

FUJITSU Storage ETERNUS Multipath Driver 3.1 User's Guide

for Oracle Solaris

This page is intentionally left blank.

Preface

This manual describes the "ETERNUS Multipath Driver" (hereinafter referred to as "Multipath Driver") software that manages the path control and connection status of the FUJITSU Storage ETERNUS Disk storage system and ETERNUS All Flash Array (hereinafter referred to as "storage system" unless stated otherwise).

The functions, installation, management, and maintenance of the Multipath Driver are explained. This manual is intended for system administrators who are connecting the storage system to an Oracle Solaris server.

Further storage system support information is available at the following website:

<http://www.fujitsu.com/global/products/computing/storage/software/eternus-mpd/>

Third Edition
July 2015

Organization

This manual is composed of the following six chapters and two appendices:

- Chapter 1 What Is Multipath Driver?
This chapter explains the features, functions, and connection patterns of the Multipath Driver.
- Chapter 2 Installation and Environment Configuration
This chapter explains how to install the Multipath Driver and configure the system environment settings.
- Chapter 3 Multipath Control Commands
This chapter explains the commands that control multipathing.
- Chapter 4 Adding or Removing Components
This chapter explains how to add and remove storage systems, logical units, and paths between servers and storage systems.
- Chapter 5 Replacing Hardware Components
This chapter explains how to replace hardware components.
- Chapter 6 Frequently Asked Questions
This chapter contains frequently asked questions.

Appendices for the following items are also provided:

- Error Messages
- Additional Information

Related Manuals

For the related manuals for each storage system, refer to the following URL:

<http://storage-system.fujitsu.com/global/doc/diskstorage/>

Acknowledgments

- Oracle and Java are trademarks or registered trademarks of Oracle Corporation in the United States and other countries.
- Screen shot(s) reprinted with permission from Microsoft Corporation.

Naming Conventions

■ Product names

Oracle Solaris might be described as Solaris, Solaris Operating System, or Solaris OS.

■ Abbreviations used in this manual

- "HBA" represents a Host Bus Adapter, which is a Solaris server-installed Fibre Channel card that connects to a storage system.
- "CA" represents a Channel Adapter installed in a storage system.
- "OVM" represents Oracle VM Server for SPARC.

■ Symbols used in this manual

The following conventions are used in this manual.



A Caution provides important or critical information to read before using a device or executing a procedure. Pay special attention to these Caution messages. Failure to do so may cause damage to the product or loss of user data.



A Note provides additional information that relates to the topic preceding or following the Note message.

High Safety

This Product is designed, developed and manufactured as contemplated for general use, including without limitation, general office use, personal use, household use, and ordinary industrial use, but is not designed, developed and manufactured for use in situations with accompanying fatal risks or dangers that, unless extremely high safety is secured, could lead directly to death, personal injury, severe physical damage or other loss (hereinafter "High Safety Required Use"), including without limitation, nuclear reaction control in nuclear facility, aircraft flight control, air traffic control, mass transport control, medical life support system, and missile launch control in weapon systems. Do not use this Product for High Safety Required Use without securing the sufficient safety level required. If you wish to use this Product for High Safety Required Use, please consult with our sales representative before such use.

Release Information

Updates

The following table shows the updates for each Multipath Driver version:

Note

The latest information is provided in "Oracle Solaris Version" found under "Downloads" at the following URL:
<http://www.fujitsu.com/global/products/computing/storage/software/eternus-mpd/>

(1/2)

Version	Modifications
3.0.0 (June 2008)	The first version.
3.0.1 (September 2009)	<ul style="list-style-type: none">• The following storage systems are supported:<ul style="list-style-type: none">- ETERNUS DX60- ETERNUS DX80- ETERNUS DX90• The following HBA driver is supported:<ul style="list-style-type: none">- LPFC 6.30g
3.0.2 (November 2009)	<ul style="list-style-type: none">• The following storage systems are supported:<ul style="list-style-type: none">- ETERNUS DX400 series- ETERNUS DX8000 series
3.0.3 (June 2011)	<ul style="list-style-type: none">• The following storage systems are supported:<ul style="list-style-type: none">- ETERNUS DX80 S2- ETERNUS DX90 S2- ETERNUS DX400 S2 series
3.0.4 (August 2011)	<ul style="list-style-type: none">• Oracle Solaris 10 8/11 is supported with the 144500-19 Kernel Patch.• The following storage systems are supported:<ul style="list-style-type: none">- ETERNUS DX60 S2- ETERNUS DX8000 S2 series

Version	Modifications
3.1.0 (April 2014)	<ul style="list-style-type: none"> • Oracle Solaris 11 is supported. • The following HBAs are supported: <ul style="list-style-type: none"> - SP1X5FAR2F - SP1X5FAS2F - SP1X5FBR2F - SP1X5FBS2F - SP1X5FAA2F - SP1X5FBA2F - SP1X7FAA2F - SP1X7FBA2F - SP1X7FAB2F - SP1X7FA2F - SP1X7FAR2F - SP1X7FAS2F - SP1X7FBR2F - SP1X7FBS2F - SE0X7F31F - SE0X7F32F - SE0X7F21F - SE0X7F22F • TPG Referrals function is supported. Refer to "ETERNUS Web GUI User's Guide (Settings) FUJITSU Storage ETERNUS DX60 S3/DX100 S3/DX200 S3, ETERNUS DX500 S3/DX600 S3 Disk storage systems, ETERNUS DX200F All Flash Array" for details. • The following storage systems are supported: <ul style="list-style-type: none"> - ETERNUS DX100 S3 - ETERNUS DX200 S3 - ETERNUS DX500 S3 - ETERNUS DX600 S3 - ETERNUS DX200F
3.1.1 (October 2014)	<ul style="list-style-type: none"> • Oracle VM guest domain is supported. • The Storage Cluster function is supported. • The following HBAs are supported: <ul style="list-style-type: none"> - SE0X7F11F - SE0X7F12F
3.1.2 (July 2015)	<ul style="list-style-type: none"> • The following storage systems are supported: <ul style="list-style-type: none"> - ETERNUS DX8700 S3 - ETERNUS DX8900 S3

Modification History

This section provides a history of modifications for the Multipath Driver.

■ 3.1.0 modifications

● Modification outline

HBAs supported by the Multipath Driver 3.0.4 or earlier cannot be used.

- Modification target
Multipath Driver

- Modification details

The following HBAs are not supported:

- GP7B8FC1
- PW008FC2
- PW008FC3
- SE0X7F11F (*1)
- SE0X7F12F (*1)

*1: Supported in the Multipath Driver 3.1.1. For details, refer to ["B.10 SE0X7F11F, SE0X7F12F" \(page 116\)](#).

- Modification effect

- HBAs for the Multipath Driver 3.0.4 or earlier cannot be used simultaneously with HBAs for the Multipath Driver 3.1.0 or later.
- Multipath Driver 3.0.4 or earlier cannot be updated to Multipath Driver 3.1.0.
- Multipath Driver 3.0.4 or earlier cannot be used simultaneously with Multipath Driver 3.1.0.

● Modification outline

Error messages have been changed.

- Modification target
Multipath Driver

- Modification details

For details, refer to ["Chapter A Error Messages" \(page 86\)](#).

- Modification effect

During an error, the message format and output may vary.

● Modification outline

The "grmpdautoconf" command functions for the Multipath Driver configuration have been changed.

- Modification target
Multipath Driver
- Modification details

Version	Modification details
3.0.4 or earlier	With the "grmpdautoconf" command, configuration of the HBA can be performed as well as configuration of the multipath.
3.1.0	The "grmpdautoconf" command can only configure the multipath. It is necessary to configure the HBA or the ssd physical device path beforehand.

- Modification effect
The procedure during a new configuration and when expanding is changed.

● Modification outline

The "mpdinfo" command is not supported.

- Modification target
Multipath Driver
- Modification details
The command is not included.
- Modification effect
The information for the devices that are displayed with "sar" or "iostat", and the information for mplb devices can be checked by using the "iompadm info" command.

■ 3.1.1 modifications

None

■ 3.1.2 modifications

● Modification outline

If the path is in an offline state during the server startup, the following error messages that are related to the path control may be output.

```
WARNING: Storage_System_Name ( e.g.:ETERNUS_DX400- 10017C) (mplb*):
        Device configuration is without redundancy.
NOTICE: mplb*: There is no redundancy of ETERNUS concurrent firmware loading.
        All paths connect to (%s).
```

For details on the error messages, refer to ["Appendix A Error Messages" \(page 86\)](#). If concurrent firmware loading (CFL) has redundancy, the action described in ["Appendix A Error Messages" \(page 86\)](#) is not required. Return the path to "online" before performing a CFL.

- Modification outline

The boot-archive is automatically updated by executing the "grmpdautoconf" command.

- Modification target
Multipath Driver

- Modification details

Version	Modification details
3.1.1 or earlier	The boot-archive is not updated.
3.1.2	The boot-archive is updated.

- Modification effect

The boot-archive is automatically updated. If automatic updates should not be performed, execute the "grmpdautoconf" command with the "-n" option.

Table of Contents

Chapter 1	What Is Multipath Driver?	15
1.1	Features	15
1.2	Connections	15
1.2.1	Connections in Normal Environments	15
1.2.2	Connections in OVM Guest Domain Environments	17
1.3	Multipath Driver Functions	20
1.4	Path Management	21
1.4.1	Path Status	21
1.4.2	Path Status Transitions	23
1.4.3	Checks and Path Status Transitions	24
1.5	Load Balancing/Failover Operation	26
1.5.1	Path Control in Normal Environments	26
1.5.2	Path Control in OVM Guest Domain Environments	32
Chapter 2	Installation and Environment Configuration	39
2.1	Operating Conditions and Installation Procedure	39
2.2	Environment Configuration	39
2.2.1	Environment Configuration Overview	40
2.2.2	Multipath Driver Environment Configuration	42
2.3	Disk Definition	44
2.3.1	Changing the Partition Configuration	44
2.3.2	Creating a File System	44
2.3.3	Registering or Changing the "/etc/vfstab" Setting	44
2.4	Assigning Virtual Disks	45
Chapter 3	Multipath Control Commands	46
3.1	Commands that Display Multipath Status	47
3.1.1	"status" Command	47
3.1.2	"info" Command	49
3.2	Commands for Replacing Hardware that Constitutes the Multipath	51
3.2.1	"change" Command	51
3.2.2	"restart" Command	52

3.3	Commands for Creating or Changing the Multipath Configuration	53
3.3.1	"grmpdautoconf" Command	53

Chapter 4 Adding or Removing Components 55

4.1	Adding a Logical Unit (LU)	55
4.2	Adding a Path or Storage System	56
4.3	Removing a Logical Unit (LU)	56
4.3.1	Stopping Applications	57
4.3.2	Removing the LU on the Storage System Side	57
4.3.3	Executing the "grmpdautoconf" Command	57
4.3.4	OVM Guest Domain Environments	58
4.4	Removing a Path	59
4.4.1	Removing the Path between the Server and the Storage System	59
4.4.2	Executing the "grmpdautoconf" Command	59
4.5	Removing a Storage System	61
4.5.1	Stopping Applications	61
4.5.2	Disconnecting the Storage System from the Server	61
4.5.3	Executing the "grmpdautoconf" Command	62
4.5.4	OVM Guest Domain Environments	63

Chapter 5 Replacing Hardware Components 64

5.1	Replacing HBAs or Server Connection Cables	64
5.1.1	How to Specify the Replacement Targets and the Set Aside Targets	64
5.1.2	Hot Swapping an HBA	66
5.1.3	Cold Swapping an HBA	66
5.1.4	Hot Swapping a Server Connection Cable	67
5.2	Replacing a Cable Connecting the Storage System and the Switch	68
5.2.1	How to Specify the Replacement Targets and the Set Aside Targets	68
5.2.2	Hot Swapping a Cable Connecting the Storage System and a Switch	70
5.3	Replacing a CA or IOB	71
5.3.1	How to Specify the Replacement Targets and the Set Aside Targets	71
5.3.2	Hot Swapping a CA or IOB	72
5.4	Replacing a CM or ROUTER	73
5.4.1	How to Specify the Replacement Targets and the Set Aside Targets	73
5.4.2	Hot Swapping a CM or ROUTER	75
5.5	Replacing a Storage System	75
5.6	Replacing a Server System Board	78
5.6.1	Hot Swapping a System Board	78
5.6.2	Cold Swapping a System Board	79

5.7	Replacing a Fibre Channel Switch or SFP	80
5.7.1	Hot Swapping a Fibre Channel Switch or SFP	80
5.7.2	Cold Swapping a Fibre Channel Switch or SFP	81
5.8	Replacing a Storage System Disk	81
Chapter 6	Frequently Asked Questions	82
<hr/>		
Appendix A	Error Messages	86
<hr/>		
A.1	Notification, Degradation, and Warning Messages	86
A.1.1	Transmission Related Messages	86
A.1.2	SCSI Sense Related Messages	87
A.2	Path Control Error Messages	89
A.3	"iompadm" Command Error Messages	92
A.4	Cluster Error Messages	96
A.5	DR Script Error Messages	99
A.6	"grmpdautoconf" Command Error Messages	100
Appendix B	Additional Information	101
<hr/>		
B.1	How to Change mplb Instance Numbers	101
B.2	Storage System Redundancy during CFL	103
B.2.1	Storage System Redundancy during a CFL for the ETERNUS6000	103
B.2.2	Storage System Redundancy during a CFL for Storage Systems other than the ETERNUS6000	105
B.3	Changing a LUN Mapping or Host Affinity	105
B.4	Troubleshooting	107
B.5	Cluster System	108
B.6	Adding Local Resources to a Cluster System	109
B.7	Performing "add device" for mplb Special Files in Non-Global Zones	111
B.8	PCI Hot Plug	112
B.9	OVM Guest Domains	113
B.10	SE0X7F11F, SE0X7F12F	116
B.11	Storage Cluster	116
Glossary		117
<hr/>		

List of Figures

Figure 1.1	Fibre Channel connection example (HBAs are connected to CAs in different CMs without Fibre Channel switches).....	16
Figure 1.2	Fibre Channel connection example (HBAs are connected to CAs in different CMs with Fibre Channel switches).....	16
Figure 1.3	Connection example when I/O domains have a multipath configuration (each path of an I/O domain is connected to a CA on a different CM).....	18
Figure 1.4	Connection example when I/O domains have a single path configuration (each I/O domain is connected to a CA on a different CM).....	19
Figure 1.5	Path status transitions	23
Figure 1.6	2-path load balancing on a storage system with CM assigned	27
Figure 1.7	4-path or more load balancing on a storage system with CM assigned	28
Figure 1.8	4-path or more load balancing on a storage system with CM not assigned.....	29
Figure 1.9	2-path failover on a storage system with CM assigned.....	30
Figure 1.10	4-path or more failover on a storage system with CM assigned, example 1	31
Figure 1.11	4-path or more failover on a storage system with CM assigned, example 2	31
Figure 1.12	Failover on a storage system with CM not assigned	32
Figure 1.13	Load balancing in OVM guest domain environments (on a storage system with CM assigned when I/O domains have a multipath configuration).....	33
Figure 1.14	Load balancing in OVM guest domain environments (on a storage system with CM assigned when I/O domains have a single path configuration).....	34
Figure 1.15	Load balancing in OVM guest domain environments (on a storage system with CM not assigned when I/O domains have a multipath configuration)	35
Figure 1.16	Load balancing in OVM guest domain environments (on a storage system with CM not assigned when I/O domains have a single path configuration)	36
Figure 1.17	Failover (when I/O domains have a multipath configuration)	37
Figure 1.18	Failover (when I/O domains have a single path configuration)	38
Figure 5.1	Failed part (HBA or server connection cable)	64
Figure 5.2	Failed part (a cable connecting the storage system and a switch)	68
Figure 5.3	Failed part (CA) and the replacement area	71
Figure 5.4	Failed part (IOB) and the replacement area.....	71
Figure 5.5	Failed part (CM) and the replacement area.....	73
Figure 5.6	Failed part (ROUTER) and the replacement area	73

Chapter 1

What Is Multipath Driver?

This chapter explains the features and operating environment of the Multipath Driver.

1.1 Features

The ETERNUS Disk storage systems and ETERNUS All Flash Array are storage systems designed to prevent the loss of data and down time due to a drive failure, thus creating a high-speed, high-capacity system with high reliability.

Installing a storage system can prevent data loss and business down time if a disk error occurs.

In addition, installing the Multipath Driver prevents business down time if a path error occurs as well as if a disk error occurs. Therefore, users can use the system safely.

If an error occurs in a path in hardware configurations where paths (HBAs, cables, CAs, etc.) are multiplexed, the Multipath Driver enables continuous operation by switching to another path.

The Multipath Driver consists of a driver to connect a storage system and an Oracle Solaris server, path status diagnostic functions, and the commands that operate the Multipath Driver.

By installing the Multipath Driver in a guest domain of OVM, a multipath can be configured with virtual disks that are assigned by multiple I/O domains.

1.2 Connections

1.2.1 Connections in Normal Environments

The Multipath Driver is used in a configuration where multiple server HBAs and multiple storage system CAs are installed and connected with multiple cables.

The Multipath Driver can also be used with switches.

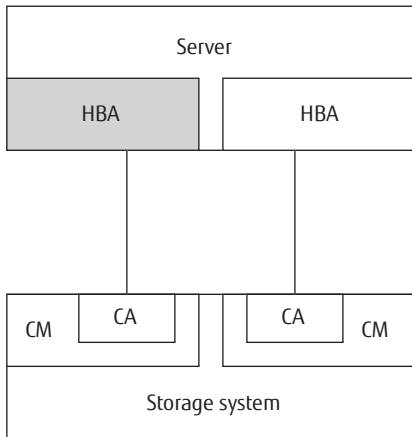
 **Caution**

In a multipath connection, all HBAs must be of the same type.

■ Fibre Channel connection example (without Fibre Channel switches)

HBA's are connected to CAs in different CMs.

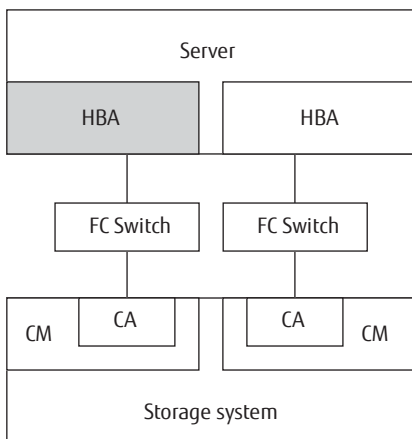
Figure 1.1 Fibre Channel connection example (HBA's are connected to CAs in different CMs without Fibre Channel switches)



■ Fibre Channel connection example (with Fibre Channel switches)

HBA's are connected to CAs in different CMs through different Fibre Channel switches.

Figure 1.2 Fibre Channel connection example (HBA's are connected to CAs in different CMs with Fibre Channel switches)



▶ **Caution**

When all of the paths to the storage system only connect to the CAs of a single CM, redundancy is not maintained.

Further storage system support information is available at the following website.

<http://www.fujitsu.com/global/products/computing/storage/software/eternus-mpd/>

1.2.2 Connections in OVM Guest Domain Environments

The Multipath Driver of a guest domain configures the multipath with virtual disks that are assigned by multiple I/O domains.

Install multiple HBAs of the I/O domain and multiple CAs of the storage system and connect the HBAs and the CAs using multiple cables. Then, configure the multipath for each I/O domain (hereinafter referred to as "multipath configuration") and assign the multipath to a guest domain as a virtual disk.

In addition, when using the Multipath Driver with guest domains only, a configuration with one path between an I/O domain and a storage system (hereinafter referred to as "single path configuration") is supported.

Both configurations can be used with switches.

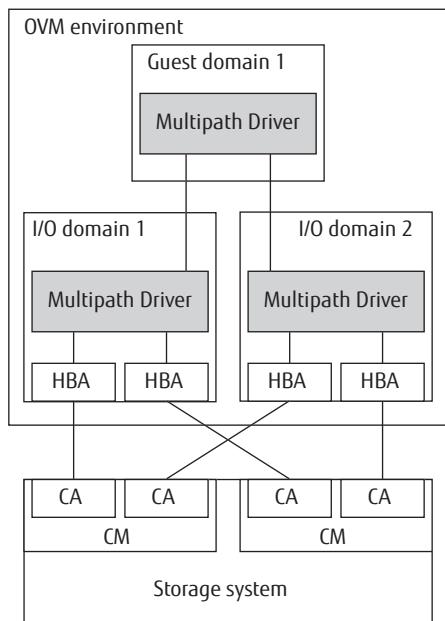
Caution

- Use the same type of HBAs.
 - The multipath configuration of the I/O domain uses the same connection type as that of a normal environment.
 - For a single path configuration of the I/O domain, each I/O domain must be connected to a CA on a different CM. If all I/O domains are connected to CAs on the same CM, redundancy is not maintained.
 - When a multipath device is used in an I/O domain, single path configurations are not supported.
 - A mix of single path configurations and multipath configurations in a single I/O domain is not supported.
 - For guest domains, a configuration where virtual disks of I/O domains in a single path configuration and virtual disks of I/O domains in a multipath configuration exist together is not supported.
-

■ Connection example when I/O domains have multipath configurations

Each path of an I/O domain connects to a CA on a different CM. This configuration can be used for connections using switches.

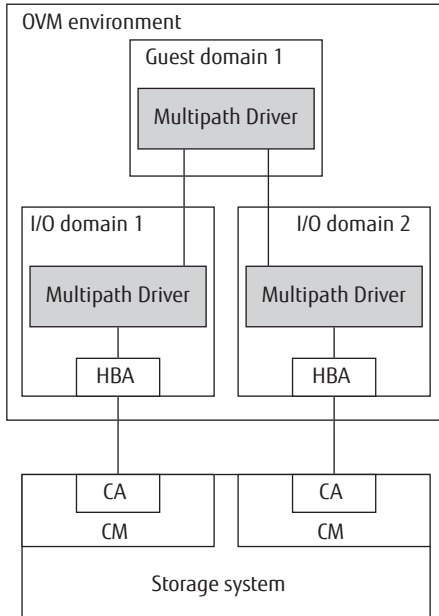
Figure 1.3 Connection example when I/O domains have a multipath configuration (each path of an I/O domain is connected to a CA on a different CM)



■ Connection example of when I/O domains have a single path configuration

Each I/O domain connects to a CA on a different CM. This configuration can be used for connections using switches.

Figure 1.4 Connection example when I/O domains have a single path configuration (each I/O domain is connected to a CA on a different CM)



Further storage system support information is available at the following website.

<http://www.fujitsu.com/global/products/computing/storage/software/eternus-mpd/>

1.3 Multipath Driver Functions

The Multipath Driver functions are described below.

- Automatic multipath configuration

After the Multipath Driver is installed in an Oracle Solaris server, when the "grmpdautoconf" command is executed while the Oracle Solaris server is connected to a storage system, the Multipath Driver automatically checks the path connection status and creates a multipath configuration.

- Load balancing

Load balancing (load distribution) is performed on all available paths, improving data accessibility. Load balancing operates differently for each storage system model. For details, refer to ["1.5 Load Balancing/Failover Operation" \(page 26\)](#).

- Failover

If a currently used path fails, the Multipath Driver automatically switches over to another normal path to access the storage system.

This allows the host application to continue operation.

Failover operates differently for each storage system model. For details, refer to ["1.5 Load Balancing/Failover Operation" \(page 26\)](#).

- Auto path check

Regular checks are performed on all paths, even when they are in a standby state, or not accessed by an application. If an unusable path is detected, this function disconnects the path and outputs a status message to the console. Reliability is increased by detecting an unusable path before a request is issued by the application.

- Auto path recovery

Regular checks are also performed on disconnected paths to reconnect those paths that have recovered from their failure.

By automatically recovering a path that is intermittently experiencing trouble, this function reinforces path redundancy and increases reliability.

There is also a function to exclude unstable paths that repeatedly cycle between a normal state and a failed state.

- Path disconnection

This function can forcibly change the status of the specified path to an offline (unusable) state.

I/O requests are not issued to the target path for preventive maintenance.

The status of the path can be returned to an online (usable) state after the maintenance is finished.

1.4 Path Management

1.4.1 Path Status

Path status is defined using a total of six states.

The states that describe the status are divided into two groups: "online" and "offline". Whether the path is "offline" or "online" is based on whether the path is available for use by user I/O request processes.

The path status can be checked by the "iompadm" command. For details on how to use the command, refer to ["3.1 Commands that Display Multipath Status" \(page 47\)](#).

The "online" state consists of the following three categories.

Online	Description
active	<ul style="list-style-type: none"> • Normal operating state. • If there are two or more active paths, load balancing is performed between them. • Display example <pre data-bbox="515 913 1466 1025" style="border: 1px solid black; padding: 5px;">/dev/rdisk/c11t500000E0D4017C80d0s2 online active block "good status [ETERNUS_DX400- 10017C-CM10-CA40-PORT00] (ssd83) "</pre>
standby	<ul style="list-style-type: none"> • Normal state. The path is in a standby state because it is connected to a non-assigned CM. • If the assigned CM (active) path fails, the standby path becomes the active path. When the assigned CM path returns to active state, this path returns to standby. • Display example <pre data-bbox="515 1196 1466 1308" style="border: 1px solid black; padding: 5px;">/dev/rdisk/c12t500000E0D4017C90d0s2 online standby block "good status [ETERNUS_DX400- 10017C-CM11-CA41-PORT02] (ssd50) "</pre>
warning	<ul style="list-style-type: none"> • I/O errors have been detected in the path, and whether or not the path can be used is currently being checked. • Warning state paths are checked every minute. • Each type of error is allowed a certain number of checks. If a path is determined to be usable before this limit is reached, its status is reverted to active or standby, otherwise its status is changed to fail. • Warning paths are not used for access if an active or standby path is available. • If no active or standby path is available, a warning path is used for access. • Display example <pre data-bbox="515 1653 1466 1765" style="border: 1px solid black; padding: 5px;">/dev/rdisk/c11t500000E0D4017C80d0s2 online active block "warning status [ETERNUS_DX400- 10017C-CM10-CA40-PORT00] (ssd83)</pre>

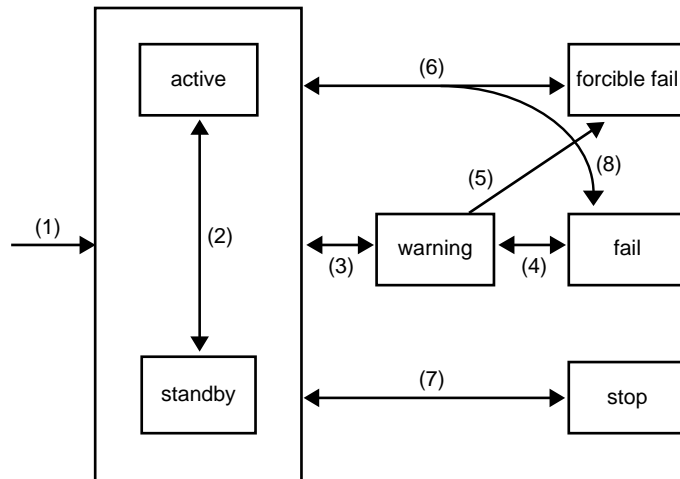
The "offline" state consists of the following three categories.

Offline	Description
fail	<ul style="list-style-type: none"> • I/O errors have been detected in the path, and the path has been judged to be unusable. • Fail paths are checked every 10 minutes for replacement of the failed parts. • If the check detects no problem, the status reverts to warning. • Use the multipath control command to instantly return the path to an active/standby (online) state after replacement of the failed parts. • Display example <pre data-bbox="515 573 1466 685" style="border: 1px solid black; padding: 5px;"> /dev/rdisk/c11t500000E0D4017C80d0s2 offline fail block "fail status [ETERNUS_DX400- 10017C-CM10-CA40-PORT00] (ssd83)</pre>
forcible fail	<ul style="list-style-type: none"> • The check result is constantly switching back and forth between normal and abnormal, so the appropriate checks cannot be performed. • Forcible fail paths are not checked. • Use the multipath control command or reboot the server to return the path to an active/standby state after replacement of the failed parts. • Display example <pre data-bbox="515 925 1466 1037" style="border: 1px solid black; padding: 5px;"> /dev/rdisk/c11t500000E0D4017C80d0s2 offline fail block "forcible fail status [ETERNUS_DX400- 10017C-CM10-CA40-PORT00] (ssd83)</pre>
stop	<ul style="list-style-type: none"> • Path has been disconnected by the multipath control command. • Stop paths are not checked. • Only the multipath control commands can return the path to an active/standby state. • In general, the stop paths are not used for access. However, if no active or standby path is available, a stop state is returned to an active state for access (except when all paths are in a stop state). • Display example <pre data-bbox="515 1312 1466 1424" style="border: 1px solid black; padding: 5px;"> /dev/rdisk/c11t500000E0D4017C80d0s2 offline stop block "stop status [ETERNUS_DX400- 10017C-CM10-CA40-PORT00] (ssd83)</pre>

1.4.2 Path Status Transitions

Path status transitions are shown in the following figure.

Figure 1.5 Path status transitions



(1) initialization→active/standby

- A multipath is automatically configured by executing the "gmpdautoconf" command. Paths connected to assigned CMs are set to active. Paths connected to non-assigned CMs are set to standby.

(2) standby↔active

- When an I/O error is detected in the active path, or when the path status changes from active to stop, the standby path becomes the active path if no other active path is available.
- When the path connected to the assigned CM is recovered and becomes the active path again, the activated standby path connected to the non-assigned CM returns to standby.

(3) active/standby↔warning

- When an I/O error occurs in the active/standby path during a normal I/O process or during a check, the state changes to warning.
- If the path is determined to be usable from the results of the check, the path transitions to active/standby.
- If the path is reconnected using the multipath control command, the path transitions to an active/standby state.

(4) warning↔fail

- During a warning, if the path is determined to be unusable from the results of the check, the path transitions to fail.
- During a failure, if the results of the check is normal, the path transitions to warning.

(5) warning→forcible fail

- When the status frequently changes between active/standby and warning, it is determined that appropriate checks cannot be performed and the status changes to forcible fail, in which the target is not checked.

(6) forcible fail \leftrightarrow active/standby

- If the path is reconnected using the multipath control command, the path transitions to an active/standby state.
- When the status frequently changes between active/standby and warning, it is determined that appropriate checks cannot be performed and the status changes to forcible fail, in which the target is not checked.

(7) active/standby \leftrightarrow stop

- When the path is disconnected by the multipath control command, the path status changes to stop.
- If the path is reconnected using the multipath control command, the path transitions to an active/standby state.

(8) active/standby \leftrightarrow fail

- When a serious failure is detected in the ordinary I/O process or active/standby check, the path status changes to fail.
- If the path is reconnected using the multipath control command, the path transitions to an active/standby state.

When the path status changes, the path connected to the assigned CM becomes active, and the path connected to the non-assigned CM becomes standby. However, if there are no active paths connected to the assigned CM, the path connected to the non-assigned CM will become active.

1.4.3 Checks and Path Status Transitions

For checks, the available three types are shown below.

Target path status	Check interval	Post check status	Transition trigger
active/standby	10 minutes	active/standby	Checked normal.
		warning	Checked failure. Error counter == 50 or 100
		fail	Checked failure. Error counter == 200
warning	1 minute	active/standby	Checked normal. Error counter == 0
		warning	Check result: 0 < Error counter < 200
		fail	Checked failure. Error counter == 200
fail	10 minutes	warning	Checked normal. Error counter == 100
		fail	Checked failure. Error counter == 200

Transitioning to another state is determined by the error counter in the warning state.

■ Check in active/standby

Active/Standby state paths are checked every 10 minutes.

When a failure is detected by the check, the status changes to either warning or fail. The error counter is determined by the failure cause.

The status changes in the same way when an error is detected during a normal I/O process. When the status transitions to fail, a message indicating that the status has changed appears. For details of this message, refer to ["A.2 Path Control Error Messages" \(page 89\)](#).

■ Check in warning

Warning state paths are checked every minute.

When a failure is detected by the check, 10 is added to the error counter. When it is normal, 10 is deducted. When the error counter after an update becomes "0", the status becomes active/standby. If it reaches 200, the status becomes fail. When the status transitions to fail, a message indicating that the status has changed appears. For details of this message, refer to ["A.2 Path Control Error Messages" \(page 89\)](#). Note that no message appears when the status transitions to active/standby.

■ Check in fail

Fail state paths are checked every 10 minutes.

When a failure is detected by the check, the status does not change. If the path is normal, the error counter is set to 100, the path status changes to warning, and a message indicating that the path has recovered appears. For details of this message, refer to ["A.2 Path Control Error Messages" \(page 89\)](#).

● Example 1

Operations when an error is detected during normal I/O process of an active path, and the error continues to occur.

Time passed	Path 1 status	Path 2 status	Operations
–	active	standby	–
0	warning (50)	active	Error in normal I/O process on Path 1. Path 1 changes to warning state, and the error counter is set by the cause of the error. (A value of 50 in this example.) Retry after setting Path 2 active for the I/O process.
1 minutes	warning (60)	active	Check detects an error. 10 added.
:	:	:	:
10 minutes	warning (150)	active	Check detects an error. 10 added.
:	:	:	:
14 minutes	warning (190)	active	Check detects an error. 10 added.
15 minutes	fail	active	Check detects an error. 10 added. Error counter reaches 200 and the status changes to fail. Path degradation message is sent to the console.

● Example 2

Auto path recovery operation after a failed path is replaced.

Time passed	Path 1 status	Path 2 status	Operations
–	standby	active	Maintenance replacement procedure completes.
0 to 10 minutes	warning (100)	active	A check at intervals of 10 minutes is performed within 10 minutes after a maintenance completion. Post replacement check gives a normal result, so status changes to warning and the error counter is set to 100.
+1 minute	warning (90)	active	Check finds normal. 10 deducted.
:	:	:	:
+5 minutes	warning (50)	active	Check finds normal. 10 deducted.
:	:	:	:

Time passed	Path 1 status	Path 2 status	Operations
+9 minutes	warning (10)	active	Check finds normal. 10 deducted.
+10 minutes	active	standby	Check finds normal. 10 deducted. Error counter reaches 0 and the status changes to active. Path 2 reverts to standby.

■ forcible fail transition

In the following cases, it is judged that accurate checks cannot be performed, so the path is changed to the forcible fail state and no longer checked. The "forcible path degrade" message is displayed on the console.

- The path stays in a warning state for one hour.
- In the hour since the last transition from active/standby to warning state, the same transition occurred 10 times.

1.5 Load Balancing/Failover Operation

1.5.1 Path Control in Normal Environments

There are two types of storage systems. The first is a storage system whose access paths are assigned to one of the two CMs (hereinafter referred to as "storage system with CM assigned"), the second is a storage system whose access paths are not assigned to a particular CM (hereinafter referred to as "storage system with CM not assigned").

In the storage system with CM assigned, paths that belong to the assigned CM are active and other paths connected to non-assigned CM are on standby.

In the storage system with CM not assigned, all paths are active, and used to access data.

Load balancing and failover operations vary according to the storage system type (whether the storage system has assigned CM or not) and number of paths.

Further information about Multipath Driver support for storage systems with CM assigned and storage systems with CM not assigned is available at the following website:

<http://www.fujitsu.com/global/products/computing/storage/software/eternus-mpd/>

■ Load balancing

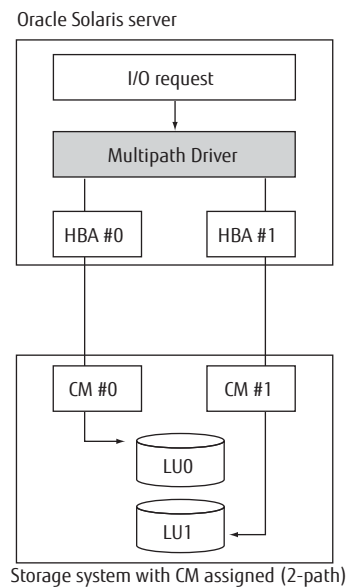
The Multipath Driver uses all active paths to perform load balancing (load distribution) to improve accessibility.

● Storage systems with CM assigned (2-path connections)

For a storage system with CM assigned (2-path), one path is connected to each controller. The total number of paths is two.

Of the two paths, the path connected to the assigned CM is active and the path connected to the non-assigned CM is on standby. If there is only one active path, load balancing is not performed.

Figure 1.6 2-path load balancing on a storage system with CM assigned



- Storage systems with CM assigned (4-path or more connections)

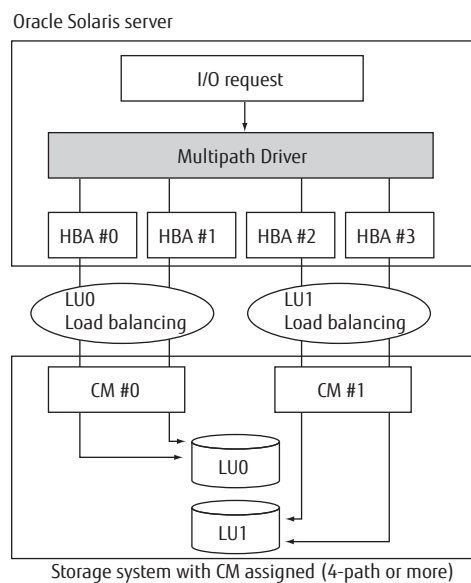
For a storage system with CM assigned that has a 4-path configuration, two paths are connected to each CM, for a total of four paths. The two paths connected to the assigned CM are active, and load balancing is performed on these two paths.

For an 8-path configuration, four paths are connected to each of the two CMs, for a total of eight paths. The four paths connected to the assigned CM are active, and load balancing is performed on these four paths.

Caution

The load balancing function is optimized for use on a storage system with CM not assigned.

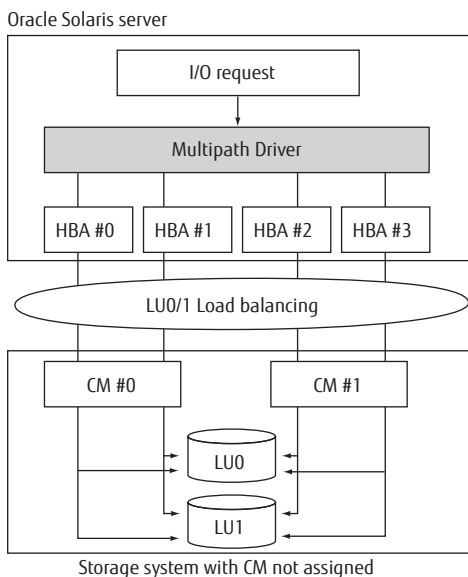
Figure 1.7 4-path or more load balancing on a storage system with CM assigned



- Storage systems with CM not assigned

For a storage system with CM not assigned, there is no concept of standby path, so all connected paths are active, and load balancing is performed over all paths. The maximum number of paths that the Multipath Driver can control is eight.

Figure 1.8 4-path or more load balancing on a storage system with CM not assigned



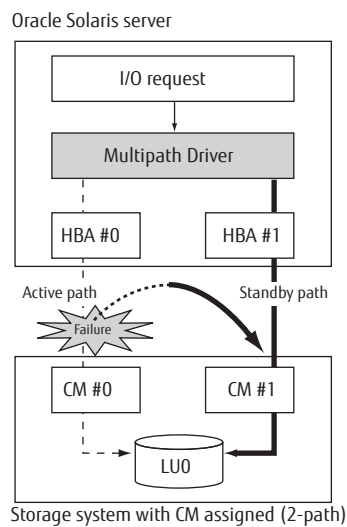
■ Failover

When the Multipath Driver detects that an active path has failed, it switches to another normal path and executes the I/O request processes. This can prevent the host application from shutting down unnecessarily.

● Storage systems with CM assigned (2-path connections)

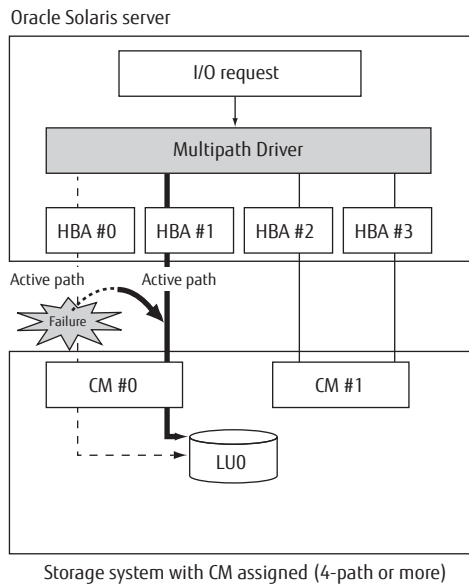
When an active path becomes unavailable, the standby path is activated to continue processing.

Figure 1.9 2-path failover on a storage system with CM assigned



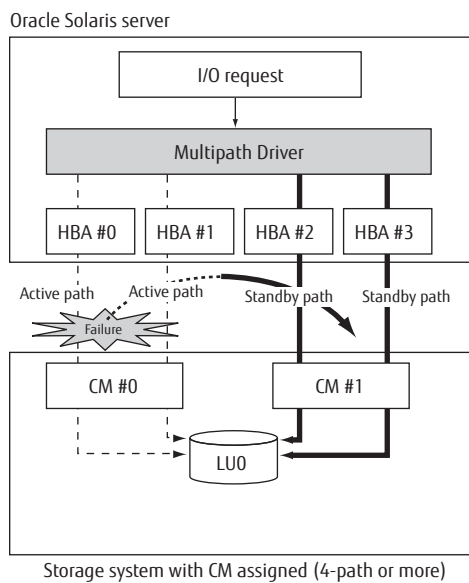
- Storage systems with CM assigned (4-path or more connections)
 - Example 1
If one of the multiple active paths becomes unavailable, processing continues on the other remaining active path.
Switching to the standby path is not performed.

Figure 1.10 4-path or more failover on a storage system with CM assigned, example 1



- Example 2
If all active paths for a CM become unavailable, a standby path is then switched to active state to continue processing.
When there are two or more standby paths, load balancing is also performed.

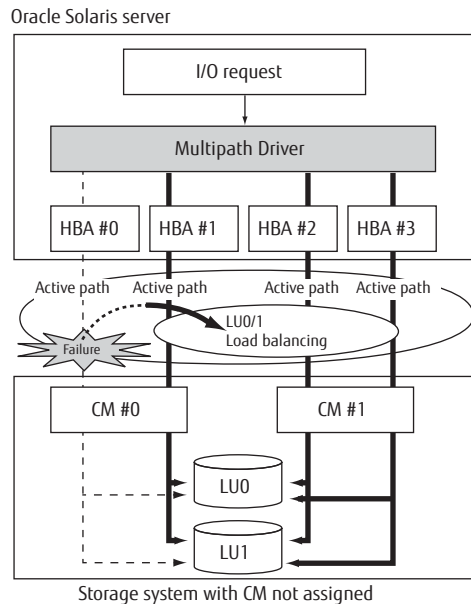
Figure 1.11 4-path or more failover on a storage system with CM assigned, example 2



- Storage systems with CM not assigned

When an active path becomes unavailable, the remaining active paths continue load balancing in a degraded state.

Figure 1.12 Failover on a storage system with CM not assigned



1.5.2 Path Control in OVM Guest Domain Environments

Operations in OVM guest domain environments are different from those in normal environments. Since the Multipath Driver operates each I/O domain and each guest domain separately, operations differ depending on the configuration.

Caution

For storage systems with CM assigned, even when all the paths on the assigned CM side are closed for the I/O domains in a multipath configuration, the paths will be used as active paths for guest domains.

- Load balancing

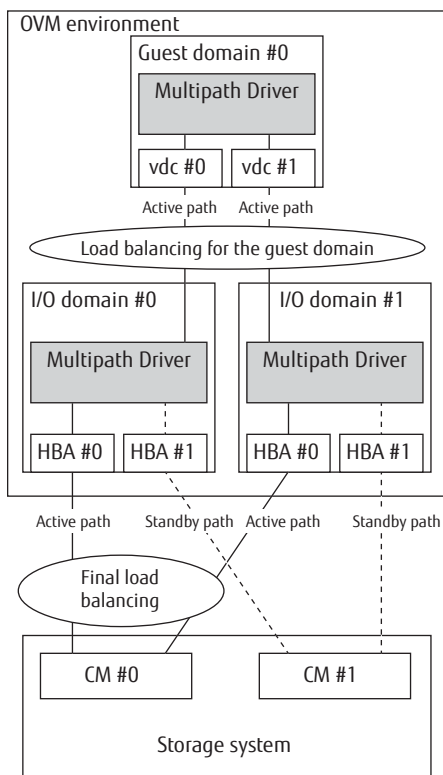
The Multipath Driver uses all active paths to perform load balancing (load distribution) to improve access performance.

- Storage systems with CM assigned (when I/O domains have a multipath configuration)

Load balancing is not performed for storage systems with CM assigned (when I/O domains have a multipath configuration) because each I/O domain only has one active path. Load balancing is performed for guest domains because the two paths are active. For I/O requests that are issued from the guest domain, a final load balancing is performed on a total of two active paths (one active path for each I/O domain).

Furthermore, when each I/O domain has two active paths in a four-path configuration, load balancing is performed for each I/O domain. In this case, a final load balancing is performed for I/O requests that are issued from the guest domain on a total of four active paths (two active paths for each I/O domain).

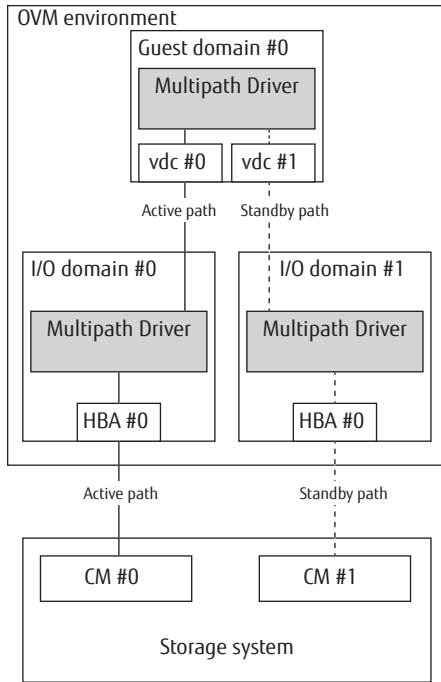
Figure 1.13 Load balancing in OVM guest domain environments (on a storage system with CM assigned when I/O domains have a multipath configuration)



- Storage systems with CM assigned (when I/O domains have a single path configuration)

For storage systems with CM assigned (when I/O domains have a single path configuration), each I/O domain connects to a different CM. Therefore, load balancing is not performed for each I/O domain because the path for I/O domain #0 is an active path and the path for I/O domain #1 is a standby path. In addition, load balancing is not performed for guest domains because only one path is active.

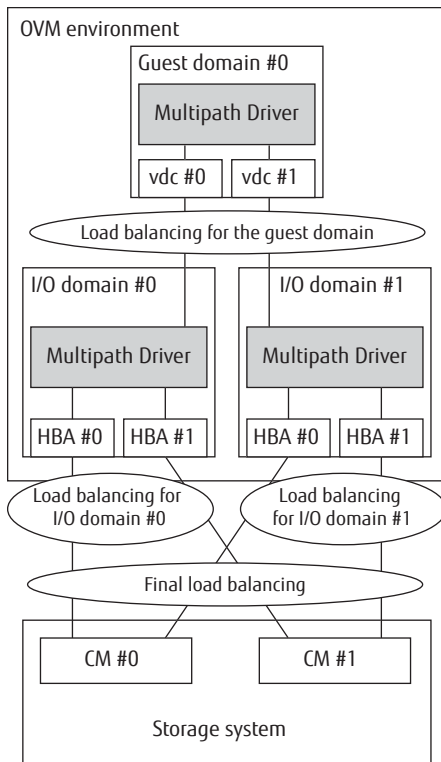
Figure 1.14 Load balancing in OVM guest domain environments (on a storage system with CM assigned when I/O domains have a single path configuration)



- Storage systems with CM not assigned (when I/O domains have a multipath configuration)

Load balancing is performed for storage systems with CM not assigned (when I/O domains have a multipath configuration) because each I/O domain and guest domain have two active paths. For I/O requests that are issued from the guest domain, a final load balancing is performed on a total of four active paths (two active paths for each I/O domain).

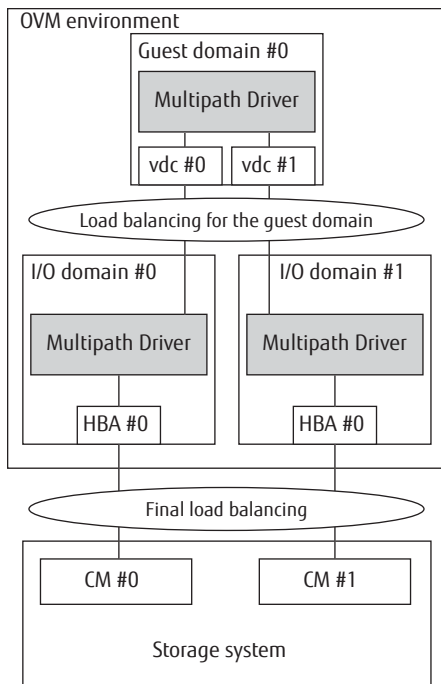
Figure 1.15 Load balancing in OVM guest domain environments (on a storage system with CM not assigned when I/O domains have a multipath configuration)



- Storage systems with CM not assigned (when I/O domains have a single path configuration)

For storage systems with CM not assigned (when I/O domains have a single path configuration), each I/O domain connects to a different CM. Therefore, load balancing is not performed for each I/O domain because there is only one active path for I/O domain #0 and I/O domain #1. Load balancing is performed for guest domains because the two paths are active. For I/O requests that are issued from the guest domain, a final load balancing is performed on a total of two active paths (one active path for each I/O domain).

Figure 1.16 Load balancing in OVM guest domain environments (on a storage system with CM not assigned when I/O domains have a single path configuration)



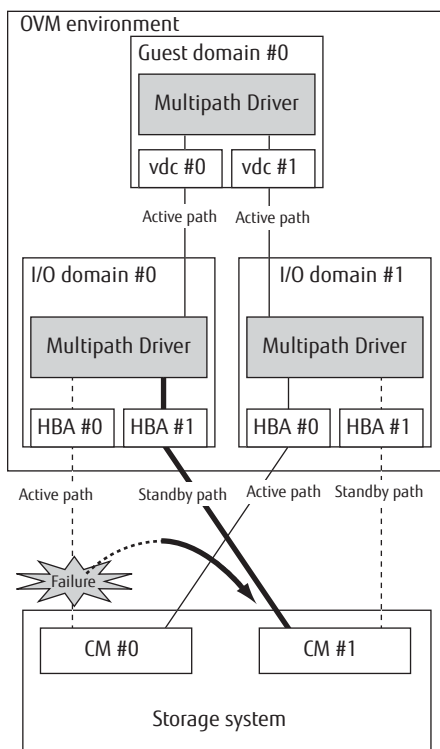
■ Failover

When the Multipath Driver detects that an active path has failed, it switches to another normal path and executes the I/O request processes. This can prevent the host application from shutting down. The location where a failover is performed differs depending on the path configuration of the I/O domains.

- When I/O domains have a multipath configuration

When an active path fails, the Multipath Driver of the I/O domain switches to the standby path to continue operations. A path is not switched by the Multipath Driver of the guest domain because the guest domain responds normally.

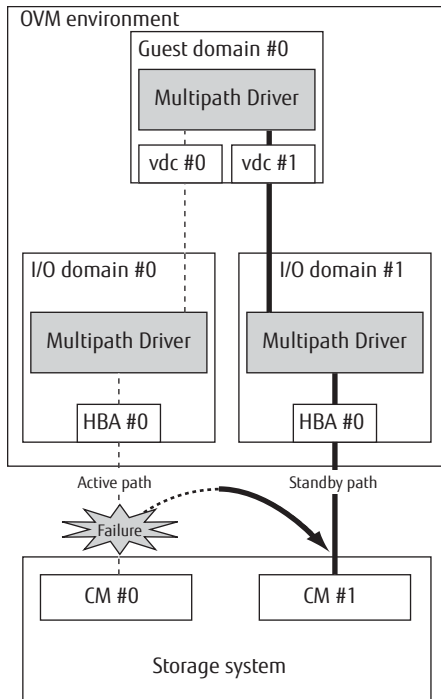
Figure 1.17 Failover (when I/O domains have a multipath configuration)



- When I/O domains have a single path configuration

When an active path fails, an I/O error message is sent to the guest domain because I/O domains have a single path configuration. After the I/O error response, the Multipath Driver of the guest domain switches to the standby path to continue operations.

Figure 1.18 Failover (when I/O domains have a single path configuration)



Chapter 2

Installation and Environment Configuration

This chapter explains how to install the Multipath Driver and configure the environment settings.

2.1 Operating Conditions and Installation Procedure

For the operating conditions and installation procedure, refer to "Software Information FUJITSU Storage ETERNUS Multipath Driver 3.1".

2.2 Environment Configuration

This section explains environment configuration for the Multipath Driver.

▶ Caution

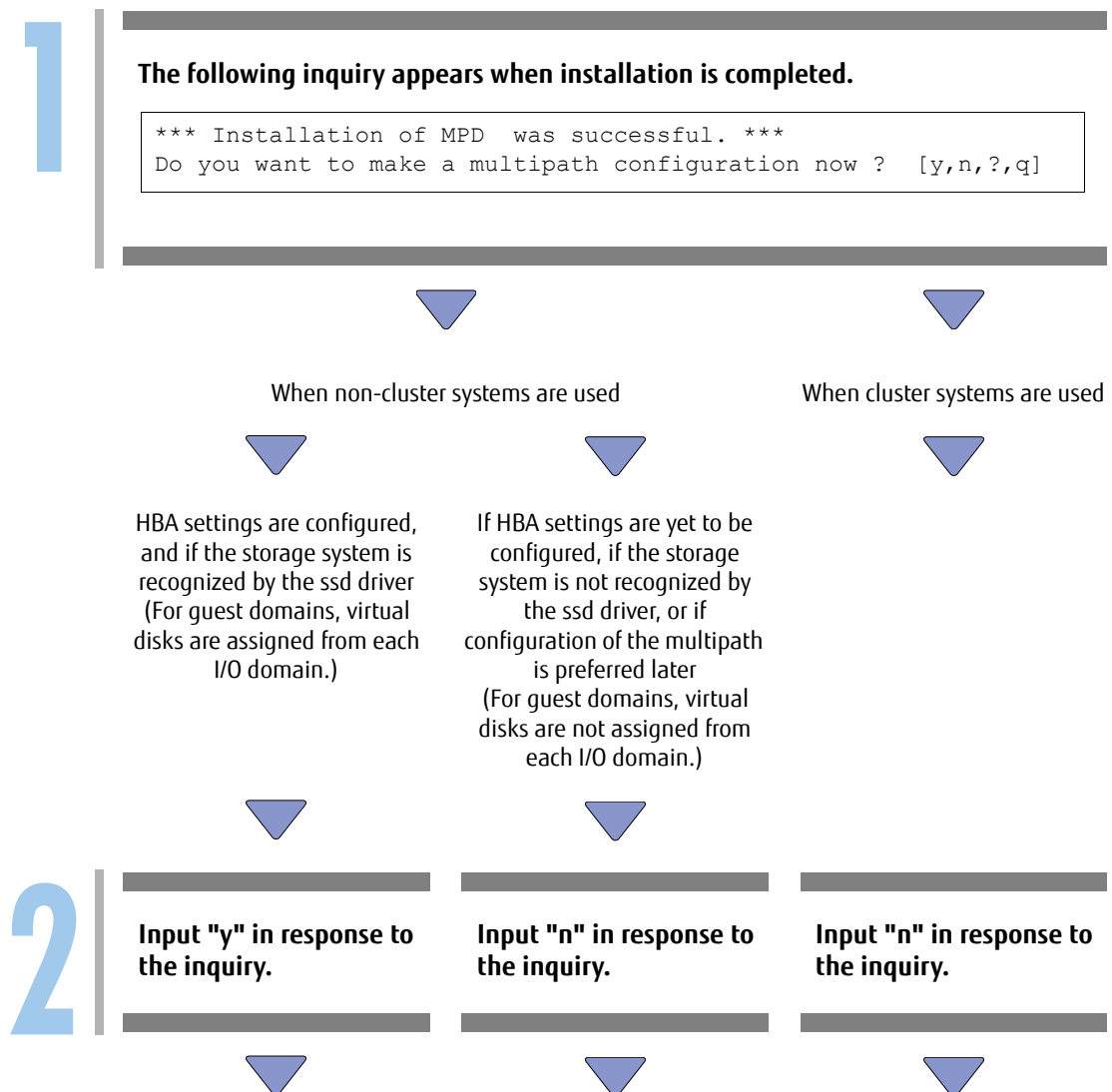
- Before beginning environment configuration, make sure that the settings of the connected storage system are correct and that the power is on.
- Before performing environment configuration, make sure that the HBA settings have been completed.
- Before beginning environment configuration, make sure the ssd driver has been recognized by the storage system.
If SAN Boot is to be used, follow the procedure given in the "SAN Boot Configuration Guide" for each server.
- When installing the Multipath Driver in a guest domain, the multipath must be configured on the I/O domains before beginning environment configuration, and virtual disks must be assigned from each I/O domain.

2.2.1 Environment Configuration Overview

2.2.1.1 Overall Flow

The "grmpdautoconf" command is used to configure the environment settings of the Multipath Driver. However, the "grmpdautoconf" command does not configure the HBA driver and the ssd driver. Refer to "FUJITSU Storage ETERNUS DX Configuration Guide -Server Connection- (Fibre Channel) for Oracle Solaris" and configure the drivers in advance. The overall workflow is as follows.

Workflow



3

Refer to "FUJITSU Storage ETERNUS DX Configuration Guide -Server Connection-(Fibre Channel) for Oracle Solaris" to configure HBA, and to ensure that the ETERNUS storage system is recognized by the `ssd` driver.

Refer to "FUJITSU Storage ETERNUS DX Configuration Guide -Server Connection-(Fibre Channel) for Oracle Solaris" to configure HBA, and to ensure that the ETERNUS storage system is recognized by the `ssd` driver.

Execute the "`grmpdautoconf`" command.

Configure the cluster multipath with the automatic resource registration.



- "["B.5 Cluster System" \(page 108\)](#)

4

Configure the Multipath Driver environment.

Create a Multipath Driver configuration file and multipath management special file.

1. Configure the Multipath Driver and create the configuration definition file.
2. Check and write the disk labels.



- "["2.2.2 Multipath Driver Environment Configuration" \(page 42\)](#)

5

Define the disks.

Configure the following setting for each logical unit as required.

1. Change the partition configuration.
2. Create a file system.
3. Register or change the "/etc/vfstab" settings.



- ["2.3 Disk Definition" \(page 44\)](#)

6

For I/O domains, assign multipath devices to guest domains.



- ["2.4 Assigning Virtual Disks" \(page 45\)](#)

2.2.2 Multipath Driver Environment Configuration

2.2.2.1 Configuring the Multipath Driver and Creating the Configuration Definition File

Configure a multipath.

● Example

```
Path      : Action : Element path                               : LUN : Storage
mplb0    : new   : c11t500000E0D4017C80d0s2 c12t500000E0D4017C90d0s2 : 0 : ETERNUS_DX400- 10017C
mplb1    : new   : c11t500000E0D4017C80d1s2 c12t500000E0D4017C90d1s2 : 1 : ETERNUS_DX400- 10017C
mplb2    : new   : c11t500000E0D4017C80d2s2 c12t500000E0D4017C90d2s2 : 2 : ETERNUS_DX400- 10017C
mplb3    : new   : c11t500000E0D4017C80d3s2 c12t500000E0D4017C90d3s2 : 3 : ETERNUS_DX400- 10017C
=== Reconfigure plan ===
Multi path: 0
Create new multi path: 4
Add path for multi path: 0 (instance) : 0 (additional path)
Delete path from multi path: 0 (instance) : 0 (deleted path)
```

 **Caution**

- When this procedure is performed, the following message may appear on the execution terminal or console, but this does not mean there is a problem.
For details of this message, refer to ["Chapter 6 Frequently Asked Questions" \(page 82\)](#).

```
[Execution terminal]
Cannot unload module: mplb
Will be unloaded upon reboot.
Forcing update of mplb.conf.
```

```
[Console]
WARNING: Duplicate instance 4095 of node "mplb" ignored.
```

- Regarding the connection configuration of the storage system with the server, if there is no redundancy when performing CFL in the storage system, the Multipath Driver configuration will end normally, but the following warning message will be output.

```
[Console]
WARNING: Storage_System_Name ( e.g.:ETERNUS_DX400- 10017C) (mplb*):
Device configuration is without redundancy.
```

In this case, in order to have an environment to perform CFL in the storage system, it is necessary to change the connection configuration of the storage system with the server.

For details on the redundancy of CFL, refer to ["B.2 Storage System Redundancy during CFL" \(page 103\)](#).

The configuration definition file for the Multipath Driver is set. Before starting, the following message appears:

- Example

```
Please wait for update configuration of multipath driver ...
!!!!CAUTION: This process requires long time !!!!!
```

2.2.2.2 Checking and Writing the Disk Label

A disk label is automatically created if there is no disk label for the LU that is now accessible. The following message appears when creating a disk label:

- Example

```
Please wait ... Checking disk label in progress.
!!!!CAUTION: This process requires long time !!!!!
```

 **Caution**

When there is no disk label information, the following message may appear on the console, but this does not mean there is a problem.

```
[Console]  
WARNING: corrupt label wrong magic number
```

2.3 Disk Definition

After configuring the settings up to ["2.2 Environment Configuration" \(page 39\)](#), the logical units of the storage system can be accessed using the multipath functions by accessing the special file `"/dev/FJSVmplb/[r]dsk/mplb*s*"` from the Solaris OS environment. Configure the following settings for each logical unit as required. These procedures are the same as with internal disks. For details, refer to the "System Administration Guide: Basic Administration" or "System Administration Guide: Devices and File Systems".

2.3.1 Changing the Partition Configuration

This procedure only needs to be done if the partition configuration of the new logical unit is to be changed from the default.

 **Caution**

Specify the path name as the argument for the "format" command as follows.

```
# format /dev/FJSVmplb/rdsk/mplb*s2
```

2.3.2 Creating a File System

If needed, create a file system in the new logical unit.

2.3.3 Registering or Changing the `"/etc/vfstab"` Setting

After a file system has been created in the new logical unit, it needs to be added to `"/etc/vfstab"` to enable it to be auto-mounted.

When the configuration is changed from single path to multipath, the special file name settings in `"/etc/vfstab"` must be changed.

2.4 Assigning Virtual Disks

When configuring a multipath on a guest domain, assign the multipath that was created on the I/O domain to the guest domain as a virtual disk.

The following example shows path specifications when exporting a virtual disk back-end through a service domain.

```
# ldm add-vdsdev /dev/FJSMplb/dsk/mplb0s2 vol10@ldom3-vds0
```

For details about the procedure and the command, refer to "Oracle VM Server for SPARC Administration Guide".

Caution

- The virtual disk that is supported by the Multipath Driver of the guest domain is the only disk that the multipath device is assigned as a full disk on the I/O domain. Virtual disks that are assigned to devices other than the Multipath Driver (for example, `ssd`, `MPxIO`, or `PRIMECLUSTER GDS` devices) and virtual disks that are assigned as slices are not supported.
- Multiple host disk control operation "mhd" is not supported.
- Notes and restrictions about virtual disks are OVM compliant. For details, refer to "Oracle VM Server for SPARC Administration Guide".

Chapter 3

Multipath Control Commands

This chapter explains the commands that control multipathing. The commands can perform the following operations.

- Displaying the multipath status
- Replacing the hardware that configured the multipath
- Changing the multipath configuration

Most of the commands are implemented as subcommands of the "iompadm" command. The specifications of the "iompadm" command are described below.

[Command format]

```
/usr/opt/FJSViomp/bin/iompadm [-c mplb] [-p] subcommand [path_name]
```

(The following explanation on the "iompadm" command uses the relative path name)

[Explanation of arguments]

- Specify one of the following commands as a subcommand.
 - "status" command
 - "info" command
 - "change" command
 - "restart" command
- -c class name: Specify "mplb" when using this product.
- -p: When the device node name is displayed, the device node name on the "/devices" side is displayed at the same time.
- In a cluster configuration, if a change unit other than "a/adapter" is specified when the "change" or "restart" subcommand is executed, communication is carried out between the cluster nodes. Refer to ["3.2.1 change" Command \(page 51\)](#) for more information. The other subcommands are not automatically executed in the other nodes.
- Path name can be any one of the following three types.
 - Instance administration device node (/dev/FJSVmplb/fiomp/adm*)
This is a device node for controlling a multipath, generated when the multipath is created.
 - User access device node (= access special file)
(/dev/FJSVmplb/[r]dsk/mplb*s*)
This is a device node for transferring multipath data, generated when the multipaths are created, and used by applications.
 - Physical device node (/dev/[r]dsk/c*t*d*s2,/devices/...)
This is a device node for a drive that is part of the multipath.

3.1 Commands that Display Multipath Status

This section explains the commands that display the multipath status.

Caution

The "status" and "info" commands display the path information that the Multipath Driver manages internally. Current path connectivity is neither checked nor displayed by these commands.

For example, when the cable of a normal state path becomes disconnected, the relevant path status shown by these commands will stay "online" for as long as the path status information managed by the Multipath Driver remains unchanged.

The path status information managed by the Multipath Driver will be changed when an I/O error or path failure is detected by the check function. Subsequent execution of these commands will then reveal the new status of the changed path information.

3.1.1 "status" Command

[Command format]

```
iompadm -c mplb [-p] status instance_administration_device_node [physical_device_node]
```

[Function overview]

This command shows the multipath status of a specified instance administration device node. Specifying a physical device node displays the status of only that physical device node.

[Usage]

This command is used to check the multipath status as need arises, such as when a mplb driver error message is displayed.

[Example]

Enter the command as follows.

```
# iompadm -c mplb status /dev/FJSVmplb/fiomp/adm0
/dev/rdsd/c11t500000E0D4017C80d0s2      online active block  "good sta-
tus                                     [ETERNUS_DX400- 10017C-CM10-CA40-PORT00] (ssd83)"

/dev/rdsd/c12t500000E0D4017C90d0s2      online standby block  "good sta-
tus                                     [ETERNUS_DX400- 10017C-CM11-CA41-PORT02] (ssd50)"

# iompadm -c mplb -p status /dev/FJSVmplb/fiomp/adm0 /dev/rdsd/
c12t500000E0D4017C90d0s2
/dev/rdsd/c12t500000E0D4017C90d0s2      online standby block  "good sta-
tus                                     [ETERNUS_DX400- 10017C-CM11-CA41-PORT02] (ssd50)"
->
/devices/pci@0,600000/pci@0/pci@8/pci@0,1/fibre-channel@1/fp@0,0/
ssd@w500000e0d4017c90,0:c,raw
```

[Output format]

```
physical_device_node, status 1, status 2, status 3, "message,  

    [storage_system_controller_identifier] (ssd_instance)"  

[-> device_path]
```

[Arguments]

Argument	Description
physical_device_node	Displays the logical unit device name (/dev/rdisk/c*t*d*s) that constitutes the multipath. The controller number of the device name (c*) cannot be changed, because it is automatically assigned by the OS. Displays the device name of the virtual disk (/dev/rdisk/c*t*s*) for guest domains.
status 1	Displays online, offline, or unconfigured. <ul style="list-style-type: none"> • online = path is active • offline = path is not active, not in use • unconfigured = path has not been configured for multipath control
status 2	Displays active, standby, stop, fail, or disconnected. <ul style="list-style-type: none"> • active = path is operating normally • standby = path is on standby • stop = path is temporarily stopped by a command • fail = path has failed for some reason • disconnected = path has been disconnected from multipath control by the Dynamic Reconfiguration function
status 3	Displays whether external access to the physical device nodes is prohibited (block) or permitted (unblock).
message	Displays the disconnected state (warning), or the state where automatic diagnosis is not performed (forcible fail). <ul style="list-style-type: none"> • warning = the online access path has been temporarily disconnected • forcible fail = a diagnosis cannot be performed
ssd_instance	Displays the instance number of ssd. Displays the instance number of the virtual disk (vdc) for guest domains.
device_path	If the -p option is used, the device path will be displayed in the "/devices" portion of the output.

The storage system controller identifier is described below.
 The identifier is the character string that identifies all items from the connected storage system to the port location of the channel adapter (controller). The output format is "storage system name" - "storage system identification number" - "group number" - "change unit number" - "adapter port number". The following section explains each item.
 Further information about the relationship between the storage system controller identifier and the physical locations of the items is available on the following website:

<http://www.fujitsu.com/global/products/computing/storage/software/eternus-mpd/>

Argument	Description
storage_system_name	Displays the storage system model name and the magic code (number for identifying the storage system).
storage_system_identification_number	Displays the unique serial number (manufacturing number) assigned to each storage system.
group_number (*1)	When replacing a group module, all the physical device nodes (or paths) that have the same storage system name, identification number, and group number are affected by the replacement. Suspend access to all paths in the group before starting the replacement. Further information about group modules for storage systems is available at the following website: http://www.fujitsu.com/global/products/computing/storage/software/eternus-mpd/
change_unit_number (*1)	When hot swapping a controller (channel adapter), all the physical device nodes (or paths) that have the same storage system name, identification number, group number, and change unit number are affected by the hot swap. Suspend access to all paths in the group before starting the replacement.
adapter_port_number (*1)	Displays the channel adapter port number of the port that is actually connected by cables on the storage system side.

*1: Not displayed for guest domains when I/O domains have a multipath configuration.

[Message list]

The following table shows the messages that are displayed and their meanings.

	Displayed message	Description
1	good status	In an online state
2	warning status	In a warning state
3	forcible fail status	In a forcible failed state
4	fail status	In a failed state
5	offline status	In a state where the host affinity and the LUN mapping are changed and a LUN that is different from the configured LUN is shown.
6	stop status	In a stopped state
7	disconnect status	In a state where the path is disconnected via Dynamic Reconfiguration or PCI Hot Plug.
8	no such device or address	In an offline state at start-up

3.1.2 "info" Command

[Command format]

```
iompadm -c mplb info [instance_administration_device_node]
```

[Function overview]

Executing this command provides information on the specified instance administration device nodes. If the instance administration device node is omitted, information will be displayed for all instances.

[Usage]

This command is used to show information regarding the entire instance administration device node.

[Example]

```
# iompadm info /dev/FJSVmplb/fiomp/adm0
IOMP: /dev/FJSVmplb/fiomp/adm0
Element:
    /dev/rdisk/c11t500000E0D4017C80d0s2      online active block  "good
status      [ETERNUS_DX400- 10017C-CM10-CA40-PORT00] (ssd83) "
    /dev/rdisk/c12t500000E0D4017C90d0s2      online standby block  "good
status      [ETERNUS_DX400- 10017C-CM11-CA41-PORT02] (ssd50) "
Node:
    /dev/FJSVmplb/rdisk/mplb0s0
    /dev/FJSVmplb/rdisk/mplb0s1
    /dev/FJSVmplb/rdisk/mplb0s2
    /dev/FJSVmplb/rdisk/mplb0s3
    /dev/FJSVmplb/rdisk/mplb0s4
    /dev/FJSVmplb/rdisk/mplb0s5
    /dev/FJSVmplb/rdisk/mplb0s6
    /dev/FJSVmplb/rdisk/mplb0s7
Function:
    MPmode=true
    AutoPath=true
    Block=true
    NeedSync=false
```

[Arguments]

Argument	Description
IOMP:	Displays the instance administration device node name
Element: Path 0 status Path 1 status : Path N status	Displays each physical device node constituting the multipath that is controlled by the above instance administration device node, and the status of the physical device node. For details of each path display, refer to the section on the "status" command.
Node: User access device node name 1 User access device node name 2 : User access device node name N	Displays all of the access path names of the paths that access the controlling multipath for the above instance administration device node.
Function: MPmode=true AutoPath=true Block=true NeedSync=false	This argument is fixed for the Multipath Driver.

3.2 Commands for Replacing Hardware that Constitutes the Multipath

3.2.1 "change" Command

[Command format]

```
iompadm -c mplb change [change_unit] target_physical_device_node
```

[Function overview]

This command is used in preparation for hardware hot swapping.

This command disconnects all of the paths that use the hardware to be replaced, which is determined by the change unit and the target physical device node, and stops all related I/O processes.

For the target physical device node, specify the physical path name of the element displayed by the "iompadm info" command or the path that is displayed in the message.

[Usage]

The procedure for hot swapping hardware is described below.

- (1) Execute the "change" command to stop all I/O processes on the paths used by the hardware that is to be replaced.
- (2) Hot swap the hardware.
- (3) Execute the "restart" command to restart all I/O processes on the paths used by the newly replaced hardware.

[Example]

```
# iompadm -c mplb change c /dev/rdisk/c11t500000E0D4017C80d0s2
```

[Change unit specification]

Argument	Description
a/adapter	Use this unit when replacing a cable between an HBA and a switch or SFP connected to the cable.
c/controller (default) (*1)	Use this unit when replacing a cable directly connecting an HBA and CA, or a cable between a CA and a switch, or a SFP connected to the cable. Necessary processes are performed on other nodes in a cluster configuration.
cu/controllerunit (*1)	Use this unit when replacing a channel adapter. Necessary processes are performed on other nodes in a cluster configuration.
g/groupmodule (*1)	Use this unit when replacing a group module. Necessary processes are performed on other nodes in a cluster configuration.

*1: Not supported for guest domains when I/O domains have a multipath configuration.

When the replacement target hardware is used by multiple servers in a cluster configuration and the change unit is other than "a" (adapter), communication is carried out between the cluster nodes. The status of the paths between each node and the target hardware automatically transitions.

- For cluster-sharing disks, this command only needs to be executed on one node of the cluster. Execution on all cluster nodes is not required.
- For non-sharing disks and guest domains in a cluster configuration, this command needs to be executed on all nodes.
- For a single (non-cluster) configuration, this command needs to be executed on all connected servers.

Further information about the relationship between change units and each storage system is available on the following website:

<http://www.fujitsu.com/global/products/computing/storage/software/eternus-mpd/>

[Caution]

- If a change unit is not specified, the process is performed as if "c" had been specified.
- When the path status transitions to "stop" by execution of this command, the stop state is canceled after the server is restarted. Similar to a normal restart, the path status transitions to online active/standby if there are no path errors, or offline fail if a path error is found.
- When PRIMECLUSTER is used, the regular diagnostic function may cause a stop state path to transition to fail state. In this case, if the relevant path has no connection error, the path will recover automatically following Multipath Driver checks. To prevent path status transition, refer to the PRIMECLUSTER manual and stop the PRIMECLUSTER diagnostic function. Restart the diagnostic function after maintenance work is completed.
- For guest domains when I/O domains have a multipath configuration, only "a/adapter" change unit is supported. Do not specify other change units. Doing so may block all paths.
- If all paths become blocked as shown below, the "change" command terminates with an error during the I/O processing. Stop the I/Os before performing the "change" command.

Examples:

- When the "change" command is performed with I/O domains of single path configurations
- When changing the online path for a two-path configuration that is in a fail and online state

3.2.2 "restart" Command

[Command format]

```
iompadm -c mplb restart [change_unit] target_physical_device_node
```

[Function overview]

This command is used to perform post-processing after hardware hot swapping.

This command restores all of the paths that use the hardware to be replaced, which is determined by the change unit and the target physical device node.

In addition, this command can be used when manually changing the path status from warning to online.

[Usage]

The procedure for hot swapping hardware is described below.

- (1) Execute the "change" command to stop all I/O processes on the paths used by the hardware that is to be replaced.
- (2) Hot swap the hardware.

- (3) Execute the "restart" command to restart all I/O processes on the paths used by the newly replaced hardware.

[Example]

```
# iompadm -c mplb restart c /dev/rdsk/c11t500000E0D4017C80d0s2
```

[Change unit specification]

Refer to ["3.2.1 "change" Command" \(page 51\)](#).

[Caution]

- When the path has a problem and the command is terminated with an error, the path status is changed to "offline fail".
- Due to the status of the fcp driver, there are cases when a restart does not restore the I/O processing. In that situation, execute "cfgadm -al" or "devfsadm", and then perform a restart again.
- For guest domains when I/O domains have a multipath configuration, only "a/adapter" change unit is supported. Do not specify other change units.

3.3 Commands for Creating or Changing the Multipath Configuration

3.3.1 "grmpdautoconf" Command

[Command format]

```
/usr/sbin/grmpdautoconf [ -d ] [ -n ] [log_file]
```

[Function overview]

This command is used to create and remove a multipath.

[Usage]

This command can be used when:

- Creating a multipath during initial installation
- Creating a multipath after adding a path, or logical unit (LU)
- Removing a path, logical unit (LU), or storage system

[Explanation of arguments]

Argument	Description
log_file_name	Specify the output destination for internal logs. This argument can be omitted. If omitted, "/var/opt/FJSMplb/grmpdautoconf.log\$" will be used for the destination (\$ is 1 or 2).
-d	Automatically recognizes paths, logical units (LUs), and storage systems that have been removed, and deletes them.
-n	Does not automatically update the boot-archive after changing the configuration.

[Usage]

Refer to ["2.2 Environment Configuration" \(page 39\)](#) or ["Chapter 4 Adding or Removing Components" \(page 55\)](#).

[Caution]

- Data saved to the target logical units before the multipath is created will remain usable after the multipath has been created. Similarly, data on a logical unit is not lost even when the multipath configuration is changed or deleted.
- Disk labels are not destroyed by a multipath configuration being created, changed, or deleted.
- Other multipath device numbers are not changed by the multipath configuration being released. For example, when both mplb0 and mplb1 exist, releasing mplb0 should have no effect on mplb1.

Chapter 4

Adding or Removing Components

This chapter provides the procedures for adding and removing storage systems, LUs, and paths between storage systems and servers.

Caution

- Before starting an addition process, use the "iompadm info" command to check that the existing devices are in normal state.
If the "grmpdautoconf" command is executed in an abnormal situation, the command may be terminated with an error.
- When processing an addition, if the added LU is not recognized by the ssd driver, use the "cfgadm -al" command, the "disks" command, or the "devfsadm" command so that the added LU is recognized by the ssd driver. After the added LU is recognized by the ssd driver, start the process.
- Before starting a reduction process, check that the device not to be reduced is in normal state.

4.1 Adding a Logical Unit (LU)

Create a logical unit (LU) and add it to the storage system.
For details on adding LUs, refer to the storage system manual.

For guest domains, assign a virtual disk from each I/O domain. For information on how to assign virtual disks, refer to ["2.4 Assigning Virtual Disks" \(page 45\)](#).

Configure the settings on the server side.
Execute the "cfgadm" command so that the LUN is recognized in the ssd driver. Using the "grmpdautoconf" command, the added LUs can be recognized without shutting down the server.

Execute the "cfgadm" command. For guest domains, this execution is not necessary.

```
# /usr/sbin/cfgadm -al
```

- When used as a shared disk for PRIMECLUSTER:
The cluster will implement the automatic resource registration. Refer to the cluster manual for more details.
- When not used as a shared disk for PRIMECLUSTER:
Execute the "grmpdautoconf" command.

```
# /usr/sbin/grmpdautoconf
```

4.2 Adding a Path or Storage System

Add the hardware first.

For guest domains, assign a virtual disk from each I/O domain. For information on how to assign virtual disks, refer to ["2.4 Assigning Virtual Disks" \(page 45\)](#).

Caution

The following functions are required to add an adapter without shutting down the server.

- Dynamic Reconfiguration function
- PCI Hot Plug function

Configure the settings on the server side.

Using the "grmpdautoconf" command, the added paths can be incorporated without shutting down the server.

- When used as a shared disk for PRIMECLUSTER:
The cluster will implement the automatic resource registration. Refer to the cluster manual for more details.
- When not used as a shared disk for PRIMECLUSTER:
Execute the "grmpdautoconf" command.

```
# /usr/sbin/grmpdautoconf
```

Caution

When the single path configuration of an I/O domain is changed to a multipath configuration by adding a path, guest domains must be rebooted and the "grmpdautoconf" command must be executed for the guest domains.

4.3 Removing a Logical Unit (LU)

Using the "grmpdautoconf" command, the LU to be removed can be recognized automatically and deleted. The LU to be removed does not need to be specified. Before executing the command, the LU must be rendered unrecognizable by the server.

For OVM guest domain environments, refer to ["4.3.4 OVM Guest Domain Environments" \(page 58\)](#).

Caution

- If the target that is to be removed is being used as a PRIMECLUSTER or PRIMECLUSTER GDS resource, it must be deregistered as a resource before it is removed.
- Remove the LU in a guest domain while I/O domains are running. While I/O domains are stopped, the LU cannot be removed from guest domains.

Use the following procedure when removing a LU:

4.3.1 Stopping Applications

Stop all applications that might access the LU that is to be removed.
If the LU has been mounted, unmount it. Any `/etc/vfstb` entries must also be deleted.
When the application is being used for ZFS, disconnect the LU with the "zpool export" command, or delete the LU with the "zpool destroy" command.

4.3.2 Removing the LU on the Storage System Side

Render the LU to be removed unrecognizable by the server. Remove the LU on the storage system side. Refer to the storage system manual for details of how to remove a LU.

4.3.3 Executing the "grmpdautoconf" Command

Execute the "grmpdautoconf" command with the "-d" option.

● Example

```
# grmpdautoconf -d
Path      : Action      : Element path                               : LUN : Storage
mplb20 : destroy : c12t2141000B5D6A0299d16s2 c11t2140000B5D6A0299d16s2 : 16 : E4000- 6A0299
=== Reconfigure plan ===
  Destroy multi path : 1
  Delete path from multi path: 0 (instance)

Do you want to continue the multipath remove?

  Yes ---> hit the 'y' key
  No ---> hit the 'n' key

y
```

 **Caution**

Even after this procedure is performed, the special file names under `/dev/[r]dsk/` will remain. Delete them, for example, by using the Solaris "devfsadm" command, if necessary.

4.3.4 OVM Guest Domain Environments

Remove LUs according to the following procedure in OVM guest domain environments.

Procedure

- 1** Stop the applications in the guest domains.
Refer to ["4.3.1 Stopping Applications" \(page 57\)](#) for more information.
- 2** Remove the LU on the storage system side.
Refer to ["4.3.2 Removing the LU on the Storage System Side" \(page 57\)](#) for more information.
- 3** Execute the "grmpdautoconf" command in the guest domains.
Refer to ["4.3.3 Executing the "grmpdautoconf" Command" \(page 57\)](#) for more information.
- 4** Release the virtual disks that are provided to the guest domains by the I/O domains.
For details, refer to "Oracle VM Server for SPARC Administration Guide".
- 5** Execute the "grmpdautoconf" command in the I/O domains.
Refer to ["4.3.3 Executing the "grmpdautoconf" Command" \(page 57\)](#) for more information.

End of procedure

4.4 Removing a Path

Using the "grmpdautoconf" command, the path to be removed can be recognized automatically and deleted. The path to be removed does not need to be specified. Before executing the command, the path to be removed must be rendered unrecognizable by the server.

Caution

Remove the path in a guest domain while I/O domains are running. While I/O domains are stopped, the path cannot be removed from guest domains.

Use the following procedure when removing a path:

4.4.1 Removing the Path between the Server and the Storage System

Remove the path between the server and the storage system. Remove the path by using the zoning settings or by physically disconnecting the cable. Note that paths do not need to be switched to stop state using the "iompadm change" command. Stopped paths are exempt from removal.

4.4.2 Executing the "grmpdautoconf" Command

Execute the "grmpdautoconf" command with the "-d" option. For Fibre Channel connections, a new Fibre Channel driver configuration definition file is output that does not contain the removed path's information.

Example

```
# grmpdautoconf -d
Path   : Action : Element path           : LUN : Storage
mplb0  : del    : c11t2140000B5D6A0299d0s2 : 0   : E4000- 6A0299
mplb1  : del    : c11t2140000B5D6A0299d1s2 : 1   : E4000- 6A0299
mplb2  : del    : c11t2140000B5D6A0299d2s2 : 2   : E4000- 6A0299
=== Reconfigure plan ===
  Destroy multi path : 0
  Delete path from multi path: 3 (instance)

Do you want to continue the multipath remove?

  Yes ---> hit the 'y' key
  No  ---> hit the 'n' key

y
```

Caution

If the path redundancy is lost after a path is removed, a confirmation message appears.

● Example

```
# grmpdautoconf -d
Path      : Action : Element path      : LUN : Storage
mplb0    : del    : c11t2140000B5D6A0299d0s2 : 0 : E4000- 6A0299
mplb1    : del    : c11t2140000B5D6A0299d1s2 : 1 : E4000- 6A0299
mplb2    : del    : c11t2140000B5D6A0299d2s2 : 2 : E4000- 6A0299
=== Reconfigure plan ===
  Destroy multi path : 0
  Delete path from multi path: 3 (instance)

mplb0 : It will be a configuration without the redundancy.
mplb1 : It will be a configuration without the redundancy.
mplb2 : It will be a configuration without the redundancy.

Do you want to continue the multipath remove?

  Yes ---> hit the 'y' key
  No  ---> hit the 'n' key

y
```

 Caution

When the multipath configuration of an I/O domain is changed to a single path configuration by removing a path, guest domains must be rebooted and the "grmpdautoconf" command must be executed for the guest domains.

4.5 Removing a Storage System

Using the "grmpdautoconf" command, the storage system to be removed can be recognized automatically and deleted. The storage system to be removed does not need to be specified. Before executing the command, the storage system to be removed must be rendered unrecognizable by the server. For OVM guest domain environments, refer to "[4.5.4 OVM Guest Domain Environments](#)" (page 63).

Caution

- The server must be rebooted after a storage system has been removed in order to activate the HBA definition file.
- If the target that is to be removed is being used as a PRIMECLUSTER or PRIMECLUSTER GDS resource, it must be deregistered as a resource before it is removed.
- Remove the storage system in a guest domain while I/O domains are running. While I/O domains are stopped, the storage system cannot be removed from guest domains.

Use the following procedure when removing a storage system:

4.5.1 Stopping Applications

Stop all applications that might access the storage system that is to be removed. If the storage system has been mounted, unmount it. Any /etc/vfstsb entries must also be deleted. When the application is being used for ZFS, disconnect the storage system with the "zpool export" command, or delete the storage system with the "zpool destroy" command.

4.5.2 Disconnecting the Storage System from the Server

Disconnect the storage system from the server by using the zoning settings or by physically removing all cables connecting the storage system and the server. Note that paths do not need to be switched to stop state using the "iompadm change" command. Stopped paths are exempt from removal.

4.5.3 Executing the "grmpdautoconf" Command

Execute the "grmpdautoconf" command with the "-d" option. For Fibre Channel connections, a new Fibre Channel driver definition file is output that does not contain the removed storage system's information.

● Example

```
# grmpdautoconf -d
Path   : Action   : Element path                                     : LUN : Storage
mplb0  : destroy  : c11t2141000B5D6A0299d0s2 c12t2141000B5D6A0299d0s2 : 0   : E4000- 6A0299
mplb1  : destroy  : c11t2141000B5D6A0299d1s2 c12t2141000B5D6A0299d1s2 : 1   : E4000- 6A0299
mplb2  : destroy  : c11t2141000B5D6A0299d2s2 c12t2141000B5D6A0299d2s2 : 2   : E4000- 6A0299
=== Reconfigure plan ===
  Destroy multi path : 3
  Delete path from multi path: 0 (instance)

Do you want to continue the multipath remove?

  Yes ---> hit the 'y' key
  No  ---> hit the 'n' key

y
```

 **Caution**

Even after this procedure is performed, the special file names under "/dev/[r]dsk/" will remain. Delete them, for example, by using the Solaris "devfsadm" command, if necessary.

4.5.4 OVM Guest Domain Environments

Remove storage systems according to the following procedure in OVM guest domain environments.

Procedure

- 1** Stop the applications in the guest domains.
For details, refer to ["4.5.1 Stopping Applications" \(page 61\)](#).
- 2** Disconnect the storage system from the server.
For details, refer to ["4.5.2 Disconnecting the Storage System from the Server" \(page 61\)](#).
- 3** Execute the "grmpdautoconf" command in the guest domains.
For details, refer to ["4.5.3 Executing the "grmpdautoconf" Command" \(page 62\)](#).
- 4** Release the virtual disks that are provided to the guest domains by the I/O domains.
For details, refer to "Oracle VM Server for SPARC Administration Guide".
- 5** Execute the "grmpdautoconf" command in the I/O domains.
For details, refer to ["4.5.3 Executing the "grmpdautoconf" Command" \(page 62\)](#).

End of procedure

Chapter 5

Replacing Hardware Components

This chapter explains how to replace hardware components.

Before replacing a hardware component, the Multipath Driver must be used to "set aside" any paths that will be disrupted by the replacement procedure. To "set aside" means to change the state of the paths to offline. Perform "set aside" to gain the following benefits:

- Checks whether redundancy is maintained.
Checks the paths on standby that will be used after the replacement to determine whether or not they are blocked. This is particularly necessary in cluster configurations and for preventive maintenance.
- Reduces the number of console error messages.

Hardware components must be replaced by a maintenance engineer. The customer should not attempt to replace failed components. Contact a maintenance engineer.

For OVM guest domain environments, perform "set aside" on the following domains.

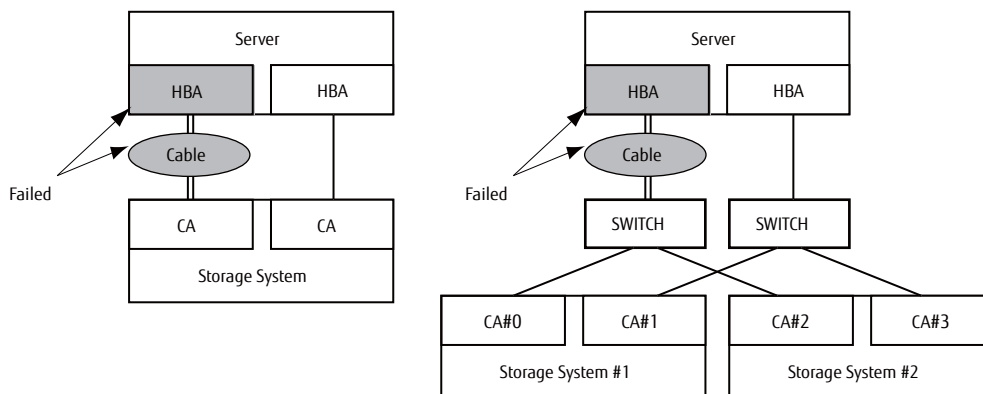
- When I/O domains have a multipath configuration, set the paths aside on the I/O domains. The "set aside" is not necessary on the guest domains.
- When I/O domains have a single path configuration, set the paths aside on the guest domains.

5.1 Replacing HBAs or Server Connection Cables

5.1.1 How to Specify the Replacement Targets and the Set Aside Targets

This section explains the procedures for replacement of the server HBA or the server connection cable, shown in the figure below.

Figure 5.1 Failed part (HBA or server connection cable)



The preventive maintenance replacement procedure described below is for the right-hand configuration above, when the following message is output to the console. All paths that run through the failed HBA or cable must be taken offline. In this example, the I/O paths that run through CA#0 and CA#2 are taken offline.

[Example console message]

Instance name is underlined.

```
/pci@0,600000/pci@0/pci@8/pci@0,1/fibre-channel@1,1/fp@0,0/ssd@w2140000b5d
6a0299,0 (ssd53):
SCSI transport failed: reason'incomplete': retrying command
```

The offline command is as follows:

```
# /usr/opt/FJSViomp/bin/iompadm -c mplb change "change_unit" "physical_device_node"
```

Use the "instance_name" (shown underlined above) from the console message as the physical device node. Search for the same "instance_name" in the output of the "iompadm info" command. Specify the matching path names (shown underlined in the output example below) as the physical device nodes.

[Example output for the "iompadm info" command]

```
# /usr/opt/FJSViomp/bin/iompadm info
IOMP: /dev/FJSVmplb/fiomp/adm0
Element:
  /dev/rdisk/c11t2140000B5D6A0299d0s2  online standby block "good status
[E4000- 6A0299-CM10-CA40-PORT01] (ssd53)"
  /dev/rdisk/c12t2141000B5D6A0299d0s2  online active block "good status
[E4000- 6A0299-CM11-CA41-PORT03] (ssd2)"
:

IOMP: /dev/FJSVmplb/fiomp/adm8
Element:
  /dev/rdisk/c11t2140000B5D6A0299d8s2  online standby block "good status
[E4000- 6A0299-CM10-CA40-PORT01] (ssd61)"
  /dev/rdisk/c12t2141000B5D6A0299d8s2  online active block "good status
[E4000- 6A0299-CM11-CA41-PORT03] (ssd10)"
:
```

Specify "adapter" as the change_unit. When "adapter" is specified, all paths that include the same controller number as that in the physical device node in addition to the paths that directly specify the physical device node will be taken offline. In the example, the /dev/rdisk/c11t2140000B5D6A0299d8s2 path for adm8, in addition to the /dev/rdisk/c11t2140000B5D6A0299d0s2 path for adm0 will be taken offline.

The following command is used to set aside a path:

```
# /usr/opt/FJSViomp/bin/iompadm -c mplb change a /dev/rdisk/c11t2140000B5D6A0
299d0s2
```

5.1.2 Hot Swapping an HBA

The following functions are required when a server HBA is to be replaced without turning off the server:

- Dynamic Reconfiguration function
- PCI Hot Plug function

Refer to the following manuals when using these functions to perform a hot swap.

- Dynamic Reconfiguration function
"Enhanced Support Facility User's Guide For Dynamic Reconfiguration"
"Enhanced Support Facility User's Guide Dynamic Reconfiguration I/O device"
- PCI Hot Plug function
Refer to "[B.8 PCI Hot Plug](#)" ([page 112](#)) for more information.

Caution

The type (model) of the replacement HBA must be the same as that of the HBA being replaced. For example, a SE0X7F31F type HBA cannot be replaced with a SE0X7F21F type HBA.

5.1.3 Cold Swapping an HBA

The following procedure describes how to replace an HBA with the server turned off.

Procedure

- 1 Turn off the server.
- 2 Replace the HBA inside the server.
- 3 If host affinity is set for the storage system and WWN zoning for the Fibre Channel switch, update the settings to use the new HBA's WWN.
Refer to the individual device manuals for details.
- 4 Turn on the server.
- 5 Execute the following command to verify that all paths are now either active or standby.

```
# /usr/opt/FJSViomp/bin/iompadm -c mplb info
```

End of procedure

Caution

- The type (model) of the replacement HBA must be the same as that of the HBA being replaced. For example, a SE0X7F31F type HBA cannot be replaced with a SE0X7F21F type HBA.
- When host affinity or WWN zoning is not correctly set in [Step 3](#), the paths will be offline in [Step 5](#). In this case, execute the "cfgadm -al" command and the "iompadm restart" command to restore the paths after correcting the settings on the storage system and the switch.

5.1.4 Hot Swapping a Server Connection Cable

The following procedure describes how to replace a cable connected to the server without turning the server off.

Procedure

- 1 Execute the following command to set aside the path.

```
# /usr/opt/FJSViomp/bin/iompadm -c mplb change a /dev/rdisk/c11t2140000B5D6A0  
299d0s2
```

- 2 Replace the cable that is connected to the server.

- 3 Execute the following command to release the path that was set aside.

```
# /usr/opt/FJSViomp/bin/iompadm -c mplb restart a /dev/rdisk/c11t2140000B5D6A0  
299d0s2
```

- 4 Execute the following command to verify that all paths are now either active or standby.

```
# /usr/opt/FJSViomp/bin/iompadm -c mplb info
```

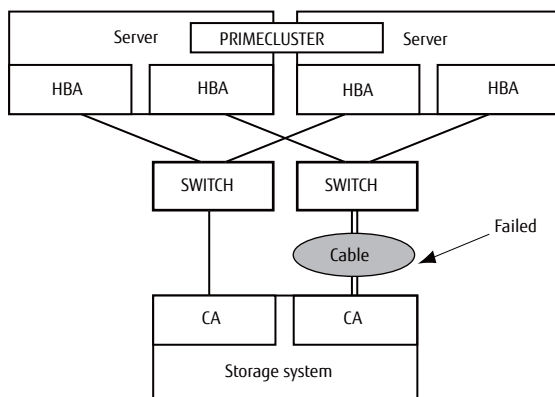
End of procedure

5.2 Replacing a Cable Connecting the Storage System and the Switch

5.2.1 How to Specify the Replacement Targets and the Set Aside Targets

This section contains the procedures for replacing a cable connecting the storage system and a switch, shown the figure below.

Figure 5.2 Failed part (a cable connecting the storage system and a switch)



The preventive maintenance replacement procedure described below is for the configuration above, when the following message is output to the main server console. All paths that run through the failed CA or cable must be taken offline.

[Example console message]

Instance name is underlined.

```
/pci@0,600000/pci@0/pci@8/pci@0,1/fibre-channel@1,1/fp@0,0/ssd@w2140000b5d  
6a0299,0 (ssd53):  
SCSI transport failed: reason'incomplete': retrying command
```

The offline command is as follows:

```
# /usr/opt/FJSViomp/bin/iompadm -c mplb change "change_unit" "physical_device_node"
```

Use the "instance_name" (shown underlined above) from the console message as the physical device node. Search for the same "instance_name" in the output of the "iompadm info" command. Specify the matching path names (shown underlined in the output example below) as the physical device nodes.

[Example output for the "iompadm info" command]

```
# /usr/opt/FJSViomp/bin/iompadm info (primary_server)
IOMP: /dev/FJSVmplb/fiomp/adm2048
Element:
  /dev/rdisk/c11t2140000B5D6A0299d0s2  online standby block  "good status
[E4000- 6A0299-CM10-CA40-PORT01] (ssd53) "
  /dev/rdisk/c12t2141000B5D6A0299d0s2  online active block  "good status
[E4000- 6A0299-CM11-CA41-PORT03] (ssd2) "
  :
# /usr/opt/FJSViomp/bin/iompadm info (secondary_server)
IOMP: /dev/FJSVmplb/fiomp/adm2048
Element:
  /dev/rdisk/c5t2140000B5D6A0299d0s2  online standby block  "good status
[E4000- 6A0299-CM10-CA40-PORT01] (ssd10) "
  /dev/rdisk/c6t2141000B5D6A0299d0s2  online active block  "good status
[E4000- 6A0299-CM11-CA41-PORT03] (ssd20) "
  :
```

Specify "controller" as the change_unit. When "controller" is specified, all paths that have the same controller identifier as that in the physical device node in addition to the paths that directly specify the physical device node will be taken offline. In the example, the /dev/rdisk/c5t2140000B5D6A0299d0s2 path for secondary server adm2048 in addition to the /dev/rdisk/c11t2140000B5D6A0299d0s2 path for primary server adm2048 will be taken offline.

The following command is used to set aside a path:

```
# /usr/opt/FJSViomp/bin/iompadm -c mplb change c /dev/rdisk/c11t2140000B5D6A0299d0s2
```

5.2.2 Hot Swapping a Cable Connecting the Storage System and a Switch

The following procedure describes how to replace a cable between the storage system and switch without turning the server and storage system off.

Procedure

- 1 Execute the following command to set aside the path.

```
# /usr/opt/FJSViomp/bin/iompadm -c mplb change c /dev/rdisk/c11t2140000B5D6  
A0299d0s2
```

- 2 Replace the cable between the storage system and the switch.

- 3 Execute the following command to release the path that was set aside.

```
# /usr/opt/FJSViomp/bin/iompadm -c mplb restart c /dev/rdisk/c11t2140000B5D6  
A0299d0s2
```

- 4 Execute the following command to verify that all paths are now either active or standby.

```
# /usr/opt/FJSViomp/bin/iompadm -c mplb info
```

End of procedure

5.3 Replacing a CA or IOB

5.3.1 How to Specify the Replacement Targets and the Set Aside Targets

This section describes the procedures for replacing the CA or the IOB. Some storage systems require the CM to be removed to allow the CA to be replaced. In that case, refer to ["5.4 Replacing a CM or ROUTER" \(page 73\)](#).

Figure 5.3 Failed part (CA) and the replacement area

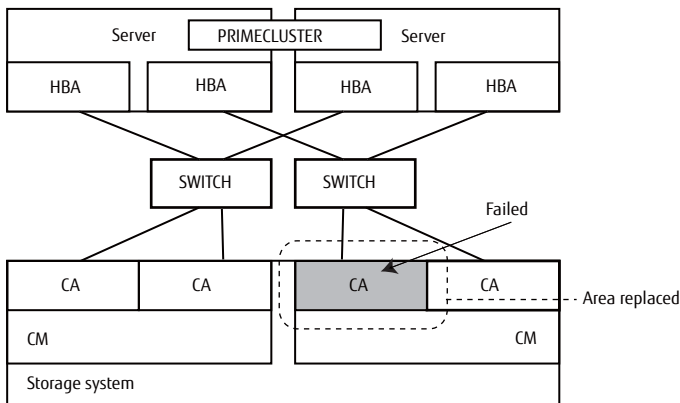
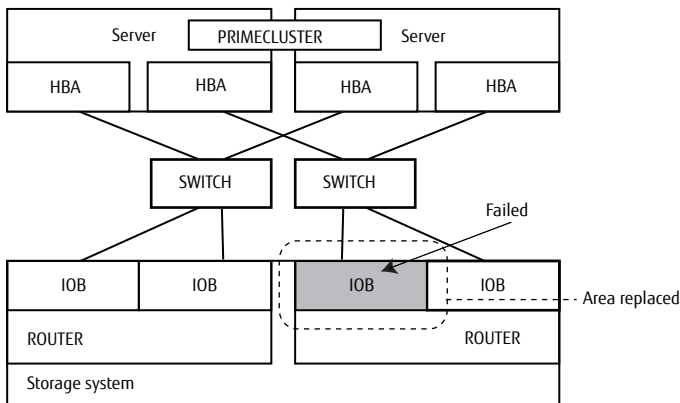


Figure 5.4 Failed part (IOB) and the replacement area



The preventive maintenance replacement procedure described below is for the configuration above, when the following message is output only once to the main server console. All paths connected to the failed CA or IOB must be taken offline.

[Example console message]

Instance name is underlined.

```
/pci@0,600000/pci@0/pci@8/pci@0,1/fibre-channel@1,1/fp@0,0/ssd@w2140000b5d
6a0299,0 (ssd53):
SCSI transport failed: reason'incomplete': retrying command
```

The offline command is as follows:

```
# /usr/opt/FJSViomp/bin/iompadm -c mplb change "change_unit" "physical_device_node"
```

Use the "instance_name" (shown underlined above) from the console message as the physical device node. Search for the same "instance_name" in the output of the "iompadm info" command. Specify the matching path names (shown underlined in the output example below) as the physical device nodes.

[Example output for the "iompadm info" command]

```
# /usr/opt/FJSViomp/bin/iompadm info (primary_server)
IOMP: /dev/FJSVmplb/fiomp/adm2048
Element:
    /dev/rdisk/c11t2140000B5D6A0299d0s2  online standby block "good status
[E4000- 6A0299-CM10-CA40-PORT01] (ssd53) "
    /dev/rdisk/c12t2141000B5D6A0299d0s2  online active block "good status
[E4000- 6A0299-CM11-CA41-PORT03] (ssd2) "
:
# /usr/opt/FJSViomp/bin/iompadm info (secondary_server)
IOMP: /dev/FJSVmplb/fiomp/adm2048
Element:
    /dev/rdisk/c5t2140000B5D6A0299d0s2  online standby block "good status
[E4000- 6A0299-CM10-CA40-PORT01] (ssd10) "
    /dev/rdisk/c6t2141000B5D6A0299d0s2  online active block "good status
[E4000- 6A0299-CM11-CA41-PORT03] (ssd20) "
:
```

Specify "controller unit" as the change_unit. When "controller unit" is specified, all paths that have same value for "number of change_unit in the controller identifier" as that in the physical device node in addition to the paths that directly specify the physical device node will be taken offline. In the example, the /dev/rdisk/c5t2140000B5D6A0299d0s2 path for secondary server adm2048 in addition to the /dev/rdisk/c11t2140000B5D6A0299d0s2 path for primary server adm2048 will be set offline.

The following command is used to set aside a path:

```
# /usr/opt/FJSViomp/bin/iompadm -c mplb change cu /dev/rdisk/c11t2140000B5D6A0299d0s2
```

5.3.2 Hot Swapping a CA or IOB

The following procedure describes how to replace a CA or IOB without turning off the server and storage system.

Procedure

- 1 Execute the following command to set aside the path.

```
# /usr/opt/FJSViomp/bin/iompadm -c mplb change cu /dev/rdisk/c11t2140000B5D6A0299d0s2
```

- 2 Replace the CA or IOB.

- 3 Execute the following command to release the path that was set aside.

```
# /usr/opt/FJSViomp/bin/iompadm -c mplb restart cu /dev/rdsk/c11t2140000B5D  
6A0299d0s
```

- 4 Execute the following command to verify that all paths are now either active or standby.

```
# /usr/opt/FJSViomp/bin/iompadm -c mplb info
```

End of procedure

5.4 Replacing a CM or ROUTER

5.4.1 How to Specify the Replacement Targets and the Set Aside Targets

This section describes the procedures for replacement of the CM or ROUTER.

Figure 5.5 Failed part (CM) and the replacement area

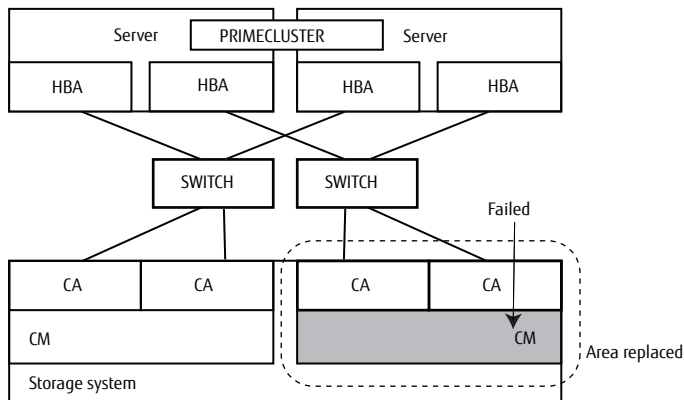
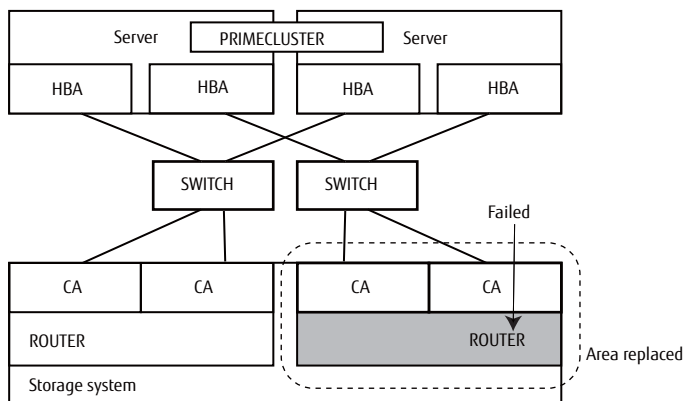


Figure 5.6 Failed part (ROUTER) and the replacement area



The preventive maintenance replacement procedure described below is for the configuration above, when the following message is output only once to the main server console. All paths connected to the failed CM or ROUTER must be taken offline. However, the ETERNUS6000 does not require those paths be set aside when replacing a CM.

[Example console message]

Instance name is underlined.

```
/pci@0,600000/pci@0/pci@8/pci@0,1/fibre-channel@1,1/fp@0,0/ssd@w214000b5d
6a0299,0 (ssd53):
SCSI transport failed: reason 'incomplete': retrying command
```

The offline command is as follows:

```
# /usr/opt/FJSViomp/bin/iompadm -c mplb change "change_unit" "physical_device_node"
```

Use the "instance_name" (shown underlined above) from the console message as the physical device node. Search for the same "instance_name" in the output of the "iompadm info" command. Specify the matching path names (shown underlined in the output example below) as the physical device nodes.

[Example output for the "iompadm info" command]

```
# /usr/opt/FJSViomp/bin/iompadm info (primary_server)
IOMP: /dev/FJSVmplb/fiomp/adm2048
Element:
  /dev/rdsk/c11t214000B5D6A0299d0s2   online standby block "good status
[E4000- 6A0299-CM10-CA40-PORT01] (ssd53)"
  /dev/rdsk/c12t2141000B5D6A0299d0s2   online active block "good status
[E4000- 6A0299-CM11-CA41-PORT03] (ssd2)"
  :
# /usr/opt/FJSViomp/bin/iompadm info (secondary_server)
IOMP: /dev/FJSVmplb/fiomp/adm2048
Element:
  /dev/rdsk/c5t214000B5D6A0299d0s2   online standby block "good status
[E4000- 6A0299-CM10-CA40-PORT01] (ssd10)"
  /dev/rdsk/c6t2141000B5D6A0299d0s2   online active block "good status
[E4000- 6A0299-CM11-CA41-PORT03] (ssd20)"
  :
```

Specify "groupmodule" as the change_unit. When "groupmodule" is specified, all paths that have same value for "number of change_unit in the controller identifier" as that in the physical device node in addition to the paths that directly specify the physical device node will be taken offline. In the example, the /dev/rdsk/c5t214000B5D6A0299d0s2 path for secondary server adm2048 in addition to the /dev/rdsk/c11t214000B5D6A0299d0s2 path for primary server adm2048 will be taken offline.

The following command is used to set aside a path:

```
# /usr/opt/FJSViomp/bin/iompadm -c mplb change g /dev/rdsk/c11t214000B5D6A0
299d0s2
```

5.4.2 Hot Swapping a CM or ROUTER

The following procedure describes how to replace a CM or ROUTER without turning off the server and storage system.

Procedure

- 1 Execute the following command to set aside the path.

```
# /usr/opt/FJSViomp/bin/iompadm -c mplb change g /dev/rdisk/c11t2140000B5  
D6A0299d0s2
```

- 2 Replace the CM or ROUTER.

- 3 Execute the following command to release the path that was set aside.

```
# /usr/opt/FJSViomp/bin/iompadm -c mplb restart g /dev/rdisk/c11t2140000B5  
D6A0299d0s2
```

- 4 Execute the following command to verify that all paths are now either active or standby.

```
# /usr/opt/FJSViomp/bin/iompadm -c mplb info
```

End of procedure

5.5 Replacing a Storage System

The procedures for replacing the storage system due to failure are explained below.

Caution

The following conditions must be met before replacing the storage system.

- When re-connecting the cables after the replacement, the cables must be connected to the same positions as before.
- The LUNs of the new storage system must be created in the same configuration as the storage system being replaced.
- In single user mode the storage system will be in a state where access is temporarily unavailable (all paths in an offline state). For that reason, do not allow access to the storage system from other applications, etc. during single user mode.

Procedure

- 1 Shut down the server.
- 2 Replace the storage system.

- 3 Boot the server to single user mode.
- 4 Mount the areas that include system files such as /var and /opt if they are not already mounted.
- 5 Execute the following command to enable the replacement destination storage system LUN to be recognized by the `ssd` driver.

```
# cfgadm -al
```

- 6 Execute the "format" command, and confirm that all of the replacement destination storage system LUNs are recognized.
- 7 Execute the following command, and configure the multipath.

```
# /usr/sbin/mplbconfig -x
*** Phase 1: checking /dev/rdisk ***
*** Phase 2: checking multi path ***
===== multipaths which body exchanged =====
      old body          new body
mplb13 : E4000 6A0170 -> E4000          6A0299
Do you want to execute this plan (y/n)?
y
*** Phase 5: result is SUCCESS ***
```

- 8 Execute the following command to confirm that all of the replacement destination storage system LUN paths are in an active or standby state. Also confirm that all of the replacement source storage system paths are in an offline state.

```
# /usr/opt/FJSViomp/bin/iompadm -c mplb info
(snip)
IOMP: /dev/FJSVmplb/fiomp/adm13
Element:
  /dev/rdsk/c13t2148000B5D6A0170d12s2      offline fail      unblock "no
such device or address [E4000- 6A0170-CM10-CA40-PORT11] (ssd122)"
  /dev/rdsk/c14t2149000B5D6A0170d12s2      offline fail      unblock "no
such device or address [E4000- 6A0170-CM11-CA41-PORT13] (ssd115)"
status      /dev/rdsk/c13t2148000B5D6A0299d12s2      online  standby block    "good
[E4000- 6A0299-CM10-CA40-PORT11] (ssd47)"
status      /dev/rdsk/c14t2149000B5D6A0299d12s2      online  active  block    "good
[E4000- 6A0299-CM11-CA41-PORT13] (ssd53)"
Node:
(snip)
```

9 Execute the following command to remove the replacement source storage system.

```
# /usr/sbin/grmpdautoconf -d
Path      : Action      : Element path      : LUN : Storage
mplb13 : del   : c13t2148000B5D6A0170d12s2 c14t2149000B5D6A0170d12s2 : 12 :
E3000- 6A0190
=== Reconfigure plan ===
  Destroy multi path : 0
  Delete path from multi path: 1 (instance)

Do you want to continue the multipath remove?

Yes ---> hit the 'y' key
No ---> hit the 'n' key
y
```

10 Execute the following command to verify that all of the replacement destination storage system paths are now either active or standby. Also confirm that the replacement source storage system paths are not displayed.

```
# /usr/opt/FJSViomp/bin/iompadm -c mplb info
(snip)
IOMP: /dev/FJSVmplb/fiomp/adm13
Element:
      /dev/rdisk/c13t2148000B5D6A0299d12s2      online  standby block  "good
status      [E4000- 6A0299-CM10-CA40-PORT11] (ssd47) "
      /dev/rdisk/c14t2149000B5D6A0299d12s2      online  active  block  "good
status      [E4000- 6A0299-CM11-CA41-PORT13] (ssd53) "
Node:
(snip)
```

11 Exit single user mode.

End of procedure

5.6 Replacing a Server System Board

5.6.1 Hot Swapping a System Board

Replacing a system board without turning the server off is only possible if the server has the Dynamic Reconfiguration function. Refer to the following manuals when using these functions to perform a hot swap.

"Enhanced Support Facility User's Guide For Dynamic Reconfiguration"

"Enhanced Support Facility User's Guide Dynamic Reconfiguration I/O device"

When Dynamic Reconfiguration that is controlled by the XSCF is performed in the SPARC Enterprise M4000/M5000/M8000/M9000, the I/O device may need to be stopped. Perform the following procedure.

Procedure

1 Disconnect the path from the Multipath Driver.

Disconnect the path on the target system board from the Multipath Driver using the "iompadm" command.

Example: If the path of the replacement target is /dev/rdisk/c11t2140000B5D6A0299d0s2

```
# /usr/opt/FJSViomp/bin/iompadm -c mplb change adapter_disconnect  
/dev/rdisk/c11t2140000B5D6A0299d0s2
```

After executing the "iompadm" command, confirm that the status of the disconnected path is "unconfigured disconnected" using the "iompadm info" command.

2 Perform Dynamic Reconfiguration and replace the system board.

3 Connect the path to the Multipath Driver.

Connect the disconnected path to the Multipath Driver using the "iompadm" command.

```
# /usr/opt/FJSViomp/bin/iompadm -c mplb restart adapter_connect  
/dev/rdisk/c11t2140000B5D6A0299d0s2
```

After executing the "iompadm" command, confirm that the status of the connected path is "online" using the "iompadm info" command.

End of procedure

5.6.2 Cold Swapping a System Board

The following procedure describes how to replace the system board with the server turned off.

Procedure

- 1 Turn off the server.
- 2 Replace the system board inside the server.
Refer to the individual server manuals for details.
- 3 Turn on the server.
- 4 Execute the following command to verify that all paths are now either active or standby.

```
# /usr/opt/FJSViomp/bin/iompadm -c mplb info
```

End of procedure

5.7 Replacing a Fibre Channel Switch or SFP

5.7.1 Hot Swapping a Fibre Channel Switch or SFP

The following procedure describes how to replace a Fibre Channel switch or an SFP installed in a Fibre Channel switch without turning off the server and storage system.

Procedure

- 1 Use the following command to confirm that the paths which are not connected to the Fibre Channel switch or SFP that is to be replaced are in "online active" or "online standby" state:

```
# /usr/opt/FJSViomp/bin/iompadm -c mplb info
```

- 2 Replace the Fibre Channel switch or SFP. No multipath commands need to be executed. Error messages may appear for paths connected to the Fibre Channel switch or SFP that is to be replaced. However, even if the path status changes to "warning" or "offline", the failover will be performed to the normal path. Error message examples are provided below.

(Example 1)

```
/pci@0,600000/pci@0/pci@8/pci@0,1/fibre-channel@1,1/fp@0,0/ssd@w2140000b5d  
6a0299,0 (ssd53):  
disk not responding to selection
```

(Example 2)

```
/pci@0,600000/pci@0/pci@8/pci@0,1/fibre-channel@1,1/fp@0,0/ssd@w2140000b5d  
6a0299,0 (ssd53):  
SCSI transport failed: reason 'tran_err': giving up
```

- 3 Execute the following command to verify that all paths are now online.

```
# /usr/opt/FJSViomp/bin/iompadm -c mplb info
```

Paths connected to the replaced Fibre Channel switch or SFP may be shown in an offline fail or online warning state. If this is the case, specify the offline fail path or the online warning path and execute the following command:

```
# /usr/opt/FJSViomp/bin/iompadm -c mplb restart a "physical_device_node"
```

(Example)

```
# /usr/opt/FJSViomp/bin/iompadm -c mplb restart a /dev/rdisk/c11t2140000B5  
D6A0299d0s2
```

End of procedure

5.7.2 Cold Swapping a Fibre Channel Switch or SFP

The following procedure describes how to replace a Fibre Channel switch or an SFP installed in a Fibre Channel switch with the server turned off.

Procedure

- 1 Turn off all the servers connected to the target Fibre Channel switch or the SFP to be replaced.
- 2 Replace the Fibre Channel switch or SFP.
- 3 Turn on the server.
- 4 Execute the following command to verify that all paths are now either in an online active or an online standby state.

```
# /usr/opt/FJSViomp/bin/iompadm -c mplb info
```

End of procedure

5.8 Replacing a Storage System Disk

The following procedure describes how to replace a storage system disk due to failure.

Procedure

- 1 Replace the disk while referring to the individual storage system manuals.
- 2 If data in the volume has been lost due to multiple failures in a RAID group, etc., use the "format" command to recreate the disk label. Refer to the man pages for details of the "format" command.

```
# format /dev/FJSVmplb/rdisk/mplb0s2
```

End of procedure

Chapter 6

Frequently Asked Questions

Category	Issue	Answer or solution
<input type="checkbox"/> Operation by commands		
	<p>The format command does not display the LUN for the storage system.</p>	<p>When setting the mplb special file, specifying the special file as the argument is necessary to execute the "format" command. Execute as shown in the following example:</p> <pre># format /dev/FJSVmplb/rdisk/mplb*s2</pre>
	<p>Even though the LUN size has been changed, the "format" command shows the old size.</p>	<p>This is due to the old disk label information remaining. Perform the following procedure. The characters in bold are lines to be entered. The format argument is the target LUN special file name.</p> <pre># format /dev/FJSVmplb/rdisk/mplb3s2 format> type AVAILABLE DRIVE TYPES: 0. Auto configure : 19. other Specify disk type(enter its number)[18]: 0 /dev/FJSVmplb/rdisk/mplb3s2: configured with capacity of 1.98GB <FUJITSU-ETERNUS_DX400-0000 cyl 254 alt 2 hd 64 sec 256> selecting /dev/FJSVmplb/rdisk/mplb3s2 [disk formatted] format> label Ready to label disk, continue? y</pre> <p>In addition, data retention is not guaranteed since the partition configuration is changed. Back up the data as necessary before changing the LUN size and then restore from the backup after changing the size and recreating the disk label.</p>
<input type="checkbox"/> Changing system configuration		
	<p>What are the procedures when converting from a single path configuration to a multipath configuration?</p>	<p>The procedure is almost the same as that shown in "2.2 Environment Configuration" (page 39). However, note the following points.</p> <p>It is not necessary to perform "2.3.1 Changing the Partition Configuration" (page 44) and "2.3.2 Creating a File System" (page 44).</p> <p>As the access path name to the disk will change, it is necessary to change the setting of applications such as /etc/vfstab.</p>
	<p>Are there methods for restricting the number of SCSI commands issued (max_throttle)?</p>	<p>Refer to "FUJITSU Storage ETERNUS DX Configuration Guide - Server Connection-" when restricting the number of SCSI commands to issue.</p>

Category	Issue	Answer or solution
<p>❑ Messages</p>		
	<p>Messages such as the below appear during operation. (Example) [ID 365881 kern.info] <FUJITSU-GR740-0321 cyl 1954 alt 2 hd 64 sec 256></p>	<p>They are messages output if the geometry information is read when starting the server or executing the "format" command, and do not indicate errors.</p>
	<p>Messages such as the below appear during operation. (Example) corrupt label wrong magic number</p>	<p>The message at left may appear if the disk label information has been lost by raw access that included cylinder 0 or by rebuilding of the RAID group and volume after replacement of a failed disk. Use the following procedure to recover the information. The characters in bold are lines to be entered. The format argument is the target LUN special file name.</p> <pre># format /dev/FJSVmplb/rdisk/mp1b3s2 format> type AVAILABLE DRIVE TYPES: 0. Auto configure : 19. other Specify disk type(enter its number)[18]: 0 /dev/FJSVmplb/rdisk/mp1b3s2: configured with capacity of 1.98GB <FUJITSU-ETERNUS_DX400-0000 cyl 254 alt 2 hd 64 sec 256> selecting /dev/FJSVmplb/rdisk/mp1b3s2 [disk formatted] format> label Ready to label disk, continue? y</pre>
	<p>Messages such as the below appear during setting of the Multipath Driver. (Example) WARNING:ssd33252:c,raw minor 0x40f22 too big for 32-bit applications</p>	<p>The maximum number of instances that can be controlled by the Solaris OS has been exceeded, and the configuration needs to be changed. Contact a Fujitsu engineer.</p>
<p>❑ Turning on the power</p>		
	<p>The storage system cannot be recognized by the server.</p>	<p>When the power of the storage system is turned on after the server is turned on, the storage system may not be recognized. Make sure to turn on the server after the storage system. Use the "iompadm info" command to check the status immediately after turning the power on. If a fail state is displayed, reboot the server.</p>
	<p>Booting the server takes a very long time to complete.</p>	<p>If the power of the storage system is turned on after the server is turned on, or if the server is turned on with no cables connecting the server and the storage system, it may take a long time for the server to complete booting because the multipath recognition process takes time. Wait until boot processing completes, or check the cable connections between the server and the storage system, and then render the disks recognizable by the server. Use the "iompadm info" command to check status after turning the server on. If the fail state is displayed, reboot the server.</p>

Category	Issue	Answer or solution
❑ Cluster configuration	<p>Messages such as below are output to the console. (Example) "FJSVcluster:Error: /etc/opt/FJSVcluster/sys/ acmaadproc. 6212.The specified configuration change procedure is already registered. (/tmp/implb/reqproc)</p>	<p>This is output according to the timing of initialization operations during cluster activation, and has no effect on system operations.</p>
	<p>The messages below are output to the console. "implbcl: error add procedure"</p>	<p>This message may appear when all the cluster nodes are restarted for the first time after the initial cluster configuration has been set up. If this message is no longer output after restarting the nodes, no further action is required. If the message appears again, contact a Fujitsu engineer.</p>
	<p>The physical device nodes(/dev/rdisk/c*t*d*s*) displayed when the "iompadm info/status" command is used vary between the cluster nodes.</p>	<p>This is not a problem. See the following example:</p> <p>Example: For Server 1 and Server 2 in a cluster configuration, the device nodes for disk 0 may be c1 and c2 for Server 1, and c3 and c4 for Server 2.</p> <ul style="list-style-type: none"> • Result when the "iompadm info" command is run on Server 1: <pre># iompadm -c implb info /dev/FJSVimplb/ fiomp/adm2048 IOMP: /dev/FJSVimplb/fiomp/adm2048 Element: /dev/rdisk/c1t2141000B5D6A0299d0s2 online active... /dev/rdisk/c2t2141000B5D6A0299d0s2 online active...</pre> <pre>Node: /dev/FJSVimplb/dsk/implb2048s0 /dev/FJSVimplb/dsk/implb2048s1 :</pre> • Result when the "iompadm info" command is run on Server 2: <pre># iompadm -c implb info /dev/FJSVimplb/ fiomp/adm2048 IOMP: /dev/FJSVimplb/fiomp/adm2048 Element: /dev/rdisk/c3t2141000B5D6A0299d0s2 online active... /dev/rdisk/c4t2141000B5D6A0299d0s2 online active...</pre> <pre>Node: /dev/FJSVimplb/dsk/implb2048s0 /dev/FJSVimplb/dsk/implb2048s1 :</pre>

Category	Issue	Answer or solution												
<p>❑ Changing the storage system settings</p>														
	How can the correspondence (LUN mapping or host affinity) between the LUNs recognized by the server and the volumes (LUN_V) in the storage system be changed or deleted?	The existing multipath needs to be released and reconfigured. For details, refer to "B.3 Changing a LUN Mapping or Host Affinity" (page 105) .												
	How can the assigned CMs of a RAID group used in the Multipath Driver be changed?	After changing the settings, reboot the server.												
<p>❑ Other</p>														
	Are there resident processes in the Multipath Driver?	<p>The two "mplbstrtrc" processes are resident processes. The main function of these resident processes is to collect I/O error information. A "ps" command output example is shown below.</p> <pre># ps -aef UID PID PPID C STIME TTY TIME CMD : root 62 1 0 13:45:25 ? 0:00 /usr/sbin/ mplbstrtrc root 63 62 0 13:45:25 ? 0:00 /usr/sbin/ mplbstrtrc :</pre>												
	How can the log file size and number of generations be determined?	<p>The log file size and the number of generations that can be stored are shown below.</p> <table border="1"> <thead> <tr> <th>Log file name</th> <th>Size</th> <th>No. of generations</th> </tr> </thead> <tbody> <tr> <td>/var/opt/FJSMplb/multi-path-config.log*</td> <td>2MB</td> <td>2</td> </tr> <tr> <td>/var/opt/FJSMplb/mplbtrc.*</td> <td>1MB</td> <td>3</td> </tr> <tr> <td>/var/opt/FJSMplb/grmpdautoconf.log*</td> <td>1MB</td> <td>2</td> </tr> </tbody> </table>	Log file name	Size	No. of generations	/var/opt/FJSMplb/multi-path-config.log*	2MB	2	/var/opt/FJSMplb/mplbtrc.*	1MB	3	/var/opt/FJSMplb/grmpdautoconf.log*	1MB	2
Log file name	Size	No. of generations												
/var/opt/FJSMplb/multi-path-config.log*	2MB	2												
/var/opt/FJSMplb/mplbtrc.*	1MB	3												
/var/opt/FJSMplb/grmpdautoconf.log*	1MB	2												

Appendix A

Error Messages

This appendix explains the error messages that are displayed by the Multipath Driver.

A.1 Notification, Degradation, and Warning Messages

For this version, transmission path messages and SCSI sense messages that were displayed for Multipath Driver 3.0.4 and earlier are not displayed.

A.1.1 Transmission Related Messages

The following messages are not displayed for this version.

WARNING: Disk pathname (e.g. /pci@1d,4000/fibre-channel@1/mplbt@10,0 (mplbt16))

SCSI transport failed: reason'incomplete': retrying command

WARNING: Disk pathname (e.g. /pci@1d,4000/fibre-channel@1/mplbt@10,0 (mplbt16))

disk not responding to selection

WARNING: Disk pathname (e.g. /pci@1d,4000/fibre-channel@1/mplbt@10,0 (mplbt16))

SCSI transport failed : reason 'tran_err' : giving up

WARNING: Disk pathname (e.g. /pci@1d,4000/fibre-channel@1/mplbt@10,0 (mplbt16))

SCSI transport failed : reason 'reset' : XXXXXXXX

WARNING: Disk pathname (e.g. /pci@1d,4000/fibre-channel@1/mplbt@10,0 (mplbt16))

SCSI transport failed: reason 'aborted' : XXXXXXXX

WARNING: Disk pathname (e.g. /pci@1d,4000/fibre-channel@1/mplbt@10,0 (mplbt16))

SCSI transport failed : reason ' timeout' : XXXXXXXX

A.1.2 SCSI Sense Related Messages

The following messages are not displayed for this version.

**WARNING: Disk pathname (e.g. /pci@1d,4000/fibre-channel@1/mplbt@10,0 (mplbt16))
information: giving up**

**WARNING: Disk pathname (e.g. /pci@1d,4000/fibre-channel@1/mplbt@10,0 (mplbt16))
FUJITSU XXXXX SK: 3,ASC: 0xB,ASCQ: 0xC,FRU: 0xD
information: Medium error**

**WARNING: Disk pathname (e.g. /pci@1d,4000/fibre-channel@1/mplbt@10,0 (mplbt16))
FUJITSU XXXXX SK: 4,ASC: 0xf1,ASCQ: 0xC,FRU: 0xD
information: FRU failed**

**WARNING: Disk pathname (e.g. /pci@1d,4000/fibre-channel@1/mplbt@10,0 (mplbt16))
FUJITSU XXXXX SK: 4,ASC: 0xf2,ASCQ: 0xC,FRU: 0xD
information: FRU failed**

**WARNING: Disk pathname (e.g. /pci@1d,4000/fibre-channel@1/mplbt@10,0 (mplbt16))
FUJITSU XXXXX SK: 4,ASC: 0xB,ASCQ: 0xC,FRU: 0xD
information: Hardware error
retry count=X,block no: Y**

**WARNING: Disk pathname (e.g. /pci@1d,4000/fibre-channel@1/mplbt@10,0 (mplbt16))
FUJITSU XXXXX SK: 6,ASC: 0x11,ASCQ: 0x90,FRU: 0xA
While rebuild processing, fill the broken data into some sectors of disk.**

**WARNING: Disk pathname (e.g. /pci@1d,4000/fibre-channel@1/mplbt@10,0 (mplbt16))
FUJITSU XXXXX SK: 6,ASC: 0xf2,ASCQ: 0x88,FRU: 0xA
Finish rebuild processing, but fill the broken data into some sectors of disk.**

**WARNING: Disk pathname (e.g. /pci@1d,4000/fibre-channel@1/mplbt@10,0 (mplbt16))
FUJITSU XXXXX SK: 6,ASC: 0xf2,ASCQ: 0x8x,FRU: 0xXX
information: Operating condition changed<Z=0xXX>**

WARNING: Disk pathname (e.g. /pci@1d,4000/fibre-channel@1/mplbt@10,0 (mplbt16))
FUJITSU XXXXX SK: 6, ASC: 0xa9, ASCQ: 0x0, FRU: 0xXX

WARNING: Disk pathname (e.g. /pci@1d,4000/fibre-channel@1/mplbt@10,0 (mplbt16))
Error for Command: XXXX Error Level: A
Requested Block: B Error Block: C
Vendor: FUJITSU Serial Number: D
Sense Key: Illegal Request
ASC: 0xE(xxxxxx),ASCQ: 0xF,FRU: 0xG

WARNING: Disk pathname (e.g. /pci@1d,4000/fibre-channel@1/mplbt@10,0 (mplbt16))
Error for Command: XXXX Error Level: A
Requested Block: B ErrorBlock: C
Vendor: FUJITSU Serial Number: D
Sense Key: yyyyyy
ASC: 0xE(xxxxxx),ASCQ: 0xF,FRU: 0xG

WARNING: Disk pathname (e.g. /pci@1d,4000/fibre-channel@1/mplbt@10,0 (mplbt16))
FUJITSU XXXXX SK: 0xb,ASC: 0xB,ASCQ: 0xC,FRU: 0xD
information: Aborted command

WARNING: Disk pathname (e.g. /pci@1d,4000/fibre-channel@1/mplbt@10,0 (mplbt16))
device busy

WARNING: Disk pathname (e.g. /pci@1d,4000/fibre-channel@1/mplbt@10,0 (mplbt16))
queue full

WARNING: Disk pathname (e.g. /pci@1d,4000/fibre-channel@1/mplbt@10,0 (mplbt16))
unknown scsi status: scsi_status = A

WARNING: Disk pathname (e.g. /pci@1d,4000/fibre-channel@1/mplbt@10,0 (mplbt16))
SCSI transport failed: reason 'timeout unit attention': giving up

**WARNING: Disk pathname (e.g. /pci@1d,4000/fibre-channel@1/mplbt@10,0 (mplbt16))
information: can't get sense: giving up**

**WARNING: Disk pathname (e.g. /pci@1d,4000/fibre-channel@1/mplbt@10,0 (mplbt16))
FUJITSU XXXXX SK: 6, ASC: 0xfb, ASCQ: 0x8x, FRU: 0xXX**

**WARNING: Disk pathname (e.g. /pci@1d,4000/fibre-channel@1/mplbt@10,0 (mplbt16))
incomplete uscsi cdb=0xXX - giving up**

A.2 Path Control Error Messages

This section explains the error messages displayed by the driver that are related to the path control. These messages are displayed on the console and logged in the /var/adm/messages log file.

**NOTICE: mplb*: I/O path failed. %s
remaining online path number=%d**

Cause:

The path is blocked because a physical path error has been detected.
%s indicates the physical path name (e.g. "/pci@1f,4000/fibre-channel@0/fp@0,0/
ssd@w2149000b5d6a0299,0").
%d indicates the number of remaining online paths that can be used.

Action:

Perform the required maintenance.

**NOTICE: mplb*: I/O Lun degraded. %s
remaining online path number of this lun =%d**

Cause:

Path access has been blocked for the LUN for which a physical path error has been detected.
%s indicates the physical path name (e.g. "/pci@1f,4000/fibre-channel@0/fp@0,0/
ssd@w2149000b5d6a0299,0").
%d indicates the number of remaining online paths that can be used.

Action:

Perform the required maintenance.

WARNING: mplb*: I/O path failed, no more online path.

Cause:

The path is blocked because a physical path error has been detected. Since this was the last online path, all paths are closed and all I/O processing has stopped.

Action:

Contact a Fujitsu engineer.

WARNING: mplb*: I/O Lun degraded, no more online path of this lun.

Cause:

Path access has been blocked for the LUN for which a physical path error has been detected. Since that was the last online path for that LUN, all paths are closed and I/O processing has stopped.

Action:

Contact a Fujitsu engineer.

NOTICE: mplb*: Use of I/O path was restarted. %s

Cause:

Since the physical path has returned to a normal state, the path is incorporated.
%s indicates the physical path name (e.g. "/pci@1f,4000/fibre-channel@0/fp@0,0/ssd@w2149000b5d6a0299,0").

NOTICE: mplb*: Use of I/O Lun was restarted. %s

Cause:

Since the physical path of a specific LUN has returned to a normal state, the path is incorporated.
%s indicates the physical path name (e.g. "/pci@1f,4000/fibre-channel@0/fp@0,0/ssd@w2149000b5d6a0299,0").

NOTICE: mplb*: Use of I/O path was restarted from STOP status. %s

Cause:

As there are no other paths available for use, the indicated path shown as %s has been restored from the stopped to the online state and reused.
%s indicates the physical path name (e.g. "/pci@1f,4000/fibre-channel@0/fp@0,0/ssd@w2149000b5d6a0299,0").

Action:

The path that was taken offline by the "iompadm change" command may not be the same path as the path on which the maintenance replacement will be performed.

WARNING: storage system name (e.g.:ETERNUS DX400- 10017C) (mplb*):

Device configuration is without redundancy.

Cause:

During path recognition, a non-redundant path configuration was detected for the multipath device specified in "storage system name".

- While the defined multipath configuration should have at least two independent paths, only one path was detected.
- In a configuration with two or more independent paths, all detected paths connect to ports on the same ROUTER or CM.

Action:

Check the connections between the server and the storage system.

WARNING: mplb*: forcible path degrade. %s

Cause:

Since the check results were unstable and found to change between normal and error, good results could not be obtained for the path shown as %s, and it has been excluded from the check.
%s indicates the physical path name (e.g. "/pci@1f,4000/fibre-channel@0/fp@0,0/ssd@w2149000b5d6a0299,0").

Action:

This path cannot auto-recover as a result of the test.
Perform the maintenance replacement, then use the "iompadm" command to bring the path online.

NOTICE: mplb*: There is no redundancy of ETERNUS concurrent firmware loading.

All paths connect to (%s).

Cause:

There is no storage system redundancy to perform a CFL. %s indicates the connected CM or ROUTER number. Further information about the relationship between CM or ROUTER numbers and physical locations in the storage system is available on the following website:
<http://www.fujitsu.com/global/products/computing/storage/software/eternus-mpd/>

Action:

The storage system configuration must be changed to perform firmware loading concurrently. Remove the target storage system. Change the cable connections so that storage system redundancy will be maintained, and then recreate the multipath using the "grmpdautoconf" command.

WARNING: Disk pathname (e.g. /pci@1f,4000/fibre-channel@0/fp@0,0/ssd@w2149000b5d6a0299,0) this path was unrecognizable. -> (%s)

Cause:

The path indicated by "Disk pathname" could not be recognized at server startup.
%s indicates the disk_device_controller_id of the target storage system. For details of this information, refer to ["3.1.1 "status" Command" \(page 47\)](#).

Action:

Check the status of the storage system and of the connections between the path and the storage system.

A.3 "iompadm" Command Error Messages

This section explains the iompadm command error messages.

iompadm: cannot initialize library: Permission Denied

Cause:

Initialization failed because a user without super user privilege tried to use the command.

Action:

Start the command with super user privilege.

iompadm: cannot initialize library: No Memory

Cause:

Initialization failed due to lack of memory.

Action:

Allocate sufficient memory and retry the command.

iompadm: Too many classes specified: Invalid Arguments

Cause:

A class was specified more than once.

Action:

Check the command format.

iompadm: invalid command: Invalid Arguments

Cause:

Subcommand name is incorrect.

Action:

Check the command format.

iompadm: -%s: Invalid Arguments

Cause:

Argument is incorrect.

Action:

Check the command format.

iompadm: cannot initialize library: Invalid Path

Cause:

No plug-in library could be found. There is no valid plug-in library.
All plug-in library initialization attempts failed.

Action:

Contact a Fujitsu engineer.

iompadm: <subcommand>: Invalid Arguments

Cause:

There is an incorrect subcommand argument.

Action:

Check the subcommand format.

iompadm: <subcommand>: No Memory

Cause:

The system ran out of memory while executing the subcommand.

Action:

Allocate sufficient memory and retry the command.

iompadm: <subcommand>: Invalid Path

Cause:

An invalid path exists in the device path specified by the argument.

Action:

Specify a valid device path. Also, check the status of connections between the server and the storage system and the hardware configuration of the storage system.

iompadm: <subcommand>: Path Busy

Cause:

The argument specified in the device path is a path in use.
Or, a path under issue of a SCSI reserve command exists in the argument specified instance.

Action:

Check that the device path is not being used.
Check whether a SCSI reserve command has been issued from the application for the argument specified Instance.

iompadm: <subcommand>: Too Many Path

Cause:

The number of device paths specified by the argument exceeds the number that can be set for the class.

Action:

Reduce the number of paths.

iompadm: <subcommand>: Not Implemented

Cause:

A class cannot use the subcommand.

Action:

Use a subcommand that is supported.

iompadm: <subcommand>: Not Supported

Cause:

A transition not supported by the class was attempted.

Action:

Check the status.

iompadm: <subcommand>: IO Error

Cause:

A hardware failure was detected in a path that is included in the instance.

Action:

Check the hardware failure.

iompadm: <subcommand>: Invalid Instance

Cause:

This plug-in does not support the specified subcommand.

Action:

Use a subcommand that is supported.

iompadm: <subcommand>: Internal Error

Cause:

The configuration information is corrupted.

Action:

Contact a Fujitsu engineer.

iompadm: <string>: Invalid Instance

Cause:

The character string is not a valid instance name.

The character string is not an instance where a class specified with "-c" manages.

Action:

Enter a valid instance name that supports the class name.

iompadm: <string>: Busy

Cause:

The device referred to in this character string is in use. If the above is not the case, and <string> is <change>, due to the presence of an instance, which will cause all paths to close if the path where a hot swap is to be made is stopped, a hot swap cannot be carried out.

Or, a path under issue of a SCSI reserve command exists in the argument specified instance.

Action:

Re-execute the command for the specified device in the character string when the other process using the device is complete.

Or instead of performing a hot swap for a change case, perform a cold swap.

Check whether a SCSI reserve command has been issued from the application for the argument specified Instance.

iompadm: <string>: Class not Found

Cause:

The character string specified by -c is not supported.

Action:

Specify the correct class name.

iompadm: <string>: Invalid Path

Cause:

The character string does not indicate a valid device.
The class does not support the device that is indicated in the character string.

Action:

Check the device name and specify a device supported by the class.

iompadm: <string>: No Memory

Cause:

The system ran out of memory while checking the character string "device".

Action:

Allocate sufficient memory and retry the command.

iompadm: <string>: Invalid Arguments

Cause:

The argument character string is not valid.

Action:

Check the command format.

iompadm: <string>: IO Error

Cause:

A failure was detected in a device.

Action:

Check the hardware failure.

iompadm: no more online path. continue (yes/no)

Cause:

If the path specified by the "iompadm -c mplb change" command is cut-off, there will be no more online paths available.

Action:

Enter "yes" to continue, or "no" to cancel.

IOMPMSG : <string> (<num>-<num>)

Cause:

This is displayed with other error messages. This message is used to investigate the malfunction.

Action:

Refer to this message to deal with the other error messages that were displayed simultaneously.

A.4 Cluster Error Messages

mplbclxec: invalid arg

Cause:

Parameter specification is incorrect.

Action:

Consult with your Fujitsu engineer if this error message is displayed during execution of the "iompadm" command.

mplbclxec: cannot get self node: %d

Cause:

Self-node information could not be obtained.
There may be a problem in the PRIMECLUSTER environment.
%d indicates an error code.

Action:

Contact a Fujitsu engineer.

mplbclxec: cannot get all nodes: %d

Cause:

Information on node configuration could not be obtained.
There may be a problem in the PRIMECLUSTER environment.
%d indicates an error code.

Action:

Contact a Fujitsu engineer.

mplbclxec: cannot create request: %d

Cause:

Check for empty areas in the /tmp directory and see if there is anything in the /tmp/mplb path.
%d indicates an error code.

Action:

Provide free space in the /tmp directory or delete data in the /tmp/mplb path and retry the command.

mplbclxec: cannot execute other nodes: %d

Cause:

Request for command execution could not be extended to other nodes.
There may be a problem in the PRIMECLUSTER environment.
%d indicates an error code.

Action:

Contact a Fujitsu engineer.

mplbclexec: invalid response info: %node

Cause:

An error was detected in the response to the execution of a command from another node.
%node indicates the node where the error occurred.

Action:

Contact a Fujitsu engineer.

mplbclexec: detected error: %node

Cause:

The execution of a command failed in another node.
%node indicates the node where command execution failed.

Action:

Observe the instructions given in the error message for the executed command following this message.

mplbclexec: response timeout: %node

Cause:

Time-out occurred while waiting for a response from the execution of the command in another node.
The node has stopped.
%node indicates the node where the error occurred.

Action:

Node stoppage does not affect the execution of "iompadm" command change/restart.
If the node is running, check the system status of the node and contact a Fujitsu engineer if required.

mplbcl: not add procedure.

Cause:

The MPLB procedure could not be registered because the cluster system setting is incomplete.

Action:

This is not a problem if the cluster system is in the configuration process.
If this message is displayed after cluster system configuration, contact a Fujitsu engineer.

mplbcl: cannot add procedure.

Cause:

The cluster MPLB procedure could not be registered because the commands or procedures required for registering it could not be found.
There may be a problem in the system environment.

Action:

Contact a Fujitsu engineer.

mplbcl: error add procedure.

Cause:

Registration of the cluster MPLB procedure failed. There may be a problem in the PRIMECLUSTER environment.

Action:

This message may appear when all the cluster nodes are restarted for the first time after the initial cluster configuration has been set up. If this message is no longer output after restarting the nodes, no further action is required. If the message appears again, contact a Fujitsu engineer.

mplbclproc: cannot get self node: %d

Cause:

A remote execution on the node failed because the self-node information could not be obtained. There may be a problem in the PRIMECLUSTER environment. %d indicates an error code.

Action:

Contact a Fujitsu engineer.

mplbclproc: invalid request proc.

Cause:

An error was detected in information from the source of request for remote execution between nodes.

Action:

Contact a Fujitsu engineer.

mplbclproc: cannot send response: %d

Cause:

The execution result could not be returned to the source of the request for remote execution between nodes. There may be a problem in the PRIMECLUSTER environment. %d indicates an error code.

Action:

Contact a Fujitsu engineer.

A.5 DR Script Error Messages

This section describes the error messages that are displayed relative to the DR Script. The DR script is a script that is provided by the Multipath Driver to work together with Dynamic Reconfiguration.

MPLB iompadm %s1 command was abnormal end.INSTANCE=%s2 device=%s3

Cause:

The "iompadm" subcommand (%s1) aborted. The specified parameters were: Instance administration device node %s2 and physical device node %s3.

Action:

Check the status of the physical device node that configures the instance administration device node. Check if the symbol link file on the /dev side for each device node exists.

MPLB all path become offline or warning. Do you continue?

Cause:

The target path to disconnect is the last access path.

Action:

Enter "yes" to continue, or "no" to cancel. Note that if "yes" is selected, all paths will be closed and it will no longer be possible to access the storage system.

MPLB all path is in same SB.Do you continue?

Cause:

The access path has been configured redundantly in the same system board.

Action:

Enter "yes" to continue, or "no" to cancel.
Note that if "yes" is selected, all paths will be closed and it will no longer be possible to access the storage system.

A.6 "grmpdautoconf" Command Error Messages

grmpdautoconf:ERROR:018:The object to be removed is still connected or the path is in the stop status.

At removing the LUN, first of all, please remove the LUN on the disk array device.

At removing the path or removing the device, first of all, please deal as follows.

- 1. Please change the path status by the restart command when the path is in the stop status.**
- 2. Please remove the path by the methods such as setting the zoning or physically pulling out the cable.**

Cause:

The logical unit or path to be removed is still connected, or the path is in a stop state. Logical units and paths may not be removed if they are still recognizable by the server. Neither may stop state paths be targeted for deletion.

Action:

LUNs should be removed on the storage system side.
Before removing a path or storage system, use the "restart" command to change the path status if a stop state path exists. After that, perform the removal by changing the Zoning settings or disconnecting all the cables between storage system being removed and the server.

grmpdautoconf:ERROR 023:There is offline path or LUN.

- **Please execute it again after making the state online.**
- **First of all, please execute grmpdautoconf -d if you want to remove.**

Cause:

An offline path or LUN was found.

Action:

Bring all statuses online then execute the command. When performing both addition and removal, perform removal first.

grmpdautoconf:ERROR 904:System command error (mkdir).

Cause:

The user command in the parentheses (), or maintenance command failed.
Since the error message of the respective command will be output, refer to the Solaris manual.

Action:

Refer to the Solaris manual to troubleshoot the problem.

Appendix B

Additional Information

This appendix provides additional information about the Multipath Driver.

B.1 How to Change mplb Instance Numbers

The mplb instance number is the "mplb*" part of the "/dev/FJSMplb/[r]dsk/mplb*s*" special file name format.

This section describes how to change the instance numbers of an existing mplb special file to an arbitrary number. To change the instance numbers, the -o and -f options of the "mplbconfig" command are used to specify the mplb special file hereinafter referred to as a "configuration file". By editing the configuration file, the instance number can be changed to any number. The procedure is as follows:

Caution

The mplb instance numbers of mplb instances registered as PRIMECLUSTER resources must not be changed, or PRIMECLUSTER may not operate normally.

Procedure

- 1 Stop all applications that might access the mplb special file whose instance number is to be changed.
If the mplb special file has been mounted, unmount it. Any /etc/vfstab entries must also be deleted.
- 2 Execute the "mplbconfig" command with the -o option to output the current configuration information to a configuration file.

The configuration file is created with whatever name is specified.

(Example) Outputting the current configuration information to /tmp/mplb_def

```
# /usr/sbin/mplbconfig -o /tmp/mplb_def
*** Phase 1: checking /dev/rdsk ***
*** Phase 2: checking multi path ***
*** Phase 3: analyzing path ***
=== Reconfigure plan ===
Multi path : 2
Create new multi path : 0
Add path for multi path : 0 (instance) : 0 (add path)
Delete path from multi path : 0 (instance) : 0 (delete path)
#
```

- 3 Edit the configuration file. Change the existing mplb instance numbers as desired.
(Example) Changing existing instance numbers mplb0 and mplb1 to mplb10 and mplb11 (respectively). Edit the strings of the "Path" column in the configuration file, as there indicate the instance numbers.

```
# vi /tmp/mplb_def
*** mplb config file ***
Path      : Action : Element path                               : LUN : Storage
mplb0     : exist  : c13t2148000B5D6A0299d0s2 c14t2149000B5D6A0299d0s2 : 0 :
E4000-    6A0299
mplb1     : exist  : c13t2148000B5D6A0299d1s2 c14t2149000B5D6A0299d1s2 : 1 :
E4000-    6A0299
```



```
*** mplb config file ***
Path      : Action : Element path                               : LUN : Storage
Mplb10    : exist  : c13t2148000B5D6A0299d0s2 c14t2149000B5D6A0299d0s2 : 0 :
E4000-    6A0299
mplb11    : exist  : c13t2148000B5D6A0299d1s2 c14t2149000B5D6A0299d1s2 : 1 :
E4000-    6A0299
```

- 4 Execute the "mplbconfig" command with the -f option specifying the edited configuration file in order to change the mplb instance numbers. The mplb instance numbers change according to the configuration file that is specified by the -f option.

(Example) Changing the mplb instance numbers using the /tmp/mplb_def file

```
# /usr/sbin/mplbconfig -f /tmp/mplb_def
*** Phase 1: checking /dev/rdsk ***
*** Phase 2: checking multi path ***
*** Phase 3: analyzing path ***
*** Phase 3.1: update plan ***
mplb10    : exist  : c13t2148000B5D6A0299d0s2 c14t2149000B5D6A0299d0s2 : 0 :
E4000-    6A0299
mplb11    : exist  : c13t2148000B5D6A0299d1s2 c14t2149000B5D6A0299d1s2 : 1 :
E4000-    6A0299
=== Reconfigure plan ===
Multi path : 2
Create new multi path : 0
Add path for multi path : 0 (instance) : 0 (add path)
Delete path from multi path : 0 (instance) : 0 (delete path)
```

- 5 Execute the "iompadm info" command to confirm that the instance numbers have changed as specified.

```
# /usr/opt/FJSViomp/bin/iompadm -c mplb info
```

End of procedure

Caution

The "mplbconfig" command should only be used to change the mplb instance numbers or as instructed by your Fujitsu engineer.

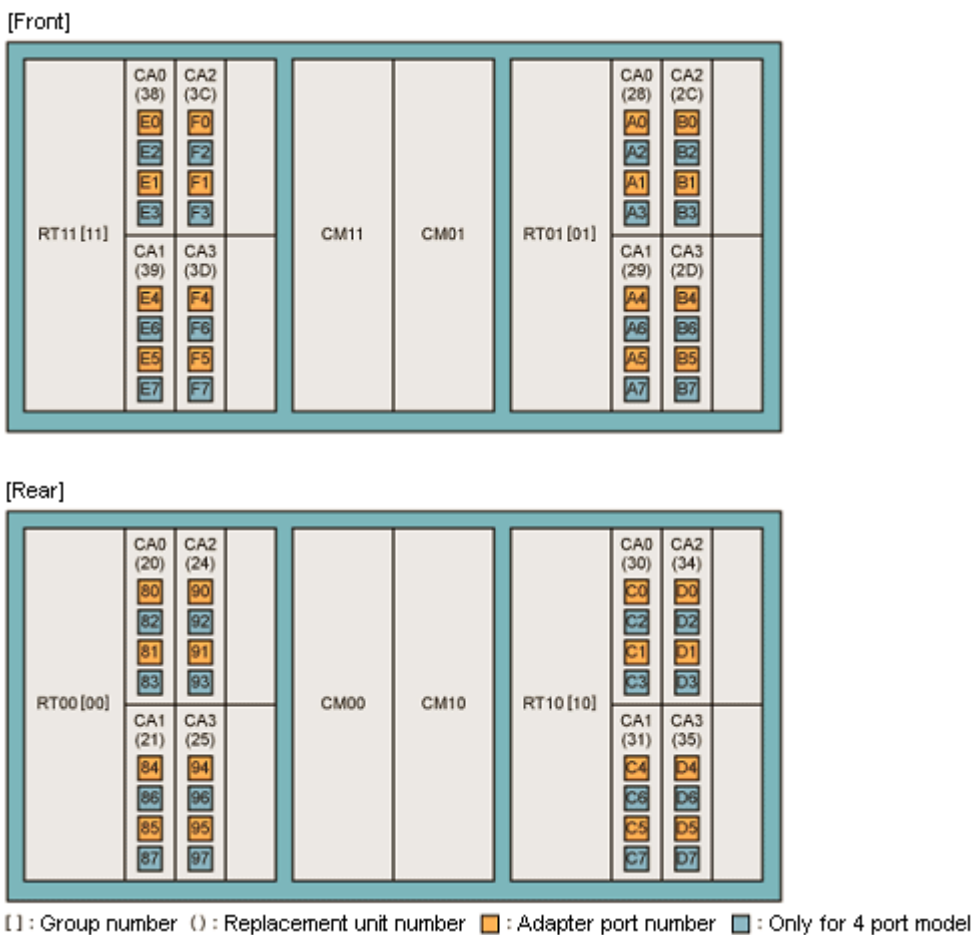
B.2 Storage System Redundancy during CFL

This section explains storage system redundancy during CFL.

B.2.1 Storage System Redundancy during a CFL for the ETERNUS6000

The ETERNUS6000 has the following configuration, with up to four ROUTERS (RT) installed.

Figure B.1 Redundant configuration during CFL



In this section, RT00 and RT01 are defined as Group 1, and RT11 and RT10 are defined as Group 2. The CFL process for the ETERNUS6000 is as follows:

- (1) The firmware installed in Group 1 (RT00 and RT01) is to be updated. Note that server access via Group 1 is temporarily disabled while the firmware is being updated.
- (2) After the firmware update for Group 1 has completed, the firmware installed in Group 2 (RT11 and RT10) is updated. Note that server access via Group 2 is temporarily disabled while the firmware is being updated.

 **Caution**

Depending on the device state of the ETERNUS6000, the firmware installed in Group 2 may be updated first.

Therefore, when the server and the ETERNUS6000 are connected in one of the following multipath configurations, access from applications will be blocked during CFL:

- All cables connected to Group 1 only
- All cables connected to Group 2 only

Since they are not redundant during CFL, if one of the above configurations is set up using the "grmpdautoconf" command, the warning message described in ["2.2.2.1 Configuring the Multipath Driver and Creating the Configuration Definition File" \(page 42\)](#) will be output during CFL.

To prevent this, the connections between the server and the ETERNUS6000 should be changed to ensure the paths are redundant during CFL. ["Reconfiguration for Path Redundancy during CFL" \(page 104\)](#) describes the necessary steps.

 **Caution**

- The steps described in ["Reconfiguration for Path Redundancy during CFL" \(page 104\)](#) are not required when cold swapping is used to replace the firmware.
- Which connection configurations provide path redundancy for CFL varies depending on the storage system model. For details of each storage system model, refer to its manuals.

■ Reconfiguration for Path Redundancy during CFL

Use the following procedure to change the connections between the server and storage system:

Procedure

- 1 Take the target storage system offline, using the procedure described in ["4.5 Removing a Storage System" \(page 61\)](#), then shut down the server.
- 2 Change the connections between the server and storage system to ensure path redundancy for CFL.
- 3 Restart the server.
- 4 Execute the "grmpdautoconf" command to configure the multipath.

End of procedure

B.2.2 Storage System Redundancy during a CFL for Storage Systems other than the ETERNUS6000

In the following connection cases, all paths would be closed during the CFL operation:

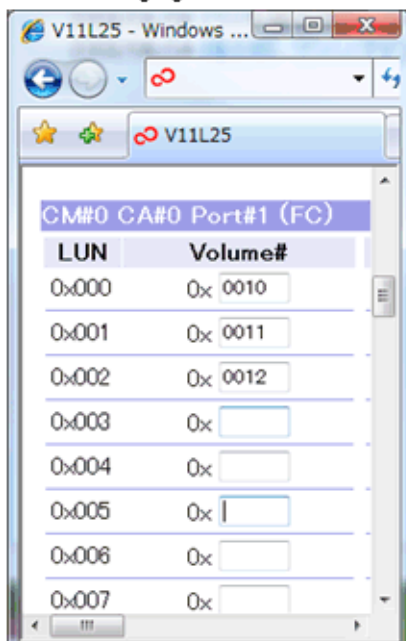
- ETERNUS DX8700 S3/DX8900 S3
When the server is connected only to CMs in group number "2*", or only to CMs in group number "1*".
- Storage systems other than the above
When the server is connected to even-numbered (group number) CMs only, or odd-numbered CMs only.

B.3 Changing a LUN Mapping or Host Affinity

To change the mapping or host affinity for the recognized LUNs after the multipath environment has been configured, the multipaths will need to be released and reconfigured. This section explains how to change ETERNUS4000 LUN mapping as an example. For situations other than this example, the same procedure is also required to change the LUN mappings. Refer to the storage system manuals for details on how to change a LUN mapping or host affinity.

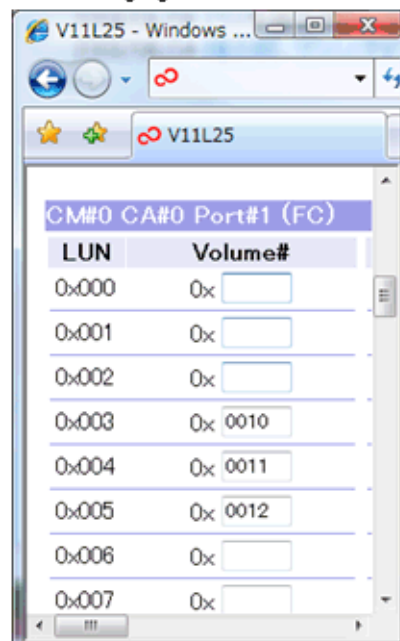
Example: Changing the LUN number from LUN0x000 – LUN0x002 to LUN0x003 – LUN0x005

Before changing



CM#0	CA#0	Port#1 (FC)
LUN	Volume#	
0x000	0x	0010
0x001	0x	0011
0x002	0x	0012
0x003	0x	
0x004	0x	
0x005	0x	
0x006	0x	
0x007	0x	

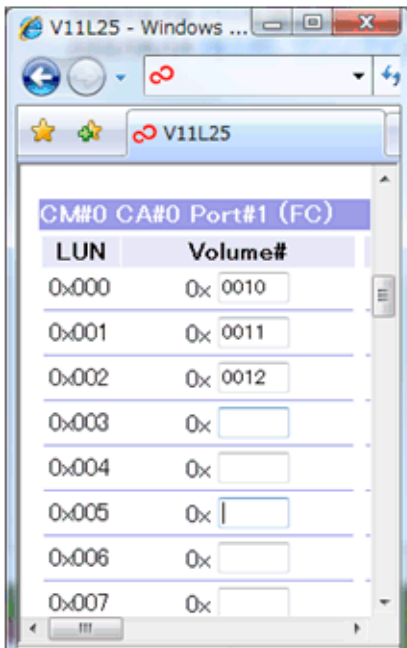
After changing



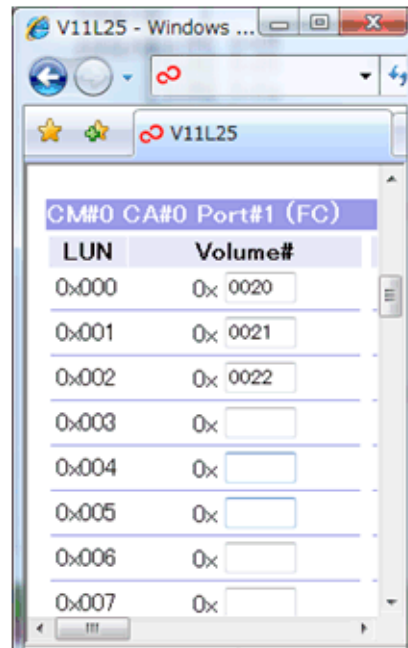
CM#0	CA#0	Port#1 (FC)
LUN	Volume#	
0x000	0x	
0x001	0x	
0x002	0x	
0x003	0x	0010
0x004	0x	0011
0x005	0x	0012
0x006	0x	
0x007	0x	

Example: Changing the mapping for LUN0x000 – LUN0x002 from 0x0010 – 0x0012 to 0x0020 – 0x0022

Before changing



After changing



The LUN for the target paths must be removed from the configuration and the multipath be reconfigured. Perform the following procedure.

Procedure

- 1 Delete the LUN mappings and host affinity settings of the target paths.
- 2 Refer to ["4.3 Removing a Logical Unit \(LU\)" \(page 56\)](#) to remove logical units with deleted configurations.
- 3 Add LUN mappings and host affinity settings after modifications.
- 4 Execute the "grmpdautoconf" command to configure the multipaths.

End of procedure

Caution

If the server is restarted after the multipaths are reconfigured but before the LUN is removed, paths whose settings have changed will be set to an "offline fail" state. If the settings for all paths connected to the storage system are changed, the following message is output:

```
WARNING: Disk_pathname  
this path was unrecognizable. -> (%s)
```

B.4 Troubleshooting

This section explains how to acquire troubleshooting information after occurrence of a Multipath Driver error.

If a Multipath Driver error occurs, collect the required diagnostic information, then contact your customer support representative. For a cluster system, collect the necessary information for all nodes in the cluster. The following two procedures can be used to collect this information:

- "mplbsnap" command
- "fjsnap" command

■ "mplbsnap" command

Only the Multipath Driver related information is collected with this command. Use the following procedure to collect the information.

Procedure

- 1 Execute the "mplbsnap.sh" command with the -a option.

```
#/opt/FJVSvmplb/bin/mplbsnap.sh -a
```

An output directory can be specified for the collected information.

```
#/opt/FJVSvmplb/bin/mplbsnap.sh -a -d /export/home/
```

The Multipath Driver related information is saved in a "mplbsnap.tar.Z" file. If another directory is not specified, this is output to the "/var/tmp" directory. If another directory is specified, the collected information file is output there.

End of procedure

■ "fjsnap" command

The "fjsnap" command is a Solaris system information tool that is included with the FJVSsnap extended support function package. This command collects both the Multipath Driver related information and unrelated information. For details of the "fjsnap" command, refer to the manuals provided with the FJVSsnap package.

B.5 Cluster System

This section provides a workflow and cautions when using a multipath configuration in a cluster system.

Workflow

1

HBA is configured in each node (or node A and node B), and the LUNs are in a recognized state via the `ssd` driver.



- ["2.2 Environment Configuration" \(page 39\)](#)



2

Automatic configuration of the cluster system is performed.

The nodes are registered as cluster resources when the multipath settings are complete.



3

In each node (or node A and node B), change the limit on the number of I/Os that can be issued simultaneously as required.

Caution

- The `grmpdautoconf` command is not used for cluster systems. The multipath is configured during the automatic configuration of the cluster.
- The controller number (`c*`) of the special files name is automatically assigned by Solaris OS. Therefore, even if under a cluster configuration, the special files names may differ between the nodes. These numbers cannot be changed. When PRIMECLUSTER is used, it is not an error that the `"c*t*d*s"` part of the special files names are different for different nodes.
- Delete any resources that are registered in the cluster system before removing or changing a storage system or a LUN.
- In a cluster system environment, if the `grmpdautoconf` command is executed by mistake, the multipath will be configured. In this case, cancel the incorrectly set multipath as detailed in ["4.3 Removing a Logical Unit \(LU\)" \(page 56\)](#) and ["4.5 Removing a Storage System" \(page 61\)](#), set the target logical unit to allow it to be recognized by the server, and then execute the `grmpdautoconf` command again to redo the multipath settings.
- When local resources (unshared disks) are added to the cluster system, refer to ["B.6 Adding Local Resources to a Cluster System" \(page 109\)](#).

B.6 Adding Local Resources to a Cluster System

When local devices are added to the nodes of a PRIMECLUSTER cluster system, make sure that each node uses unique mplb instance numbers for its local resources. Normally the PRIMECLUSTER cluster system automatically creates different instance numbers during automatic configuration, but if "grmpdautoconf" is implemented there is a chance that the instance numbers will be identical.

For example, in the following configuration both nodes (A and B) have a "mplb0" local resource.

Cluster node A	Cluster node B
mplb2048 ---> LUN_V 0x000	mplb2048 ---> LUN_V 0x000
mplb2049 ---> LUN_V 0x001	mplb2049 ---> LUN_V 0x001
mplb0 ---> LUN_V 0x002	mplb0 ---> LUN_V 0x003

This type of configuration is not supported, so the instance number "mplb0" must be changed on one of the nodes.

Cluster node A	Cluster node B
mplb2048 ---> LUN_V 0x000	mplb2048 ---> LUN_V 0x000
mplb2049 ---> LUN_V 0x001	mplb2049 ---> LUN_V 0x001
mplb0 ---> LUN_V 0x002	mplb1 ---> LUN_V 0x003

The procedure is shown below:

Procedure

- 1 Add the local resources according to the procedure in ["Chapter 4 Adding or Removing Components" \(page 55\)](#).
- 2 Execute the "iompadm info" command on each cluster node, and check the local resource mplb instance numbers used by each cluster node.
If each node uses a different set of instance numbers, no further action is required.
- 3 If the same mplb instance number is used by multiple cluster nodes, refer to ["B.1 How to Change mplb Instance Numbers" \(page 101\)](#) and assign new, unique mplb instance numbers to negate any number overlaps.

End of procedure

B.7 Performing "add device" for mplb Special Files in Non-Global Zones

The procedure is shown below:

Procedure

- 1 Create a non-global zone (example: test-zone).
- 2 Create a directory for mplb under /dev/ in the non-global zone.

```
# mkdir /export/test-zone/dev/FJSVmplb
# mkdir /export/test-zone/dev/FJSVmplb/dsk
# mkdir /export/test-zone/dev/FJSVmplb/rdisk
```

- 3 Select the target multipath device from the test-zone, and check the device major and minor numbers.

The following example is for when "mplb0s0" is added:

```
# ls -l /dev/FJSVmplb/rdisk/mplb0s0
lrwxrwxrwx  1 root  root           36 Aug 28  20:28 /dev/FJSVmplb/rdisk/
mplb
0s0 -> ../../../../devices/pseudo/mplb@0:a,raw

# ls -l /devices/pseudo/mplb@0:a,raw
crw-r-----  1 root  sys      253, 0 Aug 30  04:19 /devices/pseudo/
mplb@0:
a,raw

# ls -l /devices/pseudo/mplb@0:a
brw-r-----  1 root  sys      253, 0 Aug 30  04:19 /devices/pseudo/
mplb@0:a
```

The underlined parts are the major and minor numbers.

- 4 Create a special file under /export/test-zone/dev/FJSVmplb using the "mknod" command.

```
# mknod /export/test-zone/dev/FJSVmplb/rdisk/mplb0s0 c 253 0
# mknod /export/test-zone/dev/FJSVmplb/dsk/mplb0s0 b 253 0
```

- 5 Log in to the non-global zone and then access the mplb device.

End of procedure

B.8 PCI Hot Plug

This section explains how to perform the necessary commands and procedures for use with the Multipath Driver.

Refer to the manual for the server or Oracle Solaris for the procedure to replace HBAs using the PCI Hot Plug.

Procedure

1 Disconnect from the Multipath Driver.

Disconnect the path of the replacement target from the Multipath Driver with the "iompadm" command.

Example: If the path of the replacement target is /dev/rdisk/c11t2140000B5D6A0299d0s2

```
# /usr/opt/FJSViomp/bin/iompadm -c mplb change adapter_disconnect  
/dev/rdisk/c11t2140000B5D6A0299d0s2
```

After executing the command, confirm that the path of the replacement target is in an "unconfigured disconnected" state using the "iompadm info" command.

2 Disconnect the PCI card.

Disconnect with the "cfgadm" command.

For directions on how to use the "cfgadm" command, refer to the manual for the server or the manual for Oracle Solaris.

3 Replace the PCI card.

Replace the disconnected PCI card with the replacement part. This work is performed by a Fujitsu engineer.

4 If necessary, change the HBA WWN on the Host Affinity and Fibre Channel switch zone on the storage system after replacement.

5 Install the PCI card.

Install the card with the "cfgadm" command.

For directions on how to use the "cfgadm" command, refer to the manual for the server or the manual for Oracle Solaris.

6 Connect to the Multipath Driver.

Connect the replaced path to the Multipath Driver using the "iompadm" command.

```
# /usr/opt/FJSViomp/bin/iompadm -c mplb restart adapter_connect  
/dev/rdisk/c11t2140000B5D6A0299d0s2
```

After executing the command, confirm that the replaced path is in an "online" state using the "iompadm info" command.

End of procedure

B.9 OVM Guest Domains

Multipath operations and specifications in OVM guest domain environments differ from those in normal environments.

This section provides notes and specifications of the multipath.

■ Configuration notes

- The Multipath Driver must be configured on both I/O domains and guest domains.
- For a single path configuration of the I/O domain, each I/O domain must be connected to a CA on different CMs. If all I/O domains are connected to CAs on the same CM, redundancy is not guaranteed.
- When a multipath device is used in an I/O domain, single path configurations are not supported.
- A mix of single path configurations and multipath configurations in a single I/O domain is not supported.
- For guest domains, a mixed configuration where virtual disks of I/O domains in a single path configuration and virtual disks of I/O domains in a multipath configuration exist together is not supported.
- The virtual disk that is supported by the Multipath Driver of the guest domain is the only disk that the multipath device is assigned as a full disk on the I/O domain. Virtual disks that are assigned to devices other than the multipath (for example, `ssd`, `MPxIO`, or `PRIMECLUSTER GDS` devices) and virtual disks that are assigned as slices are not supported.
- Multiple host disk control operation "mhd" is not supported.
- Notes and restrictions about virtual disks are OVM compliant. For details, refer to "Oracle VM Server for SPARC Administration Guide".
- When a path is added or removed from I/O domains and the configuration is changed (single path configuration -> multipath configuration), the guest domains must be rebooted.
- All I/Os in guest domains are processed by the I/O domains. When processes are delayed or hung up due to overload in the I/O domains, I/Os in the guest domains are also delayed or hung up. Design I/O domains by incorporating flexibility into the performance.
- While I/O domains are stopped, components cannot be removed from guest domains.

■ Operation list for guest domains

Operations such as switching a path differ between a single path configuration and a multipath configuration of I/O domain. In addition, some operations and supported functions are different for guest domains.

Item	Single path configuration for I/O domains		Multipath configuration for I/O domains	
	I/O domain	Guest domain	I/O domain	Guest domain
Path status of the storage system with CM not assigned	active	active/active	active/active	active/active
Path status of the storage system with CM assigned	active or standby	active/standby	active/standby	active/active
Operation when an error occurs on a physical path	An I/O error message is sent to the guest domain. The path status changes to "warning".	The path is switched. The path status changes to "warning".	The path is switched. Normal response with the new path. The path status changes to "warning".	Normal response. The path status does not change.
Operation when an error occurs in an I/O domain	—	The path is switched. I/Os are issued to other I/O domains.	—	The path is switched. I/Os are issued to other I/O domains.
"iompadm info" and "status" display	Same as normal	Same as normal	Same as normal	CM-CA-PORT information is not displayed. Example: [ETERNUS_DXL-6A0299]
"c/cu/g" specification for "iompadm change/restart"	Supported	Supported	Supported	Not supported
"c/cu/g" operation for "iompadm change/restart" in a cluster configuration	Must be performed on only one node	Must be performed on all nodes	Must be performed on only one node	—

■ Live Migration

Live Migration is supported.

▶ Caution

- When performing Live Migration, virtual disk assignments must be the same for the migration source and the migration destination.
- When performing Live Migration, the number of multipaths for I/O domains must be the same for the migration source and the migration destination.

■ SPARC M10-4S Dynamic Reconfiguration

SPARC M10-4S Dynamic Reconfiguration is supported.

The procedure differs depending on whether or not I/O domains are stopped.

- When I/O domains are stopped
 - Operating the Multipath Driver is not necessary.
- When I/O domains are running
 - Before deleting virtual disks

Procedure

- 1 When the SPARC M10-4S server is replaced using the Dynamic Reconfiguration function and I/O domains exist, change the status of the virtual disks in the guest domains to the state where disconnection is possible before deletion.
Execute the following command for all the guest domains that use the target I/O domain.
Example: When the target path for the Dynamic Reconfiguration replacement is /dev/rdisk/c2d0s2

```
# /usr/opt/FJSSViomp/bin/iompadm -c mplb change adapter_disconnect /dev/  
rdisk/c2d0s2
```

After executing the command, confirm that the state of the paths that are using the target I/O domain is "unconfigured disconnected" using the "iompadm info" command.

End of procedure

- After reassigning the virtual disks

Procedure

- 1 Execute the following command in the guest domains.

```
# /usr/opt/FJSSViomp/bin/iompadm -c mplb restart adapter_connect /dev/rdisk/  
c2d0s2
```

After executing the command, confirm that the state of the paths that are using the target I/O domain is "online" using the "iompadm info" command.

End of procedure

B.10 SE0X7F11F, SE0X7F12F

To use SE0X7F11F and SE0X7F12F, a driver must be defined when configuring a new multipath or when a component is added or removed.

Caution

- "fjpfca.conf" must be set manually.
- A definition must be manually added to "sd.conf" for a LUN to be recognized.
- When an error occurs, it takes more time to switch a path than other HBAs.
- OVM guest domains are not supported.
- SAN Boot is not supported.
- For Oracle Solaris 11, SE0X7F11F and SE0X7F12F cannot be used.

The procedure is shown below:

Procedure

- 1 Install the HBA driver.
- 2 Set "/kernel/drv/fjpfca.conf".
For details, refer to the respective HBA manual.
- 3 Set "/kernel/drv/sd.conf" and make the LUN be recognized in "sd".
For details, refer to the manuals such as for Oracle Solaris.
- 4 Execute the "grmpdautoconf" command to configure the multipath.

End of procedure

B.11 Storage Cluster

When storage systems are switched with the Storage Cluster function, the status of the paths may become "warning". I/O processing is taken over without any problems. The "warning" state returns back to "online" with a diagnosis.

A

Assigned CM

The controller module that has primary control of the LUs in a storage system with CM assigned. Going through the assigned CM provides faster LU access.

C

CA (Channel Adapter)

A storage system I/O adapter connected to the server device. The CA controls transmission and reception of data.

CM (Controller Module)

Controller Module in the storage system. The main module for controlling the storage system.

D

Device node

Device nodes are indicated by the following types of special files.

Instance administration device node:

(/dev/FJSMplb/fiomp/adm*)

User access device node:

(/dev/[r]dsk/c*t*d*s2, /dev/FJSMplb/[r]dsk/mplb*s*)

Physical device node:

(/dev/[r]dsk/c*t*d*s2,/devices/*..)

F

Fabric connection

One of the formats for connecting between the server and the storage, using Fibre Channel interface. Fabric connection allows simultaneous transmission between multiple nodes, by going through the Fibre Channel switch.

H**HBA (Host Bus Adapter)**

I/O adapter installed on the server. Connected to the storage system, controls transmission and reception of data.

HBAs such as SCSI HBAs, Fibre Channel HBAs, and SAS HBAs are available.

Hot expansion function

Adding a component into the system without affecting tasks.

I**Instance (No.)**

Units managed by the driver. For example, one instance in the MPLB driver is indicated by one LU number.

L**LU (Logical Unit)**

Unit representing logical disks that are recognized from the server.

LUN (Logical Unit Number)

A number assigned to a logical unit.

M**MPLB driver**

It is accessed from the special file `/dev/FJSMplb/[r]dsk/mplb*s*`.

N**Non-Assigned CM**

The controller module that has secondary control of the LUs in a storage system with CM assigned. Compared to an assigned CM, access to the LUs is slower.

P**PRIMECLUSTER**

Fujitsu's cluster system product. Can be used at the same time as the Multipath Driver. Configures a high availability system of multiple nodes that are connected in a dedicated network. High availability is achieved by taking over tasks through automatically switching from nodes that are down to other nodes.

R

Router

The module that controls the transmission interface between each of the modules (between CM-CA, for example) in a storage system with CM not assigned.

S

SAN (Storage Area Network)

High-speed network that is connected to the server and storage system. It is independent from the LAN, prevents degradation of LAN functions when large volumes of data are being handled.

SFP (Small Form Factor Pluggable)

An input and output module used to plug into a switch port, and connect the port to the network. Hot swapping is possible.

Special file

Interface to access the driver.
For example, `/dev/FJSMplb/[r]dsk/mplb*s*`, `/dev/[r]dsk/c*t*d*s2`, `/devices/...`

ssd driver

Standard Solaris OS SCSI disk driver.

Switch

An intermediary device that is necessary when configuring a network using Fibre Channel. It can run multiple clusters of data transmission simultaneously.

Z

Zoning function

Zones are divided according to Fibre Channel switch port names or World Wide Names (WWNs). This function limits access between separate zones.

FUJITSU Storage ETERNUS Multipath Driver 3.1 User's Guide
for Oracle Solaris

P2S0-0063-03ENZO

Date of issuance: July 2015
Issuance responsibility: FUJITSU LIMITED

- The content of this manual is subject to change without notice.
- This manual was prepared with the utmost attention to detail. However, Fujitsu shall assume no responsibility for any operational problems as the result of errors, omissions, or the use of information in this manual.
- Fujitsu assumes no liability for damages to third party copyrights or other rights arising from the use of any information in this manual.
- The content of this manual may not be reproduced or distributed in part or in its entirety without prior permission from Fujitsu.


FUJITSU