FUJITSU Software
PRIMECLUSTER
Global File Services

Configuration and Administration Guide 4.5

Linux

Preface

Global File Services is called "GFS" in the document.

This manual explains the functions, settings, and operations of GFS.

Target Readers

This manual is intended for all the users operating the products with the GFS Shared File System (e.g. PRIMECLUSTER.)

To read this manual, readers will require a general knowledge of UNIX and Linux(R) operating systems.

Because GFS Shared File System uses the functions of PRIMECLUSTER, readers will also need knowledge of cluster control as used in a PRIMECLUSTER system.

Because GFS Shared File System uses the functions of PRIMECLUSTER Global Disk Services, readers will also need knowledge of shared volume as used in a PRIMECLUSTER Global Disk Services.

Organization

This manual is organized as below.

Chapter 1 File System Functions

This chapter describes the functions and features of the GFS Shared File System.

Chapter 2 File System Structure

This chapter describes the structure of the GFS Shared File System.

Chapter 3 Failure Recovery

This chapter describes the failure recovery functions of the GFS Shared File System.

Chapter 4 File System Design

This chapter describes the main features of the operational design of the GFS Shared File System.

Chapter 5 Management Partition

This chapter describes the management partition.

Chapter 6 Starting and Exiting the Management View

This chapter describes the topology, preparation, and operation to use the GFS Management View.

Chapter 7 Operation Management View Screen Elements

This chapter describes the screen elements of the GFS Management View.

Chapter 8 Management Partition Operations (GUI)

This chapter describes how to operate the management partition by the GFS Management View.

Chapter 9 Management Partition Operations (Command)

This chapter describes how to operate the management partition by commands.

Chapter 10 File System Operations (GUI)

This chapter describes how to operate the GFS Shared File System by the GFS Management View.

Chapter 11 File System Operations (Command)

This chapter describes how to operate the GFS Shared File System by commands.

Chapter 12 File System Management

This chapter describes the procedures for managing the GFS Shared File System using basic commands.

Chapter 13 File System Backing-up and Restoring

This chapter describes how to backing up and restoring data in the GFS Shared File System.

Chapter 14 Tuning

This chapter describes how to use a variety of utilities to optimize and make effective use of the GFS Shared File System.

Chapter 15 Migration to the GFS Shared File System

This chapter describes how to migrate from existing file systems to the GFS Shared File Systems.

Appendix A List of Messages

This appendix describes GFS Shared File System messages.

Appendix B Reference Manual

This appendix describes management commands of the GFS Shared File System.

Appendix C Troubleshooting

This appendix describes how to collect troubleshooting information and take corrective action in the event of a GFS Shared File System failure.

Appendix D Startup Scripts and Startup Daemons, and Port Numbers in GFS [RHEL6]

This appendix provides the explanations on scripts and daemons that are started by GFS, and the port numbers being used.

Appendix E systemd Services and Startup Daemons, and Port Numbers in GFS [RHEL7]

This appendix provides the explanations on systemd services and startup daemons in GFS, and the port numbers being used.

Appendix F Release information

This appendix describes the main contents for change of this manual.

Glossary

The glossary defines the terms related to the GFS Shared File System.

Related documentation

Please refer to the following manuals if necessary:

- PRIMECLUSTER Concepts Guide

- PRIMECLUSTER Installation and Administration Guide

- PRIMECLUSTER Web-Based Admin View Operation Guide

- PRIMECLUSTER Cluster Foundation (CF) Configuration and Administration Guide

- PRIMECLUSTER Reliant Monitor Services (RMS) with Wizard Tools Configuration and Administration Guide

- PRIMECLUSTER Global Disk Services Configuration and Administration Guide

- PRIMECLUSTER Global Link Services Configuration and Administration Guide: Redundant Line Control Function

- PRIMECLUSTER Messages

- FJQSS (Information Collection Tool) User's Guide

Note

A related document of PRIMECLUSTER includes the following documents besides the above-mentioned manuals.

- PRIMECLUSTER Software Release Guide and Installation Guide
  The software release guide and installation guide appended to each product of PRIMECLUSTER.
  The data is stored on DVD of the product. For information about the file name, see "Product introduction."

Manual Printing

Use the PDF file to print this manual.

Adobe Reader is required to read and print this PDF file. To get Adobe Reader, see Adobe Systems Incorporated's website.

Online Manuals

To reference the online manual, use the Cluster management server to register the user name in user group wvroot, clroot, cladmin, or clmon.

📚 See
....
To enter a user group and meaning thereof, see "PRIMECLUSTER Web-Based Admin View Operation Guide."
....

Notational Conventions

Notation

Prompts

Command line examples that require system administrator (or root) rights to execute are preceded by the system administrator prompt, the hash sign (#). Command line examples that do not require system administrator rights are preceded by a dollar sign ($).

Manual page section numbers

Section No. of manual page

Section numbers of the manual page appear in brackets after the commands of Linux(R) operating system and PRIMECLUSTER. Example: cp(1)

The keyboard

Keystrokes that represent nonprintable characters are displayed as key icons such as [Enter] or [F1]. For example, [Enter] means press the key labeled *Enter*; [Cntl]+[B] means hold down the key labeled *Cntl* or *Control* and then press the [B] key.

Italic type

Italic types represent variables which need to be replaced by specific numerical values or character strings.

Description on OS

[RHEL6] represents information essential for using Red Hat Enterprise Linux 6.

[RHEL7] represents information essential for using Red Hat Enterprise Linux 7.

Symbol

Material of particular interest is preceded by the following symbols in this manual:

🅿 Point
....
Describes the contents of an important point.
....

📑 Note
....
Describes the points the users should pay close attention to.
....

📖 Information
....
Provides useful information related to the topic.
....

📚 See
....
Provides manuals for users' reference.
....

Trademarks

- UNIX is a registered trademark of The Open Group in the United States and other countries.

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

- Linux is a registered trademark of Linus Torvalds.

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

- Microsoft, Windows, and Internet Explorer are registered trademarks of Microsoft Corporation in the United States and other countries. All other trademarks are the property of their respective owners.

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

Copyright (c) 1983-1989 Portions may be derived from Berkeley BSD system, licensed from the U. of CA.

Export Controls

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

Acknowledgement

This product includes software developed by the University of California, Berkeley and its contributors.

Requests

- No part of this document may be reproduced or copied without permission of FUJITSU LIMITED.

- The contents of this document may be revised without prior notice.

Revision History

Revision	Location	Edition
Changed "Response" of the following message: ERROR: sfcfsrm:0005: can not load *module*	A.3.1 sfcfsrm script	2.1
Added notes on using the Mount namespace function.	1.8.26 Notes on using the Mount namespace function [RHEL7]	2.2

Contents

Chapter 1 File System Functions

This chapter describes the functions and features of the Global File Services (hereinafter GFS) Shared File System.

1.1 File System Overview

The GFS Shared File System allows simultaneous access to it from multiple Linux systems in an integrated manner. (A file system of this type is referred to as a shared file system. A file system such as ext3 that will be used only within a **node** is referred to as a local file system.)

With GFS Shared File System commercial systems can enjoy high reliability and high performance without changing applications code due to Application Program Interface (API) compatibility with ext2/ext3.

The GFS Shared File System maintains data consistency even if file data is updated from multiple nodes. So, applications on different nodes can deliver application data without changing applications code.

It is suitable for demands for high-availability in file systems due to non-disruptive file operation on another node even if one node fails.

The GFS Shared File System has the following functions:

- Simultaneous shared access from multiple nodes to files or file systems

- Maintaining consistency for **file data** referenced or updated from multiple nodes

- File access using a file cache on each node

- Continuous file operation on other nodes if one node fails while maintaining file system consistency

- High-speed file system recovery function

- High-speed I/O processing by contiguous block assignment to areas in a file

- Support of multi-partition to implement I/O processing load distribution

- Support of multi-partition to implement the extension of file system size without rebuilding file system

- GUI-based file system operation using a Web browser

Like the ext2/ext3 file system, the following application interface is commonly available:

- **64-bit file system interface**

- Advisory file lock

See
. . . .
For the functions that are unavailable with the GFS Shared File System, see "1.8.2 Unavailable functions."
. . . .

1.1.1 Basic hardware configuration for use of file system

The following basic configuration is required to build the GFS Shared File System:

- Shared disk device, which provides simultaneous access between **node**s sharing the file system

- One or more NICs for the cluster interconnect to monitor internodes communication failures.

- Management Board (for PRIMEQUEST)

- Remote Service Board (for PRIMERGY)

- GUI display personal computer

- NIC for public LAN

Figure 1.1 Basic hardware configuration
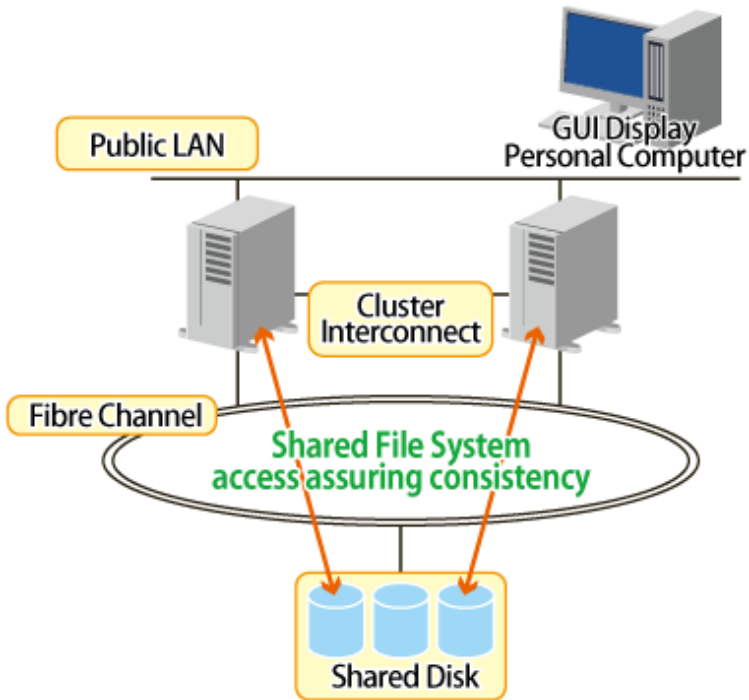
1.1.2 File system sharing with PRIMEQUEST

The PRIMEQUEST can be used for multiple nodes by splitting one frame into multiple partitions. Using the GFS Shared File System in this configuration allows a file system to be shared between nodes.

![Note]

Partitions as used here refer to nodes split into multiple nodes through logical partitioning of the system.

Figure 1.2 File system sharing with PRIMEQUEST
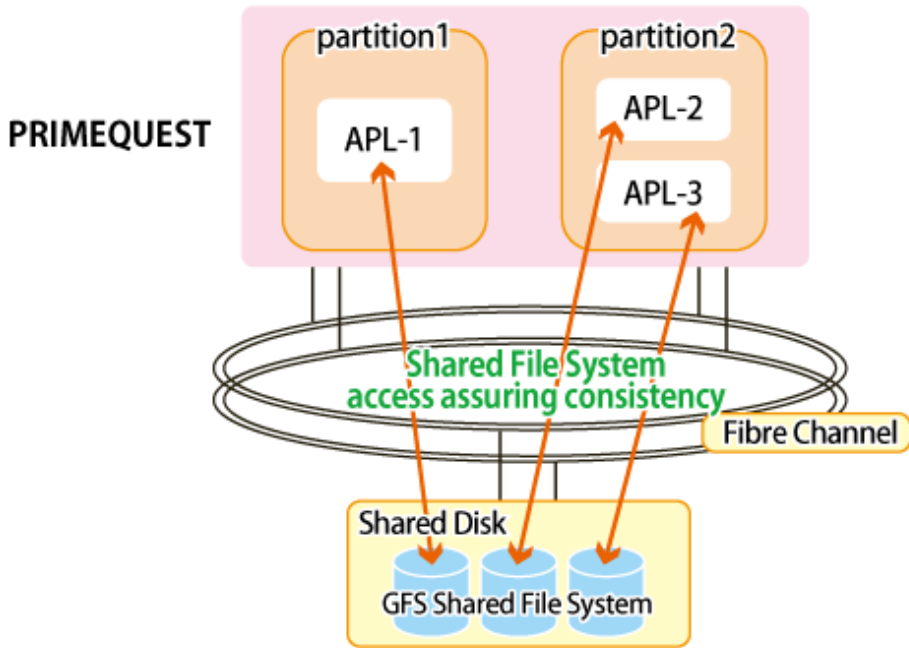
1.2 Simultaneous Shared Access Maintaining Coherency

The GFS Shared File System provides to access and update a file system on a shared disk device from multiple nodes simultaneously. The GFS Shared File System maintains consistency in updating from multiple nodes to files or file systems.

In this way, applications on multiple nodes can modify data mutually and read the latest data. A conventional UNIX file system API makes such operations available.

Figure 1.3 Distributed execution of file related application
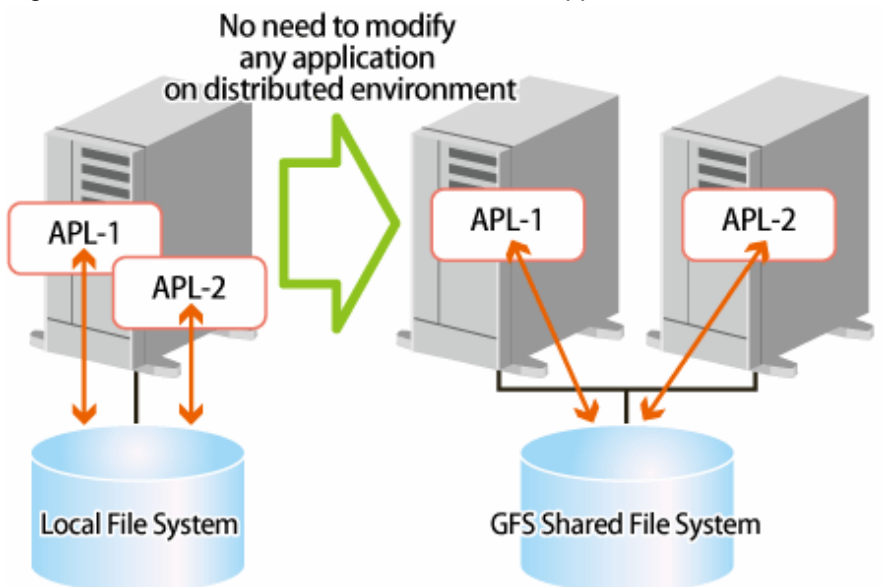
1.3 High Availability

The GFS Shared File System allows continuous access to file systems if a node failure or disk block failure occurs.

1.3.1 Uninterrupted operation in case of node failure

If one **node** fails and the GFS Shared File System is being used from multiple nodes, the other nodes can continue file accesses. The file system data retained by the failing node is automatically recovered in consistency within the GFS Shared File System from the other nodes. In other words, application programs operating on the other nodes can continue without a file system operation error.

See

For more information on the uninterrupted operation function, see "Chapter 3 Failure Recovery."

Figure 1.4 Uninterrupted operation in case of node failure

GFS : Global File Services
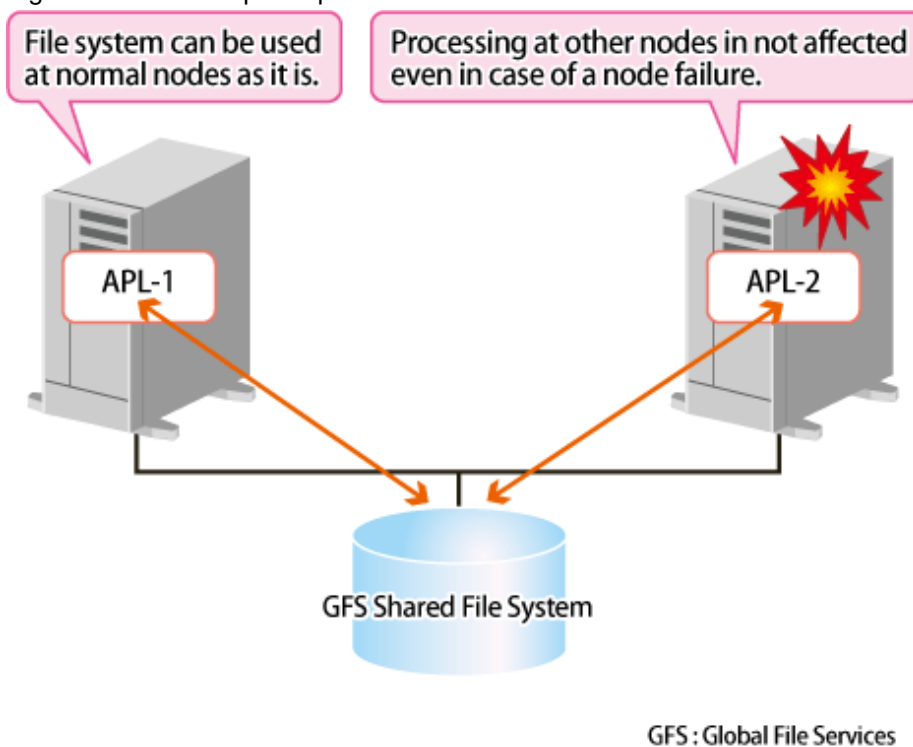
1.3.2 High-speed file system recovery

In case of a **node** failure, **fsck**(8) must be executed for an ordinary file system to recover consistency. In most file systems, it is necessary to inspect all of the **meta-data** for the systems in order to recover consistency. If a node failure occurs, considerable time may be required before the file system actually becomes available.

The GFS Shared File System records operations that have changed the file system structure in an area called updated log. These operations include file creation and deletion. Using the **data** in the **update log area** allows the file system to recover from a system failure in less than a minute.

When recovering from a system failure, the GFS Shared File System retrieves the **update log** in the recovery process. The file system then decides whether to invalidate or complete the file system operation that was done, and reflects the result. The file system structure can then be mounted and used without fully checking it.

As mentioned in "1.3.1 Uninterrupted operation in case of node failure", the GFS Shared File System operating on multiple **node**s does not require execution of **sfcfsck**(8) because recovery processing of consistency is performed automatically in case of a node failure.

1.3.3 Area reassignment in case of disk block failure

The GFS Shared File System automatically assigns another disk block to a new **meta-data area** if a disk block hardware failure occurs. Assignment allows continuous file system processing if a disk failure occurs only on a specific block.

1.4 Performance

The GFS Shared File System has the following functions for implementing file system access at high speed:

1.4.1 Data access performance

The GFS Shared File System allows access to the file system on the **shared disk device** from multiple **node**s. Conventional distributed file systems have transferred file system **data** from management servers to clients that have actually issued an access request via a LAN in the network. However, the GFS Shared File System accesses the disk directly from a request node, reducing network load and speeding up response time for a read or write request in comparison with NFS.

1.4.2 Contiguous block assignment

The GFS Shared File System assigns contiguous blocks to **file data** to take better advantage of batch I/O processing to improve the file system performance. The GFS Shared File System manages area on the **extent-base** for assigning contiguous blocks. For the range of blocks that can be acquired contiguously, the file system manages area assignment using a file offset, a start block number, and the number of blocks used.

The file system also takes into consideration the possibility of contiguous block assignment when extending the file.

Following picture shows a block assignment example that a file has three extents which locate 0, 64 megabytes, 96 megabytes file offset and 64 megabytes, 32 megabytes, 4 megabytes extent length respectively.

Figure 1.5 Contiguous block assignment

File Offset	0MB	64MB	96MB
Partition Number	1	1	0
Start Block Number	20155	9966	2024
Number of Blocks	8192	4096	512
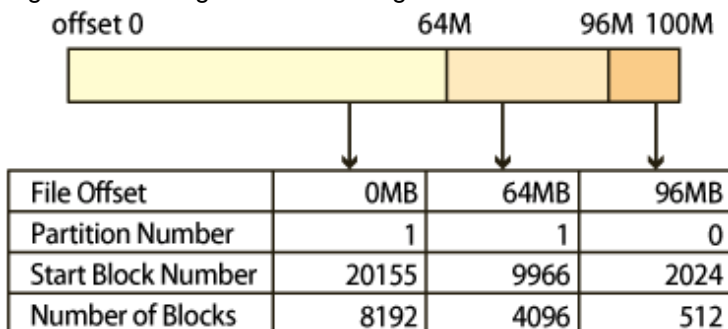
Empty file data areas are also managed in the extent base to implement high-speed assignment of optimum empty areas.

1.4.3 Multi-partition configuration

The GFS Shared File System can unify **partition**s into one file system.

With the GFS Shared File System and addition of partitions, you can easily solve the problem of data insufficiency.

In multi-partition configuration, a **round robin allocation** system allows that different files can be allotted to different partitions, thus improving file system performance by distributing the I/O loads to different disks.

1.4.4 Individualized meta-cache management

The GFS Shared File System has the original cache management for **meta-data**. Many conventional file systems have managed meta-data in a uniform manner. However, the GFS Shared File System provides individualized management of **i-node**, directory block, and indirect extent management information areas on disks, taking into consideration the characteristics of the access. As a result, it improves the cache-hit ratio and reducing resources used.

1.5 Extendibility

The GFS Shared File System is capable of easily extending its data partitions by specifying free disk partitions, providing a quick solution if the file system capacity is insufficient.

1.5.1 File system extension

The GFS Shared File System can expand a file system by just adding partitions. File capacity insufficiency can be resolved without complicated operations for backup and/or re-creation of the file system.

Figure 1.6 File system extension
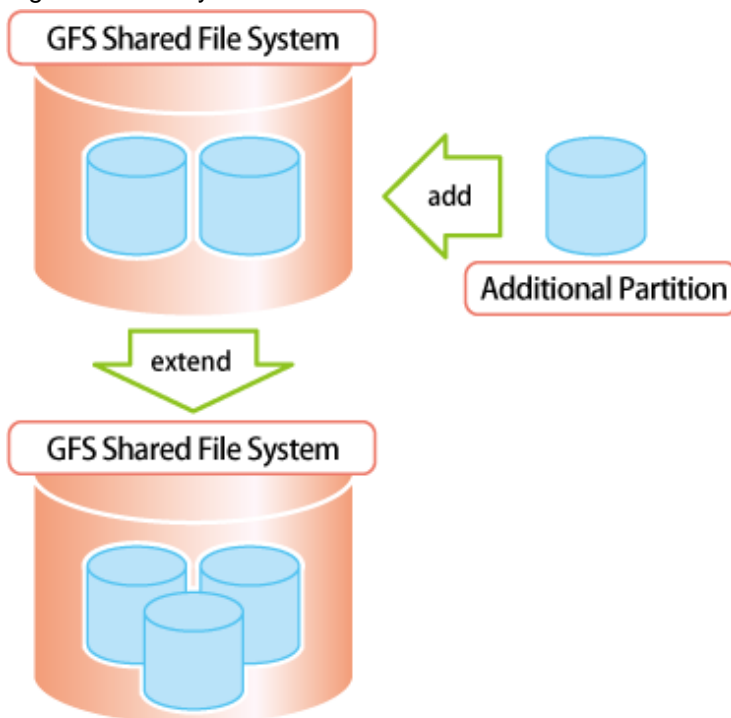
GFS : Global File Services

1.6 Operability

1.6.1 GUI

The GFS Shared File System uses a GUI that allows a Web browser to be used to create, delete, operate, and modify its data and monitor its status.

1.7 Upper Limits of the File System

The following table lists the upper limits of the GFS Shared File System on a file system basis.

Table 1.1 Upper limits of file system

Item	Upper limit
Maximum file system capacity	1 terabyte - 1 kilobyte
Maximum partitions size	1 terabyte
Maximum file size	1 terabyte - 512 byte
Maximum number of sharable nodes	2 nodes
Maximum directory size	2 gigabytes - 1 kilobyte
Maximum number of partitions consisting of a file system	32 partitions
Maximum number of i-nodes per file system	16 mega
Maximum number of directories per file system	1 mega
Maximum number of concurrently open files	5000 files (*1)

*1 Maximum number of files can be opened at a time on one node.

The maximum number of GFS Shared File Systems in a cluster system is 10.

🔶 Note

We recommend you to lessen the number of the GFS Shared File Systems simultaneously used in one cluster system as much as possible. When you use many GFS Shared File Systems, please allot a sufficient number of CPUs to relevant nodes, and carry out system verification in advance.

1.8 Notes

Below are notes on using the GFS Shared File System.

1.8.1 Relationships with Other Components of PRIMECLUSTER

Cluster Interconnect Protocol (CIP) must have been set up, as it is used for the **sfcfrmd** daemon of the GFS Shared File System.

📘 See

For details to set up CIP, see "PRIMECLUSTER Cluster Foundation (CF) Configuration and Administration Guide."

Global Disk Services (hereinafter GDS) is used to create logical volumes for use by the GFS Shared File System.

For details to set up GDS, see "PRIMECLUSTER Global Disk Services Configuration and Administration Guide."
....

The GFS Shared File System must work with the other nodes in the same cluster and must recognize that each node works correctly. When a node downs, it turns to be the state called LEFTCLUSTER. If there is a node with this state, it is impossible to change the management state of the GFS Shared File System. To keep the GFS Shared File System operational whenever a node downs, set up the Shutdown Facility so that the LEFTCLUSTER state will turn into the DOWN state automatically.

1.8.2 Unavailable functions

1.8.2.1 Functions not provided by the GFS Shared File System

The GFS Shared File System explained in this guide does not provide the following functions:

- Use as a root file system, **/usr**, **/var**, **/opt**

- To use as a mount point

- quota function

- ACL function

- Asynchronous I/O function

- File sharing with other nodes by an NFS

- Setting of hard link for a directory

- Direct I/O function

- Mandatory file lock

- File creation and access of the following file types:

    - Block special files

    - Character special files

    - Named pipes

    - Sockets

- The strict file terminal management in the case of **write**(2) for **open**(2) file for which specified **O_APPEND** compete with **write**(2) from other nodes

- Guaranteeing the atomicity of the write for the following operations from other nodes, when data written by **write**(2), **writev**(2), **pwrite**(2) is stored in several blocks.
write(2), **writev**(2), **pwrite**(2), **truncate**(2), **ftruncate**(2), **creat**(2), **open**(2) (with **O_TRUNC**)

- Changing **PROT_READ** and **PROT_WRITE** using **mprotect**(2)

- Enabling lease between nodes using **fcntl**(2) with **F_SETLEASE** or **F_GETLEASE**.

- Using **remap_file_pages**(2) or **mmap**(2) with **MAP_POPULATE**

- **mmap**(2) from another node to a writable file opened by **open**(2)

- Writable **open**(2) from another node to a file mapped by **mmap**(2)

- **open**(2) from another node to a **mmap**(2) file for which **MAP_SHARED** and **PROT_WRITE** are specified

- **mmap**(2) which specifies **MAP_SHARED** and **PROT_WRITE** from another node to a **open**(2) file

- Assurance of data consistency when **mmap**(2) is executed for **open**(2) files from other nodes

- Assurance of data consistency when **open**(2) is executed for **mmap**(2) files from other nodes

- Using **fallocate**(2)

- Mounting function using **mount**(8)

- Unmounting function using **umount**(8)

1.8.3 Service ports used by the GFS Shared File System

In the GFS Shared File Systems, TCP service ports within the range of **9200** to **9263**, and **9300** are reserved.

If these port numbers conflict with those for other applications, change the entry port number beginning with **sfcfs-** and **sfcfsrm** in the **/etc/services** file. If you need to change the port numbers, be aware that the port numbers beginning with **sfcfs-** and **sfcfsrm** must be the same on all the nodes.

1.8.4 Notes on use of logical volumes of GDS

GDS is volume management software designed to improve the availability and operation manageability of information stored on disk units of the Storage Area Network (SAN). Various kinds of access controls can be performed on the logical volumes of GDS to protect data against damage due to unauthorized access.

See
. . . .
For the operations of disk classes to which logical volumes of GDS belong, refer to the description of Class operation of "Operation" "GDS Management View" in "PRIMECLUSTER Global Disk Services Configuration and Administration Guide."
For the operations of logical volumes of GDS, refer to the description of Volume operation of "Operation" "GDS Management View" in "PRIMECLUSTER Global Disk Services Configuration and Administration Guide."
. . . .

1.8.4.1 Notes on setting of logical volumes of GDS

To use a logical volume of GDS in the GFS Shared File System, make the following settings:

- The disk class to which the logical volume belongs is of type: "shared".

- All nodes sharing the GFS Shared File System are specified in the scope of the disk cluster to which the logical volume belongs.

- When the attribute lock mode of the logical volume is **lock=off**, the volume is started automatically.

- The attribute access mode of the logical volume is set as read-write.

To use the GFS Shared File Systems, logical volumes of GDS must be **ACTIVE**. If they are **STOP**, all accesses to the shared file systems will be prohibited.

Note
. . . .
If you create the GFS Shared File System on GDS-striped logical volumes, set the width of the stripe group to 256 blocks or more.
. . . .

📑 Note

Do not use online volume expansion for the logical volume of GDS that is including GFS Shared File System. Because the copy of the super block and of the partition configuration information of the GFS Shared File System cannot be read when the size of the logical volumes of GDS is changed by the online volume expansion, the GFS Shared File System cannot be used.

If you want to expand the GFS Shared File System, add an unoccupied logical volume partition, or create a new logical volume and add it.

1.8.4.2 Notes when using logical volumes of GDS that were used in a file system other than GFS

If the logical volume of GDS had been used for a file system other than GFS before, clear 1 MB from the beginning of it with **dd**(1) to zero before using it for a GFS Shared File System.

The execution example is shown below.

```
# dd bs=1M count=1 if=/dev/zero of=/dev/sfdsk/gfs01/dsk/volume01
```

📑 See

For details of **dd**(1), see the online manual page.

1.8.5 Notes on unmounting

You must use **sfcumount**(8) or **sfcumntgl**(8) to unmount a GFS Shared File System.

If the GFS Shared File System is unmounted with **umount**(8) by mistake, the file system cannot be mounted or unmounted. Restoration of such a file system can be performed in the manner described in "C.3.3 Action to Be Taken When Unmounted by umount(8)."

1.8.6 Notes on updatedb command

Function to retrieve the location of the file at high speed is provided in Linux as the **locate** command. The **updatedb** command is periodically executed to update the file name database used by this function. The **updatedb** command retrieves the file on each file system, and updates the file name database.

📑 See

For details about the **updatedb** and **locate** commands, see the following online manual pages.

- [RHEL6] **updatedb**(1L) and **locate**(1L)

- [RHEL7] **updatedb**(8) and **locate**(1L)

While the **updatedb** command is executed to look for GFS Shared File System, file access performance can deteriorate considerably. To evade it, it is recommended to set up not to register the files on GFS Shared File System to the file name database.

📑 Note

The file on the GFS Shared File System cannot be retrieved by the **locate** command when setting up it.

So as not to register the files on the GFS Shared File System to the file name database, add "**sfcfs**" to PRUNEFS parameter (specifies the file system excluded from the retrieval object by the **updatedb** command) defined in **/etc/updatedb.conf**. It must be set up on all nodes which share the GFS Shared File System. The following shows an example of setting.

```
# cat /etc/updatedb.conf
PRUNE_BIND_MOUNTS = "yes"
PRUNEFS = "9p afs anon_inodefs auto autofs bdev binfmt_misc (text omitted) tmpfs ubifs udf
```

```
usbfs sfcfs"
PRUNENAMES = ".git .hg .svn"
PRUNEPATHS = "/afs /media /mnt /net /sfs /tmp /udev /var/cache/ccache /var/lib/yum/yumdb /var/
spool/cups /var/spool/squid /var/tmp"
```

When the files on the GFS Shared File System is registered to the file name database, deterioration in the file access performance can be reduced while the **updatedb** command operates by setting as follows.

- Set the execution time of the **updatedb** command to when GFS Shared File System is accessed fewer according to the user operation.

- Adjust the execution time of the **updatedb** command to avoid executing the **updatedb** command simultaneously by the nodes which shared GFS Shared File System.

Usually, the **updatedb** command is executed by the file that exists on the **/etc/cron.daily** directory. The execution time of **/etc/cron.daily** is specified in **/etc/anacrontab**. Change the settings of **/etc/anacrontab** as necessary.

See
. . . .
For details of the contents of **/etc/anacrontab**, see **anacrontab**(5).
. . . .

1.8.7 Notes on system time change

Do not perform the following operation while a GFS Shared File System is working.

- Change the system time by the **date**(1) command etc.

If the system time is changed by the **date**(1) command, the file system will be blockaded. The monitoring mechanism on a GFS Shared File System might not be operated normally due to a sudden change of the system time.

To resume the system operation, mount again after unmounting the file system.

1.8.8 Notes on operation stop

Be sure to unmount all the GFS Shared File System in a node, before stopping the operation of a GFS Shared File System using the following command.

[RHEL6]

```
# /etc/init.d/sfcfsrm stop <Enter>
```

[RHEL7]

```
# systemctl stop fjsvgfsfsrm2.service <Enter>
# systemctl stop fjsvgfsfsrm.service <Enter>
# systemctl start fjsvgfsfsrm2.service <Enter>
```

When the command mentioned above is executed with GFS Shared File System mounted, the mount information of GFS Shared File System remains, so mounting or unmounting the file system becomes impossible. In this case, reboot the node to restarting the operation of GFS Shared File System.

1.8.9 Deleting a file being created from the other node

When you are handling GFS Shared File Systems, do not execute **unlink**(2) or **rename**(2) for a file being created by **open**(2) with **O_CREAT** or by **creat**(2), from the other node.

If you execute **unlink**(2) or **rename**(2) for a file being created by **open**(2) with **O_CREAT** or by **creat**(2) from the other node, **unlink**(2) or **rename**(2) will return an exit code, but **open**(2) or **creat**(2) might encounter an **ESTALE** error.

Accordingly, commands (*1) that are using the **open**(2) or **creat**(2) might also return an error code.

*1 Creating a file with **touch**(1) returns an error code of 1.

1.8.10 Notes on file owner change

If you need to change the file owner (*1) after changing the file mode (*2) to the same file, perform both operations on the same node.

*1 File owner change using a command such as **chown**(1), **chown**(2), or **fchown**(2)

*2 File mode change using a command such as **chmod**(1), **chmod**(2), or **fchmod**(2)

If you perform file owner change after file mode change for the same file from the other node, the file might go back to the previous mode. You can prevent this problem by acquiring (*3) updated file information before changing file owner.

*3 File information acquisition using a command such as **stat**(1), **stat**(2), or **fstat**(2)

1.8.11 Notes on ioctl(2) with FIONREAD

If you need to execute **ioctl**(2) with **FIONREAD** after changing the file size (*1) to the same file, perform both operations on the same node.

*1 File size change using a command such as **write**(2), **truncate**(2), or **ftruncate**(2)

If you execute the **ioctl**(2) after changing the file size for the same file from the other node, the file might calculate the prefetch quantity based on the old file size. You can prevent this problem by acquiring (*2) updated file size before executing the command.

*2 File information acquisition using a command such as **stat**(1), **stat**(2), or **fstat**(2)

1.8.12 Notes on fcntl(2) with F_SETFL

If you need to execute **fcntl**(2) with **F_SETFL** and change the file owner (*1), perform both operations on the same node.

*1 File owner change using a command such as **chown**(1), **chown**(2), or **fchown**(2)

If you execute **fcntl**(2) with **F_SETFL** and change the file owner on each of nodes, **fcntl**(2) might check permission and change the open flag, with the old file owner information.

If you need to execute **fcntl**(2) with **F_SETFL** and change the file owner on each of nodes, please update the file information (*2) before executing **fcntl**(2).

*2 File information acquisition using a command such as **stat**(1), **stat**(2), or **fstat**(2)

1.8.13 Notes on fcntl(2) with F_SETLEASE

If you need to execute **fcntl**(2) with **F_SETLEASE** and change the file owner (*1), perform both operations on the same node.

*1 File owner change using a command such as **chown**(1), **chown**(2), or **fchown**(2)

If you execute **fcntl**(2) with **F_SETLEASE** and change the file owner on each of nodes, **fcntl**(2) might check permission and set a file lease, with the old file owner information.

If you need to execute **fcntl**(2) with **F_SETLEASE** and change the file owner on each of nodes, please update the file information (*2) before executing **fcntl**(2).

*2 File information acquisition using a command such as **stat**(1), **stat**(2), or **fstat**(2)

1.8.14 Notes on rename(2)

The following describes notes for executing **rename**(2).

- If a directory is specified for the place from which a file is moved, **rename**(2) must not be used for the directory's parent directory.

  If **rename**(2) is executed by mistake, file operation may hang. If this occurs, reboot a node where the file hangs and recover the failed file.

- When changing a file name in the same directory, specify the name so that absolute path and relative path will not coexist in newpath and oldpath arguments of **rename**(2).

1.8.15 Notes on file lock function

This section describes notes on file lock function that is used in the GFS Shared File System.

1.8.15.1 Unavailable functions

The GFS Shared File System does not support the following functions:

- File lock between nodes to files other than regular file

- Deadlock check function between nodes with another file system

1.8.15.2 Notes caused by Linux kernel specification

The following functions might not operate normally because of the Linux kernel specification.

- Setting and getting of file lock when SEEK_END is specified

  When you execute the operation that sets or gets a file lock with **SEEK_END** to the same file after the operation that changes the size of the file by **write**(2), **truncate**(2) or the **ftruncate**(2), etc, execute the operation on the same node.
  When operations to the same file are executed on a different node, the file lock operation might be done with the old file size.
  When it wants to execute above-mentioned two operation on a different node, the issue can be evaded by getting the latest size of the file by **stat**(1), **stat**(2), and the **fstat**(2), etc. before file lock operation with **SEEK_END**.

- The setting of file lock between the process, its child processes and the threads that it generated

  File operation might hang up when the following operation competes for the same file.

    1. The setting operation of file lock by thread that process(A) generated

    2. The setting operation of file lock by process(A)

    3. The terminating operation of process(B) that process(A) generated or
       The **close**(2) operation in process(B) for the file that is target of file lock setting in 1 or 2

  Or

    1. The setting operation of file lock by thread that process(A) generated

    2. The setting operation of file lock by the child process(C) that process (A) generated

    3. The terminating operation of process(C) or
       The **close**(2) operation in process(A) for the file that is target of file lock setting in 1 or 2

When the file operation hangs up, it is possible to recover from the state which has hung up by terminating by **kill**(1) the process which has hung up.

1.8.15.3 Effect of file lock on system performance

If many file locks are set on a GFS Shared File System, the system performance can be affected as follows:

- Loads on the memory will be increased by increase in amount of memory used by MDS.

- Loads on CPU will be intensive by the file lock processing.

- If there are many processes that wait for the file lock, and the following operations are executed, loads in the network might be temporarily increased.

    1. The file lock is released. or

    2. The node where the process setting file lock exists goes down.

- If the recovery procedure is executed when many file locks are set, it will take time and fail.

1.8.16 Notes on starting and stopping a node

Please note the following points when you start the node.

- In RHEL7.1 or later version environment, the following message may be output. However, it does not affect the starting operation.

```
kernel: Request for unknown module key 'FUJITSU Software: Fujitsu BIOS DB FJMW
Certificate: Hexadecimal, forty-digit' err -11
```

Please note the following points when you stop the node.

- Before stopping a node, please stop all the applications on a node accessing GFS Shared File System.

- Use **shutdown**(8) to stop a node. If **reboot**(8), **halt**(8), or **poweroff**(8) is used to stop a node, GFS may not work correctly.

- If you stop plural nodes including the node that operates the primary MDS, stop the node that operates the primary MDS at the end. Otherwise, it might take time to stop the nodes.

 Information
....
If the GFS Shared File System is unmounted by **sfcumntgl**(8) on all nodes before the node is stopped when all shared nodes of the GFS Shared File System are stopped, the order of stopping the node need not be considered.
....

1.8.17 Notes on starting and stopping CF by GUI

Please note the following points when you start CF by GUI.

- In RHEL7.1 or later version environment, the following message may be output. However, it does not affect the starting operation.

```
kernel: Request for unknown module key 'FUJITSU Software: Fujitsu BIOS DB FJMW
Certificate: Hexadecimal, forty-digit' err -11
```

- In degenerated operation in which a part of cluster nodes is stopped, CF cannot be started by GUI. In this case, reboot the node. The following message will be output when the node is rebooted.

```
WARNING: sfcfsrm:5001: Starting the sfcfrmd daemon was suspended because quorum does not
exist
```

See "C.3.4 Corrective action when the sfcfrmd daemon is not started" to take corrective action.

Please note the following points when you stop CF by GUI.

- If you want to stop CF on the plural nodes including the node that operates the primary MDS, stop CF on the node that operates the primary MDS at the end. Otherwise, it might take time to stop CF.

 Information
....
If the GFS Shared File System is unmounted by **sfcumntgl**(8) on all nodes before CF is stopped when all CF of a shared node of the GFS Shared File System is stopped, the order of stopping CF need not be considered.
....

1.8.18 Notes on opening a file on the GFS Shared File System

The number of files that can be concurrently open on one node at one time is limited in the GFS Shared File System. The maximum number is 5000 files per file system. When **open**(2) is attempted for the file on the GFS Shared File System exceeding the maximum number, **open**(2) will fail and **ENFILE** will be returned as an error number.

1.8.19 Notes on writing file data on the GFS Shared File System

The GFS Shared File System uses extent based area management. If a file has many extents, some of extents are stored as indirect extent management information in V-data area.

In this case, an allocation of new blocks could induce a rebuilding of indirect extent management information. It would consume many V-data area. To prevent V-data starvation, GFS prevents an allocation of new block when V-data area usage exceeds 80%. In this case, a system-call (e.g. **write**(2)) which attempts to allocate a new block fails with **ENOSPC**.

Deleting unused files or directories or moving them into other file systems helps you to reduce V-data area usage.

See

For information on how to check usage of V-data area, see "B.1.1 sfcdf(8) Display the GFS Shared File System use status and configuration information."

1.8.20 Notes on use of GFS in environments where SELinux is enabled

In environments where SELinux is enabled, the security context of a file on the GFS Shared File System cannot be changed.

1.8.21 Notes on cache of the file system

When accessing a large number of files continuously, it takes a lot of time to access files and it may affect the performance and operation of the entire system for the following reasons:

- It takes time to get new page cache with increase of unused page cache.

- If the release process runs to keep the number of i-node cache of the meta-data server below the limit value, the system load increases.

Consider the following corrective actions:

- Before and after processing, you can clear all caches for GFS if you unmount target file systems on all nodes and then mount them. By doing this, you can prevent the increase of the processing time due to unused cache.

- By dividing the process accessing a large number of files continuously into small steps and setting the processing intervals with **sleep**(3), the load of the entire system may be reduced and the performance of other processing that runs concurrently may be improved.

- If accessing total of 50000 files or more on all nodes continuously, it is recommended to release dentry caches periodically to discard GFS page cache.

  Release unused dentry caches according to the following procedure.

  1. Check the usage status of dentry caches.

  ```
  # /bin/cat /proc/sys/fs/dentry-state <Enter>
  411155 391674 45 0 0 0
  ```

  The second field indicates the number of unused dentry caches.

  2. Release unused dentry caches.

  ```
  # /bin/echo 2 > /proc/sys/vm/drop_caches <Enter>
  ```

Note

Caches other than the GFS file system are released with the command.

Information

You can periodically check and release unused dentry caches by using **cron**(8).

1.8.22 Notes on using the Migration function in a KVM environment

In a system where the Migration function is used in a KVM environment, set the communication timeout value of the GFS Shared File System to 630 seconds or more.

It prevents a GFS Shared File System from being blockaded by detecting a communication timeout during the Migration.

 See
....
For details on how to set a communication timeout value and the notes, see "14.1.2 Communication timeout value."
....

1.8.23 Notes on lseek(2) with SEEK_DATA or SEEK_HOLE

When **lseek**(2) with SEEK_DATA or SEEK_HOLE is executed to a file on the GFS Shared File System, the behavior is as follows:

- When the specified offset does not exceed the end of the file

  - SEEK_DATA:

    Set the file offset to the specified offset.

  - SEEK_HOLE:

    Set the file offset to the end of the file.

- When the specified offset exceeds the end of the file

  lseek(2) fails (return code: -1, errno: ENXIO).

 Information
....
- Specify SEEK_DATA and SEEK_HOLE in the third argument of **lseek**(2).

- Specify the offset in the second argument of **lseek**(2).
....

1.8.24 Notes on updating OS to RHEL7.1 or later version

Please note the following points when OS is updated to RHEL7.1 or later version.

When using GFS in an environment where SELinux is enabled, due to the specification change in OS, the SELinux security context of a file on the GFS Shared File System displayed by the following operations is changed:

- When the following command is executed to display the list of the GFS Shared File System (if the previous OS version is RHEL7.0)

  - mount(8) without option specified

- When the following commands are executed to display the security context of the file on the GFS Shared File System

  - dir(1) -Z, --context

  - ls(1) -Z, --context

  - vdir(1) -Z, --context

  - find(1) -printf "%Z"

  - stat(1) without option specified, -c %C, --format=%C, --printf=%C

Here are samples of the command execution and display:

- mount(8)

  [RHEL7.0]

```
# mount | grep /mnt/gfs01 <Enter>
/dev/sfdsk/gfs/dsk/gfs01 on /mnt/gfs01 type sfcfs
(rw,relatime,context=system_u:object_r:file_t:s0)
```

[RHEL7.1 or later version]

```
# mount | grep /mnt/gfs01 <Enter>
/dev/sfdsk/gfs/dsk/gfs01 on /mnt/gfs01 type sfcfs
(rw,relatime,context=system_u:object_r:unlabeled_t:s0)
```

- ls(1)

[RHEL6 and RHEL7.0]

```
# ls -Z /mnt/gfs01/file01 <Enter>
-rw-r--r--. root root system_u:object_r:file_t:s0      /mnt/gfs01/file01
```

[RHEL7.1 or later version]

```
# ls -Z /mnt/gfs01/file01 <Enter>
-rw-r--r--. root root system_u:object_r:unlabeled_t:s0 /mnt/gfs01/file01
```

1.8.25 Notes on RHOSP environment

In Red Hat OpenStack Platform (hereinafter RHOSP) environment, set the communication timeout value of the GFS Shared File System to 630 seconds or more.

When a node fails in RHOSP environment, this setting prevents the GFS Shared File System from being blockaded by detecting a communication timeout during MDS failure recovery.

 See

For details on how to set a communication timeout value and the notes, see "14.1.2 Communication timeout value."

1.8.26 Notes on using the Mount namespace function [RHEL7]

Please note the following points when you use the Mount namespace function.

- Do not mount or unmount the GFS Shared File System in a process whose mount namespace is unshared.

- Do not unmount the GFS Shared File System if there is a process which mounts it as private mount.

  Use the following method to check if there is a process which mounts the GFS Shared File System as private mount.

```
# grep "\- sfcfs" /proc/*/mountinfo | egrep -v "shared:|master:"
#
```

  If nothing is output, there is no process which mounts the GFS Shared File System as private mount.

Chapter 2 File System Structure

This chapter describes the structure of the GFS Shared File System.

2.1 Disk Structure of File System

This section describes the disk structure of the GFS Shared File System.

As illustrated below, the disk structure of the GFS Shared File System consists of the following elements:

- **Super block**

- **Partition configuration data**

- **Meta-data area** (area management data, i-node area, V-data area)

- **Update log area**

- **File data area**

Figure 2.1 Disk structure of the GFS Shared File System
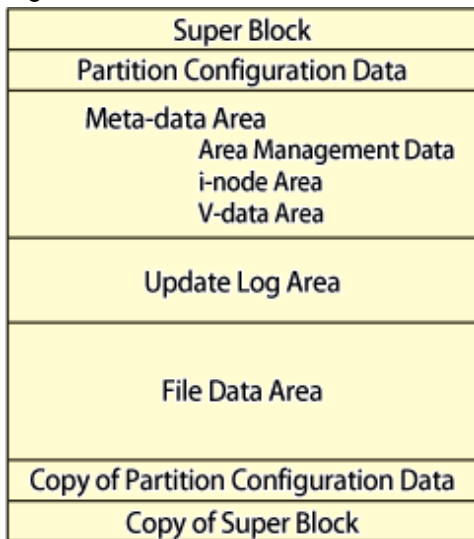
The GFS Shared File System manages the V-data area in units of 1,024 bytes and the file data area in unit of 8,192 bytes.

Conventional file systems have decentralized meta-data areas on their disks. However, the GFS Shared File System centralizes **meta-data** to improve meta-data access performance. Besides, localizing updated portions upon recovery of the file system allows the update log area to shorten the recovery time.

2.1.1 Super block

Super block contains a file system type, creation and update dates, label data, size and layout of the file system, and a data history of empty areas.

The copy of super block at the end of the file system exists to allow continuous processing if a disk device block failure occurs in the area where the super block is stored.

2.1.2 Partition configuration data

The partition configuration information contains the following information:

- What partitions are used to configure a file system. A GFS Shared File System can consist of multiple partitions.

- Area allocation in the partition

In order to expand the file system configuration and change the shared device information, the partition area will be changed.

Similar to the super block, the partition configuration data is important for operation of the file system and is designed to have a partition copy at the end of the file system and to be resistant to disk device block failures.

2.1.3 Meta-data area

This section describes the area of meta-data management. The meta-data area is an area for storing meta-data, which exists only in a **representative partition**.

2.1.3.1 Area management data

Area management data includes the allocation information of **i-node**, **V-data** and **file data**.

File data allocation information is managed in **extent-base** while combining a part to include extent information with a list managed with a degree of continuity.

When a file system is created with **sfcmkfs**(8), an area of the fixed size is allocated for management information, i-nodes and V-data.

2.1.3.2 i-node area

The i-node area is a data structure containing file data. The pointer to extent information including the file data and the file type, length, owner, group IDs, and access right are stored in each i-node. There is one i-node for each file.

2.1.3.3 V-data area

The V-data area includes a directory block, a symbolic link path name, and indirect extent management information. An area is obtained from this area if necessary. The V-data area is managed in units of 1,024 bytes for efficient use of directory blocks.

2.1.4 Update log area

The **update log area** stores a history of file system structure changes for high-speed file system recovery. This area is maintained as a cyclic log.

The update log area contains data about the following processes:

- Acquiring or releasing an **i-node**

- Updating i-node data

- Allocating or releasing an extent, which is a group of contiguous file system data blocks handled as a single unit

- Acquiring or releasing V-data

- Updating V-data

The GFS Shared File System assures that data is written in the **V-data area** before the file system structure is updated. If a system failure occurs, file system consistency is restored by using **sfcfsck**(8) to either invalidate or re-execute unprocessed changes on the file system. Also, system failure recovery processing is conducted. Only changes to the file system structure are recorded in the update logs; file-data is not recorded.

2.1.5 File data area

The file data area is an area for storing file data. In these areas, 8192 bytes areas are defined as the minimum block, and blocks are managed for contiguous allocation in a file area.

2.1.6 Partition configuration

A GFS Shared File System can consist of single partition or multiple partitions. In the multi-partition configuration, multiple partitions are allocated to single file system.

Meta-data area and update log area are gathered by one partition and posted. The partition posted meta-data area is called the representative partition.

In a single partition configuration, a meta-data area, an update log area, and a file data area are allocated to one partition.

Figure 2.2 Single partition configuration
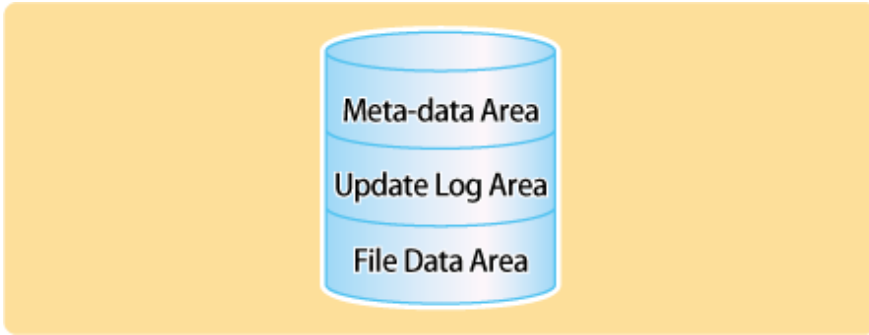
In the multi-partition configuration, the file data area can hold multiple partitions. Also, the file system can be created by detaching the file data area from the meta data area.

Typical examples of multi-partition configurations are shown on the following.

1. Adding a file data area to another partition. (file data area addition)

   Figure 2.3 File data area addition
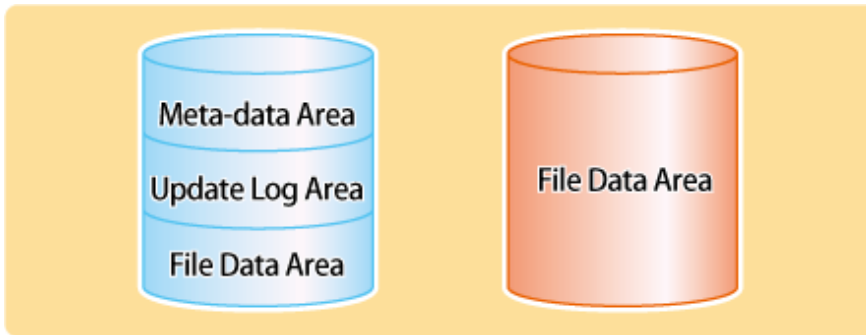
2. Partition configuration in which the file data area is separated from the representative partition.

   Figure 2.4 File data area separation
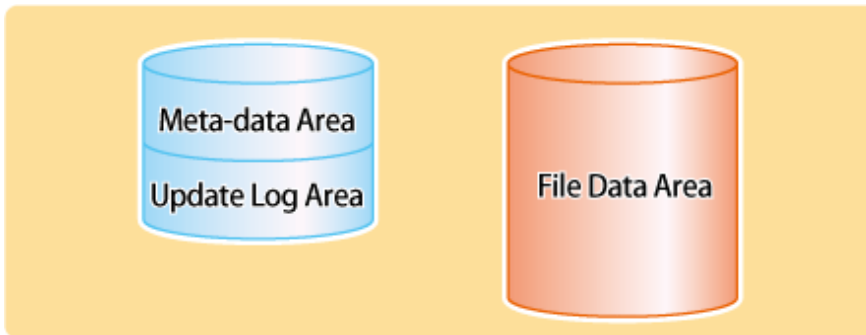
In a multi-partition configuration, **super block** and **partition configuration data** are maintained in all partitions.

2.2 Component configuration

The component configuration of the GFS Shared File System is illustrated below.

The GFS Shared File System consists mainly of three components:

1. **MDS** (Meta-data server)
   The file system server function component (sfcfsmg) of the GFS Shared File System

2. **AC** (Access client)
   The file system client function component (kernel component AC) of the GFS Shared File System

3. Monitoring daemons (sfcfrmd, sfcprmd, sfcfsd, sfchnsd, sfcpncd)
   The control function component of MDS, AC and composition component of GFS Shared File System.

Figure 2.5 Component composition
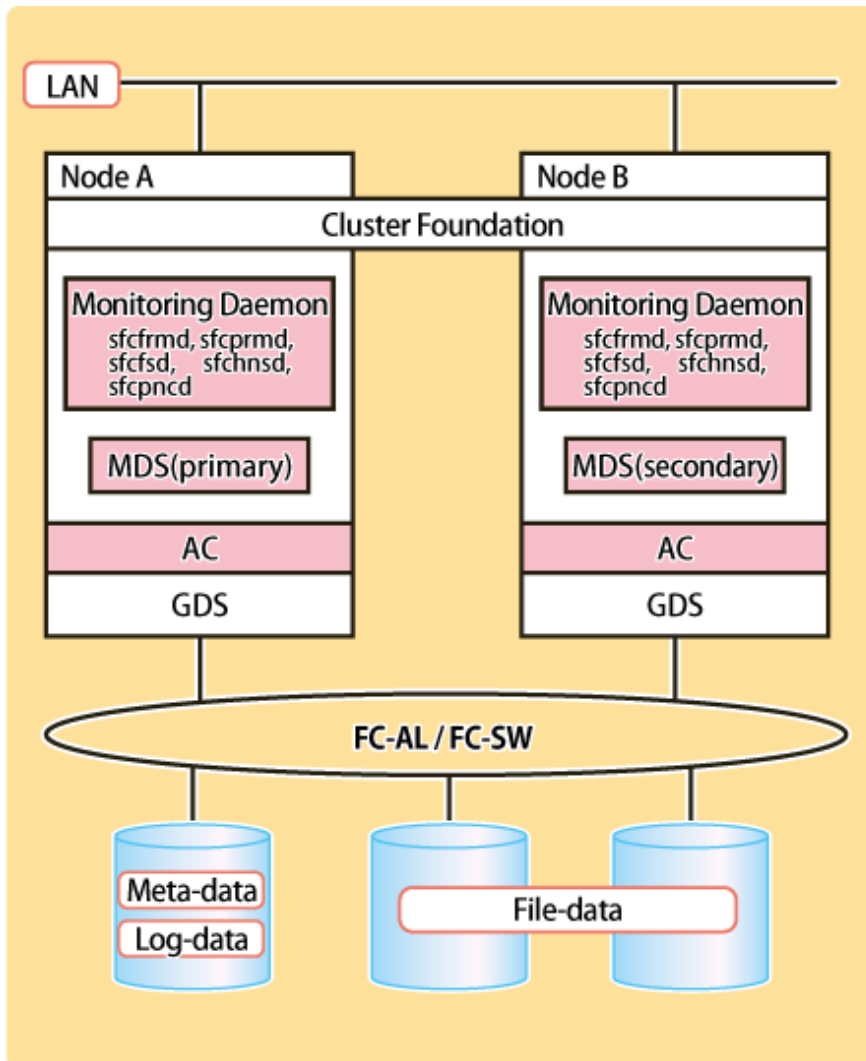
GDS : Global Disk Services

2.2.1 MDS (Meta Data Server)

The MDS is a server daemon that manages the meta-data of the GFS Shared File System and operates as a user process. There are two processes of primary MDS and secondary MDS for each file system. Two nodes of those that shares a file system are predefined as **MDS nodes** and MDS operates on one of these nodes. The MDS nodes are only the two nodes of the primary and secondary MDS node.

The secondary MDS is used for standby and runs on a different node from the primary MDS. However, when only one of MDS nodes is active, the secondary MDS does not run. The MDS uses multithreading in order to execute processing requests from **AC** in parallel.

The MDS has the following major functions:

- Token management

- Meta-data area management

- Update log area management

- File data area management

- File system free area management

- AC node management

2.2.2 AC (Access Client)

This client processes requests from an application that accesses the GFS Shared File System. It operates within the kernel. Access clients exist on all the nodes where the GFS Shared File System is mounted. The AC allows cache management of meta-data and file data within the client.

2.2.3 Monitoring daemon

- sfcfrmd

  Receives a request from the sfcfsd or the commands and provides the communication paths of commands-sfcfsd, commands-sfcprmd and inter-node. Configuration database management and the management partition I/O functions are also provided.

- sfcprmd

  Used to monitor the startup, stop, or failure of the processes that shares the GFS Shared File System. When the configuration process starts or executes another daemon or command, a request is transmitted to sfcprmd, and then sfcprmd executes a daemon or a command as required.

  The process from which a request is transmitted is referred to as the client process, while the process to be started from sfcprmd as required is referred to as the target process.

  The target process is monitored by sfcprmd and, when the process terminates normally or abnormally, the status is posted to the client process.

- sfcfsd

  Provides MDS/AC control functions such as the GFS Shared File System operation start/stop, failover of MDS.

- sfchnsd

  Used to provide functions such as node down event reception, domain-wide lock, membership information acquisition, and node state acquisition.

- sfcpncd

  Performs live monitoring of sfcprmd or sfcfrmd that is to be monitored. When the process to be monitored abnormally end (process down), this daemon makes system go down in order to maintain consistency of file systems.

2.3 File Access and Token Management

Each **AC** caches file data to improve file system performance. Since it is necessary to manage discard operation of cache to ensure consistency in each file system, the GFS Shared File System uses file-exclusive tokens (referred to hereafter as tokens) for management.

A token guarantees that the AC can access files. The token contains information that indicates how it accesses a particular file. An AC must have a token of a file when the AC accesses the file. Two types of tokens are provided: one that ensures **meta-data** access and another that ensures **file data** access.

- Meta-data access and token management

  The meta-data token required to access i-node and V-data.

- File data access and token management

  Tokens are used to access file data in ordinary files. The two types of available tokens (write authority and read authority) are managed in units of logical blocks. Such operations as simultaneously reading from the same block and simultaneously reading from/writing to different blocks in the same file can be performed simultaneously by multiple shared nodes without degrading performance. The file data token is required to access ordinary files.

Chapter 3 Failure Recovery

This chapter describes the following failure recovery functions of the GFS Shared File System:

- **MDS** Failure Recovery

- **AC** degeneration

3.1 MDS Failure Recovery

If a **MDS** fails by node down, the GFS Shared File System automatically performs failure recovery to continue operation of the file system.

The following provides an overview of the operations performed if either the primary MDS or secondary MDS fails.

3.1.1 Automatic recovery for primary MDS failures

There is one case in which the standby secondary **MDS** becomes the primary MDS to continue operation. This case occurs if both the primary **MDS** and secondary **MDS** are set up as operable and the primary MDS fails by node down. In this case, the following process is performed:

1. Switching the secondary MDS to the primary MDS

   The secondary MDS operates as the primary MDS in place of the primary MDS that has failed.
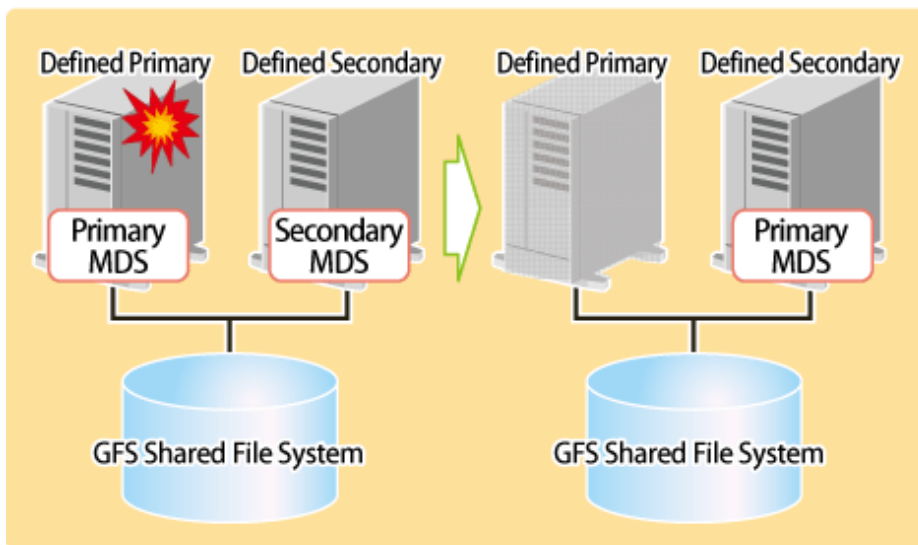
GFS : Global File Services

2. Replaying the update log

   The new primary MDS replays the update log to ensure consistency of the file system.

3. Resuming processing after MDS switching

   The AC sends a processing request to the new primary MDS. Until that request is accepted, all access to the file system is blocked.

4. Restarting the failed MDS

   Usually, the MDS is restarted as the secondary MDS.
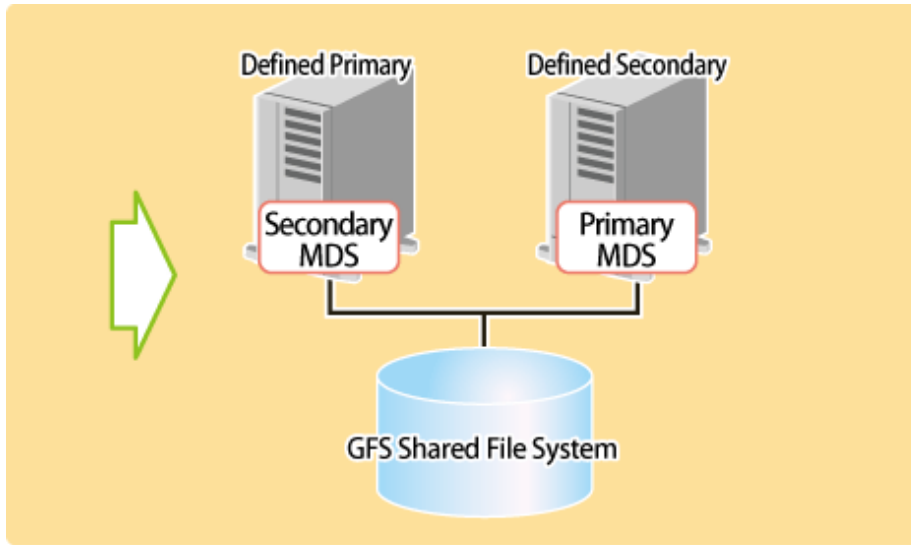
3.1.2 Automatic recovery for secondary MDS failures

If the secondary **MDS** fails because of a node failure, the primary MDS will not be affected. It will continue operation without the secondary MDS. In this case, the following process is performed:

1. Restarting the failed MDS

   The MDS is restarted as a secondary MDS.
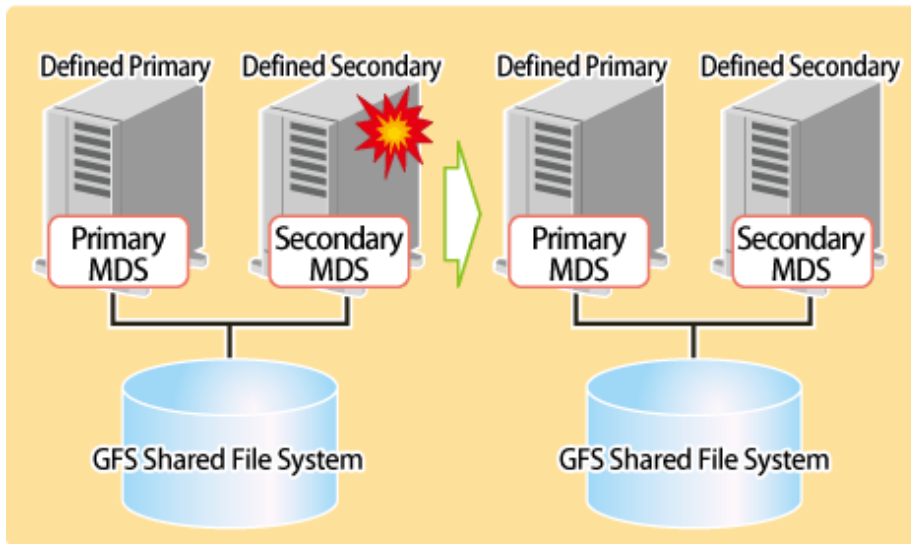
GFS : Global File Services

2. Making the restarted secondary MDS effective

   The AC performs the processing necessary for recognizing the restarted secondary MDS.

3.1.3 Automatic recovery in case primary MDS goes down during primary MDS operation only

If the primary **MDS** goes down in a state in which the MDS nodes are degraded only during primary MDS operation, the following processing is performed automatically.

1. Restarting the failed MDS

   The MDS is started as the primary MDS again.

2. Replaying the update log

   The new primary MDS replays the **update log** to ensure the consistency of the file system.

3. Resuming processing after MDS switching

   The AC makes a processing request to the new primary MDS.

3.2 AC degeneration

If the GFS Shared File System is mounted, MDS will hold the information on the AC. When the node to which AC operates is downed, MDS cancels the information on the AC automatically, and continuation of management of the file system in a normal node is performed automatically.

Chapter 4 File System Design

This chapter describes the main features of the operational design of the GFS Shared File System.

4.1 Mount and unmount opportunity of the file system at the time of system starting and ending

In the GFS Shared File System, a setup for mounting and unmounting automatically at the time of system starting and ending differs from that of ext2, ext3, etc.

The mount options field of **/etc/fstab** must contain noauto in the GFS Shared File System. A setup of the automatic mount at system startup is determined by setting a parameter "mount options" to "noatrc" in **/etc/fstab** or not.

- When noatrc is not described in the mount options field of **/etc/fstab**

    The automatic mounting is performed on system startup.

    The mounting is performed at the following timings:

    - [RHEL6] At the time of executing the startup script **S57sfcfsrm** in run levels 2, 3, and 5

    - [RHEL7] At the time of starting the **fjsvgfsfsrm** service

- When noatrc is described in the mount options field of **/etc/fstab**

    The automatic mounting is not performed on system startup.

GFS Shared File Systems are automatically unmounted when the system is shut down. The unmounting is performed at the following timings:

- [RHEL6] At the time of executing the stop script **K36sfcfsrm**

- [RHEL7] At the time of stopping **fjsvgfsfsrm** service

Note

It is necessary to set the management partition and the GDS volumes where partitions of the GFS Shared File System that will automatically be mounted are allocated to automatically become **ACTIVE** during system startup.

See

For more information about GDS, see "PRIMECLUSTER Global Disk Services Configuration and Administration Guide."

4.2 Disk Design

This section describes the disk design to create a GFS Shared File System.

GFS Shared File System requires GDS logical volumes for not only a file system but also a management partition (to manage the important information for GFS Shared File System operation).

The number of required GDS logical volumes is shown in the following table.

Table 4.1 Number of required GDS logical volumes on GFS Shared File System

Usage	Number of GDS logical volumes	Remarks
Management partition	1	Over 40 megabytes of volume size is necessary.
File system	1 to 32 for each file system	The maximum number of file system is 10.

GDS : Global Disk Services

4.3 System Design

This section describes the following items that must be considered in creating or using the GFS Shared File System:

- Effects of file system operation and system load

- Node layout

- Disk layout

- LAN selection

- For future expansion

4.3.1 Effects of file system operation and system load

To maintain strict file system consistency, the GFS Shared File System processing can significantly slow down during the following operations:

- File access

  - Open and close a file for *each* successive read or write, without putting all I/O together within one open and close.

  - Sequential write with a small record size (less than 1 megabyte). It can cause a fragmentation of a file data.

  - Deleting 1000 or more open files at once.

  If one or more of the above situations occur, you can improve file system performance by changing open/close frequency or timing of the target file, or optimizing I/O size.

- Confliction between nodes, and directory access.

  - Frequently referencing a file from other nodes that is frequently updated from one or more nodes.

  - Frequently updating the same block in the same file from two or more nodes.

  - Frequently creating or deleting files or directories in the same directory from multiple nodes, and repeatedly monitoring directory updates with **readdir**(2) or **stat**(2).

  - When an **ls**(1) command with the **-l** option and a command requiring attributes of files in the directory such as **cp**(1), **du**(1), and **tar**(1) are both issued in a directory containing more than 10000 files.

If one or more of the above situations occurs, you can improve file system performance by changing the monitoring or access frequency, or dividing up files into separate directories.

CPU load or I/O load may be concentrated on the MDS node that manages the file system meta-data operates. A heavy load here indicates that operations, which require updating of the file system meta-data, such as file creation, deletion, or extension, are frequently being done. In such cases, file system throughput may improve by optimizing the layout of the meta-data area and update log area. An alternative is to increase CPU throughput of the MDS node operates.

4.3.2 Node layout

In the GFS Shared File System, up to two nodes can share one file system simultaneously. Therefore, file system processing can continue even if the node (on which MDS manages file system meta-data) fails. For such operation, you must select two nodes on which MDS can operate and set these nodes as **MDS node**s. From the MDS nodes, select the primary MDS node (on which the primary MDS usually operates), the secondary MDS node (on which the secondary MDS usually operates).

For details on the effects of MDS node settings, see "3.1 MDS Failure Recovery."

As previously described in "4.3.1 Effects of file system operation and system load," the following should be taken into consideration:

- CPU and I/O loads accompanying updating of the meta-data of the file system that occur on the MDS node operates

- CPU throughput, I/O throughput, memory capacity, and swap area

When activating the secondary MDS, activation of the primary MDS may take the time for a while.

4.3.3 Disk layout

In the GFS Shared File Systems, area of a disk configuring the file system consists of the following area:

- **meta-data area**

- **update log area**

- **data area**

The primary MDS references or updates two areas: the meta-data area and the update log area. The **AC** references or updates the data area.

A GFS Shared File System can consist of a single partition or multiple partitions. The file system with multiple partitions can improve I/O performance.

For example, in an environment where bottlenecks of intensive file data access will reduce the I/O performance of the file system, load balancing with access control will be enabled by configuring the file data area with multiple partitions. Also, the file data area is detached from the representative partition, and the meta data area and update log area can be detached from the file data area. This will increase the performance and throughput of file data I/O.

For more information about available partition configurations, see "2.1.6 Partition configuration."

4.3.4 LAN selection

In the GFS Shared File System, the **MDS** communications needed to obtain **meta-data** and to maintain consistency are processed through the LAN.

Set up the LAN path while keeping the following in mind.

- LAN traffic volume and LAN load

- Redundant Configuration Definition for a LAN fault

The GFS Shared File System has been designed in order not to increase the LAN load. But, the load intensity can be high under certain operation conditions. When a higher load is imposed on the LAN due to another processing, the more response time may be consumed in accessing the GFS Shared File System.

The LAN load status can be checked with the following commands:

- [RHEL6] **netstat**(8)

- [RHEL7] **ip**(8)

The use of the following is recommended when the load on a file system is large or when the response time is more important than other factors.

- High-speed LAN

- Private LAN

In the GFS Shared File System, specific nodes blockade the file systems if an error occurs in the communications between MDS and node AC. This is done because AC determines file system processing cannot be continued. To avoid problems in the communication paths, it is recommended to establish a multiplex communication path using Redundant Line Control Function of Global Link Services (hereinafter GLS).

See

For more information about the selection and settings of GLS redundant line switching mode, see "PRIMECLUSTER Global Link Services Configuration and Administration Guide: Redundant Line Control Function."

4.3.5 For future expansion

When a GFS Shared File System is created with **sfcmkfs**(8), the file system is configured, considered for future expansion by default. Check the parameter values in the following table and see if they are sufficient for the future expansion. For example, the partition addition function cannot add partitions exceeding the **-o maxvol** parameter of **sfcmkfs**(8). In this case, to expand the file system, the current data must be backed up and then deleted, a new file system must be created, and the data must be restored.

In preparation for file system capacity expansion for future needs, create a file system, estimating the maximum size for expansion.

Table 4.2 Parameter that you should confirm for future expansion

sfcmkfs(8) parameter	Default	Meaning
maxdsz	File data area size at specification	Maximum **file data area** size
maxnode	16	Maximum number of sharing nodes
maxvol	16	Maximum number of partitions

4.4 Backup Design

This section describes the hardware required to back up and restore file systems, and also describes how to back up and restore them in the GFS Shared File System.

- Required hardware

  A tape unit or hard disk unit is used to back up the GFS Shared File System.

- Backing-up and restoring procedure

  The GFS Shared File System backs up and restores file systems with the following methods:

  - Backing up and restoring the entire file system: **dd**(1)

  - Backing up and restoring the file system on a file basis: **cpio**(1L) or **tar**(1)

If the GFS Shared File System consists of multiple partitions, backing-up and restoring should be done on a **partition** basis. If tape devices and hard disk devices are configured, considered for a load distribution control of assigning I/O transactions, backing-up and restoring tasks can be distributed on all the nodes, and this will reduce the time required to back up and restore each partition.

See

For more information about backing-up and restoring in the GFS Shared File System, see "Chapter 13 File System Backing-up and Restoring."

Chapter 5 Management Partition

This chapter describes purposes and functions of management partitions.

For information on how to operate management partitions, see the following;

- "Chapter 8 Management Partition Operations (GUI)"

- "Chapter 9 Management Partition Operations (Command)"

5.1 Management Partition

Management partitions provide the important information for GFS Shared File System operation. Also, it is used for the "sfcfrmd" daemon to maintain data integrity. The management partition contains the following information:

1. Information of each node constituting the GFS Shared File System.

2. Information linking to a shared node and a shared device in each GFS Shared File System.

3. MDS and AC allocation information for each GFS Shared File System.

The node information of 1 must be set up using **sfcsetup**(8). The information of 2 and 3 can be set up, modified and added with **sfcmkfs**(8), **sfcadm**(8), **sfcnode**(8) and **sfcadd**(8). The **sfcgetconf**(8) command is available for backing-up the information within the management partition.

Figure 5.1 Configuration of the connection between nodes and the management partition
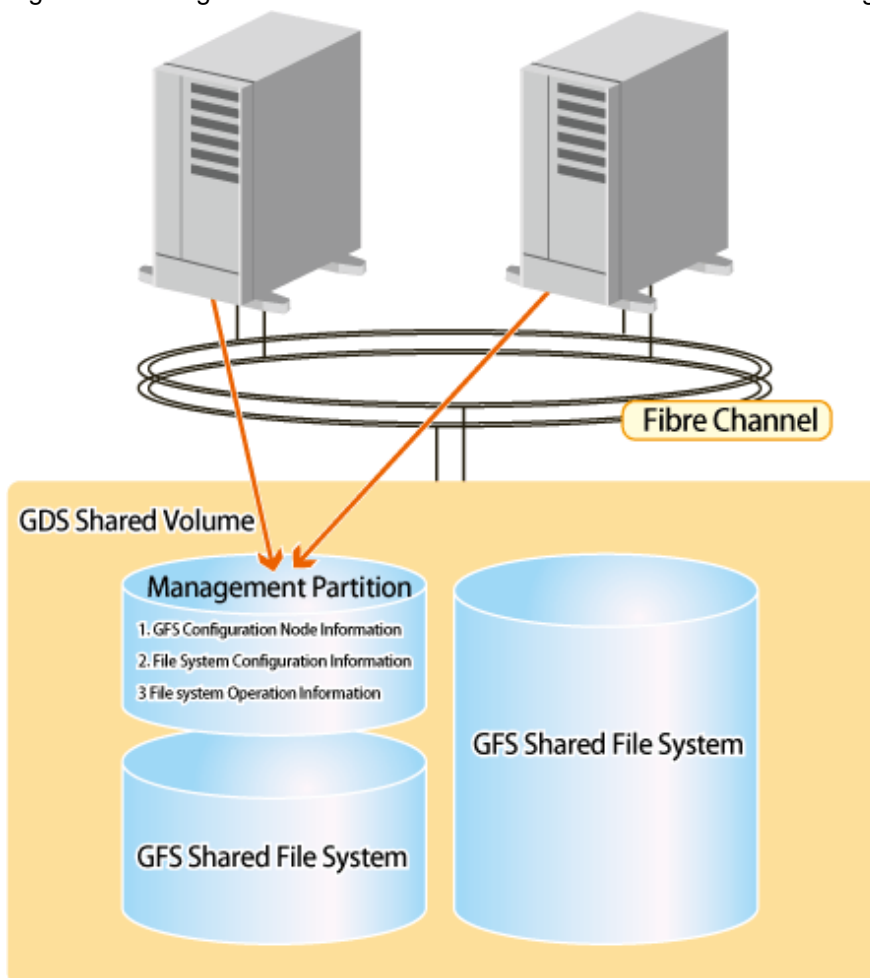
GDS : Global Disk Services
GFS : Global File Services

5.2 Resources Required for the Management Partition

The GFS Shared File System requires, in addition to the volumes that shares the file system, one **GDS** Shared volume that is exclusively for the management partition for each cluster system. Even when using two or more file systems, the number of required management partitions is one.

Specify a volume size of at least 40 megabytes for the management partition.

See
. . . .
For details on how to set up the GDS Shared disk system, see "PRIMECLUSTER Global Disk Services Configuration and Administration Guide."
. . . .

Note
. . . .
A setup of the volume which creates the management partition should set to "**shared**" the type of a disk class with which volume belongs. Also, all the nodes that share the GFS Shared File System must be specified as the **scope** of the disk class to which the volume belongs. Moreover, the settings must be such that the volume is automatically activated when the node is started.

Set "**no**" to [Exclusive use] of the disk class use including the volume where the management partition exists in.

If a disk class that includes volumes with management partitions is set in a GDS resource of a cluster application by mistake, the volume will not enter **ACTIVE** automatically during node startup. Therefore, the GFS Shared File Systems does not work.
. . . .

5.3 Notes on Management

If a failure occurs in the management partition, the GFS Shared File System within a cluster system stops. If I/O to the management partition fails, GFS may panic the node with an I/O error in order to maintain the consistency of the file system. To overcome this, it is recommended to mirror the volumes within the management partition.

Also, prior to changing the configuration of the file system, make a backup of the management partition information with **sfcgetconf**(8).

See
. . . .
For details on backup of the management partition, see "9.6 Backup of the management partition information" and "9.7 Restoring of the management partition information."
. . . .

Chapter 6 Starting and Exiting the Management View

GFS Management View operates under a Web environment.

As long as you have a web environment, you will have access via a browser from any location.

This chapter describes how to start and exit the GFS Management View.

For information about the following items, all of which are necessary for using the GFS Management View, see "PRIMECLUSTER Web-Based Admin View Operation Guide":

- Topology

- Preparation

- Starting Screen

- Operating Screen

Note that, because the GFS Shared File System is subject to the following functions, they must be usable prior to starting file system operation:

- Global Disk Services

- Cluster Foundation

For details on each function, see the following:

- "PRIMECLUSTER Global Disk Services Configuration and Administration Guide"

- "PRIMECLUSTER Cluster Foundation (CF) Configuration and Administration Guide "

When startup of the GFS Management View is enabled, you can operate the management partition with the GUI. For details, see the following:

- "Chapter 8 Management Partition Operations (GUI)"

If the management partition has already been created, the file system can be operated with the GUI. For details, see the following:

- "Chapter 10 File System Operations (GUI)"

If the management partition or the file system is to be operated from the command line, you do not have to make any Management View settings.

6.1 Starting Web-Based Admin View

If all preparations are complete, start Web-Based Admin View using the following procedures.

1. Take any of the following procedures to start Web-Based Admin View.

   - Startup from the command (if using Java Web Start) (recommended)

     1. Press <Win>+<R> key on the client and the [Run] dialog box opens.

     2. Input the **javaws** command with the format below and access to the management server.

     ```
     javaws http://host-name:port-number/
     ```

   - Startup from the Web browser (if using Java Web Start)

     1. Start the Web browser on the client.

     2. Specify the URL with the format below and access to the management server.

     ```
     http://host-name:port-number/
     ```

3. When using Microsoft Edge browser, click the <Open> button at the notification bar of file download completion which is displayed at the lower part of the browser.

When the notification bar of file download confirmation is displayed, click the <Save> button or the <Save As> button to save the file and then click the <Open> button.

📙 Note

- A downloaded file name may be something other than **WebStart.jnlp**.

- When starting Web-Based Admin View by taking this procedure, a tab remains in the Web browser after starting the Web-Based Admin View screen and the user login screen. Closing this tab will not cause any problems because it does not operate with Web-Based Admin View after starting the screen.

- Startup from the Web browser (If using Java Plug-in)

1. Start the Web browser on the client.

2. Specify the URL in the following format and access to the management server.

```
http://host-name:port-number/Plugin.cgi
```

host-name:

Specify the IP address or the host name for the client to connect to the primary or secondary management server.

port-number:

Specify "**8081**". If the port number is changed, specify the changed port.

📘 See

- For information about a client that starts the Web browser, see "PRIMECLUSTER Web-Based Admin View Operation Guide."

- For information on how to change the http port number, see "Change http port number" of "PRIMECLUSTER Web-Based Admin View Operation Guide."

📙 Note

If Web-Based Admin View is not started even when a host name of the management server is specified for *host-name*, specify an IP address of the management server directly.

2. If the Web-Based Admin View is started, the user login screen will appear as follows.

Figure 6.1 Login screen
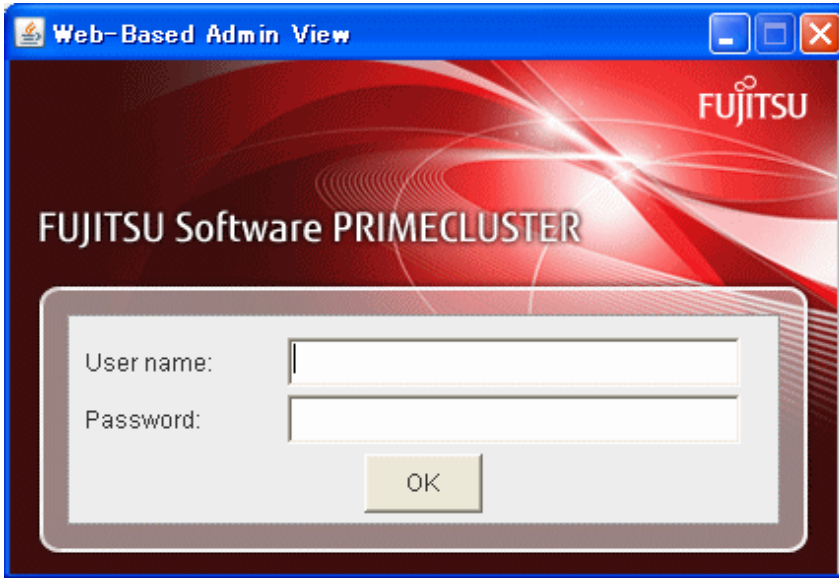
Enter user name and password for the management server then click the <OK> button.

 Note
....

- The user login screen cannot be closed by clicking the <X> button of upper-right corner in the screen.

- Either of the pop-up warning below may be displayed before the user login screen is displayed. These pop-up warnings are verification messages when the signed Java application is run with an unrestricted access privileges. These messages will not cause any problems. Press the <Run> button to continue the operation.

Security Warning

Do you want to run this application?

⚠️

Name: Web-Based Admin View

Publisher: FUJITSU LIMITED

Location: http://192.168.10.1:8081

Running this application may be a security risk

Risk: This application will run with unrestricted access which may put your computer and personal information at risk. The information provided is unreliable or unknown so it is recommended not to run this application unless you are familiar with its source

Unable to ensure the certificate used to identify this application has not been revoked.

More Information

Select the box below, then click Run to start the application

☑ I accept the risk and want to run this application   [Run]   [Cancel]

3. When the user authentication processing is completed, the top menu of the Web-Based Admin View will appear.

6.2 Web-Based Admin View Top Menu

After starting Web-Based Admin View has started, the [Web-Based Admin View operation Screen] appears.

Web-Based Admin View allows you to operate and manage GFS, set environment from the WWW screen.

Top menu refers to the Web-Based Admin View operation menu.

6.2.1 Web-Based Admin View Operation Menu Functions

The Web-Based Admin View operation menu supports the following functions related with the GFS Shared File System.

Table 6.1 Web-Based Admin View Operation Menu Description

Menu	Description
Global File Services	Set and operates GFS Shared File System.

GFS : Global File Services

🖐 Note

- Note that the contents of the Web-Based Admin View operation menu vary depending on the installed product.

- When a dialog is displayed because of an error in Web-Based Admin View, the picture on the right side of the Web-Based Admin View top menu also turns red. If other screens hide the dialog, clicking the red picture brings the dialog box to the front. To find some errors occurred surely, the picture on the right side of the Web-Based Admin View top menu must be always displayed.

6.2.2 Web-Based Admin View Tool Menu Functions

For information about the Web-Based Admin View Tool Menu, see "PRIMECLUSTER Web-Based Admin View Operation Guide."

6.3 Starting GFS Management View

Select the [Global File Services] Menu on the Web-Based Admin View operation menu, the GFS Management screen (hereafter
referred to as the main screen) opens.

From the main screen, you can create the management partition, create a file system, change a file system's configuration, and delete a file
system.

 See

....

For more information, see "Chapter 8 Management Partition Operations (GUI)" and "Chapter 10 File System Operations (GUI)."

....

Figure 6.2 GFS Management Main Screen

If there is more than one node, [Select Node] Screen appears on top of the main screen. Select the desired node. The selected node can
be changed from the main screen.

Figure 6.3 Select Node Screen
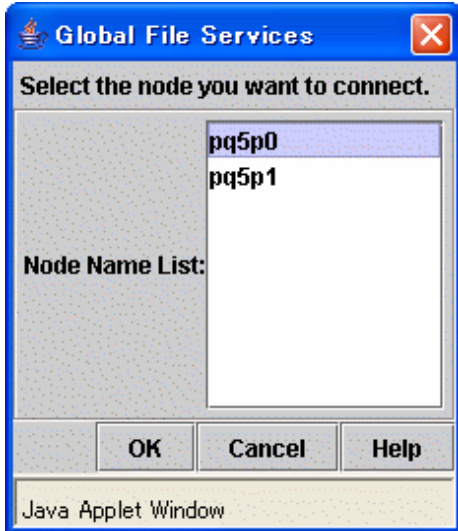
6.4 Exiting GFS Management View

To exit the GFS Management View, click <Exit> button on the [General] menu. The following message will appear.

Figure 6.4 Exit Screen

If you click <Yes> button, the Web-Based Admin View Screen (top menu) is displayed.

If you click <No> button, the main screen is displayed again.

6.5 Exiting Web-Based Admin View

Exit Web-Based Admin View, as follows:

1. Select the <Logout> button on the top menu.

2. The user login screen is displayed. Take either of the following steps while the user login screen remains displayed:

   - If Web-Based Admin View is started by using Java Web Start

     - Select <X> button of the upper right corner of the Web-Based Admin View menu screen.

   - If Web-Based Admin View is started by using Java Plug-in

     - Exit the Web browser.

     - Switch the screen automatically by specifying a new URL or other bookmark.

     - Click <Back> of the Web browser.

     - Select [Close tab] of the Web browser.

Note

If the user login screen continues to be displayed

The user login screen will remain displayed briefly after the Web browser exits; however, no additional action is required to close the user login screen as it disappears after a short period of time.

Chapter 7 Operation Management View Screen Elements

This chapter describes the screen elements of the GFS Management View.

7.1 Screen Configuration

Main screen

Select Global File Services from Web-Based Admin View, and the screen below appears.

From this main screen, you can create a file system, change a file system's configuration and delete a file system.

Screen configuration of main screen is shown below.

Figure 7.1 GFS Management Screen (main screen)
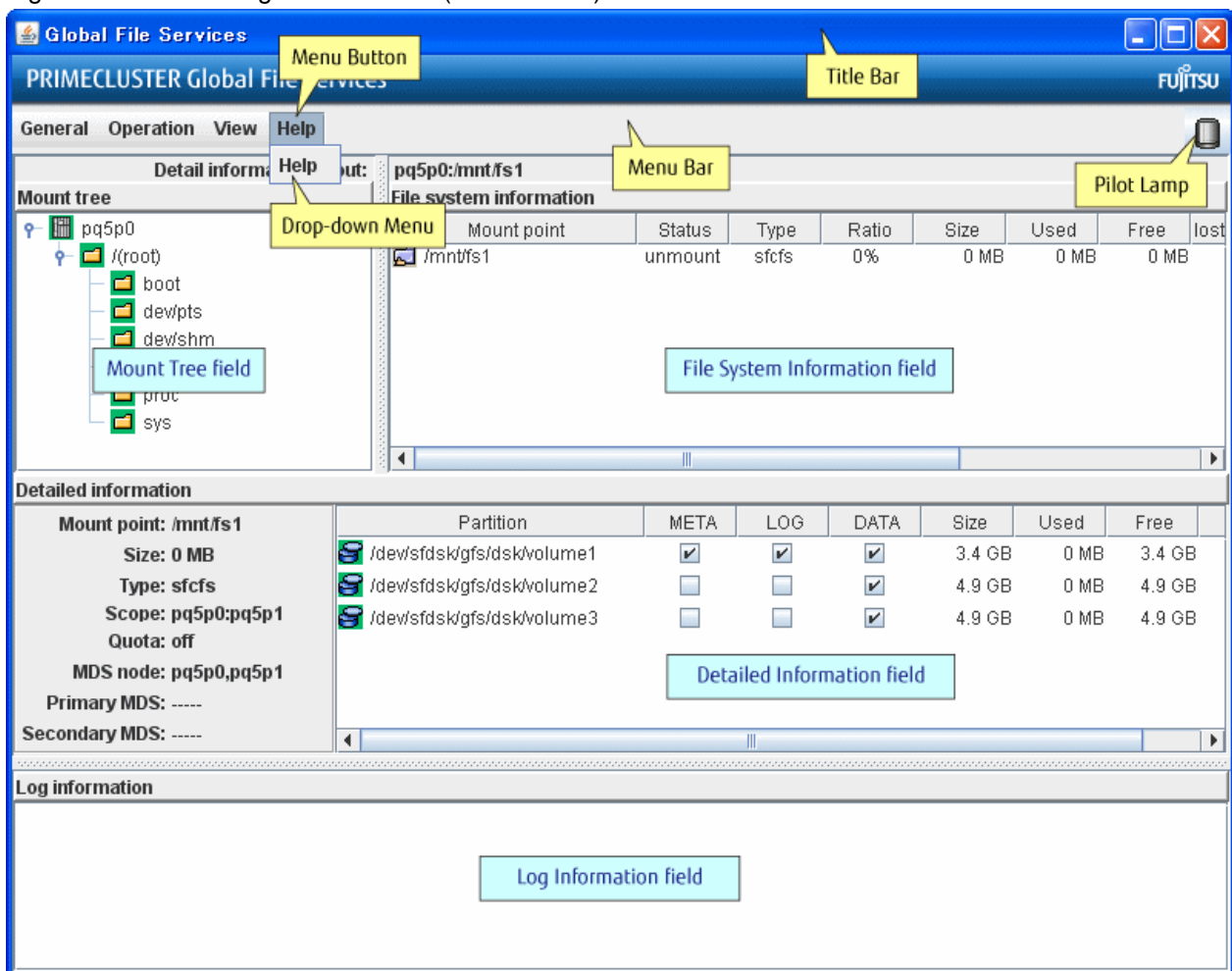
Mount Tree field

File systems accessible from the node selected with [Select Node] in the [General] menu on the main screen can be displayed in the tree structure. The tree displays only file systems listed in **/etc/fstab**.

When the GFS Management screen appears, the [Select Node] screen appears first to select target node.

File System Information field

Displays the file systems for the node/directory selected from the "Mount Tree field". If the file system is unmounted state, 0 is displayed in the size.

Each object has an icon representing the object type. If the file system type is sfcfs, icons are color-coded so that the status of the objects can be determined at a glance.

Detailed Information field

When the type of the file system selected in the "Mount Tree field" or the "File System Information field" is sfcfs, that file system's detailed information such as the placement of the MDS and the state of quota are displayed.

If the file system is unmounted state, 0 is displayed in the size in a left column.

The size of the data area is displayed in the size in the column of each right partition. Therefore, the size column is not displayed as for the partition without the data area.

Log Information field

Displays messages concerning the GFS Shared File System daemon program. The node name that the message was displayed is added at the head of each message.

Title Bar

Displays screen title [Global File Services].

Menu Bar

Display the menu buttons.

Menu Button

Allow you to control the objects selected on screen.

There are <General>, <Operation>, <View> and <Help>.

Drop-down Menu

When a menu button from the Menu Bar is selected, a Drop-down Menu appears.

Pilot Lamp

Shows the status of monitored objects.

The lamp can indicate the following status.

Table 7.1 Types of Pilot Lamp

Pilot Lamp	Status	Meaning
(Gray, lit up)	Normal	-
(Red, blinking)	Abnormal	The file system is abnormal. (Unavailable)
(Red, lit up)	Abnormal	When red blinking warning lamp is single-clicked
(Yellow, blinking)	Alarm	The utilization rate of the file system exceeds the threshold or the file system is abnormal at other node.
(Yellow, lit up)	Alarm	When yellow blinking warning lamp is single-clicked

When a field cannot be fully displayed, move a mouse cursor to that part that is not being displayed clearly. A popup display will appear.

In a pop-up list of the scope, information on the nodes sharing the file system is displayed in the node (host) name format. When several host names are specified, they are displayed in parentheses.

7.2 Menu Configurations and Functions

Each menu button has a Drop-down Menu; you can operate the selected object on screen. This section explains the menu configuration and functions.

7.2.1 General

Figure 7.2 General menu

Select Node

Select the node you want to operate, and click <OK> button. Only one node can be selected at a time.

If you do not need selecting node, click <Cancel> button.

Figure 7.3 [General]: [Select Node] Screen
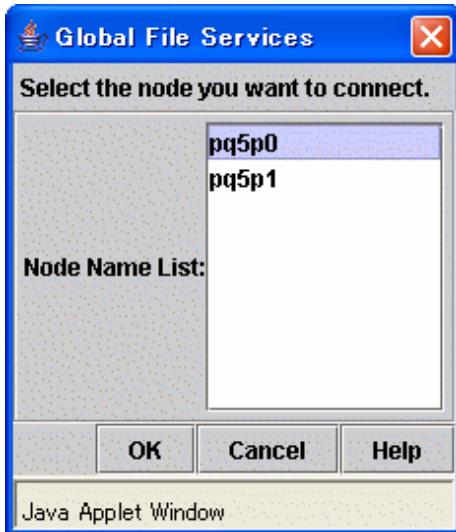
Exit

Exits Global File Services.

Figure 7.4 [General]: [Exit] Screen

7.2.2 Operation

Figure 7.5 Operation menu
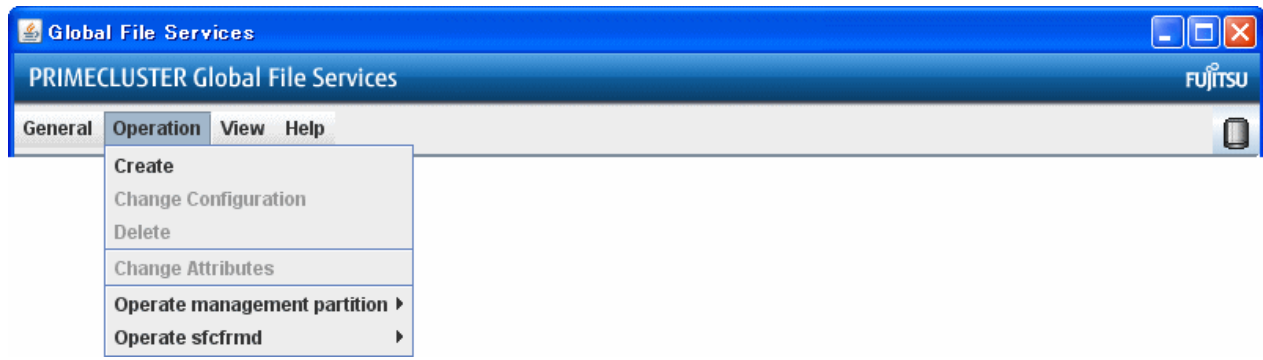
Create

Creates a file system.

See

For details, see "10.2.1 Creating a file system."

Figure 7.6 [Operation]: [Create] Screen
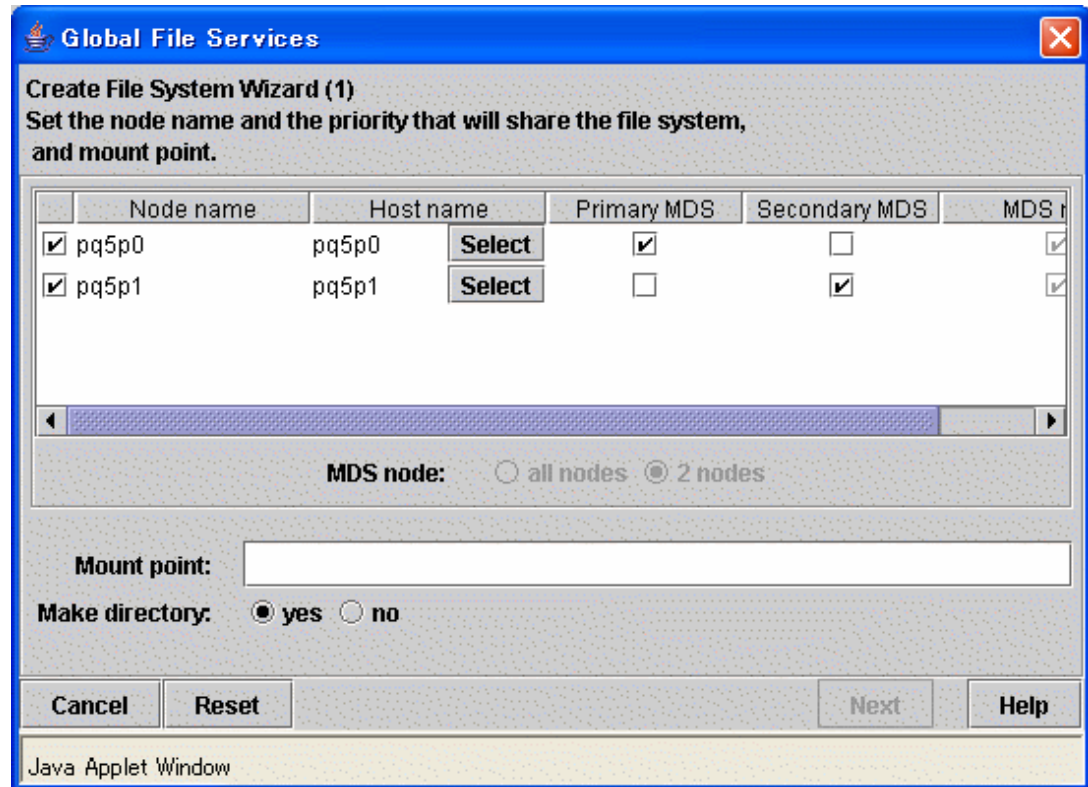
Change Configuration

Changes the partition configuration of a file system.

See

For details, see "10.3.2 Changing the file system configuration (partition addition)."

Figure 7.7 [Operation]: [Change Configuration] Screen
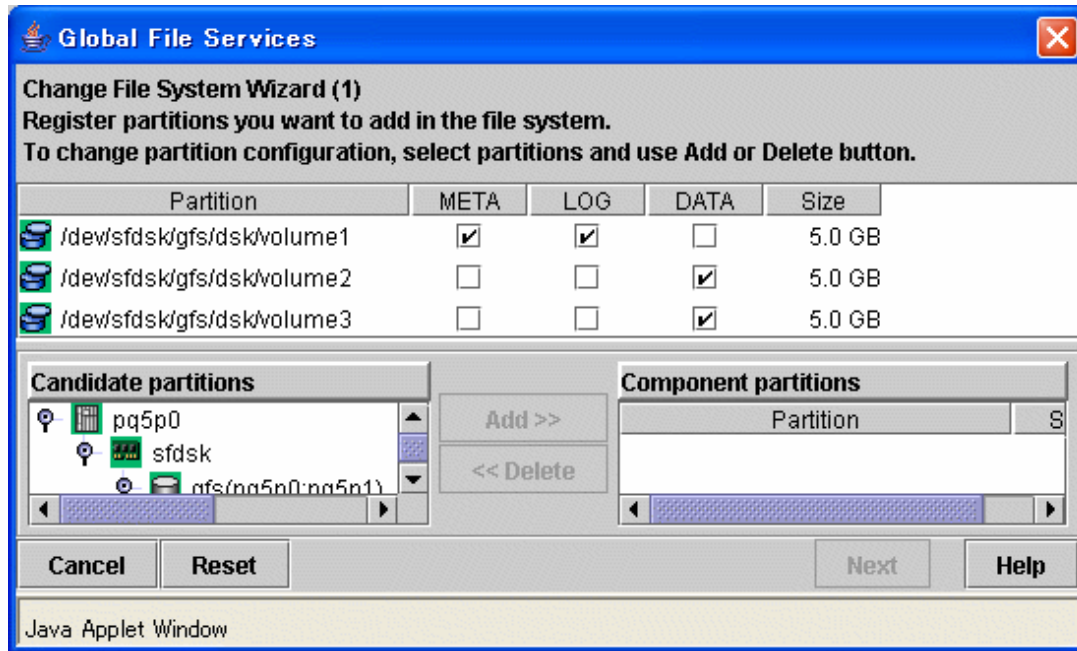
Delete

Deletes a file system.

🗂 **See**

. . . .

For details, see "10.4.1 Deleting the file system."

. . . .

Figure 7.8 [Operation]: [Delete] Screen
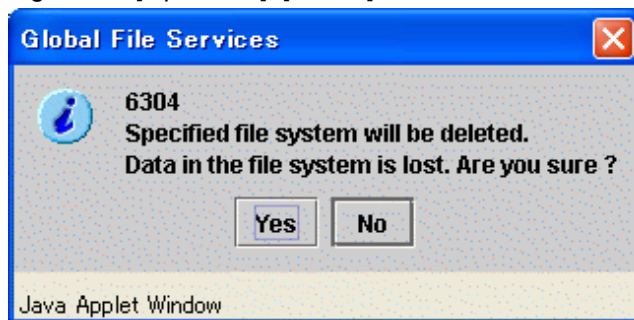
Change Attributes

Changes the mount information, share information, and detailed information for a file system.

🗂 **See**

. . . .

For details, see "10.3.1 Changing the file system attributes" and "10.3.3 Changing the shared node information."

. . . .

Figure 7.9 [Operation]: [Change Attributes] Screen
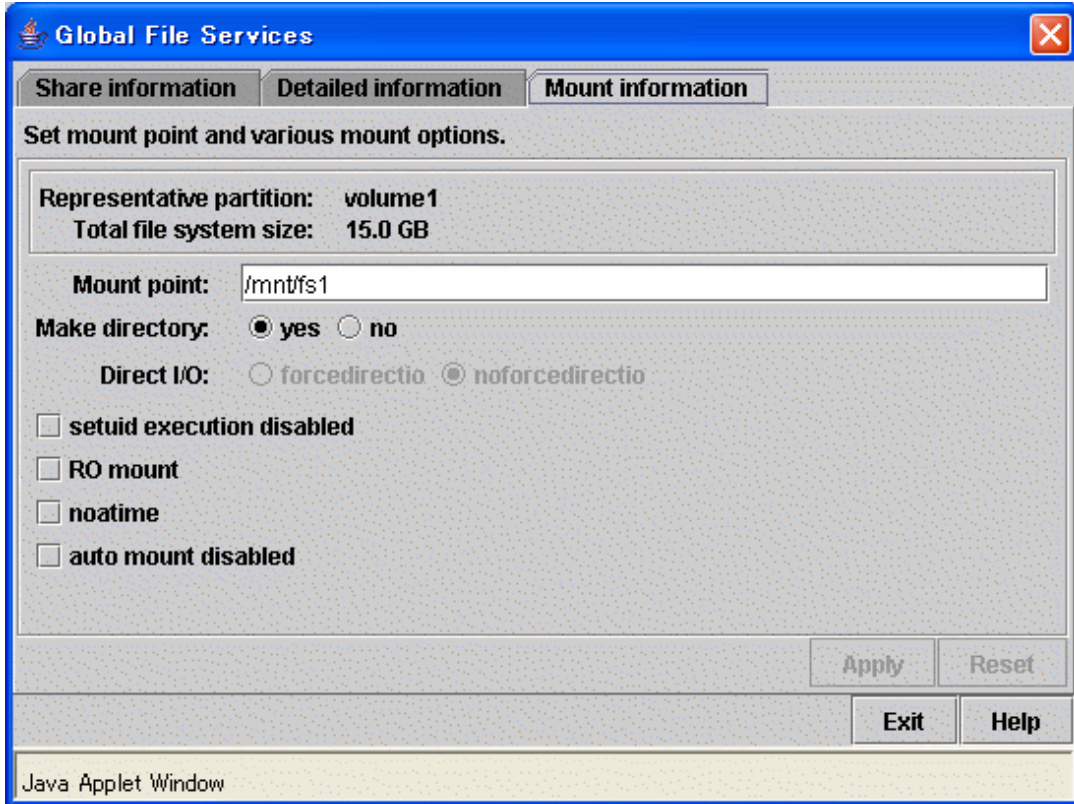
Operate management partition

You can create the management partition, register and add node information.

![See icon] **See**
. . . .
For details, see "8.2 Creating the management partition" and "8.3 Adding node configuration information to the management partition."
. . . .

Figure 7.10 [Operation]:[Operate management partition]:[Create] Screen 1
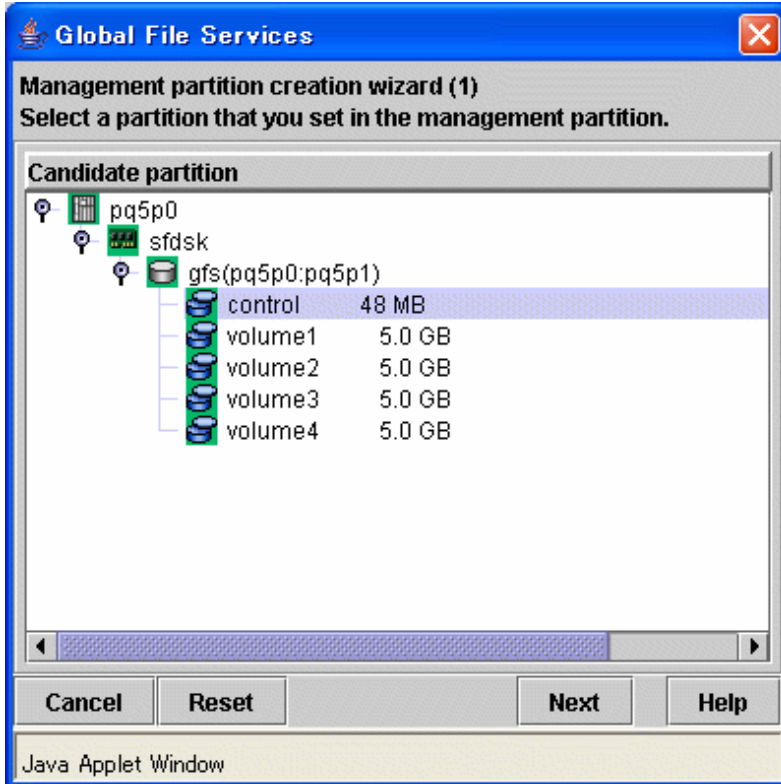
Figure 7.11 [Operation]:[Operate management partition]:[Create] Screen 2
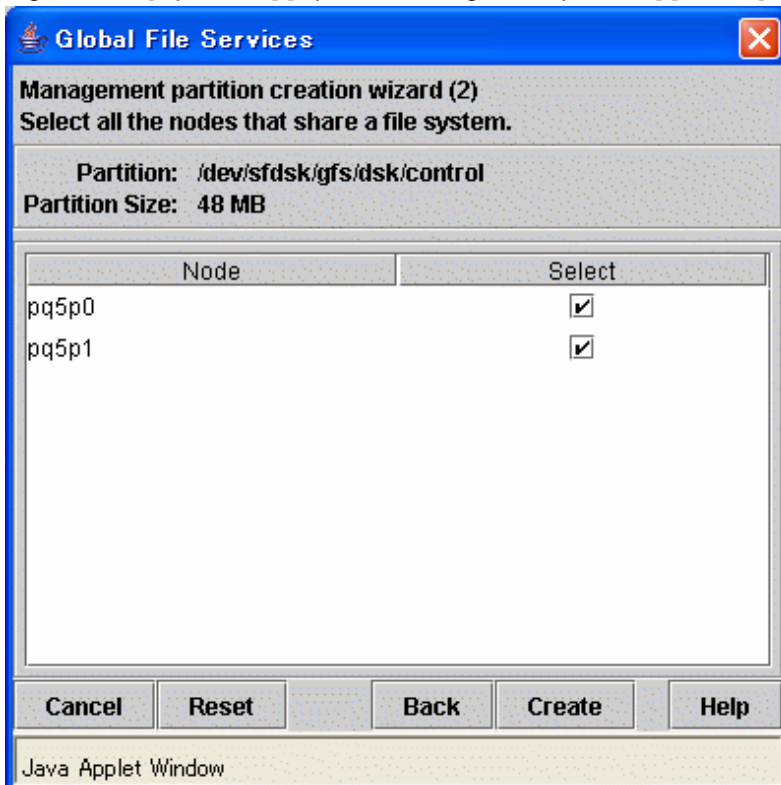
Figure 7.12 [Operation]:[Operate management partition]:[Add node] Screen
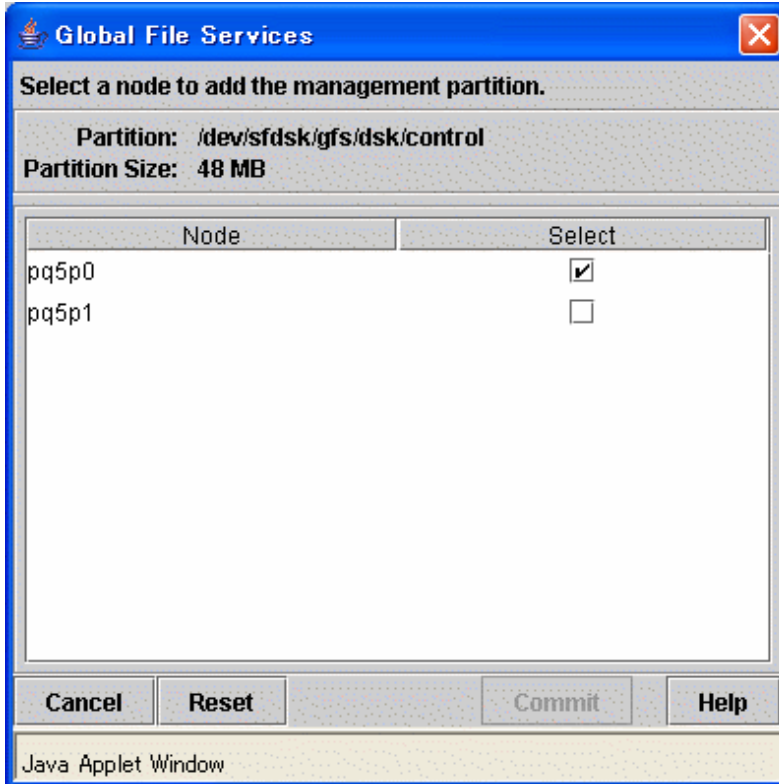
Operate sfcfrmd

sfcfrmd daemon can be started or stopped.

![See] **See**

. . . .

For details on the operation, see "8.3.4 Starting sfcfrmd daemon", "8.3.2 Stopping sfcfrmd daemon."

. . . .

Figure 7.13 [Operation]: [Operate sfcfrmd]:[Start] Screen
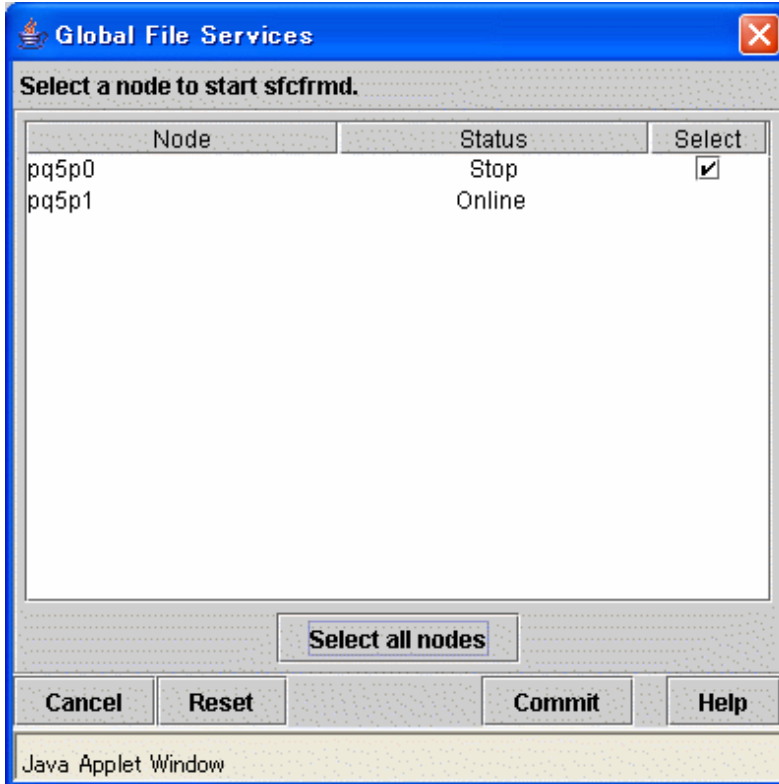
Figure 7.14 [Operation]: [Operate sfcfrmd]:[Stop] Screen
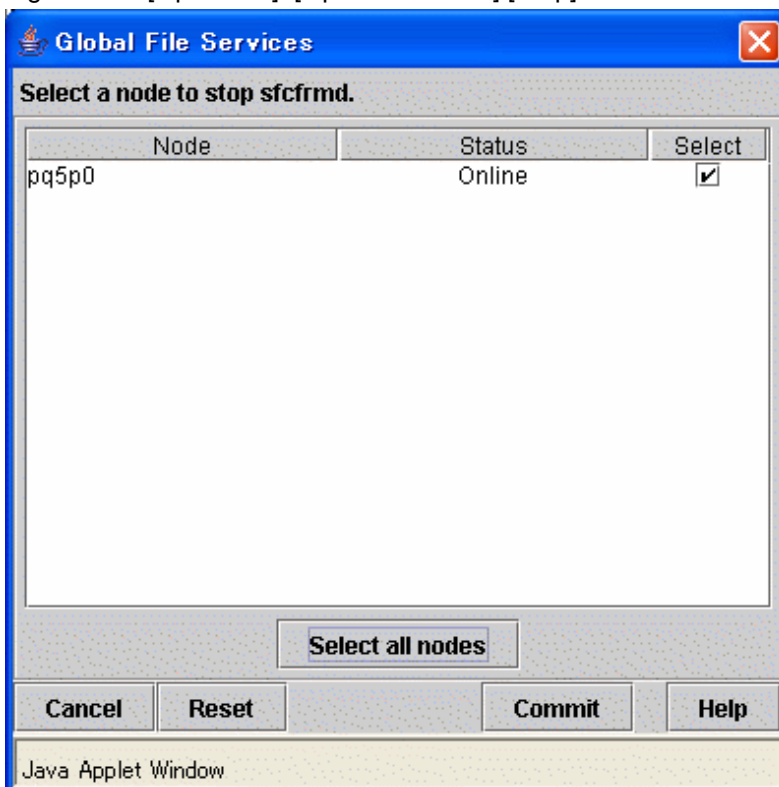
7.2.3 View

Figure 7.15 View menu
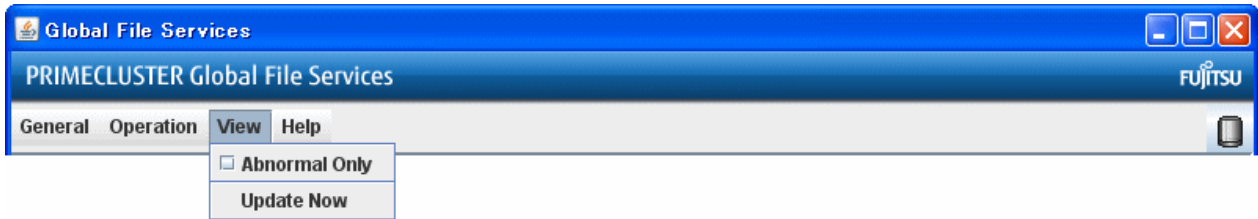
Abnormal Only

Only displays file systems with abnormalities.

Update Now

Display the latest file system information.

7.2.4 Help

Figure 7.16 Help menu

Help

Displays help information.

7.3 Icon Types and Object Status

GFS Management View uses icons to show object types and status.

The status and icons are shown below.

1. Node

Icon	Status	Meaning
	-	-

2. Adapter

Icon	Status	Meaning
	-	-

3. File system (sfcfs)

Icon	Status	Meaning
(Green)	Normal	The file system is working normally on all nodes.
(Yellow)	Alarm	The usage rate of the file system exceeds the threshold or an error occurred at another node.
(Red)	Abnormal	An error occurred in the local node.

Icon	Status	Meaning
(Blue)	In transition	The file system is in mounted or unmounted.
(Light brown)	Inactive	The file system is in unmounted.

4. File system (ext2, ext3, etc.)

Icon	Status	Meaning
(Green)	Normal	-
(Red)	Abnormal	-

5. Physical Disk

Icon	Status	Meaning
	-	-

6. Partition

Icon	Status	Meaning
	-	-

Note

File system status

The GFS Shared File System (file system type: sfcfs) is displayed as normal until it is accessed, and then the abnormality is detected. Even if the file system is unmounted, the icon shows the status.
When the file system type is ext2, ext3, etc., the status is not shown under an unmounted status.

Chapter 8 Management Partition Operations (GUI)

This chapter describes how to operate the management partition by the GFS Management View.

Note that, because the GFS Shared File System is subject to the following functions, they must be usable prior to starting file system operation:

- Global Disk Services

- Cluster Foundation

See
. . . .

For details on each function, see the following:

- "PRIMECLUSTER Global Disk Services Configuration and Administration Guide"

- "PRIMECLUSTER Cluster Foundation (CF) Configuration and Administration Guide "
. . . .

To execute the management partition operation, the GFS Management View settings must have been made in advance. For details, see the following:

- " Chapter 6 Starting and Exiting the Management View"

When you want to create a file system immediately after you create a management partition, see the following:

- " Chapter 10 File System Operations (GUI)"

- " Chapter 11 File System Operations (Command)"

8.1 Flow of Operations

This section describes the flow of Management Partition operations of the GFS Management View.

8.1.1 Creating the management partition

The following figure describes the flow of operations for creating a file system. For details on operations in the following figure, see "8.2 Creating the management partition."

Figure 8.1 Operation flow for creating the management partition
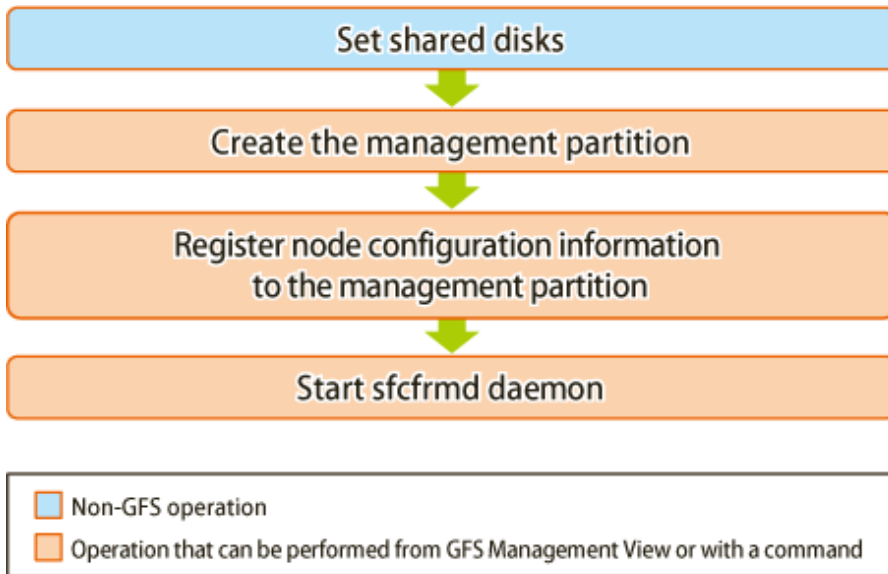
8.1.2 **Adding node configuration information to the management partition**

The following figure describes the flow of operations for adding node configuration information to the management partition. For details on operations in the following figure, see "8.3 Adding node configuration information to the management partition."

Figure 8.2 Operation flow for adding node configuration information to the management partition
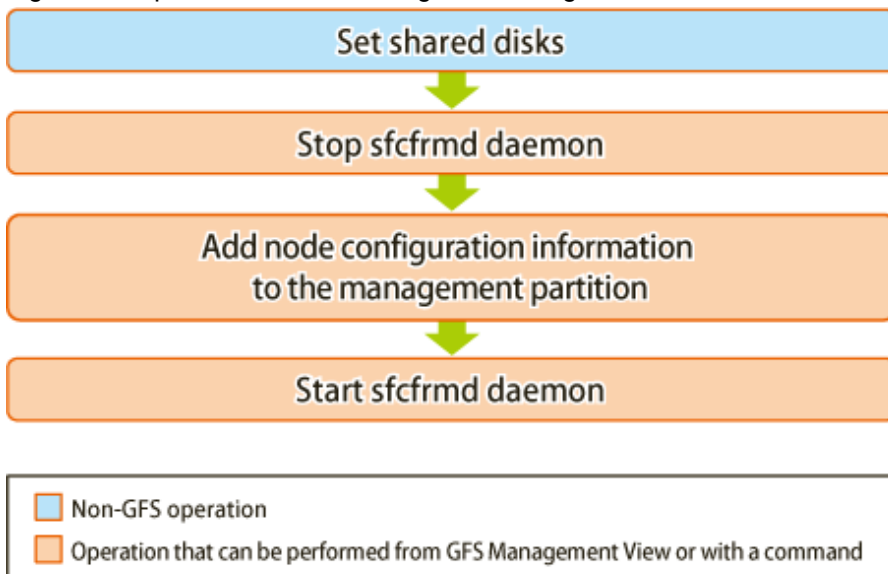
8.2 **Creating the management partition**

This section describes how to create the management partition.

Note

Before creating the management partition, confirm that a cluster partition error has not occurred. If a cluster partition error has occurred, fix the cluster partition error first.

See

For information on how to check whether a cluster partition error has occurred and, how to fix a cluster partition error, see "C.3.4 Corrective action when the sfcfrmd daemon is not started."

8.2.1 Setting shared disks

The device used as a management partition of a GFS Shared File System needs to fulfill the following conditions.

- Size should be 40 megabytes or more. (At least 40 megabytes will be required for the management partition.)

- It is the GDS logical volume and a status should be "**ACTIVE**." (Check with the **sdxinfo**(8))

- It is shared from all the nodes that use a file system. (Check with the **sdxinfo**(8))

See

For details on the **sdxinfo**(8), see "Appendix B.1.6 sdxinfo" in "PRIMECLUSTER Global Disk Services Configuration and Administration Guide."

GDS logical volume used as management partition should be set up as follows.

- Set up the type of a disk class with which GDS logical volume belongs as shared.(Change GDS class attributes)

- Set all the nodes that share GFS Shared File System as the scope of the disk class to which logical volume belongs.(Change GDS class attributes)

See

For information on logical volume operation, see the relevant item in "Operation" "GDS Management View" in "PRIMECLUSTER Global Disk Services Configuration and Administration Guide."

8.2.2 Creating the management partition, registering node configuration information and starting sfcfrmd daemon

Creating the management partition and registering node configuration information can be done after selecting [Operation]:[Operate management partition]:[Create] on the GFS Management View.

Figure 8.3 Management partition creation wizard (1)
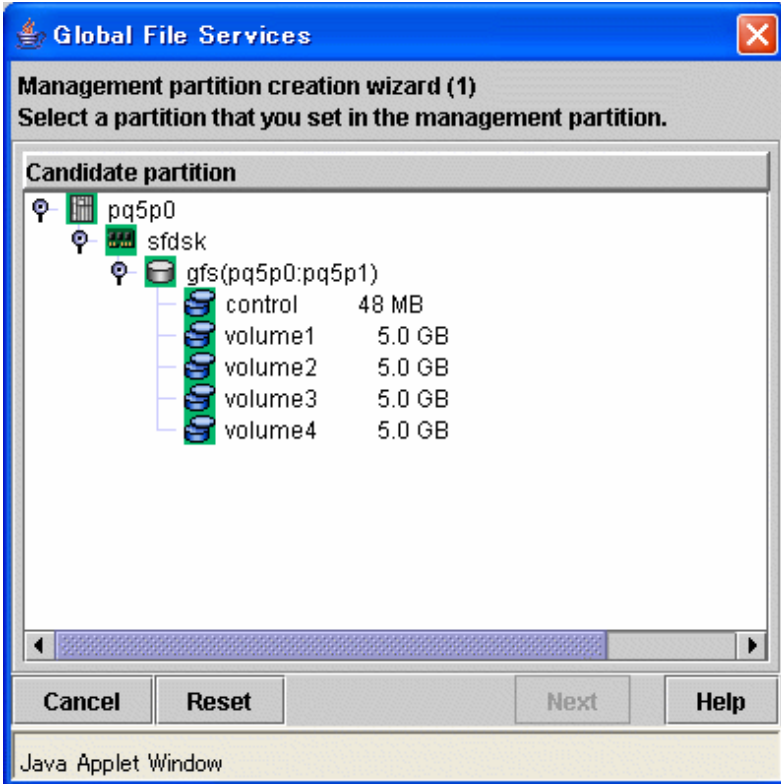
In the [Management partition creation wizard (1)] screen, select partition used as management partition.

In [Candidate partition], partitions that can use as management partition at the node executing management partition creation are displayed. If the partition used as management partition is selected with the check mark, it becomes possible to click the <Next> button.
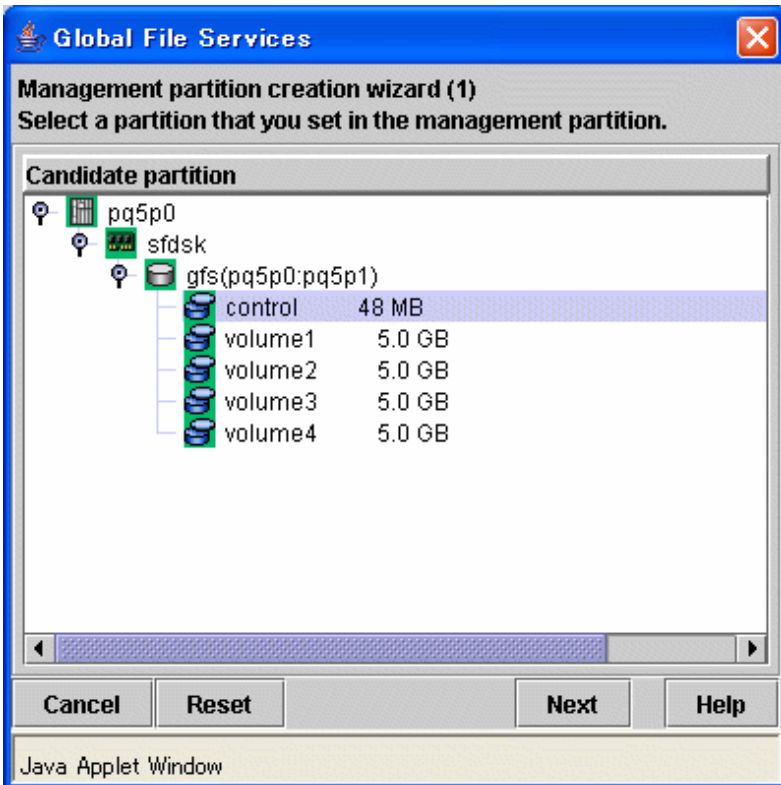
After selecting the partition use as management partition from [Candidate partition], it will go to [Management partition creation wizard (2)] screen if you click <Next> button.

To stop the creation of the management partition, click the <Cancel> button.

To restore the selection state to the default value, click the <Reset> button.

📒 Note
....
Before attempting to create the management partition, first stop sfcfrmd daemon.
....

📒 Note
....
Partitions that are used as management partitions cannot be initialized from the GFS Management View. Please do with **sfcsetup**(8). For details, see "9.2.2 Initializing of the management partition."
....

Figure 8.4 Management partition creation wizard (2)
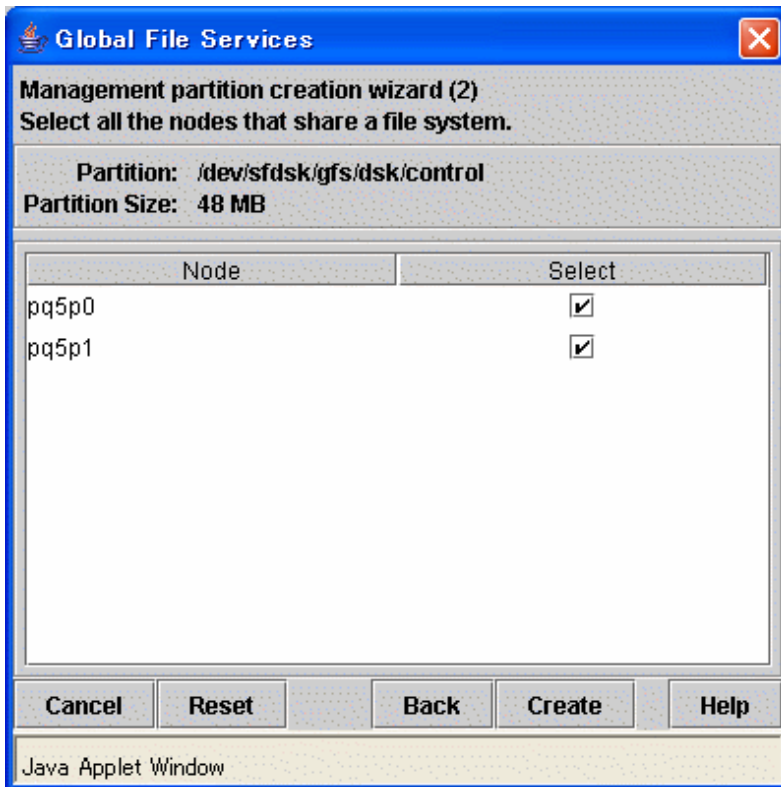
In the [Management partition creation wizard (2)] screen, select (with the check mark) nodes to register into management partition selected in the [Management partition creation wizard (1)] screen.

In [Node], the node included in the scope of the disk class to which the GDS logical volume selected as a management partition at the [Management partition creation wizard (1)] screen belongs is displayed. By the default, [Select] of all nodes are checked.

📒 Note
....
The node selected as a connecting node in [Select Node] screen appearing at the time of starting the GFS Management View needs to surely register as management partition. Therefore, the check mark of [Select] cannot be deselected.
....

To stop creation processing of the management partition, click the <Cancel> button.

To restore the selection state to the default value after change [Select] value, click the <Reset> button.

Click <Back> button when returning to the screen that selects the partition creating as a management partition.

When the check state of [Select] button of a node registering with a management partition is checked and the <Create> button is clicked, the following messages will be displayed.
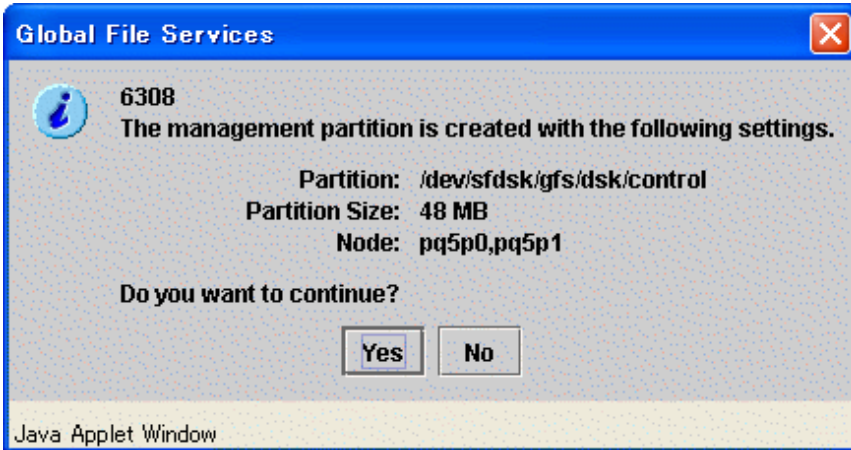
Processing will be interrupted, if you click the <No> button of the displayed message.

Creating management partition is started, if you click <Yes> button. It cannot be interrupted after creating management partition is started.
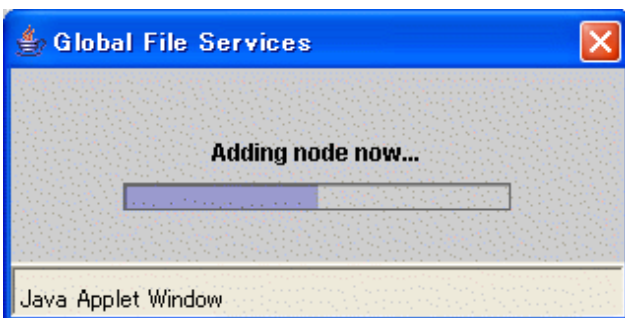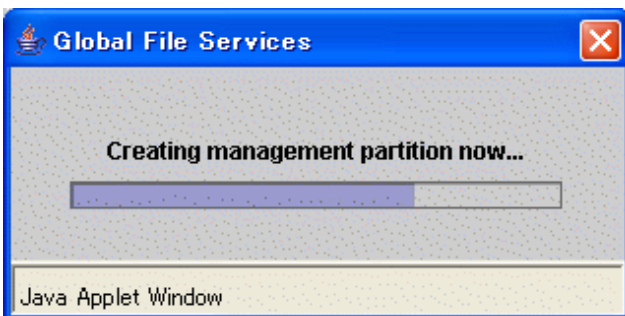
Start sfcfrmd daemon on the node registered with management partition.
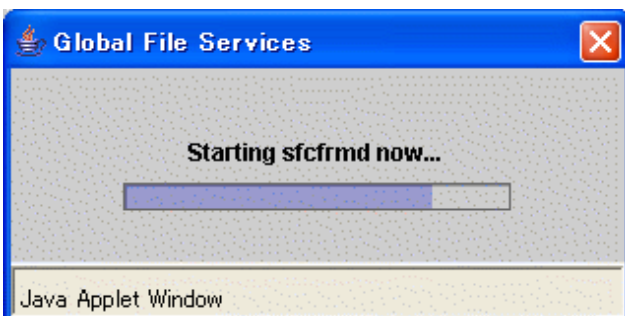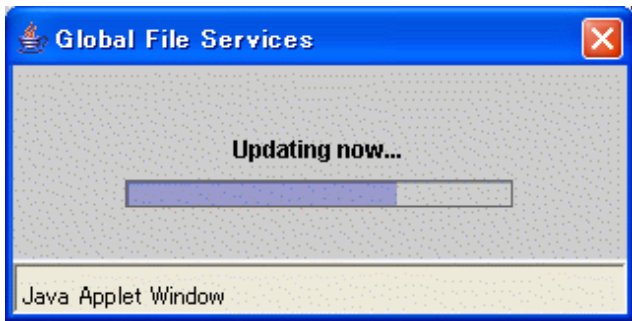
The following message appears and management partition creation is completed.

8.3 Adding node configuration information to the management partition

This section describes how to add node configuration information to the management partition.

🔰 Note

....

You cannot add a configuration node while the sfcfrmd daemon is started.

....

🔰 Note

....

Before adding node configuration information to the management partition, confirm that a cluster partition error has not occurred. If a cluster partition error has occurred, fix the cluster partition error first.

....

📘 See

....

For information on how to check whether a cluster partition error has occurred and, how to fix a cluster partition error, see "C.3.4 Corrective action when the sfcfrmd daemon is not started."

....

8.3.1 Setting shared disks

The node added as configuration node needs to be included in the scope of disk class to which GDS logical volume shared by all nodes registered to management partition belongs.

📘 See

....

For details, see "8.2.1 Setting shared disks."

....

8.3.2 Stopping sfcfrmd daemon

You can stop sfcfrmd daemon by selecting [Operation]:[Operate sfcfrmd]:[Stop] on the GFS Management View.

You must stop sfcfrmd daemon in all nodes when node configuration information is added to management partition.

![Note icon] Note
....
Unmount all GFS Shared File Systems before stopping sfcfrmd daemon.
....

![Information icon] Information
....
If node configuration information is added to management partition without stopping sfcfrmd daemon, sfcfrmd daemon under running is stopped automatically under processing of adding node configuration information to management partition.
....

Figure 8.5 sfcfrmd daemon stop node selecting dialog box
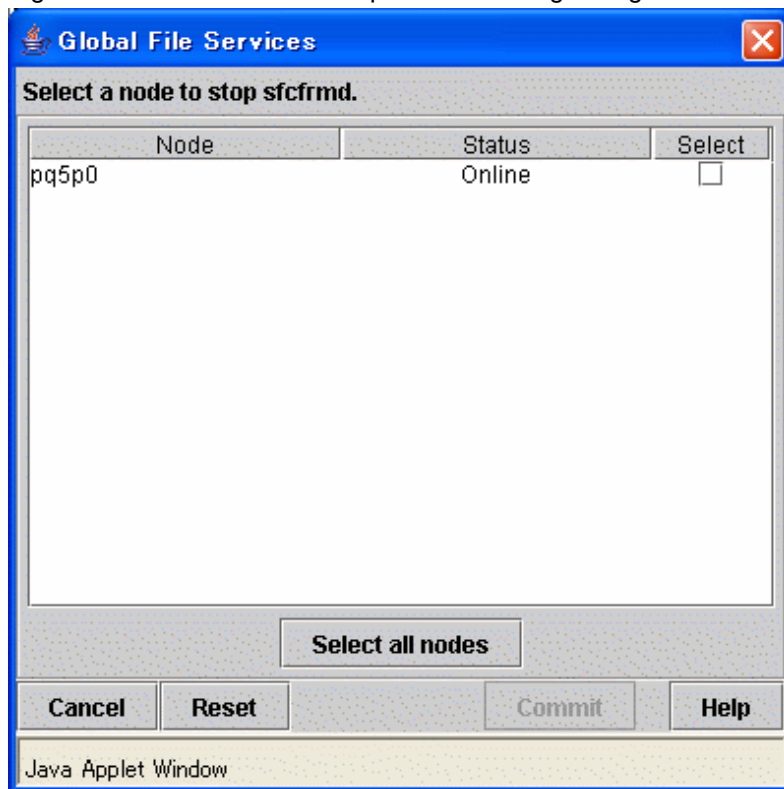
In the [sfcfrmd daemon stop node selecting dialog box], select the check mark of node stopping sfcfrmd daemon.

In [Node], all nodes registered to management partition are displayed. If you select [Select] of node stopping sfcfrmd daemon, <Commit> button can be clicked.
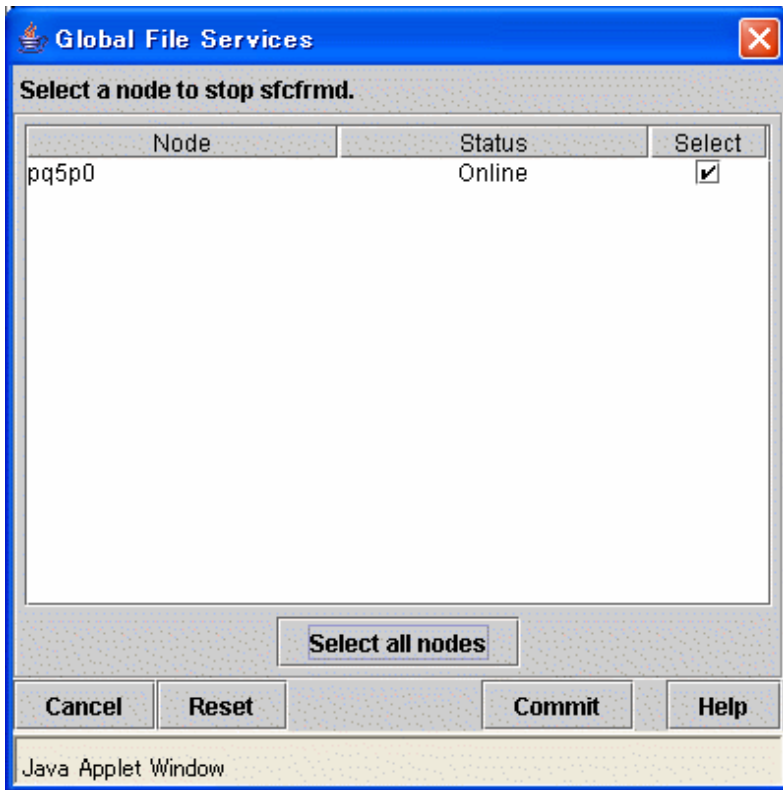
If you click the <Select all nodes> button, all the nodes will be checked.

To restore the selection state to the default value, click the <Reset> button.

To cancel stopping sfcfrmd daemon, click the <Cancel> button.

When you click <Commit> button, the following messages will be displayed.
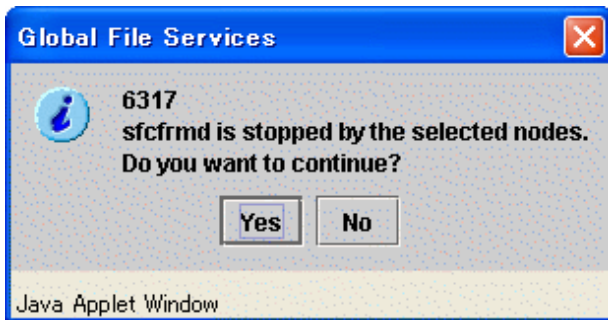
Processing will be interrupted, if you click the <No> button of the displayed message.

Stopping sfcfrmd daemon is started, if you click <Yes> button. It cannot be interrupted after stopping sfcfrmd daemon is started.
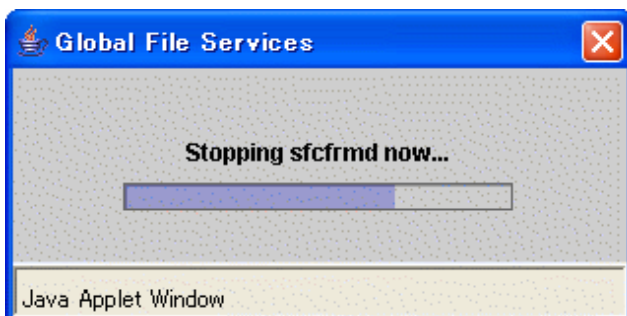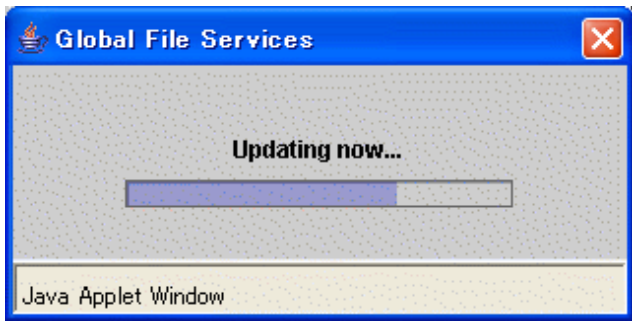
The following message appears and stopping sfcfrmd daemon is completed.
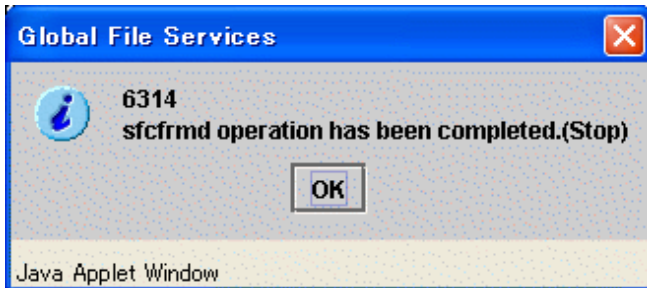
8.3.3 Adding node configuration information to the management partition

Node configuration information can be added to the management partition by selecting [Operation]:[Operate management partition]:[Add node] on the GFS Management View.

If the node which newly added to management partition is selected as a connecting node in [Select Node] screen appearing at the time of starting the GFS Management View, following message will appear when [Operation]:[Operate management partition]:[Add node] menu is selected.
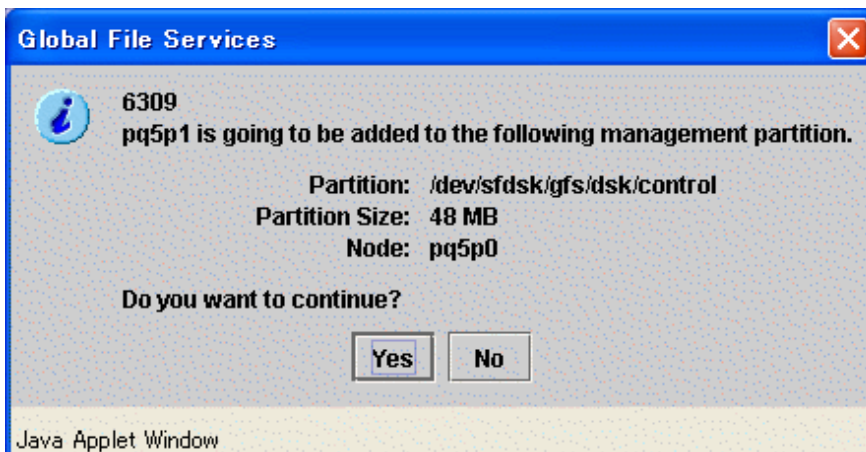
Processing will be interrupted, if you click the <No> button of the displayed message.

It will go to the screen that selects the node added to a management partition if a <Yes> button is clicked.

If the node which is already registered into the management partition is selected as a connecting node in [Select Node] screen appearing at the time of starting the GFS Management View, message does not appear and it will go to the screen that selects the node added to a management partition.

Figure 8.6 Management partition - Node selection dialog box
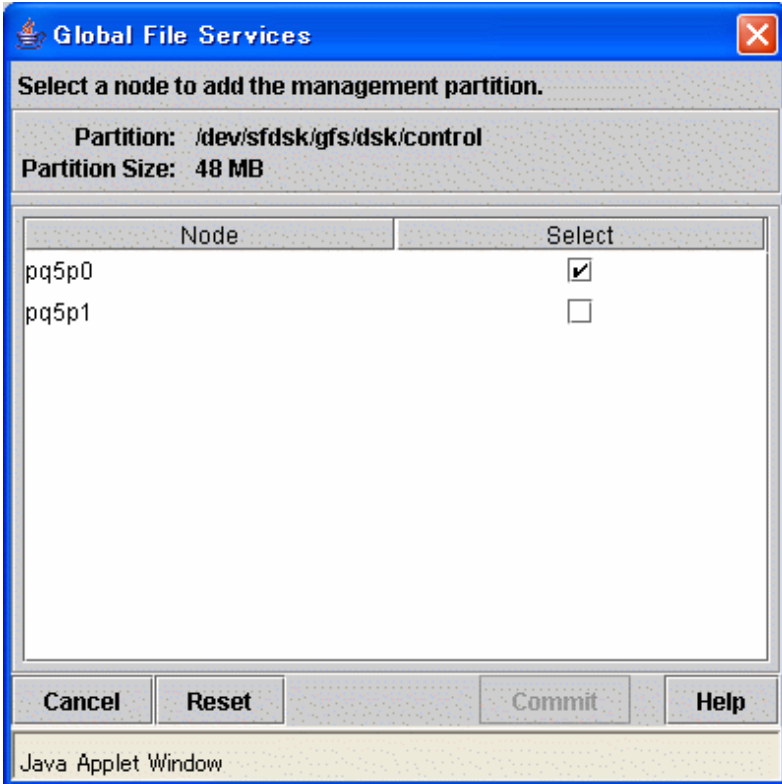
In the [Management partition - Node selection dialog box], select the node that adds the node configuration information to management partition.

In [Node], all nodes included in the scope of the disk class to which the GDS logical volume used as a management partition belong are displayed. By the default, it checks as follows.

- When the node which is already registered into the management partition is selected as a connecting node in [Select Node] screen appearing at the time of starting the GFS Management View.

  [Select] of the node registered into management partition is checked.

- When the node which newly added to management partition is selected as a connecting node in [Select Node] screen appearing at the time of starting the GFS Management View.

  [Select] of the node registered into management partition and the node connected in [Select Node] screen is checked.

If [Select] of node added to management partition is selected with the check mark, it becomes possible to click the <Commit> button.
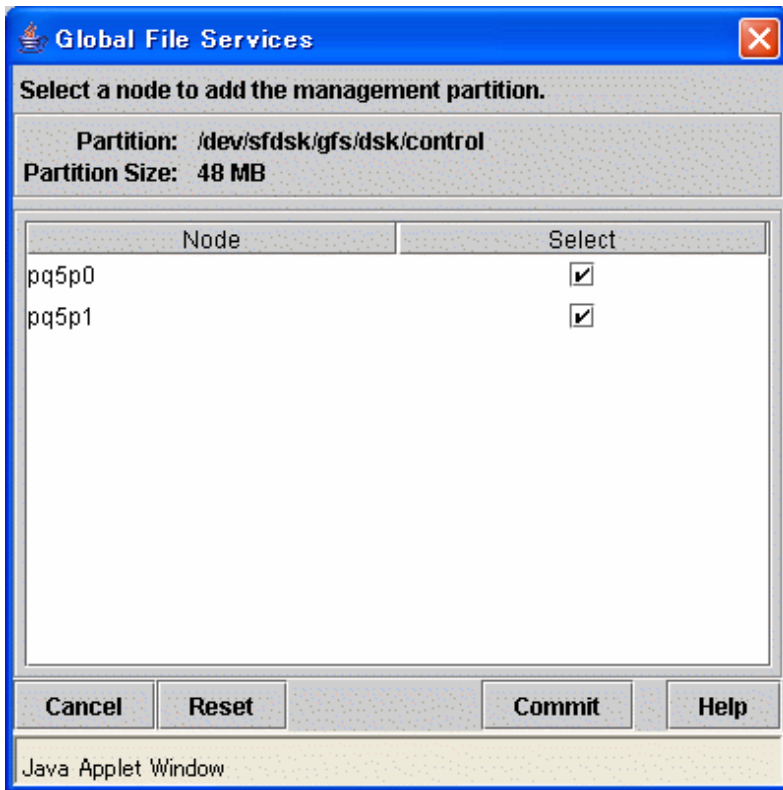
To stop to add nodes to the management partition, click the <Cancel> button.

To restore the selection state to the default value, click the <Reset> button.

After selecting (with the check mark) [Select] of added node and the <Commit> button is clicked, the following messages will be displayed.
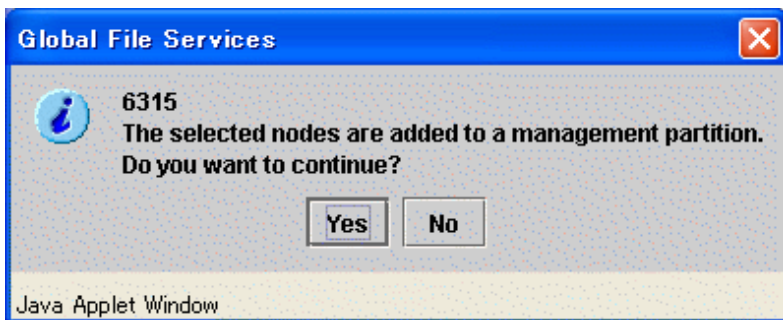
Processing will be interrupted, if you click the <No> button of the displayed message.

Adding node to management partition is started, if you click <Yes> button. It cannot be interrupted after adding node to management partition is started.
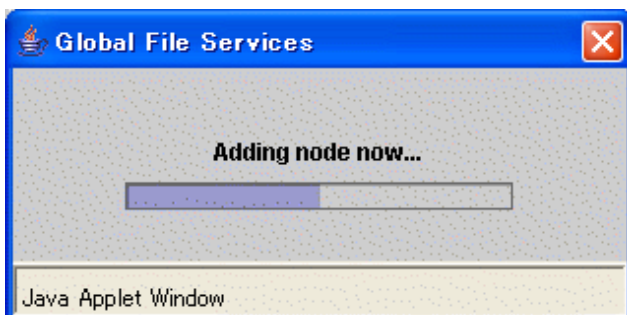
sfcfrmd daemon of the node newly added to management partition is started.
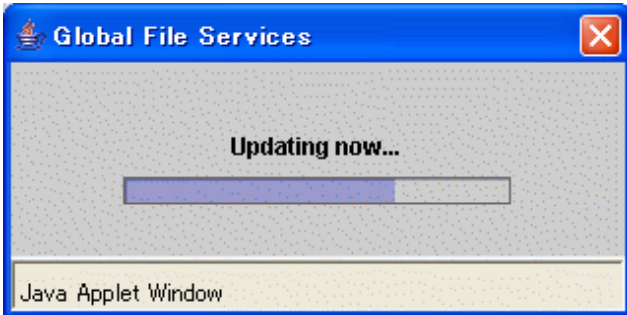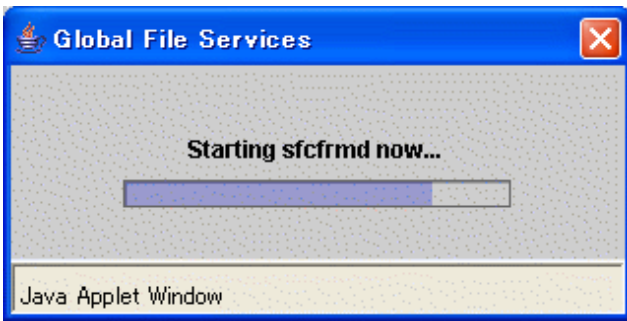
The following message appears and addition of the node to management partition is completed.
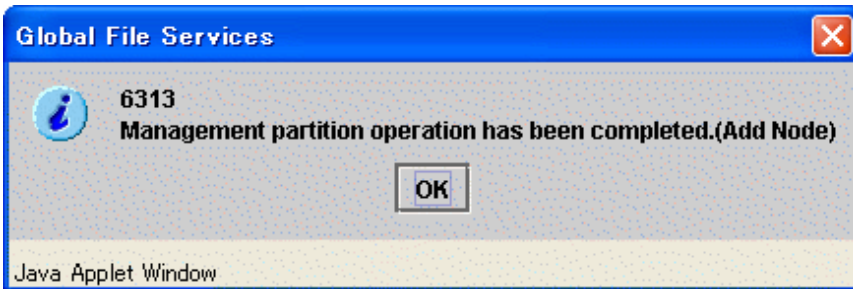
📝 Note

. . . .

Since the GFS Management View does not support the deletion of node configuration information from the management partition, you cannot clear the check mark corresponding to a node that is already registered.

. . . .

8.3.4 Starting sfcfrmd daemon

You can start sfcfrmd daemon by selecting [Operation]:[Operate sfcfrmd]:[Start] on the GFS Management View.

You must start sfcfrmd daemon in all nodes after node configuration information is added to management partition.

📝 Note

. . . .

To start sfcfrmd daemon, you must register the node configuration information to the management partition in advance.

Only when you stop sfcfrmd daemon manually before adding the node configuration information to management partition, start sfcfrmd daemon.

. . . .

In the [sfcfrmd daemon start node selection dialog box], select node starting sfcfrmd daemon.

Figure 8.7 sfcfrmd daemon start node selection dialog box
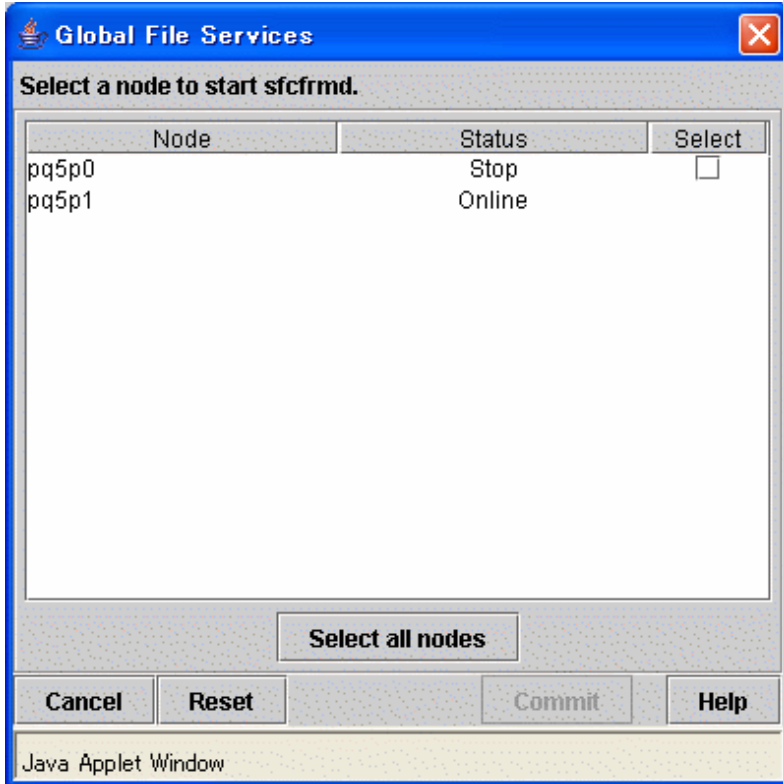
In [Node], all nodes registered to management partition are displayed. If you mark [Select] of node starting sfcfrmd daemon, <Commit> button can be clicked.
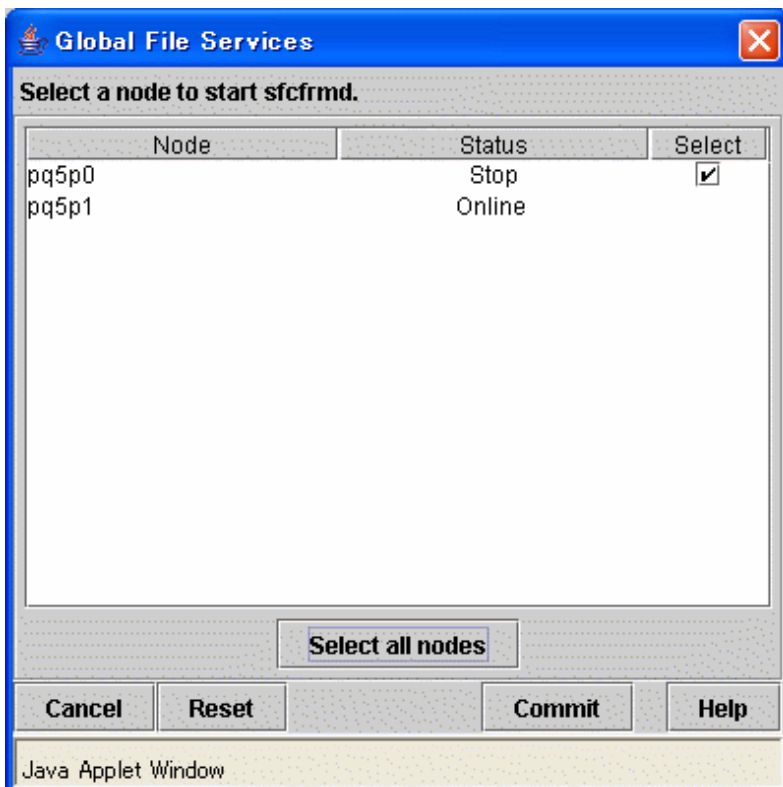
If you click the <Select all nodes> button, all the nodes will be checked.

To cancel starting sfcfrmd daemon, click the <Cancel> button.

To restore the selection state to the default value, click the <Reset> button.

When you click <Commit> button, the following messages will be displayed.
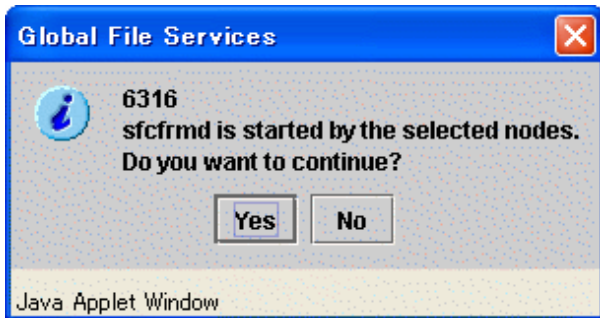
Processing will be interrupted, if you click the <No> button of the displayed message.

Starting sfcfrmd daemon is started, if you click <Yes> button. It cannot be interrupted after starting sfcfrmd daemon is started.
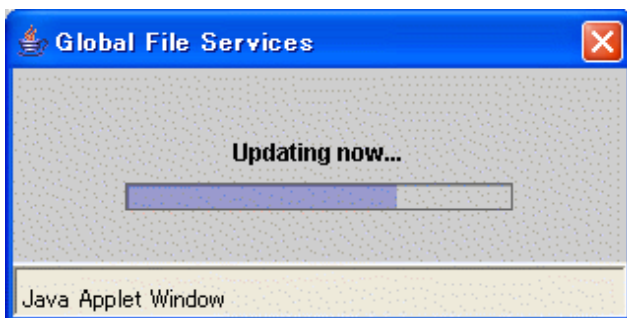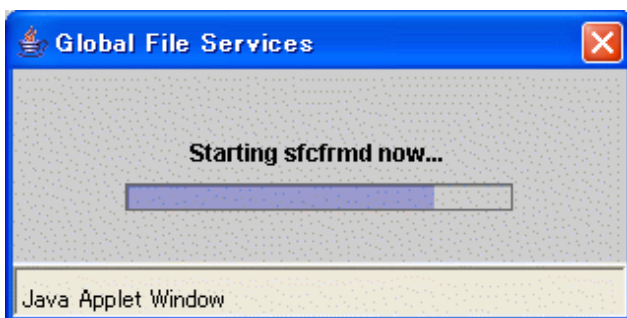
The following message appears and starting sfcfrmd daemon is completed.

Chapter 9 Management Partition Operations (Command)

This chapter describes how to operate the management partition by commands.

Note that, because the GFS Shared File System is subject to the following functions, they must be usable prior to starting file system operation:

- Global Disk Services

- Cluster Foundation

See

....

For details on each function, see the following:

- "PRIMECLUSTER Global Disk Services Configuration and Administration Guide"

- "PRIMECLUSTER Cluster Foundation (CF) Configuration and Administration Guide"

....

If you want to create a file system immediately after creating a management partition, see either of the following:

- "Chapter 11 File System Operations (Command)"

- "Chapter 10 File System Operations (GUI)"

9.1 Flow of Operations

This section describes the flow of operations in management partition.

9.1.1 Creating the management partition

The following figure describes the flow of creating the management partition.

For details on operations in the following figure, see "9.2 Creating of the management partition."

Figure 9.1 Operation flow for creating the management partition
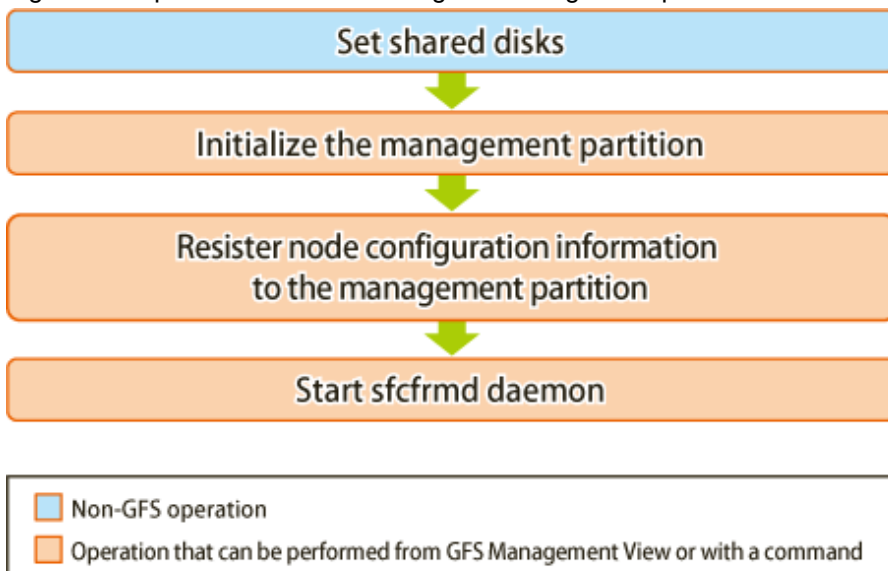
GFS : Global File Services

9.1.2 Adding node configuration information to the management partition

The following figure describes the flow of operations for addition of node configuration information to the management partition.

For details on operations in the following figure, see "9.3 Adding node configuration information to the management partition."

Figure 9.2 Operation flow for addition of node configuration information to the management partition
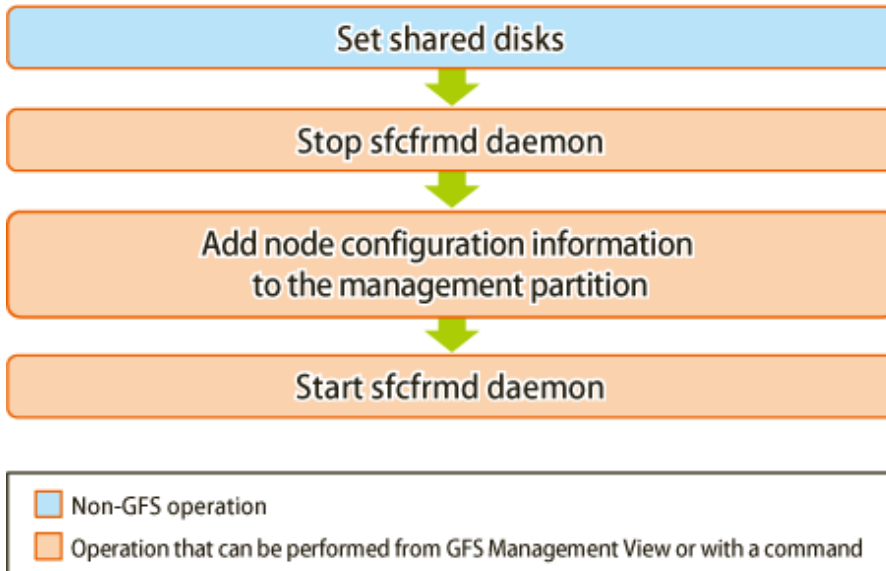
9.1.3 Deleting node configuration information from the management partition

The following figure describes the flow of operations for deleting node configuration information from the management partition.

For details on operations in the following figure, see "9.4 Deleting node configuration information from the management partition."

Figure 9.3 Operation flow for deleting node configuration information from the management partition
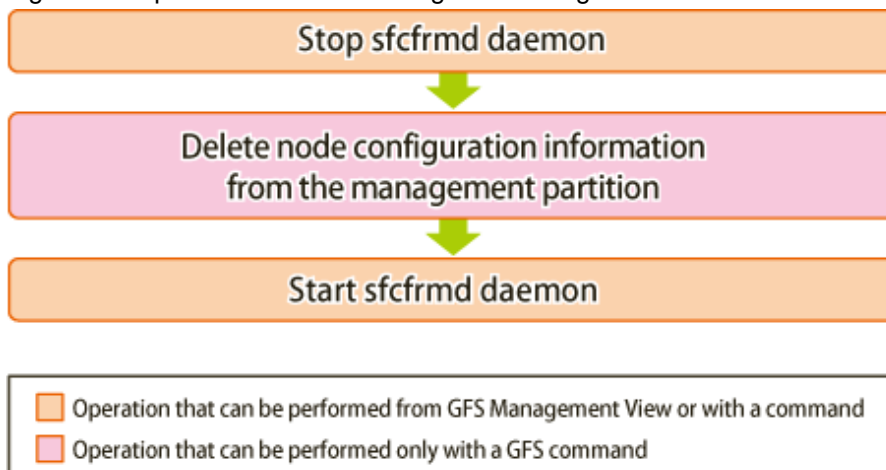
9.1.4 Changing the sfcfrmd daemon's startup mode registered in the management partition

The following figure describes the flow of operations for changing the sfcfrmd daemon's startup mode registered in the management partition.

For details on operations in the following figure, see "9.5 Changing the sfcfrmd daemon's startup mode registered in the management partition."

Figure 9.4 Operation flow for changing the sfcfrmd daemon's startup mode registered in the management partition
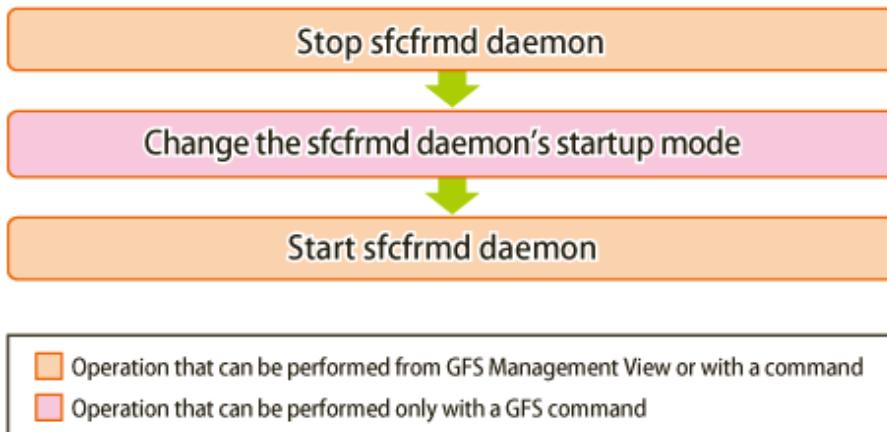
9.1.5 Backup of the management partition information

The following figure describes the flow of operations for backup of the management partition information.

For details on operations in the following figure, see "9.6 Backup of the management partition information."

Figure 9.5 Operation flow for backup of the management partition information
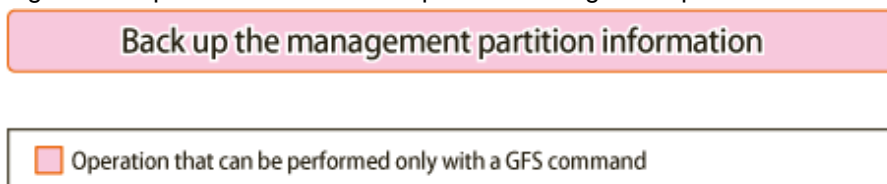
9.1.6 Restoring of the management partition information

The following figure describes the flow of operations for restoring of the management partition information.

For details on operations in the following figure, see "9.7 Restoring of the management partition information."

Figure 9.6 Operation flow for restoring of the management partition information
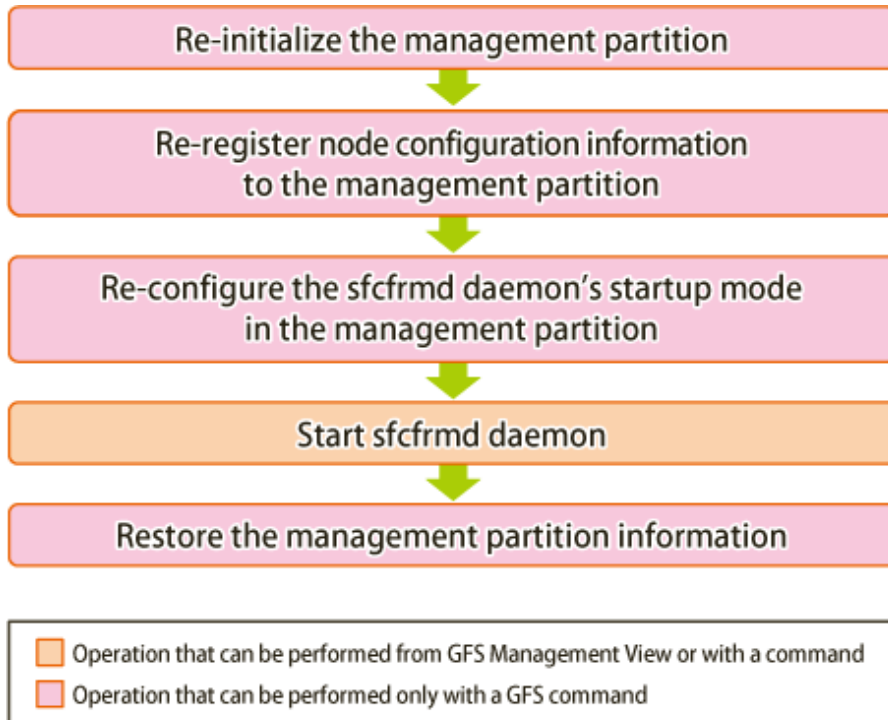
9.2 Creating of the management partition

This section describes how to create the management partition information.

📙 Note
....
Before creating of the management partition, confirm that a cluster partition error has not occurred. If a cluster partition error has occurred, fix the cluster partition error first.
....

📘 See
....
For information on how to check whether a cluster partition error has occurred and, how to fix a cluster partition error, see "C.3.4 Corrective action when the sfcfrmd daemon is not started."
....

9.2.1 Setting Shared disks

The management partition of the GFS Shared File System is created on the GDS logical volume of the shared device.

📘 See
....
For information about the shared device, see "8.2.1 Setting shared disks."
....

9.2.2 Initializing of the management partition

Use the **sfcsetup(8)** command with the **-c** option to initializing the management partition.

Note
....

To initialize the management partition, sfcfrmd daemon must have already been stopped on the cluster system.
....

The example of initializing /dev/sfdsk/gfs/dsk/control as a management partition is shown as follows.

```
# sfcsetup -c /dev/sfdsk/gfs/dsk/control <Enter>
```

Note
....

Please confirm whether to meet the following requirements beforehand in the case re-initialized the management partition.

- All the GFS Shared File Systems must be deleted.

- All the node information must not be registered in the management partition.
....

When a partition that is already initialized as the management partition is to be initialized again, execute the command with the **-f** option specified, as shown below:

```
# sfcsetup -c -f /dev/sfdsk/gfs/dsk/control <Enter>
```

Note
....

Initialization of a management partition sets the sfcfrmd daemon's startup mode to **wait**. Use the **sfcsetup(8)** command with the **-m** option to change the sfcfrmd daemon's startup mode.
....

9.2.3 Registering node configuration information to the management partition

sfcsetup(8) command with the **-a** option registers node configuration information to the management partition.

An example of registering node configuration information to the management information is shown below.

Note
....

Register node configuration information on all the nodes on which will be shared GFS Shared File Systems.
....

1. Registering a node in the management partition

```
# sfcsetup -a /dev/sfdsk/gfs/dsk/control <Enter>
```

2. The path name of the management partition that has been set up can be confirmed by executing the **sfcsetup(8)** command with the **-p** option specified.

```
# sfcsetup -p <Enter>
/dev/sfdsk/gfs/dsk/control
```

3. Registered node configuration information can be confirmed by executing **sfcsetup(8)** with no option specified.

```
# sfcsetup <Enter>
HOSTID      CIPNAME      MP_PATH
80380000    sunnyRMS     yes
80380001    moonyRMS     yes
```

9.2.4 Starting sfcfrmd daemon

The sfcfrmd daemon is started by all nodes so that operation may begin. Use the **sfcfrmstart(8)** command to start sfcfrmd daemon.

 See
....
For details on **sfcfrmstart(8)**, see sfcfrmstart(8).
....

 Note
....
To start sfcfrmd daemon, node configuration information must have been registered in the management partition.
....

Execute the command as follows on the node on which sfcfrmd daemon is to be started.

```
# sfcfrmstart <Enter>
```

9.3 Adding node configuration information to the management partition

This section describes how to add node configuration information to the management partition.

 Note
....
Before adding node configuration information to the management partition, confirm that a cluster partition error has not occurred. If a cluster partition error has occurred, fix the cluster partition error first.
....

 See
....
For information on how to check whether a cluster partition error has occurred and, how to fix a cluster partition error, see "C.3.4 Corrective action when the sfcfrmd daemon is not started."
....

9.3.1 Setting Shared disks

It is necessary to set nodes that will be added as configuration nodes in the scope of the disk class to which the GDS logical volume belongs. This GDS logical volume is shared among the nodes that are registered in the management partition.

9.3.2 Stopping sfcfrmd daemon

The sfcfrmd daemon should be stopped on all nodes in the clustered system to add node configuration information to the management partition. Use **sfcfrmstop(8)** to stop sfcfrmd daemon.

Note

Please unmount all GFS Shared File Systems before stopping a sfcfrmd daemon.

Execute the command as follows on all the nodes in the cluster system.

```
# sfcfrmstop <Enter>
```

9.3.3 Adding node configuration information to the management partition

sfcsetup(8) command with the **-a** option adds node configuration information to the management partition.

An example of adding node configuration information to the management information is shown below.

Perform the following operations at the adding node.

1. Add node configuration information to the management partition.

    ```
    # sfcsetup -a /dev/sfdsk/gfs/dsk/control <Enter>
    ```

2. The path name of the management partition that has been set up can be confirmed by executing the **sfcsetup(8)** command with the **-p** option specified.

    ```
    # sfcsetup -p <Enter>
    /dev/sfdsk/gfs/dsk/control
    ```

3. Added node configuration information can be confirmed by executing **sfcsetup(8)** with no option specified.

    ```
    # sfcsetup <Enter>
    HOSTID      CIPNAME      MP_PATH
    80380000    sunnyRMS     yes
    80380001    moonyRMS     yes
    ```

9.3.4 Starting sfcfrmd daemon

The sfcfrmd daemon is started by all nodes so that operation may begin.

9.4 Deleting node configuration information from the management partition

This section describes how to delete node configuration information from the management partition.

9.4.1 Stopping sfcfrmd daemon

The sfcfrmd daemon should be stopped on all nodes in the clustered system to deleting node information from the management partition.

See
....
For details of stopping sfcfrmd daemon, see "9.3.2 Stopping sfcfrmd daemon."
....

9.4.2 Deleting node configuration information from the management partition

When you delete node configuration information from the management partition, you will execute **sfcsetup(8)** with -d option. As a result, it is deleted node information that executed **sfcsetup(8)**, from the management partition. And it is deleted the management partition from executed **sfcsetup(8)** node.

Note
....
Please confirm whether to meet the following requirements beforehand when you delete node configuration information from the management partition.

- The GFS Shared File System that includes the shared node to be deleted does not exist.
....

On the node that is to be deleted, execute the command as shown below:

```
# sfcsetup -d <Enter>
```

9.4.3 Starting sfcfrmd daemon

The sfcfrmd daemon is started by all nodes so that operation may begin.

See
....
For details of starting sfcfrmd daemon, see "9.2.4 Starting sfcfrmd daemon."
....

9.5 Changing the sfcfrmd daemon's startup mode registered in the management partition

This section describes how to change the sfcfrmd daemon's startup mode.

9.5.1 Choosing the sfcfrmd daemon's startup mode

In the GFS Shared File System, monitoring daemons such as sfcfrmd, sfcprmd, sfcfsd, sfchnsd, and sfcpncd monitor the cluster state. The monitoring daemons also ensure that only a single primary MDS is running on a given cluster at one time, so file system access is consistent on all nodes.

If the state of some of the nodes in the cluster cannot be confirmed due to a cluster partition error, the node that is operating the primary MDS cannot be determined because a quorum does not exist. In this case, the GFS Shared File System services suspended startup of the sfcfrmd daemon, in order to ensure data consistency.

There are two available sfcfrmd daemon's startup modes following:

- wait

  When a node is started up, the startup of the sfcfrmd daemon will be suspended until a cluster quorum exists, and also startup of nodes will be suspended at the same time.
  And, when the CF is started up by the GUI, the startup of the sfcfrmd daemon is given up if a quorum does not exist.

- wait_bg

  When a node is started up or the CF is started up by the GUI, the startup of the sfcfrmd daemon will be suspended in the background until it can confirm that a cluster quorum exists. Node startup or CF startup will continue.

When the value is omitted (When initializing management partition), the default is "**wait**".

If the GFS Shared File System is used for cluster applications, please choose **wait**.
If the GFS Shared File System is not used for cluster applications, you can choose **wait_bg**. In this case, the cluster applications will be started without waiting for the GFS Shared File System become usable. Also, the time required for system startup can be reduced.

Note

- If you start some of the nodes after stopping all the nodes in the cluster, the state of nodes that are not operated in a cluster system cannot be confirmed. So, the startup of the sfcfrmd daemon is suspended until all nodes are operated in a cluster system.

- If **wait** is set and you want to start up CF by the GUI after stopping CF on all the nodes, see "12.6 How to start up CF by the GUI when a GFS Shared File System is used."

Information

When a node's startup is suspended, you can login to system using the network.

9.5.2 Stopping sfcfrmd daemon

To change the sfcfrmd daemon's start mode registered in the management partition, the sfcfrmd daemon must be stopped on all the cluster nodes.

See

For details of stopping sfcfrmd daemon, see "9.3.2 Stopping sfcfrmd daemon."

9.5.3 Changing the sfcfrmd daemon's startup mode

To change the sfcfrmd daemon's startup mode, use the **-m** option of **sfcsetup(8)**.

How to change the sfcfrmd daemon's startup mode to **wait_bg** is explained below.

1. Confirm the current the sfcfrmd daemon's startup mode.

   ```
   # sfcsetup -m <Enter>
   wait
   ```

2. Change the sfcfrmd daemon's startup mode.

   ```
   # sfcsetup -m wait_bg <Enter>
   ```

3. Confirm the sfcfrmd daemon's startup mode after change.

```
# sfcsetup -m <Enter>
wait_bg
```

9.5.4 Starting sfcfrmd daemon

Start the sfcfrmd daemon on all the nodes.

See
. . . .

For details of starting sfcfrmd daemon, see "9.2.4 Starting sfcfrmd daemon."
. . . .

9.6 Backup of the management partition information

This section describes how to back up the management partition information.

9.6.1 Backup of the management partition information

Use the **sfcgetconf(8)** command to make a backup of the management partition information.

The method of backup of the management partition information is explained below.

```
# sfcgetconf _backup_file_ <Enter>
```

See
. . . .

For details on **sfcgetconf(8)**, see sfcgetconf(8).
. . . .

Note
. . . .

To back up management partition information with **sfcgetconf(8)**, the sfcfrmd daemon should be operating by the execution node. If sfcfrmd daemon is not operating, start it by using **sfcfrmstart(8)**. For details, see "9.2.4 Starting sfcfrmd daemon."

Make a backup of the management partition information, after changing the configuration of the file system by **sfcmkfs(8)**, **sfcadd(8)**, **sfcadm(8)**, **sfcnode(8)**.
. . . .

sfcgetconf(8) generate a shell script _backup_file_:

The contents of the shell script are explained below.

```
# cat _backup_file_ <Enter>
#!/bin/sh
# This file is made by:
#    sfcgetconf _backup_file_
# Thu Mar 14 09:23:04 2011
#---- fsid : 1 ----
# MDS primary   (port) : host1 (sfcfs-1)
# MDS secondory (port) : host2 (sfcfs-1)
# MDS other            :
# AC       : host1,host2
# options :
# device  : /dev/sfdsk/gfs01/dsk/volume01
sfcadm -m host1,host2 -g host1,host2 -p sfcfs-1,sfcfs-1 /dev/sfdsk/gfs01/dsk/volume01
...
```

9.7 Restoring of the management partition information

This section describes how to restore of the management partition information.

9.7.1 Re-initializing the management partition

If a disk failure occurs in the GDS volume that was being used as the management partition, initialize the management partition in the GDS volume in which the disk unit was replaced, after you replace the disk unit.

See
. . . .
For details of initializing of the management partition, see "9.2.2 Initializing of the management partition."
. . . .

9.7.2 Re-registering node configuration information to the management partition

Re-register node configuration information to the re-initialized management partition.

See
. . . .
For details of registering node configuration information to the management partition, see "9.2.3 Registering node configuration information to the management partition."
. . . .

9.7.3 Re-configuring the sfcfrmd daemon's startup mode in the management partition

Re-configure the sfcfrmd daemon's startup mode in the management partition.

Note
. . . .
The registered sfcfrmd daemon's startup mode is **wait** when the management partition is initialized. If another sfcfrmd daemon's startup mode is selected, this operation is required.
. . . .

See
. . . .
For details of changing the sfcfrmd daemon's startup mode, see "9.5.3 Changing the sfcfrmd daemon's startup mode."
. . . .

9.7.4 Starting sfcfrmd daemon

The sfcfrmd daemon is started by all nodes so that operation may begin.

See
. . . .
For details of starting sfcfrmd daemon, see "9.2.4 Starting sfcfrmd daemon."
. . . .

9.7.5 Restoring of the management partition information

The shell script generated by **sfcgetconf(8)** in "9.6.1 Backup of the management partition information" is executed by the restoring of management partition information.

The method of restoring of the management partition information is explained below.

```
# sh _backup_file_ <Enter>
get other node information start ... end
```

Chapter 10 File System Operations (GUI)

This chapter describes how to operate the GFS Shared File System by the GFS Management View.

The structure of this chapter is the same as that of "Chapter 11 File System Operations (Command)." For details about command operation, see "Chapter 11 File System Operations (Command)."

The management partition settings must have been made before you create a file system.

For information on the outline of the management partition, see the following:

- "Chapter 5 Management Partition"

For details on how to set up the management partition, see the following:

- "Chapter 8 Management Partition Operations (GUI)"

- "Chapter 9 Management Partition Operations (Command)"

Start file system operation after you have created a file system. For details on file system operation, see the following:

- "Chapter 12 File System Management"

10.1 Flow of Operations

This section explains the flow of file system operations of the GFS Management View.

10.1.1 Creation

The following figure shows the basic design flow for creating a file system in the GFS Shared File System.

Figure 10.1 Operation flow for creating file system
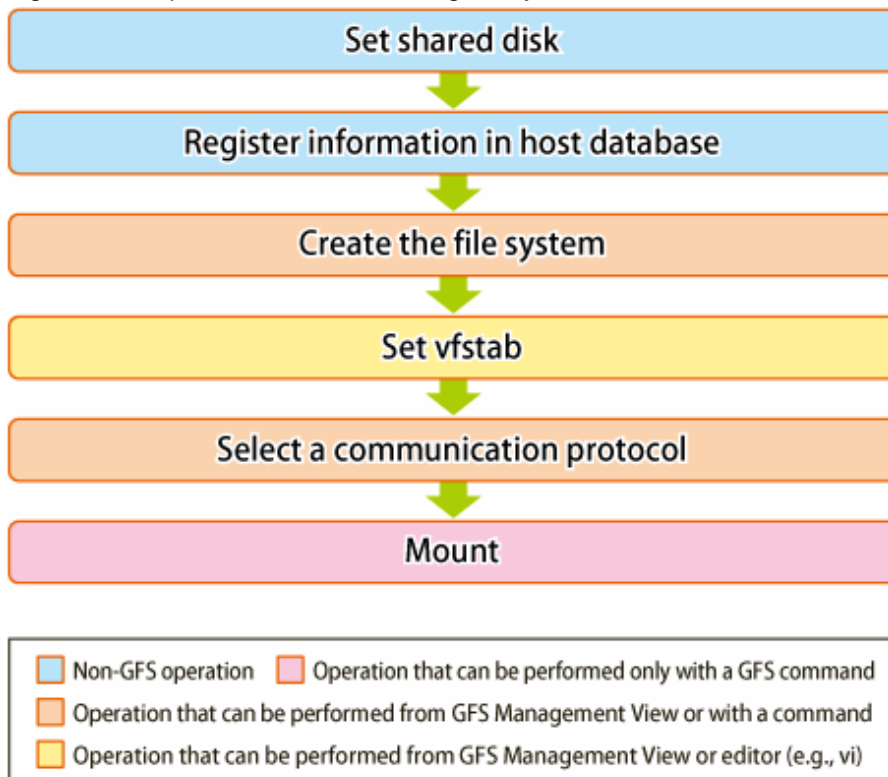
10.1.2 Change (in file system attributes)

Using the GFS Management View, the following file system attributes can be changed:

- Mount information

- File system tuning

The following figure shows the basic design flow for changing the file system attributes in the GFS Shared File System.

Figure 10.2 Operation flow for change in file system attributes
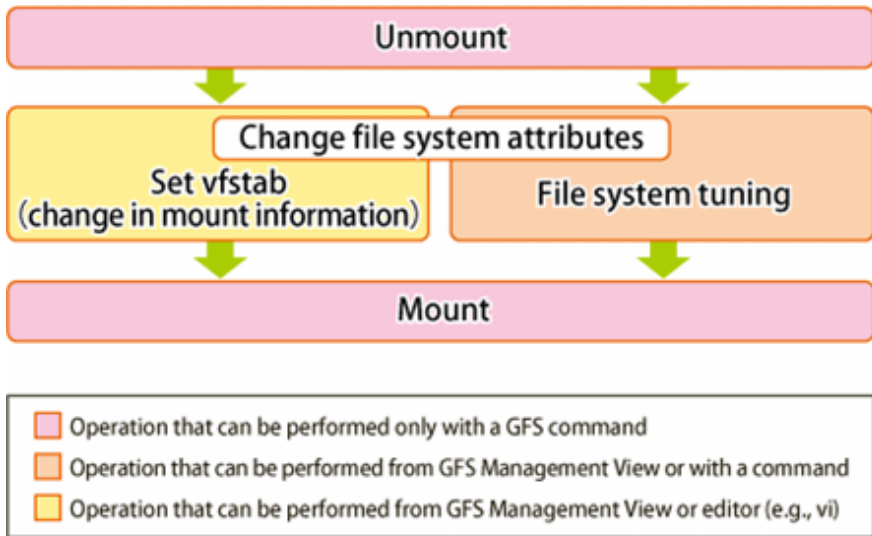
GFS : Global File Services

10.1.3 Change (partition addition)

Using the GFS Management View, the following configuration changes can be made:

- Addition of a file data partition

The following figure shows the basic design flow for changing the file system configuration in the GFS Shared File System.

Figure 10.3 Operation flow for partition addition
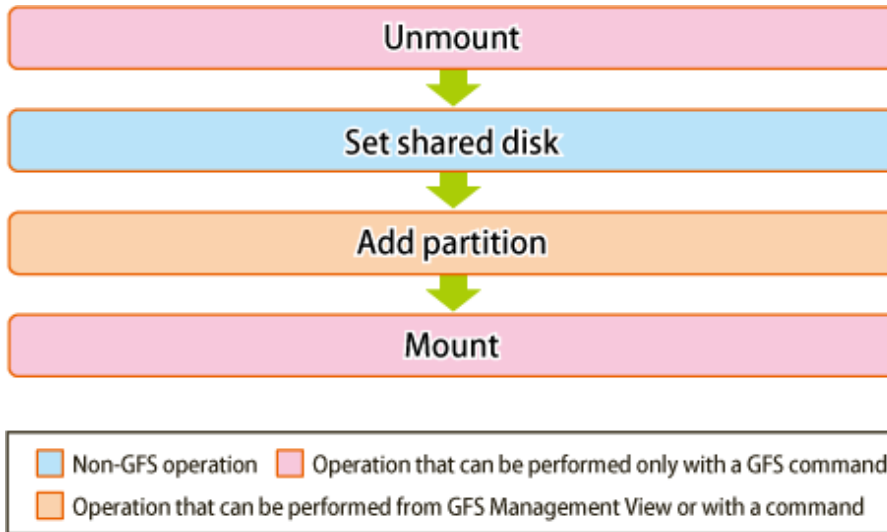
See
....
For details about GUI operation in the above figure, see "10.3.2 Changing the file system configuration (partition addition)."
For details about command operation in the above figure, see "11.4 Change (partition addition)."
....

10.1.4 Change (shared node information)

The following describes the flow of operations to change information about nodes sharing a file system. Information about nodes sharing a file system can be changed only when the file system is unmounted on all nodes.

Figure 10.4 Operation flow for changing shared node information
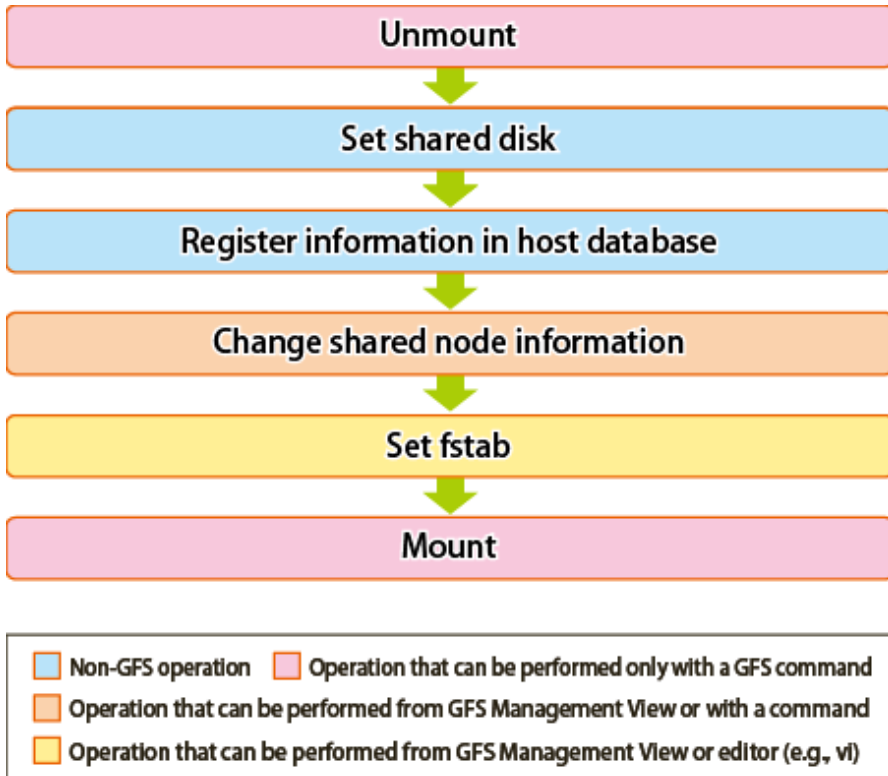
GFS : Global File Services

🗂 See
....
For details about GUI operation in the above figure, see "10.3.3 Changing the shared node information."
For details about command operation in the above figure, see "11.5 Change (shared node information)."
....

10.1.5 Deletion

The following describes the flow of operations to delete a file system. A file system can be deleted only when the file system is unmounted on all nodes.

Figure 10.5 Operation flow for deleting a file system
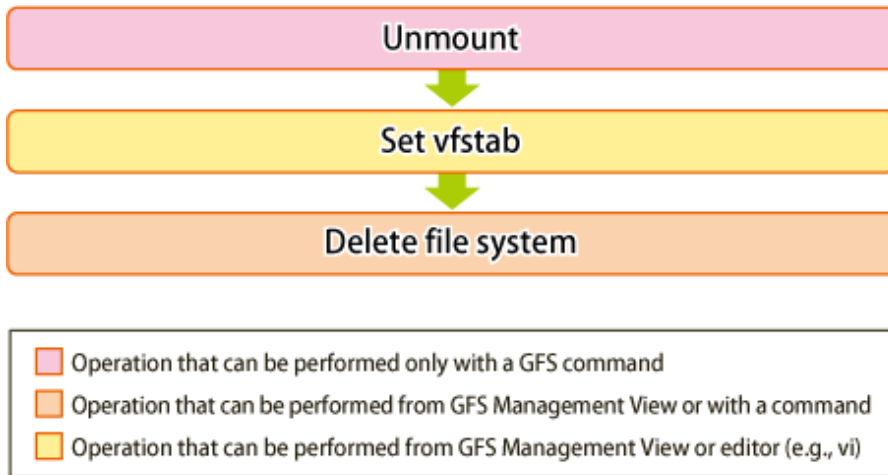
GFS : Global File Services

![See icon] **See**
....
For details about GUI operation in the above figure, see "10.4 Deletion."
For details about command operation in the above figure, see "11.8 Deleting."
....

10.2 Creation

This section explains how to create a file system.

It is easy to create a file system because the file system creation operation uses a wizard.

To start the file system creation wizard, Click [Create] on the [Operation] drop-down menu.

![Note icon] **Note**
....
Note of Operation

- Information that has already been input becomes invalid and the default value corresponding to information newly set on the previous screen is set when <Next> button is selected after <Back> button is selected on the wizard screen and the operation that exerts the influence on the following screens is done.

- The information changed on the Management View is updated on the main screen by clicking [Update Now] on the [View] drop-down menu.
....

10.2.1 Creating a file system

The following describes the procedure for creating a file system.

(1) Setting a shared disk device

The GFS Shared File System uses logical volumes of GDS as shared devices. The status of the logical volumes of GDS must be **ACTIVE** on each node.

![Note icon] **Note**
....
GDS logical volume should be set up as follows.

- Set up the type of a disk class with which GDS logical volume belongs as shared.(Change GDS class attributes)

- Set all the nodes that share GFS Shared File System as the scope of the disk class to which logical volume belongs.(Change GDS class attributes)

. . . .

See

When creating a GFS Shared File System on the logical volume of GDS that was used in the file system other than GFS, see "1.8.4.2 Notes when using logical volumes of GDS that were used in a file system other than GFS."

See

For GDS disk class operations, see the applicable items under "Operation" "GDS Management View" in "PRIMECLUSTER Global Disk Services Configuration and Administration Guide."
For the logical volume operations of GDS, see the applicable items under "Operation" "GDS Management View" in "PRIMECLUSTER Global Disk Services Configuration and Administration Guide."

(2) Registering information in host database

MDS communicates with AC by using the IP address corresponding to the host name of the node that shares the GFS Shared File System. Therefore, the host name of the shared node and the IP address are registered in host database (for example: /etc/hosts).

Note

- The IP address type of the shared node needs to match the communication protocol selected in "(9) Selecting a communication protocol." If they are not matched, the GFS Shared File System cannot be mounted.

- Register the host name of the shared node as the canonical host name (for /etc/hosts, the host name described in the left end).

- The following two IP address types are available for the GFS Shared File System:

  - Unique Local Unicast Address

  - Global Unicast Address

- The setting of host database must be the same on all the shared nodes. If it is not, MDS down recovery fails and it may trigger a file system blockade or system panic.

See

For details on /etc/hosts, see hosts(5) of the online manual page.

(3) Setting share information

Click [Create] on the [Operation] menu and the [Create File System Wizard] screen appears. It is shown in the following "Figure Create File System Wizard (1)."

In the [Create File System Wizard (1)] screen, set the node information and enter the mount point.

Figure 10.6 Create File System Wizard (1)
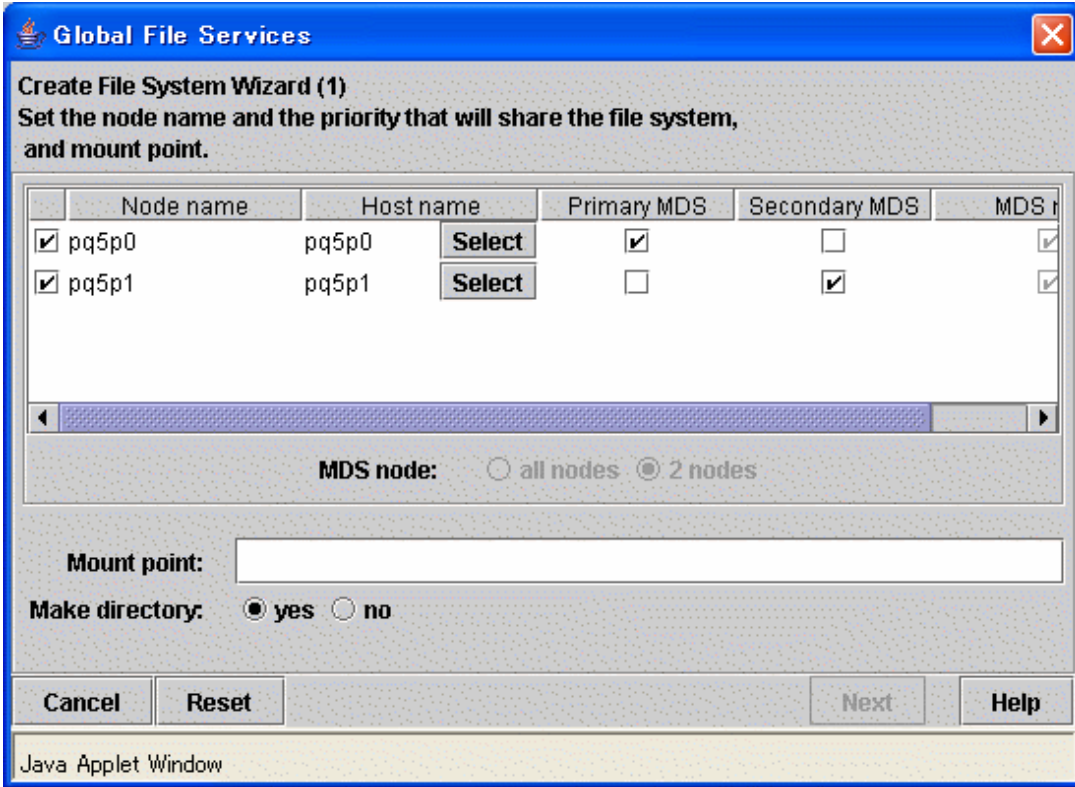
Selecting node names

In the [Node name], select the names of the sharing nodes. Be sure to select two nodes. It is not possible to deselect the local node.

Selecting host names

In order to select host names other than the host names displayed, click the <Select> button then select the LAN host that will be used for each node in the following [Host Name Selection Wizard] screen. If you do not click the <Select> button, the displayed host name will be used.

Figure 10.7 Host Name Selection Wizard
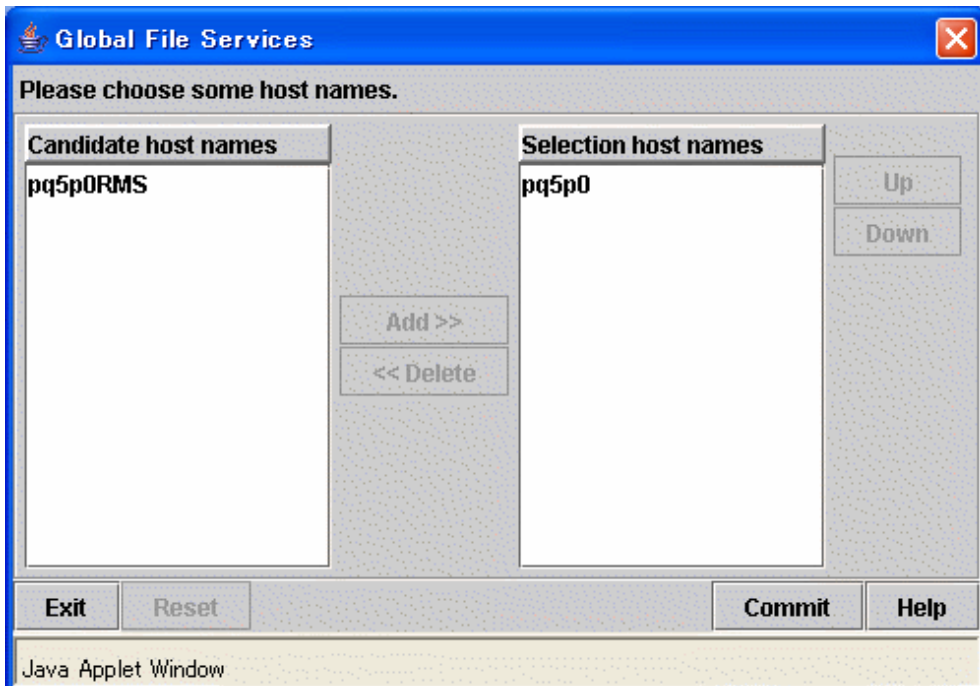
Please choose the host name of LAN used from the left-hand side [Candidate host names]. Then please click the <Add> button in order to add to the right-hand side [Selection host names]. Note that only one host name can be selected at a time in [Selection host names]. If you want to change [Selection host names], delete the host name being selected in [Selection host names].

Selecting the primary MDS and secondary MDS

Specify the nodes on which the [Primary MDS] and [Secondary MDS] will be started. A single, unique node must be selected for each MDS.

Setting mount point and creating a directory

Specify the full path name of a mount point in the [Mount point] field. If <yes> button is selected for [Make directory], a directory is created with the following attributes:

Owner: root
Group: sys
Access right: 775

After setting the area information items, click the <Next> button to proceed to the next screen which is shown in the following "Figure Create File System Wizard (2)."

To restore the settings to their default values, click the <Reset> button.

To cancel the file system creates operation, click the <Cancel> button.

(4) Selecting a partition

After completing the MDS configuration, the register partition screen will be displayed. It is shown in the following "Figure Create File System Wizard (2)."

Figure 10.8 Create File System Wizard (2)
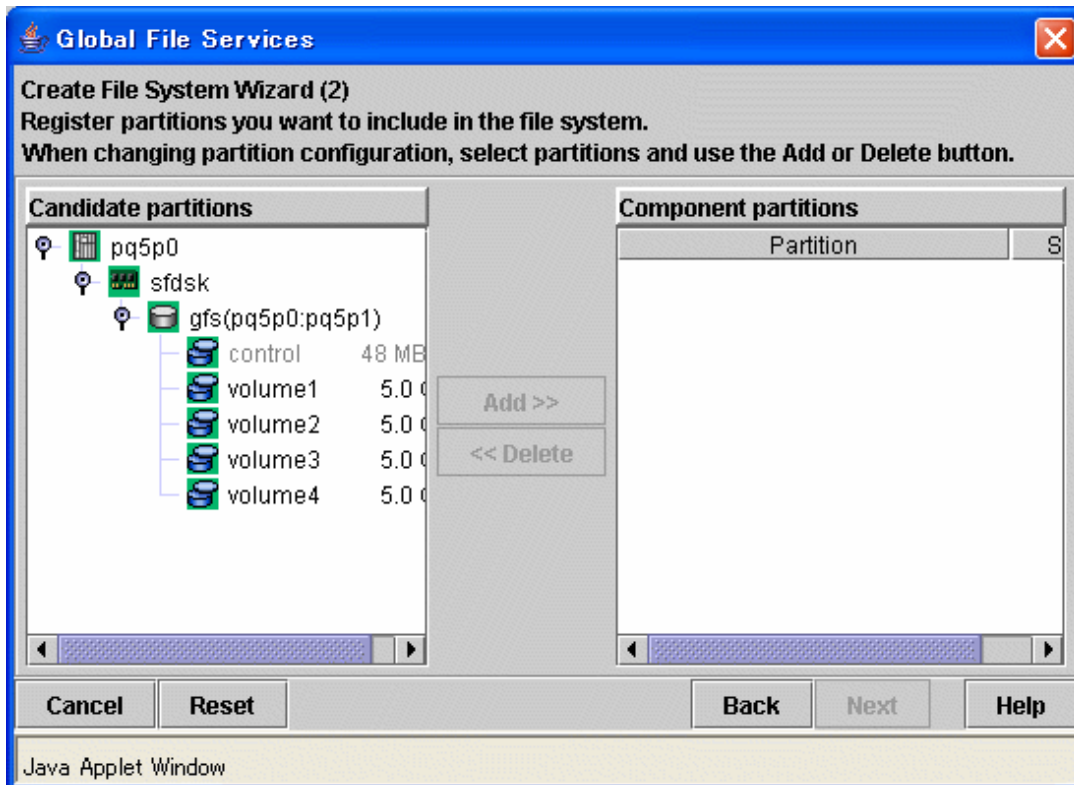
Select the partition from [Candidate partitions] field, and click the <Add> button. More than one partition can be selected at a time. However, cannot select a partition that is a part of existing file system or a management partition.

After selecting a partition, click the <Next> button to continue. The next screen is shown in "Figure Create File System Wizard (3)."

To cancel all selections of the partition currently displayed on [Component partitions] field, click the <Reset> button.

To return to the [Create File System Wizard (1)] screen, click the <Back> button.

To cancel the file system creates operation, click the <Cancel> button.

Note

If the GDS logical volume is not **ACTIVE**, it cannot be selected as a configuration partition. If it needs to be selected as a configuration partition, bring the logical volume into **ACTIVE**.

(5) Setting partition information

In the [Create File System Wizard (3)] screen, select the area (META/LOG/DATA) to be assigned to each partition selected in the [Create File System Wizard (2)].

The partition to which META is assigned becomes a representative partition.

See

For an explanation of the combinations of areas (META/LOG/DATA) assigned to partitions, see "2.1.6 Partition configuration."

Figure 10.9 Create File System Wizard (3)
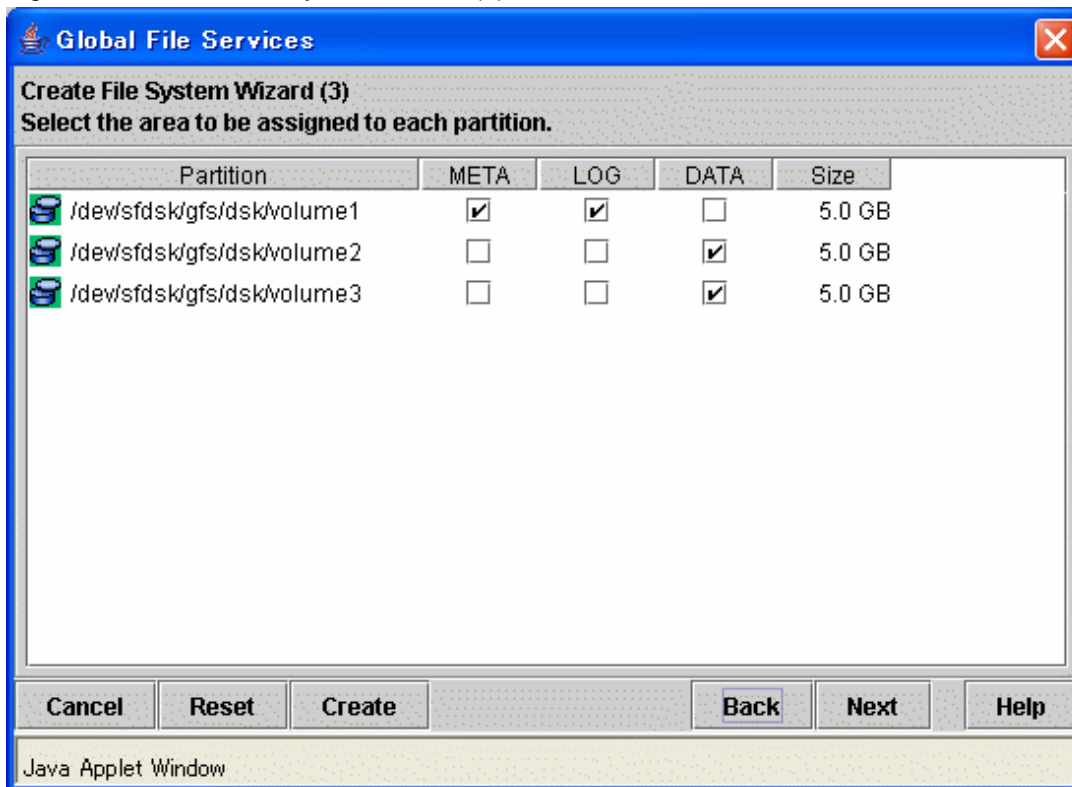
After setting the area information items, click the <Next> button to proceed to the next screen which is shown in the following "Figure Create File System Wizard (4)."

To restore the settings to their default values, click the <Reset> button.

To return to the [Create File System Wizard (2)] screen, click the <Back> button.

To cancel the file system creation operation, click the <Cancel> button.

If you do not need to change extended information, detailed information and mount information, click the <Create> button to create the file system.

(6) Setting extended information

In the [Create File System Wizard (4)] screen, the [extended information] values should be set such that they allow for future expansion of the file system and configuration changes.

Figure 10.10 Create File System Wizard (4)
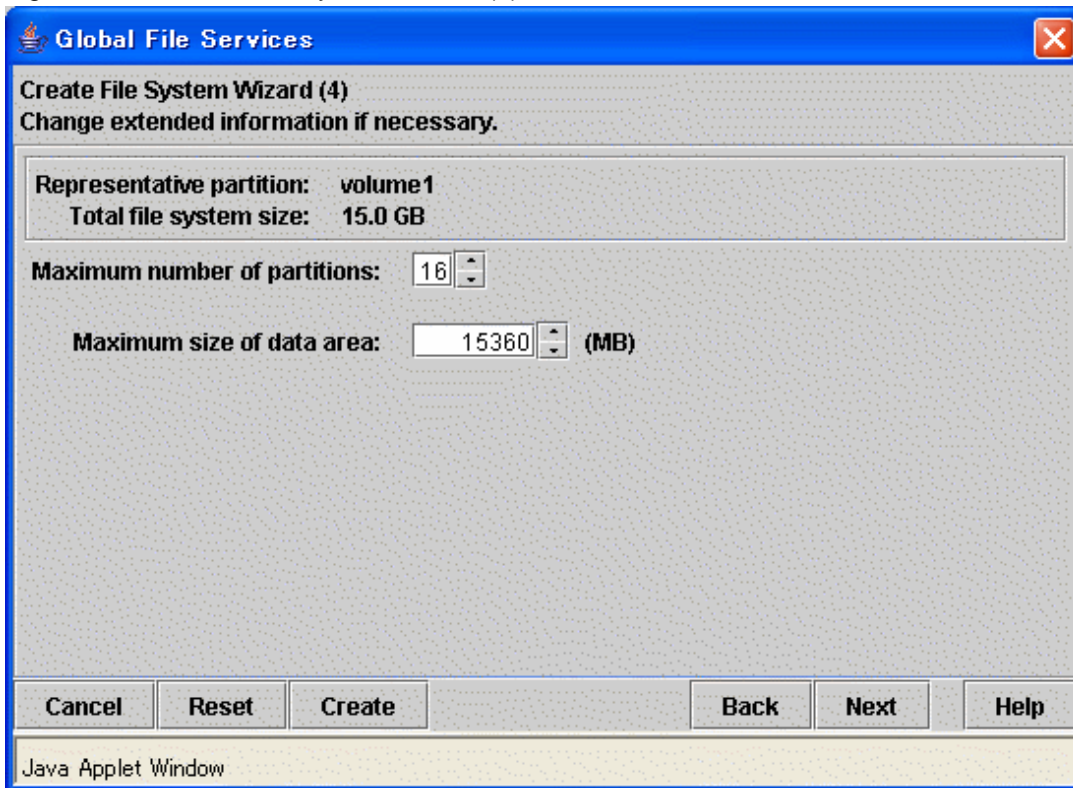
![Note icon] **Note**

. . . .

Set the maximum number of partitions and the maximum size of data area after adequately considering future data area expansion for the file system to be created. For details, see "4.3.5 For future expansion."

. . . .

![See icon] **See**

. . . .

The above parameter for expansion setup corresponds to specific **sfcmkfs**(8) options of the GFS Shared File System as follows:

- [Maximum number of partitions]: **maxvol**=n

- [Maximum size of data area]: **maxdsz**=n

For details about the parameters, see sfcmkfs(8).

. . . .

When the extended information items are correct, click the <Next> button to proceed. The next screen is shown in the following "Figure Create File System Wizard (5)."

To restore the settings to their default values, click the <Reset> button.

To return to the [Create File System Wizard (3)] screen, click the <Back> button.

To cancel the file system creation operation, click the <Cancel> button.

If you do not need to change the detailed information and mount information, click the <Create> button to create the file system.

(7) Setting detailed information

In the [Create File System Wizard (5)], set the detailed information.

Figure 10.11 Create File System Wizard (5)
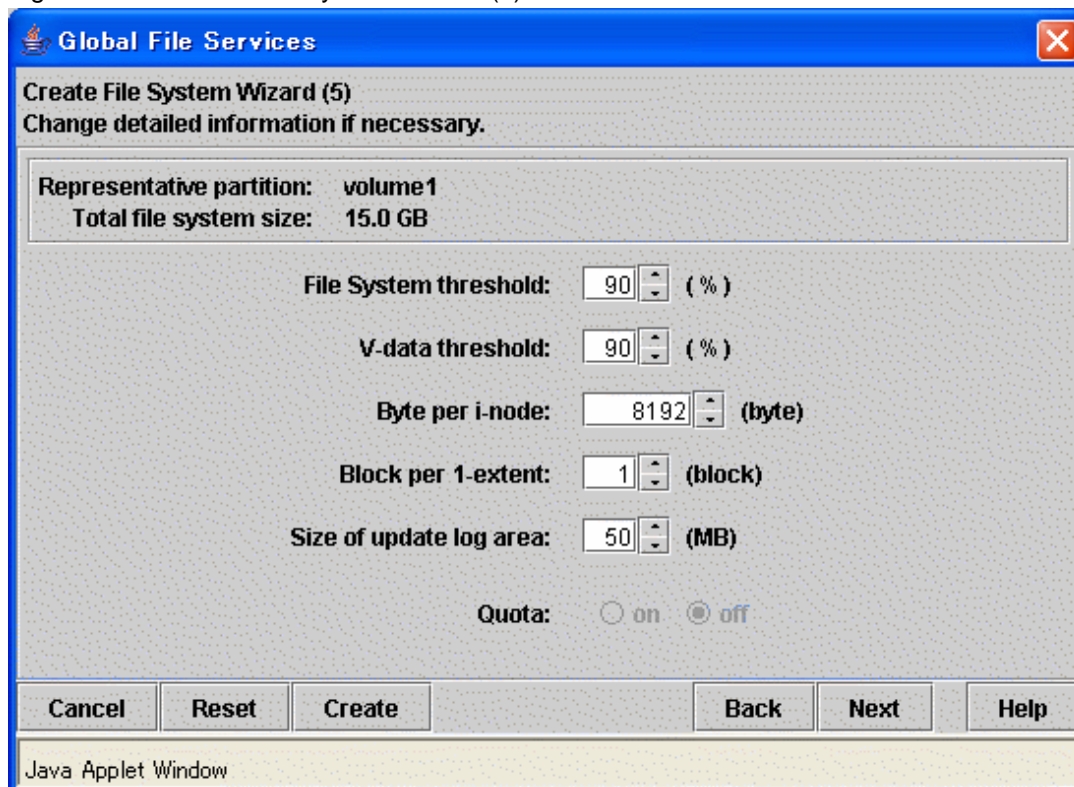
See

. . . .

The above parameter for expansion setup corresponds to specific **sfcmkfs**(8) options of the GFS Shared File System as follows:

- [File System threshold]: **free**=*n*

- [V-data threshold]: **mfree**=*n*

- [Byte per i-node]: **nbpi**=*n*

- [Block per 1-extent]: **nblkpext**=*n*

- [Size of update log area]: **logsz**=*n*

For details about the parameters, see sfcmkfs(8).

. . . .

Note

. . . .

The [File System threshold] is the value obtained by subtracting from 100 the minimum percentage of free area in the file system specified by the **-o free** option of **sfcmkfs**(8).
The [V-data threshold] is the value obtained by subtracting from 100 the minimum percentage of free area in the V-data area specified by the **-o mfree** option of **sfcmkfs**(8).

. . . .

When the detailed information items are correct, click the <Next> button to continue. The next screen is shown in the following "Figure Create File System Wizard (6)."

To restore the settings to their default values, click the <Reset> button.

To return to the [Create File System Wizard (4)] screen, click the <Back> button.

To cancel the file system creation operation, click the <Cancel> button.

If you do not need to change the mount information, click the <Create> button to create the file system.

(8) Setting mount information

In the [Create File System Wizard (6)] screen, set the mount information.

Figure 10.12 Create File System Wizard (6)
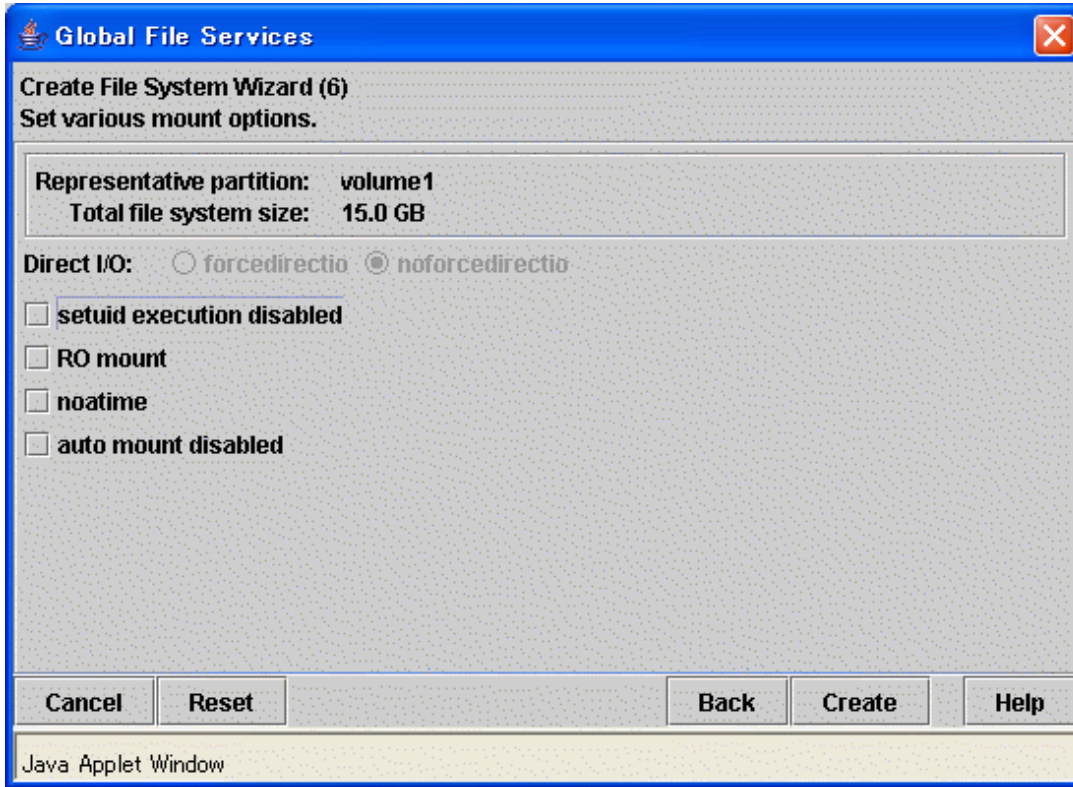
![See]

The above each mount option corresponds to mount option of **sfcmount(8)** as follows:

- [setuid execution disabled]: **nosuid**

- [RO mount]: **ro**

- [noatime]: **noatime**

- [auto mount disabled]: **noatrc**

For details about the parameters, see sfcmount(8).

![Note]

The mount option **strictatime** cannot be specified with the GFS Management View. When specifying **strictatime**, directly edit **/etc/fstab** of all the nodes that share a file system after clicking the <Create> button.

After setting the mount information, click the <Create> button to create the file system.

To restore the settings to their default values, click the <Reset> button.

To return to the [Create File System Wizard (5)] screen, click the <Back> button.

To cancel the file system creation operation, click the <Cancel> button.

📒 **Note**

. . . .

If you create a file system by clicking the <Create> button, the GFS Management View will automatically add an entry to **/etc/fstab**.

. . . .

(9) Selecting a communication protocol

After completing a file system creation, determine either IPv6 or IPv4 to be used as a communication protocol between MDS and AC.

The default (a communication protocol to be registered when creating a file system) is IPv4.

The setting of a communication protocol is performed in [Detailed information] of the [File System Attributes] screen started by clicking [Change Attributes] on the [Operation] menu of the GFS Management View. When using IPv4, this setting is not required.

📒 **Note**

. . . .

- A communication protocol can be set for each file system; however, we recommend that users set the same communication protocol for all the file systems so that it can be easily managed.

- Select the communication protocol that matches the IP address type of the shared node registered in host database in "(2) Registering information in host database." If you have selected a communication protocol that does not match it, the GFS Shared File System cannot be mounted.

. . . .

📒 **See**

. . . .

For details of setting a communication protocol, see "10.3.1 Changing the file system attributes" in this manual.

. . . .

(10) Mounting the file system

Mount the file system.

📒 **See**

. . . .

For information on how to mount the file system using a command, see "11.2.6 Mount."

. . . .

10.3 Change

The GFS Management View provides three change operations:

- Changing the file system attributes

- Changing the file system configuration (partition addition)

- Changing the shared node information

📒 **Note**

. . . .

Note of Operation

- The information changed on the Management View is updated on the main screen by clicking [Update Now] on the [View] drop-down menu.

- If the sfcfs information is not available, the file system cannot be changed. Take action according to the output error message.

. . . .

10.3.1 Changing the file system attributes

The following describes the procedure for changing file system attributes using the GFS Management View.

Changing the file system attributes includes changing the mount information.

(1) Unmounting a file system

Before changing the file system attribute, unmount the file system on all the nodes.

🛠 See
....
For information on how to unmount the file system using a command, see "11.3.1 Unmount."
....

(2) Selecting a file system

From the [Mount tree] or [File system information] field in the main screen, select the file system you want to change. Only file systems of the **sfcfs** type can be selected.

Click [Change Attributes] on the [Operation] menu, and the [File System Attributes] screen appears.

From the [File System Attributes] screen, the [Mount information], [Share information], and [Detailed information] can be selected using the screen switching tabs.

(3) Changing the mount information

In the [Mount information], the mount information can be changed.

To change the settings, change the corresponding parameters.

To restore the settings to their previous values, click the <Reset> button.

To execute the mount information change operation, click the <Apply> button.

Specify the full path name of a mount point in the [Mount point] field. If [yes] is selected for [Make directory], a directory is created with the following attributes:

Owner: root
Group: sys
Access permissions: 775

Figure 10.13 Mount information for file system attributes
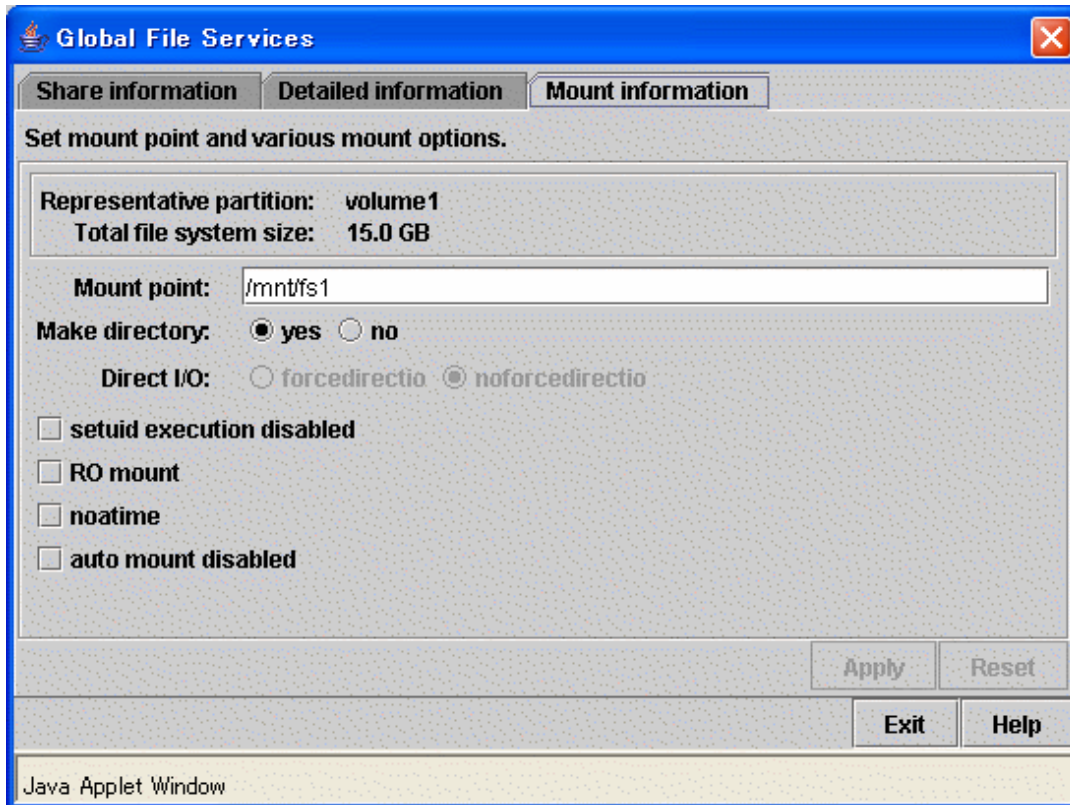
See

The above each mount option corresponds to mount option of **sfcmount(8)** as follows:

- [setuid execution disabled]: **nosuid**

- [RO mount]: **ro**

- [noatime]: **noatime**

- [auto mount disabled]: **noatrc**

For details about the parameters, see sfcmount(8).

Note

- If you change the mounting information by clicking the <Apply> button, the GFS Management View will automatically add the information in **/etc/fstab**.

- The mount option **strictatime** cannot be specified with the GFS Management View. When specifying **strictatime**, directly edit **/etc/fstab** of all the nodes that share a file system after clicking the <Apply> button.

(4) Changing detailed information

In the [Detailed information], tune the file system.

To change the settings, change the corresponding parameters.

To restore the settings to their previous values, click the <Reset> button.

To execute the detailed information change operation, click the <Apply> button.

Figure 10.14 Detailed information for file system attributes
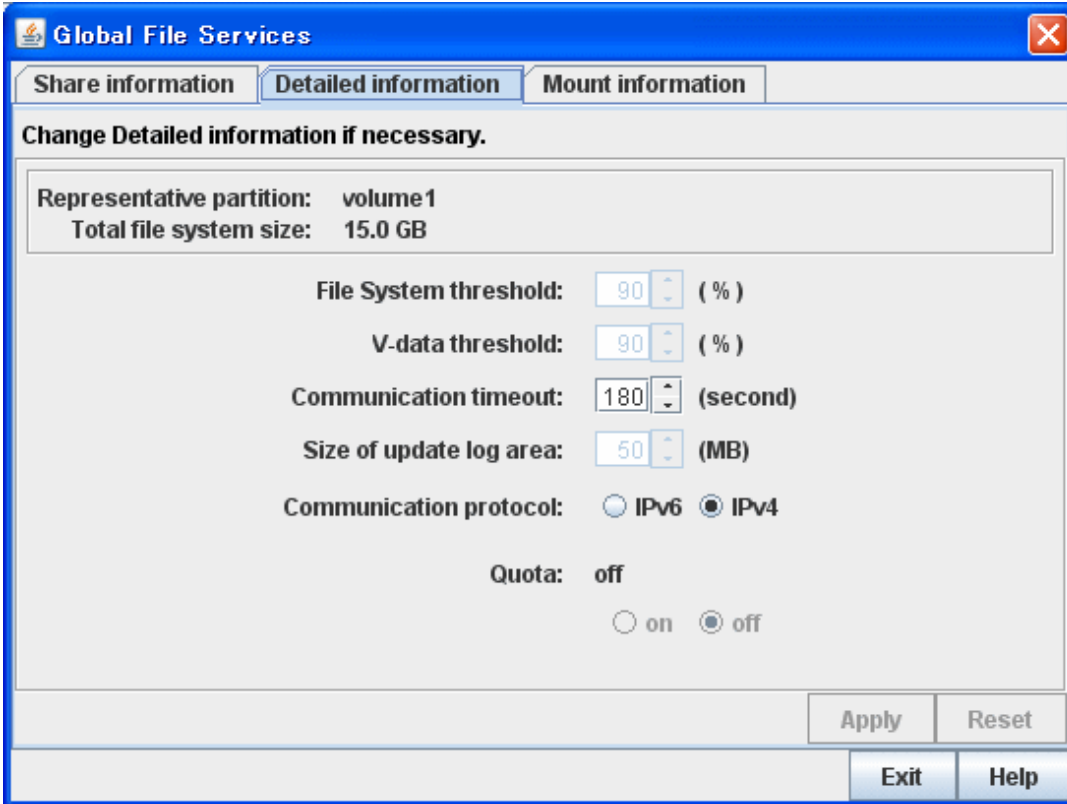
See

....

The variable parameter of detailed information corresponds to the tuning parameter of the GFS Shared File System as follow:

- [Communication timeout]: **CON_TIMEOUT**

- [Communication protocol]: **SFCFS_IPV6**

For details about the tuning parameter, see "14.1.2 Communication timeout value" and "14.1.3 Communication protocol" in this manual.

....

(5) Mounting a file system

After the completion of the file system attribute change operation, a file system is in mounted.

See

....

For information on how to mount the file system using a command, see "11.2.6 Mount."

....

10.3.2 Changing the file system configuration (partition addition)

This section explains the procedure for adding file data partitions.

You can easily change file system configuration using a wizard.

To start the Change File System Configuration Wizard, click [Change Configuration] on the [Operation] drop-down menu.

Note

....

Note of Operation

Information that has already been input becomes invalid and the default value corresponding to information newly set on the previous screen is set when <Next> button is selected after <Back> button is selected on the wizard screen and the operation that exerts the influence on the following screens is done.

(1) Unmounting a file system

Before changing the file system configuration, unmount the file system on all the nodes.

See

For details about command, see "11.3.1 Unmount."

(2) Setting shared disk device

A partition that will be added to the GFS Shared File System must be a logical volume of a GDS shared class. Also, the GDS logical volume must be **ACTIVE**.

Note

GDS logical volume should be set up as follows.

- Set up the type of a disk class with which GDS logical volume belongs as shared.(Change GDS class attributes)

- Set all the nodes that share GFS Shared File System as the scope of the disk class to which logical volume belongs.(Change GDS class attributes)

See

When adding the logical volume of GDS, which was used in a file system other than GFS, as the partition of the GFS Shared File System, see "1.8.4.2 Notes when using logical volumes of GDS that were used in a file system other than GFS."

See

For GDS disk class operations, see the applicable items under "Operation" "GDS Management View" in "PRIMECLUSTER Global Disk Services Configuration and Administration Guide."
For GDS logical volume operations, see the applicable items under "Operation" "GDS Management View" in "PRIMECLUSTER Global Disk Services Configuration and Administration Guide."

(3) Selecting a file system

From the [Mount tree] or [File system information] field in the main screen, select the file system to be changed. Only file systems of the **sfcfs** type can be selected.

Click [Change Configuration] on the [Operation] menu, the [Change File System Configuration Wizard (1)] screen appears.

(4) Selecting a partition

From the [Candidate partitions] field, select the partition to be added and click the <Add> button. More than one partition can be selected at a time. Note that those partitions that are already in use for file systems cannot be selected.

To cancel all selections of the partition currently displayed on [Component partitions] field, click the <Reset> button.

Figure 10.15 Change File System Configuration Wizard (1)
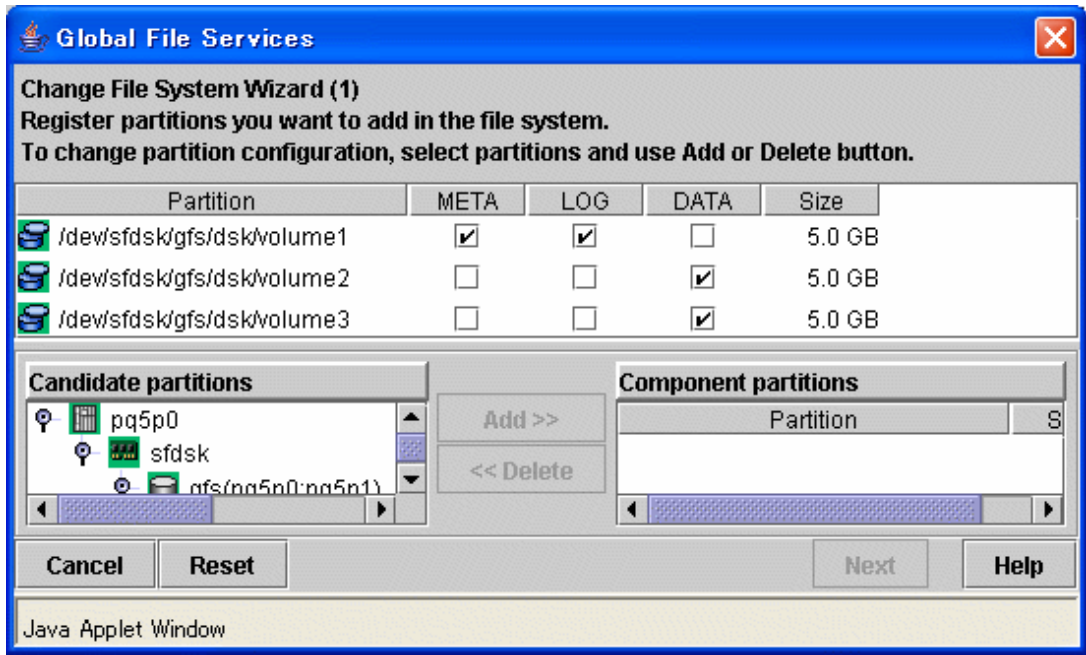
After selecting the partitions, click the <Next> button to continue. The next screen is shown in the following "Figure Change File System Configuration Wizard (2)."

To cancel the file system configuration change operation, click the <Cancel> button.

(5) Checking partition information

In the [Change File System Configuration Wizard (2)] screen, confirm the file system configuration including adding partition.

See
. . . .
For an explanation of the combinations of areas (META/DATA) assigned to partitions, see "2.1.6 Partition configuration."
. . . .

Note
. . . .
The added partition is automatically allocated to the file data area.
. . . .

Figure 10.16 Change File System Configuration Wizard (2)
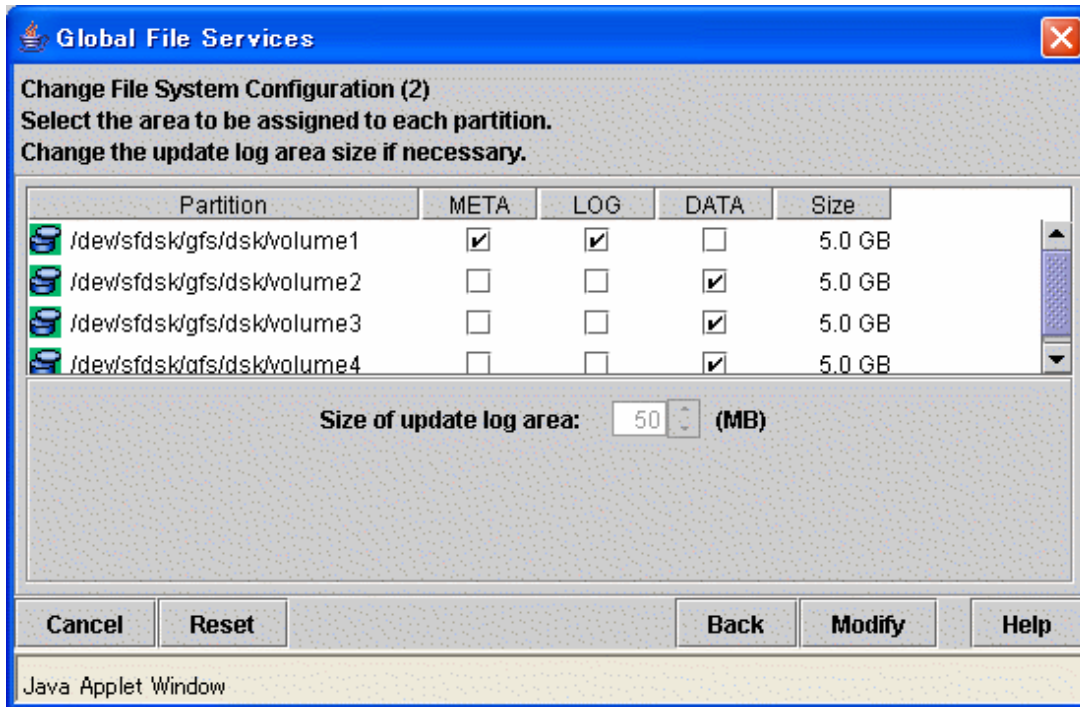
After confirming the information items, click the <Modify> button to execute the file system configuration change operation.

To return to the [Change File System Configuration Wizard (1)], click the <Back> button.

To cancel the file system configuration change operation, click the <Cancel> button.

(6) Mounting a file system

After completion of the file system attribute change operation, mount the file system if the file system is unmounted.

See
. . . .
For details about command, see "11.2.6 Mount."
. . . .

10.3.3 Changing the shared node information

The following describes the procedure for changing sharing-node information using the GFS Management View.

(1) Unmounting a file system

Before changing the shared node information, unmount the file system on all the nodes.

See
. . . .
For details about command, see "11.3.1 Unmount."
. . . .

(2) Selecting a file system

From the [Mount tree] or [File system information] field in the main screen, select the file system you want to change. Only file systems of the **sfcfs** type can be selected.

Click [Change Attributes] on the [Operation] menu and go to [Share information] screen using the screen-switching tab.

(3) Changing share information

In the [Share information], the information on the nodes sharing the file system can be changed.

To change the settings, change the corresponding parameters.

To restore the settings to their previous values, click the <Reset> button.

Figure 10.17 Share information for file system attributes
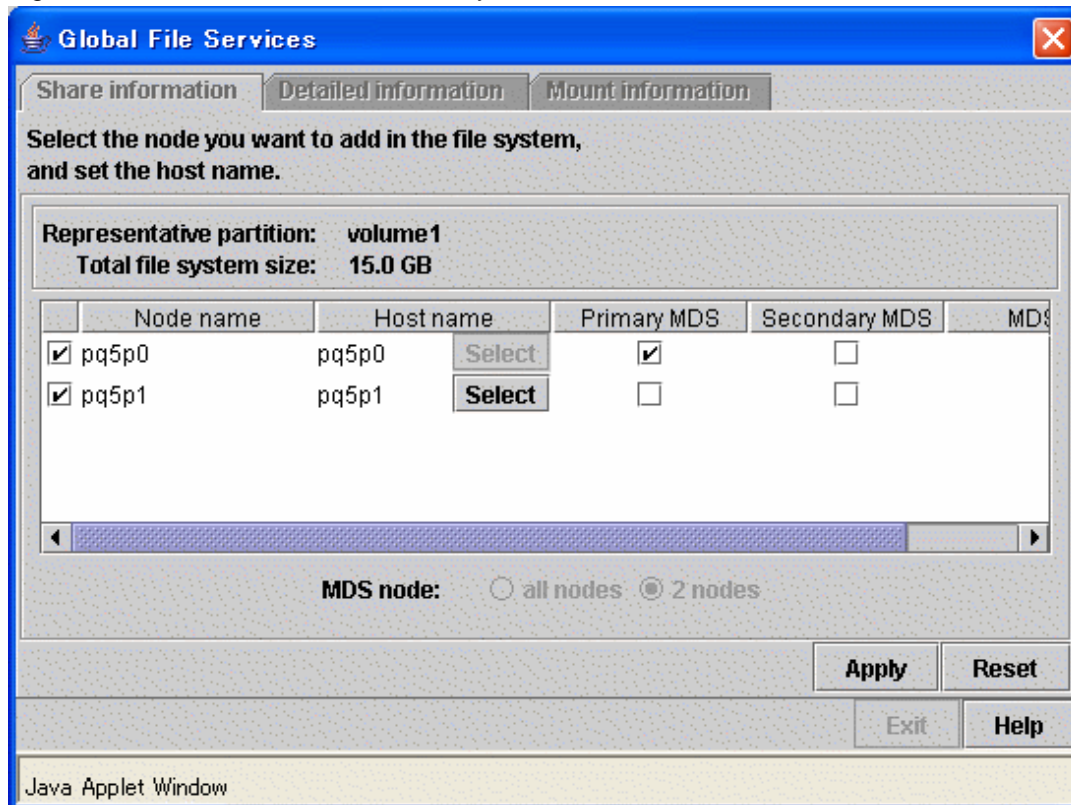
Selecting node names

In the [Node name] column, select the names of the sharing nodes. Be sure to select two nodes. It is not possible to deselect the local node, the primary MDS or the secondary MDS.

Selecting a host name

In order to select host names other than the host names displayed, click the <Select> button. Then select the LAN host that will be used for each node in the following [Host Name Selection Wizard] screen.

If you do not click the <Select> button, the displayed host name will be used.

The host name for each node already selected cannot be changed. Delete the host name being selected, and select the LAN host from the candidate hosts then add it to the hosts for selection.

Figure 10.18 Host Name Selection Wizard
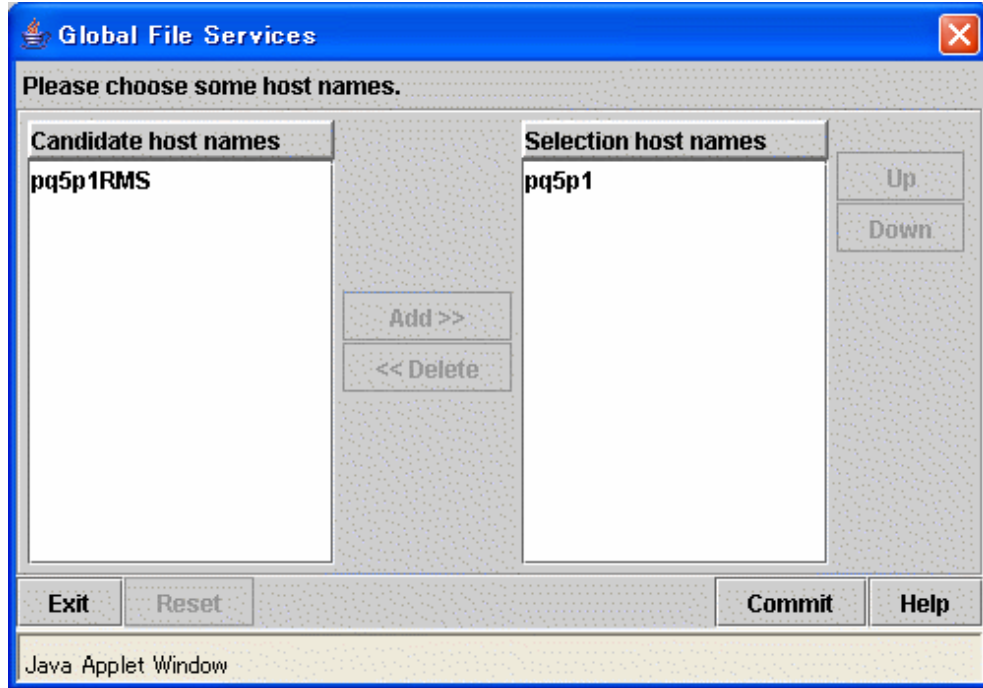
Note

Primary MDS and Secondary MDS

The [Primary MDS] and [Secondary MDS] cannot be changed. If a change is necessary, the file system must be deleted and re-created.

(4) Confirming the sharing-node information change operation

To execute the sharing-node information change operation, click the <Apply> button.

Note

If you change sharing-node information by clicking the <Apply> button, the GFS Management View will automatically add and change the entry in **/etc/fstab**.

(5) Mounting a file system

After completion of the file system attributes change operation, mount the file system.

See

For details about command, see "11.2.6 Mount."

10.4 Deletion

This section explains how to delete a file system.

Note

Note of Operation

- The information changed on the Management View is updated on the main screen by clicking [Update Now] on the [View] drop-down menu.

- If the sfcfs information is not available, the file system cannot be changed. Take action according to the output error message.

10.4.1 Deleting the file system

The following describes the procedure for deleting a file system using the GFS Management View.

(1) Unmount the file system

Unmount the file system on all nodes before deletion.

See
For detail on the unmount procedure, see "11.3.1 Unmount."

(2) Selecting the file system

From the [Mount tree] or [File system information] field in the main screen, select the file system to be deleted. Only file systems of the **sfcfs** type can be selected.

(3) Deleting the file system

Click [Delete] on the [Operation] menu and the following warning message appears.

To delete the file system, click the <Yes> button.

To cancel the deletion of the file system, click the <No> button.

Figure 10.19 File system deletion warning message
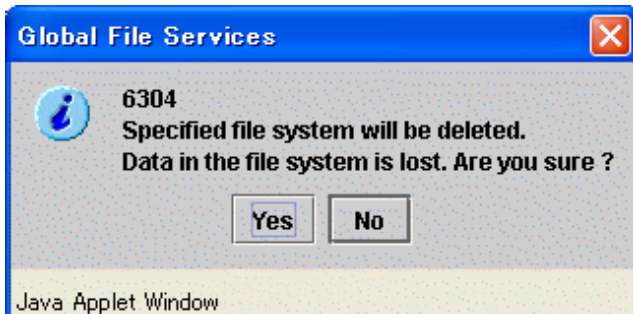
Note
If you delete the file system by clicking the <Yes> button, the GFS Management View will automatically delete the entry from **/etc/fstab**.

Chapter 11 File System Operations (Command)

This chapter describes how to operate the GFS Shared File System by commands.

The structure of this chapter is the same as that of "Chapter 10 File System Operations (GUI)." For details about GUI operation, see "Chapter 10 File System Operations (GUI)."

The management partition settings must have been made before you create a file system.

For information on the outline of the management partition, see the following:

- "Chapter 5 Management Partition"

For details on how to set up the management partition, see the following:

- "Chapter 8 Management Partition Operations (GUI)"

- "Chapter 9 Management Partition Operations (Command)"

Start file system operation after you have created a file system. For details on file system operation, see the following:

- "Chapter 12 File System Management"

11.1 Flow of Operations

This section describes the flow of operations in the GFS Shared File System.

11.1.1 Creation

The following figure shows the basic design flow for creating a file system in the GFS Shared File System.

Figure 11.1 Operation flow for creating file system
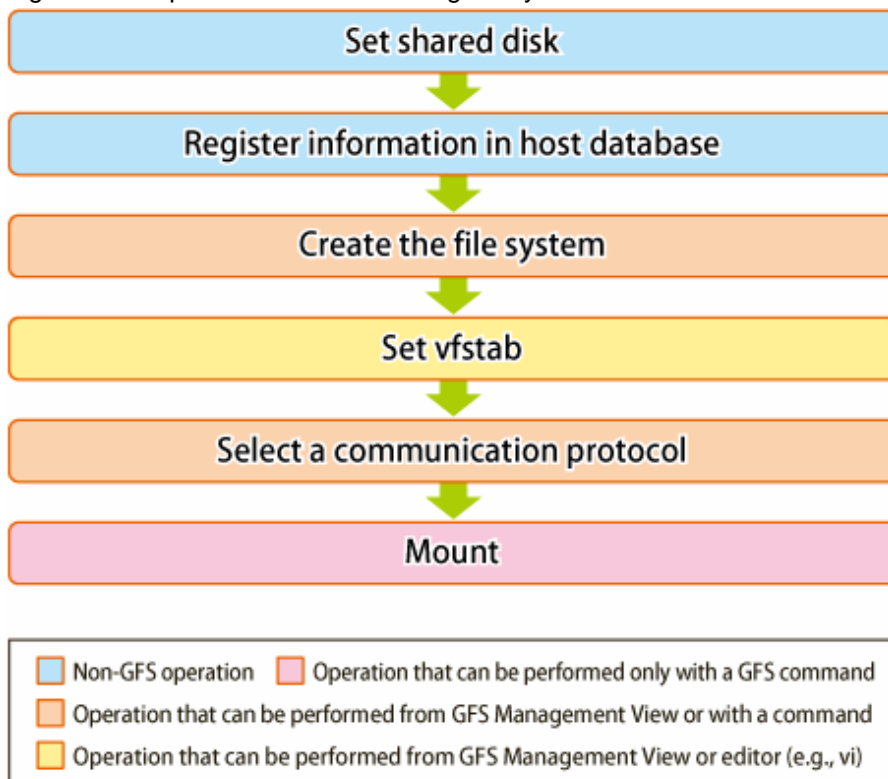
11.1.2 Change (in file system attributes)

In the GFS Shared File System, the following file system attributes can be changed:

- Mount information

- File system tuning

The following figure shows the basic design flow for changing the file system attributes in the GFS Shared File System.

Figure 11.2 Operation flow for change in file system attributes
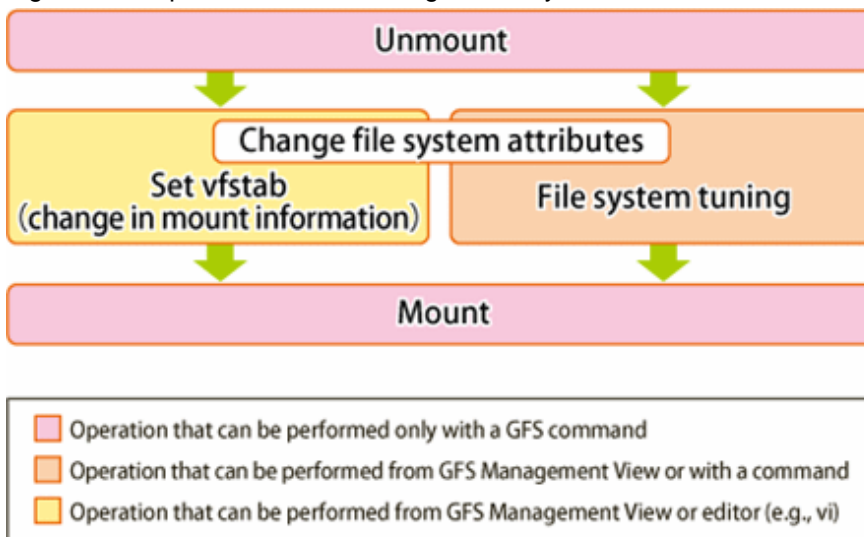
11.1.3 Change (partition addition)

In the GFS Shared File System, the following information can be changed:

- Addition of file data partitions

The following figure shows the basic design flow for changing the file system configuration in the GFS Shared File System.

Figure 11.3 Operation flow for partition addition
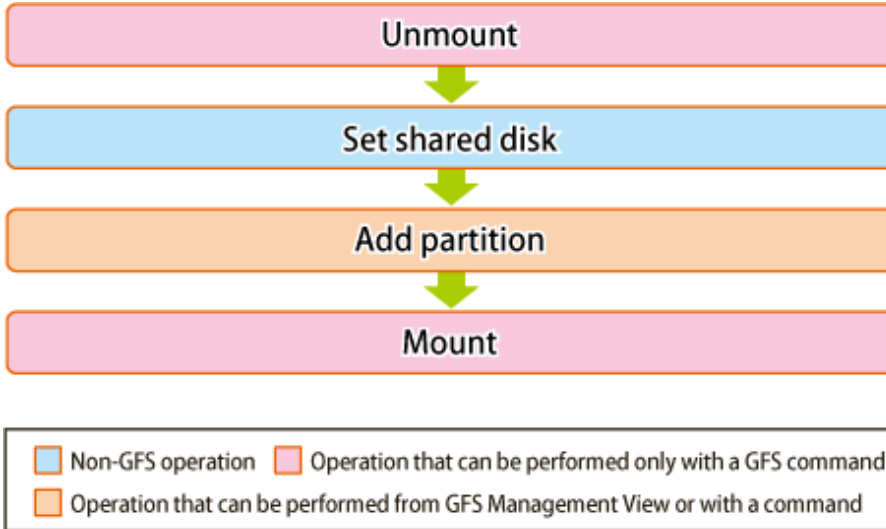
See

For details about command operation in the above figure, see "11.4 Change (partition addition)."
For details about GUI operation in the above figure, see "10.3.2 Changing the file system configuration (partition addition)."

11.1.4 Change (shared node information)

In the GFS Shared File System, the following change of shared node information can be operated:

- Adding shared node information

- Deleting shared node information

- Updating shared node information

The following figure shows the basic design flow for changing the file system shared node information in the GFS Shared File System.

Figure 11.4 Operation flow for changing the shared node information
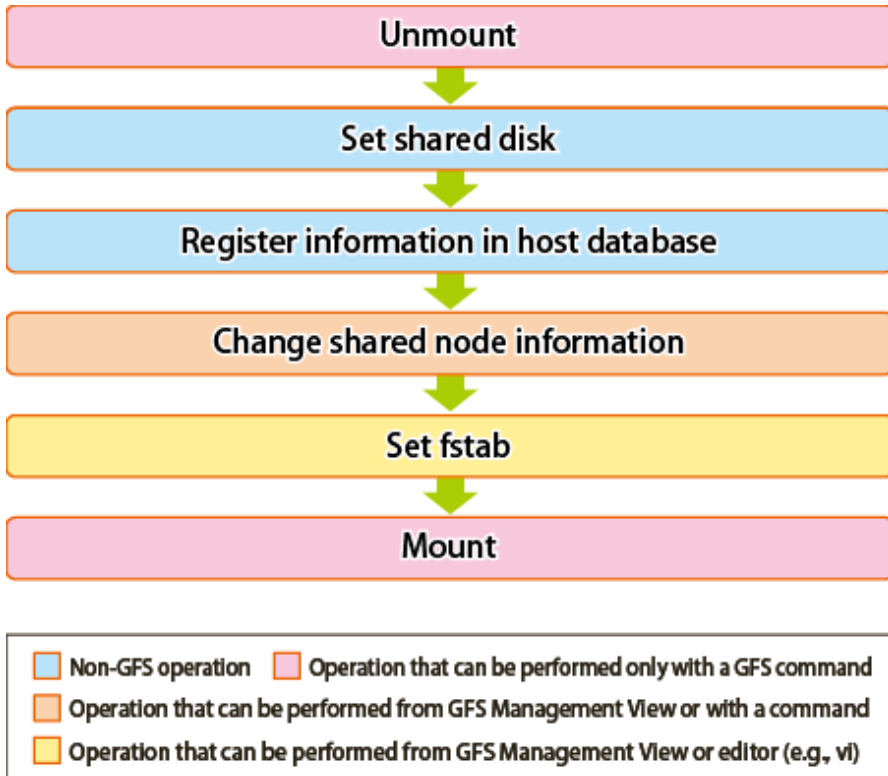
Figure 11.4 Operation flow for changing the shared node information

![Flowchart: Unmount → Set shared disk → Register information in host database → Change shared node information → Set fstab → Mount]

Legend:
- ▢ Non-GFS operation
- ▢ Operation that can be performed only with a GFS command
- ▢ Operation that can be performed from GFS Management View or with a command
- ▢ Operation that can be performed from GFS Management View or editor (e.g., vi)

GFS : Global File Services

🗃 See
....
For details about command operation in the above figure, see "11.5 Change (shared node information)."
For details about GUI operation in the above figure, see "10.3.3 Changing the shared node information."
....

11.1.5 Change (re-creating a file system)

The following figure shows the basic design flow for re-creating the file system in the GFS Shared File System.

Figure 11.5 Operation flow for re-creating a file system
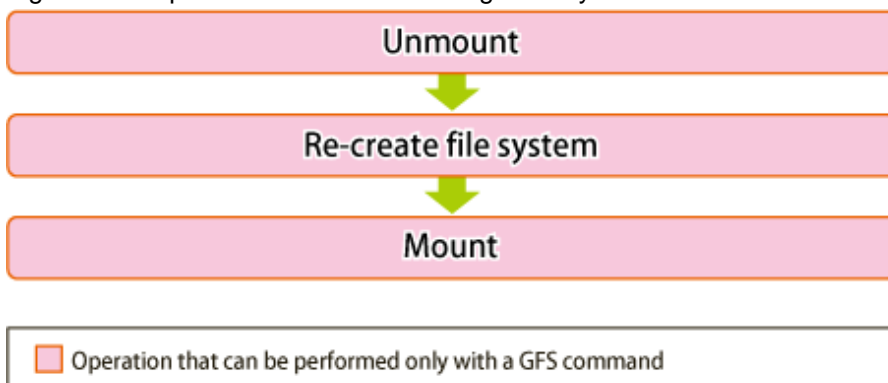
![Flowchart: Unmount → Re-create file system → Mount]

Legend:
- ▢ Operation that can be performed only with a GFS command

GFS : Global File Services

11.1.6 Change (MDS operational information)

The following figure shows the basic design flow for changing the MDS operational information in the GFS Shared File System.

Figure 11.6 Operation flow for changing the MDS operational information
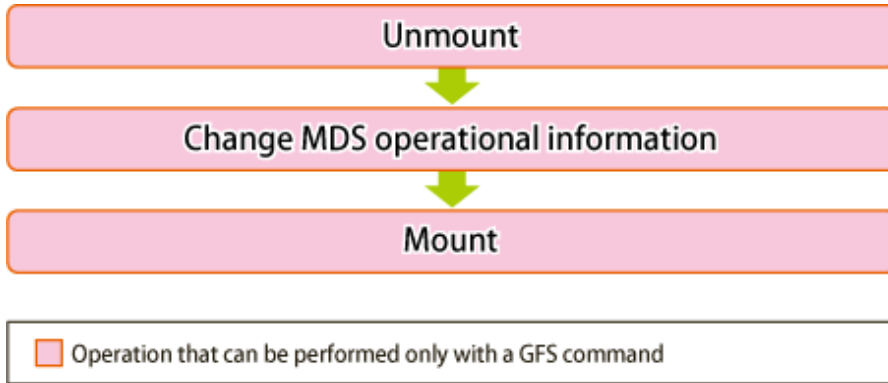
11.1.7 Deletion

The following figure shows the basic design flow for deleting the file system in the GFS Shared File System.

Figure 11.7 Operation flow for deleting a file system
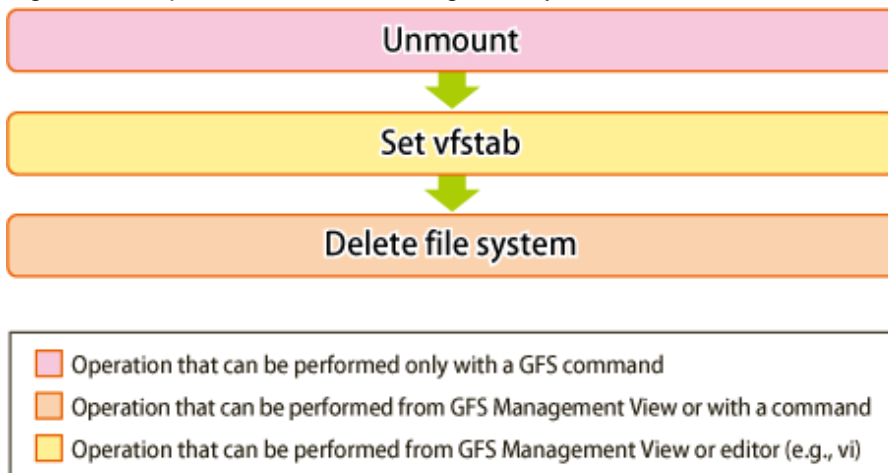
11.2 Creation

This section describes the operations from GFS Shared File System creation to operation.

1. Setting up a shared disk

2. Registering information in host database

3. File system creation

4. Setting up /etc/fstab

5. Selecting a communication protocol

6. Mount

7. File system status check

Also, the notes are provided when the GFS Shared File System is created by using the partition that already used File System.

11.2.1 Setting shared disks

The GFS Shared File System uses logical volumes of GDS as shared devices. The status of the logical volumes of GDS must be ACTIVE on each node.

📒 **Note**
. . . .
GDS logical volume should be set up as follows.

- Set up the type of a disk class with which GDS logical volume belongs as shared.(Change GDS class attributes)

- Set all the nodes that share GFS Shared File System as the scope of the disk class to which logical volume belongs.(Change GDS class attributes)

. . . .

📚 **See**
. . . .
When creating a GFS Shared File System on the logical volume of GDS that was used in the file system other than GFS, see "1.8.4.2 Notes when using logical volumes of GDS that were used in a file system other than GFS."
. . . .

📚 **See**
. . . .
For GDS disk class operations, see the applicable items under "Operation" "GDS Management View" in "PRIMECLUSTER Global Disk Services Configuration and Administration Guide."
For GDS logical volume operations, see the applicable items under "Operation" "GDS Management View" in "PRIMECLUSTER Global Disk Services Configuration and Administration Guide."
. . . .

11.2.2 Registering information in host database

MDS communicates with AC by using the IP address corresponding to the host name of the node that shares the GFS Shared File System. Therefore, the host name of the shared node and the IP address are registered in host database (for example: /etc/hosts).

 Note

. . . .

- The IP address type of the shared node needs to match the communication protocol selected in "11.2.5 Selecting a communication protocol." If they are not matched, the GFS Shared File System cannot be mounted.

- Register the host name of the shared node as the canonical host name (for /etc/hosts, the host name described in the left end).

- The following two IP address types are available for the GFS Shared File System:

    - Unique Local Unicast Address

    - Global Unicast Address

- The setting of host database must be the same on all the shared nodes. If it is not, MDS down recovery fails and it may trigger a file system blockade or system panic.

. . . .

 See

. . . .

For details on /etc/hosts, see hosts(5) of the online manual page.

. . . .

11.2.3 Creating a file systems

To create a GFS Shared File System, use sfcmkfs(8). Create the GFS Shared File System on any one of the file system shared nodes.

 See

. . . .

For details on sfcmkfs(8), see sfcmkfs(8).

. . . .

11.2.3.1 Defaults of parameters used by sfcmkfs(8)

Specifies a value of parameter along use of the file system that decided beforehand as had described in "Chapter 4 File System Design."

The defaults of parameters used by sfcmkfs(8) are as follows:

Table 11.1 Defaults of parameters used by sfcmkfs(8)

Parameter	Default
Data block size	Always 8192 bytes
Minimum free space (**-o free=***n*)	10% of the file data area
Number of i-nodes (**-o nbpi=***n*)	One in an 8192-byte disk area Up to 16 mega.
Update log area size (**-o logsz=***n*)	1% of the file system size The available range is from 5 megabytes to 50 megabytes.
Meta-data area size (**-o metasz=***n*)	About 10% of the file system size

 Note

. . . .

The use ratio of the meta-data area decreases by growing of the file system size when the meta-data area size is not specified.

. . . .

11.2.3.2 Examples of creating a representative file system

Examples of creating a typical GFS Shared File System are given below.

Single partition configuration

To create a GFS Shared File System with a single partition, specify the partition configuring the file system. In this case, areas for all types of data (meta-data area, update log area, and file data area) are created in the representative partition.

The following is an example of using **sfcmkfs**(8) to create a file system with a single partition.

```
# sfcmkfs -o node=mikan,karin /dev/sfdsk/gfs01/dsk/volume01 <Enter>
   Representative partition (meta-data, log, and data): /dev/sfdsk/gfs01/dsk/volume01
   Shared hosts    : mikan, karin
   Primary MDS : mikan
   Secondary MDS : karin
```

Figure 11.8 Single partition configuration
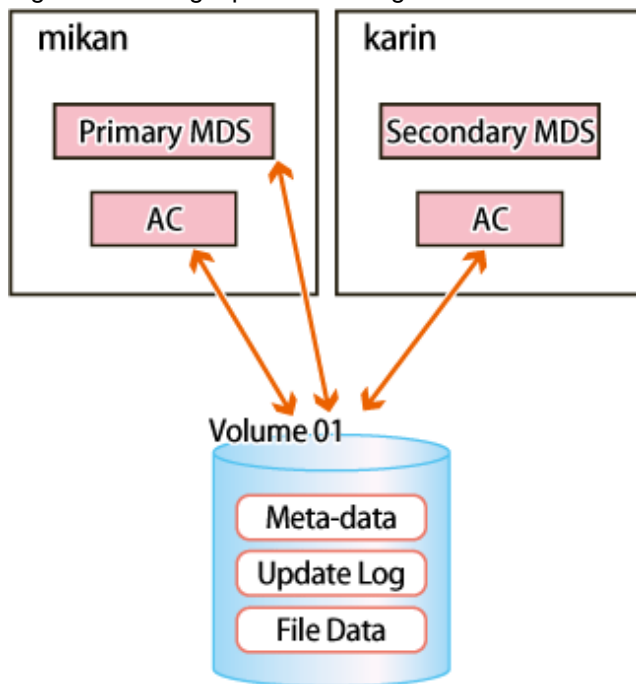
Multiple partitions configuration (Multiple file data partitions configuration)

To specify multiple partitions in the data area, specify a representative partition and file data partitions. In this case, a meta-data area, update log area, and data area are created in the representative partition.

The following is an example of using sfcmkfs(8) to create a file system with multiple file data partitions.

```
# sfcmkfs -o data=/dev/sfdsk/gfs01/dsk/volume02,data=/dev/sfdsk/gfs01/dsk/
volume03,node=mikan,karin /dev/sfdsk/gfs01/dsk/volume01 <Enter>
   Representative partition (meta-data, log, and data): /dev/sfdsk/gfs01/dsk/volume01
   Data partition:  /dev/sfdsk/gfs01/dsk/volume02, /dev/sfdsk/gfs01/dsk/volume03
   Shared hosts:  mikan,karin
   Primary MDS:  mikan
   Secondary MDS:  karin
```

Figure 11.9 Multiple data partitions configurations
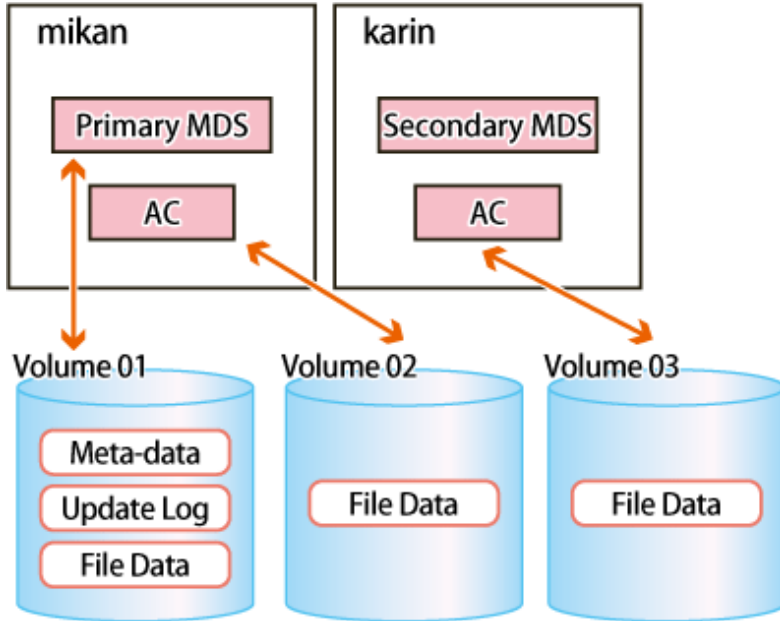
Multiple partitions configuration (Separating the data area)

If multi-partition of the file data area is specified with the "-o dataopt=y" option, representative partitions will not contain the file data area.

The following example shows how to create a file system where the file data area is not included in multiple file data partitions or representative partitions with sfcmkfs(8).

```
# sfcmkfs -o dataopt=y,data=/dev/sfdsk/gfs01/dsk/volume02,data=/dev/sfdsk/gfs01/dsk/
volume03,node=mikan,karin /dev/sfdsk/gfs01/dsk/volume01 <Enter>
  Representative partition (meta-data and log):  /dev/sfdsk/gfs01/dsk/volume01
   Data partition:  /dev/sfdsk/gfs01/dsk/volume02, /dev/sfdsk/gfs01/dsk/volume03
   Shared hosts:  mikan, karin
   Primary MDS:  mikan
   Secondary MDS:  karin
```

Figure 11.10 Separating the data area
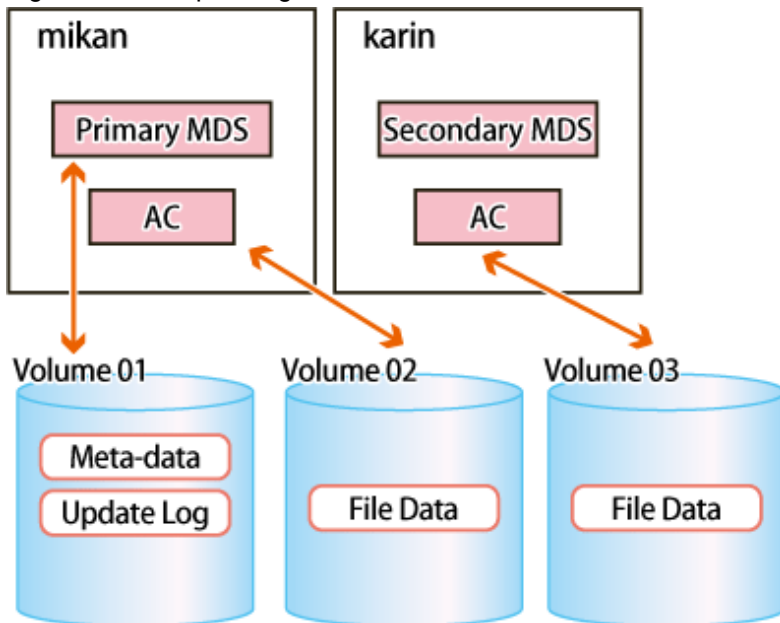
11.2.3.3 Customizing a GFS Shared File System

Parameters for customizing a file system are provided by **sfcmkfs**(8), which is used to create a GFS Shared File Systems. Typical parameters are described below.

Update log area size (-o logsz=*n*)

Specify the size of the update log area with a value from 5 to 100 megabytes. The default size of an update log area is of 1% of the file system size. If, however, 1% of the file system is less than 5 megabytes, 5 megabytes is set. Also, if the value is greater than 50 megabytes, 50 megabytes is set.

Meta-data area size (-o metasz=n)

Specifies the size of the meta-data area. The default is about 10% of the file system size. However, when the file system size is larger, the ratio of the meta-data area becomes smaller.

The minimum value that can be specified is same as default. The maximum value is minimum requirement size for managing 1-mega V-data. (The total number of V-data of the file system can be confirmed by executing sfcdf(8).) However, larger value than the size of the representative partition cannot be specified.

Maximum data area size (-o maxdsz=n)

Specify the maximum total size of the file data area when adding a file data partition as an expansion to a GFS Shared File System. If a file data partition that exceeds the maximum data space, area management information might be insufficient. If this occurs, part of free file data space might not be used. You can add a file data partition using sfcadd(8).

The maximum size is less than 1048576 megabytes (1 terabyte).

Maximum number of partitions (-o maxvol=*n*)

Specify the maximum number of partitions that may ever comprise this GFS Shared File System. The types of partitions configuring the GFS Shared File System are the representative partition, update log partition, and file data partition. The value specified here is the total of the partitions of all types. The default value is 16. Specifiable values are 1 to 32.

11.2.3.4 Setting MDS operational information

The following shows an example of creating a typical GFS Shared File System with MDS operational information specified to improve availability:

Priority settings for MDS placement

Specify the priority of placing the primary MDS and secondary MDS using the shared host name in the -o node option of the sfcmkfs(8) command. The primary MDS and secondary MDS candidates are selected in the order the shared host names are specified. The primary MDS is ultimately determined based on operation status of actual shared nodes when starting the file system operation.

The following shows an example of creating a file system using sfcmkfs(8) specifying mikan as the primary MDS candidate and karin as the secondary MDS candidate.

```
# sfcmkfs -o node=mikan,karin /dev/sfdsk/gfs01/dsk/volume01 <Enter>
   Representative partition (meta-data,log and data):  /dev/sfdsk/gfs01/dsk/volume01
   Shared hosts:  mikan, karin
   Primary MDS:  mikan
   Secondary MDS:  karin
```

11.2.4 Setting fstab

In the GFS Shared File System, it is necessary to code an entry that is to be mounted on /etc/fstab of all the nodes that share a file system. The entry to the same file system should be corresponding by all nodes that share the file system.

Entry of GFS Shared File System sets it to /etc/fstab explains as follows.

- **device to mount field**

  Specify the path name of a block special file to the representative partition of the GFS Shared File System which does the mount.

- **mount point field**

  Specify a directory becoming the mount point at the time of the abbreviation.
  Please specify the directory which there is in all nodes using file system in common.

- **file system type field**

  As a type of the file system, please specify sfcfs.

- **mount options field**

  Please specify a mount option used in a mount.
  Please specify noauto by all means.
  The following table shows the detail of the mount options.

Table 11.2 Mount options

Option	Description
atime | noatime | relatime | strictatime	Sets the method to update the last access time of a file. If **atime** or **relatime** is specified, the last access time of a file is updated when the previous last access time is the same as or older than the last modify time or the last change time. Also, when the previous last access time of a file is old for over a day, it is updated. If **noatime** is specified, the last access time of a file is not updated. If **strictatime** is specified, the last access time of a file is updated for each file access. The default is **relatime**.
atrc | noatrc	Specifies whether to mount (**atrc**) or not to mount (**noatrc**) at the node startup. The default is **atrc**.
async	Mounts in the manner of processing any I/O for the file system asynchronously. The default is **async**.
defaults	Synonymous with **rw**,**suid**,**dev**,**exec**,and **async**.
dev | nodev	Mounts making character special file and block special files available(**dev**) or unavailable (**nodev**). The default is **dev**.
exec | noexec	Mounts giving permission(**exec**) or no permission(**noexec**) to execute binary files on the file system. The default is **exec**.
rw | ro	Mounts the file system as read/write enabled(**rw**) or read only(**ro**). The default is **rw**.
suid | nosuid	Disallows a **setuid** execution. By default, allows **setuid** execution. The default is **suid**.

- **dump frequency field**

  0 Please specify.

- **fsck pass field**

  0 Please specify.

By writing the entries concerning the created file system in **/etc/fstab** on the node that does not share the file system, you can prevent the file system from being used by mistake as other file systems.

Examples of /etc/fstab entries are shown below:

Table 11.3 When mounting in rw mode at startup of the node:

Parameter	Value
device to mount	/dev/sfdsk/gfs01/dsk/volume01
mount point	/mnt/fs1
file system type	sfcfs
mount options	rw,noauto
dump frequency	0
fsck pass	0

Table 11.4 When mounting is not performed at startup of the node

Parameter	Value
device to mount	/dev/sfdsk/gfs01/dsk/volume01
mount point	/mnt/fs1
file system type	sfcfs
mount options	rw,noauto,noatrc
dump frequency	0
fsck pass	0

11.2.5 Selecting a communication protocol

Determine either IPv6 or IPv4 to be used as a communication protocol between MDS and AC, and then set it for each file system.

The default (a communication protocol to be registered when creating a file system) is IPv4.

- A communication protocol can be set for each file system; however, we recommend that users set the same communication protocol for all the file systems so that it can be easily managed.

- Select the communication protocol that matches the IP address type of the shared node registered in host database in "11.2.2 Registering information in host database." If you have selected a communication protocol that does not match it, the GFS Shared File System cannot be mounted.

When using IPv6, the setting procedure is as follows. When using IPv4, this setting is not required.

1. Set 1 to the tuning parameter SFCFS_IPV6.

```
# sfcadm -o SFCFS_IPV6=1 /dev/sfdsk/gfs01/dsk/volume01 <Enter>
```

2. Check that 1 has been set to the tuning parameter SFCFS_IPV6.

```
# sfcinfo -e /dev/sfdsk/gfs01/dsk/volume01 | grep SFCFS_IPV6 <Enter>
SFCFS_IPV6=1
```

11.2.6 Mount

If a GFS Shared File System is used, mount it on all the nodes.

To do so, execute the **sfcmntgl(8)** on any one of the nodes where the GFS Shared File System will be used.

If it is only used on one node, mount the file system using **sfcmount(8)**.

If this command is executed, sfcfsd and MDS will be started on a MDS node is configured. Then, the file system will be mounted on all the nodes.

🛈 Note
. . . .

If the file system is not added in /etc/fstab, mounting of the file system will fail. Also, sfcmntgl(8) and sfcmount(8) are not used for the mount option. If either of them is executed, the current mount option that is specified for the mount option filed of /etc/fstab will be used.
. . . .

11.2.6.1 Mount of all nodes

If the file system is mounted on all the nodes, execute sfcmntgl(8) on any one of the nodes.

🛈 See
. . . .

For details on sfcmntgl(8), see sfcmntgl(8).
. . . .

You can mount the file system on all the nodes as follows:

- When the mount point is specified:

```
# sfcmntgl /mnt/fs1 <Enter>
   Mount point: /mnt/fs1
```

- When a representative partition is specified:

```
# sfcmntgl /dev/sfdsk/gfs01/dsk/volume01 <Enter>
   Representative partition: /dev/sfdsk/gfs01/dsk/volume01
```

- When both the mount point and a representative partition are specified:

```
# sfcmntgl /dev/sfdsk/gfs01/dsk/volume01 /mnt/fs1 <Enter>
   Representative partition: /dev/sfdsk/gfs01/dsk/volume01
   Mount point: /mnt/fs1
```

If mounting fails for a node, the name of the node on which the failure occurred is reported.

11.2.6.2 Mount

If you want to mount a file system on specific nodes, use **sfcmount**(8). Then, the file system will only be mounted on the nodes where the file system will be shared.

🛈 See
. . . .

For details on **sfcmount**(8), see sfcmount (8).
. . . .

Mount the GFS Shared File System on the node as shown below:

- When the mount point is specified:

```
# sfcmount /mnt/fs1 <Enter>
    Mount point: /mnt/fs1
```

- When a representative partition is specified:

```
# sfcmount /dev/sfdsk/gfs01/dsk/volume01 <Enter>
    Representative partition: /dev/sfdsk/gfs01/dsk/volume01
```

- When both the mount point and a representative partition are specified:

```
# sfcmount /dev/sfdsk/gfs01/dsk/volume01 /mnt/fs1 <Enter>
    Representative partition: /dev/sfdsk/gfs01/dsk/volume01
    Mount point: /mnt/fs1
```

11.2.7 Checking file system status

sfcrscinfo(8) can check if a GFS Shared File System can be mounted. Execute the command on any one of the nodes where the file system will be shared:

See
For details on sfcrscinfo(8), see sfcrscinfo(8).

The following example shows how to check the mount status of the GFS Shared File System whose representative partition is /dev/sfdsk/gfs/dsk/volume01 with sfcrscinfo(8).

```
# sfcrscinfo -m /dev/sfdsk/gfs01/dsk/volume01 <Enter>
/dev/sfdsk/gfs01/dsk/volume01:
FSID  MDS/AC  STATE  S-STATE  RID-1  RID-2  RID-N  hostname
   2  MDS(P)  run    -            0      0      0  mikan
   2  AC      run    -            0      0      0  mikan
   2  MDS(S)  wait   -            0      0      0  karin
   2  AC      run    -            0      0      0  karin
```

If following states can be confirmed, the file system is operated normally.

- Either state of MDS(P) or MDS(S) is "run-".

- States of access client (AC) within the range of a shared node of the object file system are "run-" all.

11.2.8 Notes applying when the partitions of a created file system are used

When creating a GFS Shared File System using a partition of an existing file system, be aware of the following important information.

11.2.8.1 The GFS Shared File System

To use a partition that is already in use by an existing GFS Shared File System for a new GFS Shared File System, delete the existing file system and then create the new file system.

See
For information about deleting a file system, see "11.8 Deleting."

11.2.8.2 File systems other than GFS Shared File System

To use a partition that is in use by a file system other than the GFS Shared File System, unmount the file system on all nodes and then delete in accordance with the procedure for that file system. Also, delete the definition in /etc/fstab and then create the GFS Shared File System.

11.3 Change (file system attributes)

To change mount information (GFS Shared File System attribute change) or tuning the file system, perform the following tasks:

1. Unmount

2. Change the file system attributes

3. Mount

11.3.1 Unmount

To change the GFS Shared File System, unmount it on all the nodes.

First, stop using the file system by stopping applications.

Then, unmount the file system.

Check if the file system is unmounted properly on all the nodes using sfcrscinfo(8).

See
....
In order to use the process of the file system, execute fuser(1). For details about the command, see the online manual page.
For information on how to check the file system status using sfcrscinfo(8), see "11.2.7 Checking file system status."
....

11.3.1.1 Unmount of all nodes

To unmount the file system on all the nodes, use sfcumntgl(8). Execute the command on any one of the nodes where the file system is shared.

See
....
For details on sfcumntgl(8), see sfcumntgl(8).
....

- When the mount point is specified:

```
# sfcumntgl /mnt/fs1 <Enter>
   Mount point: /mnt/fs1
```

- When a representative partition is specified:

```
# sfcumntgl /dev/sfdsk/gfs01/dsk/volume01 <Enter>
   Representative partition: /dev/sfdsk/gfs01/dsk/volume01
```

If unmounting fails on a node, the name of the node on which the failure occurred is reported.

11.3.1.2 Unmount

To unmount the file system on specific nodes, use **sfcumount**(8). Execute the command on any one of the nodes where the file system is shared.

See
....
For details on **sfcumount**(8), see sfcumount(8).
....

Unmount the GFS Shared File System on the node as shown below:

- When the mount point is specified:

```
# sfcumount /mnt/fs1 <Enter>
    Mount point: /mnt/fs1
```

- When a representative partition is specified:

```
# sfcumount /dev/sfdsk/gfs01/dsk/volume01 <Enter>
    Representative partition: /dev/sfdsk/gfs01/dsk/volume01
```

![Note]**Note**

....

To unmount the GFS Shared File System, be sure to use sfcumount(8).

....

11.3.2 Change the file system attributes

11.3.2.1 Changing the mount information

To change the mount information of the GFS Shared File System, change fstab information. The parameters of mount point and mount options can be changed.

![See]**See**

....

For details of setting fstab, see "11.2.4 Setting fstab."

....

11.3.2.2 Tuning file system

In the GFS Shared File System, each file system can be tuned for the following items:

- Communication timeout value (CON_TIMEOUT)

- Communication protocol (SFCFS_IPV6)

Execute sfcadm(8) to change the above value on any one of the nodes where the file system is shared.

![See]**See**

....

For details of sfcadm(8), see sfcadm(8).

....

The following example shows how to set the timeout value to 240 seconds for the existing file system (/dev/sfdsk/gfs01/volume01) using sfcadm(8).

```
# sfcadm -o CON_TIMEOUT=240 /dev/sfdsk/gfs01/dsk/volume01 <Enter>
```

![See]**See**

....

For details about the tuning parameter, see "14.1.2 Communication timeout value" and "14.1.3 Communication protocol" in this manual.

....

11.3.3 Mount

Mount the file system after completing the GFS Shared File System changes.

See

. . . .

For details on how to mount the GFS Shared File System, see "11.2.6 Mount."

. . . .

11.4 Change (partition addition)

To add a file data partition (GFS Shared File System configuration change), perform the following tasks:

1. Unmount

2. Setting shared disk

3. Adding a file data partition

4. Mount

11.4.1 Unmount

Unmount the GFS Shared File System on all the nodes.

 See

. . . .

For details on how to unmount the GFS Shared File System, see "11.3.1 Unmount."

. . . .

11.4.2 Setting shared disk

The partition that will be added to GFS Shared File System must be GDS logical volume. The volume must be ACTIVE on each node.

 Note

. . . .

GDS logical volume should be set up as follows.

- Set up the type of a disk class with which GDS logical volume belongs as shared.(Change GDS class attributes)

- Set all the nodes that share GFS Shared File System as the scope of the disk class to which logical volume belongs.(Change GDS class attributes)

. . . .

 See

. . . .

When adding the logical volume of GDS, which was used in a file system other than GFS, as the partition of the GFS Shared File System, see "1.8.4.2 Notes when using logical volumes of GDS that were used in a file system other than GFS."

. . . .

 See

. . . .

For GDS disk class operations, see the applicable items under "Operation" "GDS Management View" in "PRIMECLUSTER Global Disk Services Configuration and Administration Guide."
For GDS logical volume operations, see the applicable items under "Operation" "GDS Management View" in the "PRIMECLUSTER Global Disk Services Configuration and Administration Guide."

. . . .

11.4.3 Partition addition

To add a partition as file data area, use sfcadd(8). Execute the command on any one of the nodes where the file system is shared.

See

For details on sfcadd(8), see sfcadd(8).

Note

If you add a partition to the file system that is not stopped properly, recovery the file system by executing the sfcfsck(8) command in advance.

How to add a file data partition (/dev/sfdsk/gfs01/dsk/volume02) to an existing file system (/dev/sfdsk/gfs01/dsk/volume01) by sfcadd(8) is shown below.

1. Check the configuration of the current file system.

```
# sfcinfo /dev/sfdsk/gfs01/dsk/volume01 <Enter>
/dev/sfdsk/gfs01/dsk/volume01:
FSID special                                   size  Type  mount
   1 /dev/sfdsk/gfs01/dsk/volume01(1e756)     14422  META  -----
   1 /dev/sfdsk/gfs01/dsk/volume01(1e756)      5116  LOG   -----
   1 /dev/sfdsk/gfs01/dsk/volume01(1e756)     95112  DATA  -----
```

2. Add a file data partition (/dev/sfdsk/gfs01/dsk/volume02) as file data area.

```
# sfcadd -D /dev/sfdsk/gfs01/dsk/volume02 /dev/sfdsk/gfs01/dsk/volume01 <Enter>
```

3. Confirm that the file data partition has been added.

```
# sfcinfo /dev/sfdsk/gfs01/dsk/volume01 <Enter>
/dev/sfdsk/gfs01/dsk/volume01:
FSID special                                   size  Type  mount
   1 /dev/sfdsk/gfs01/dsk/volume01(1e756)     14422  META  -----
   1 /dev/sfdsk/gfs01/dsk/volume01(1e756)      5116  LOG   -----
   1 /dev/sfdsk/gfs01/dsk/volume01(1e756)     95112  DATA  -----
   1 /dev/sfdsk/gfs01/dsk/volume02(1e757)     65512  DATA  -----
```

Note

To add a file data partition, be aware of the following:

- A file data partition addition is not possible if doing so would exceed the maximum number of partitions (-o maxvol=n) for the file system, as specified by sfcmkfs(8).

- When file system's file data area size after adding partition exceeds the maximum file data area size specified by sfcmkfs(8) (-o maxdsz=n), a part of free area in file data area might not be able to be used because of management region shortage.

11.4.4 Mount

After the GFS Shared File System change is completed, mount the file system.

See

For details on how to mount the GFS Shared File System, see "11.2.6 Mount."

11.5 Change (shared node information)

To add or delete a node from the nodes where the GFS Shared File System is shared, perform the following:

1. Unmount

2. Shared disk setting (When adding node information)

3. Registering information in host database (When adding node information)

4. Shared node information change

5. Setting fstab

6. Mount

🛈 Note
. . . .

Addition of a node is not allowed if it exceeds the maximum size of the GFS Shared File System. See "1.7 Upper Limits of the File System."

A node where a primary MDS and secondary MDS are configured cannot be deleted from a group of shared nodes.
. . . .

11.5.1 Unmount

To change shared node information, unmount the GFS Shared File System on all the nodes.

📖 See
. . . .
For details on how to unmount the GFS Shared File System, see "11.3.1 Unmount."
. . . .

11.5.2 Shared disk setting (When adding node information)

When adding a node to a group of nodes that share the file system, be aware of the following important information.

🛈 Note
. . . .

The node must be cluster-configured.

The node must provide the GFS Shared File System that is registered in the management partition.

It is necessary to include the adding node in the scope of the disk class of the GDS logical volumes that are used as management partition and GFS Shared File System.
. . . .

📖 See
. . . .
For information on how to add a node to a cluster system, see "Adding a new node to CF" of "PRIMECLUSTER Cluster Foundation (CF) Configuration and Administration Guide."
For information on how to operate a disk class to which the GDS logical volume belongs, see "Operation" "GDS Management View" in "PRIMECLUSTER Global Disk Services Configuration and Administration Guide."
For information on how to add a node to a group of nodes that provide the GFS Shared File System, see "8.1.2 Adding node configuration information to the management partition" or "9.1.2 Adding node configuration information to the management partition."
. . . .

11.5.3 Registering information in host database (When adding node information)

When adding a shared node, register its host name and IP address in host database (for example: /etc/host).

1. Register the host names and IP addresses of all the shared nodes in host database of the shared nodes to be added. You can check the host names of the existing shared nodes with the -**n** option of **sfcinfo(8)**.

```
# sfcinfo -n /dev/sfdsk/gfs01/dsk/volume01 <Enter>
/dev/sfdsk/gfs01/dsk/volume01:
FSID  hostID    status  hostname
   1  80a4f75b  STOP    sunny
```

2. Resister the host name and IP address of the shared node to be added in host database of the existing shared nodes.

📖 Note

Register the IP address that matches the communication protocol between MDS and AC. You can check the communication protocol with -**e** option of **sfcinfo(8)** on the existing shared node.

```
# sfcinfo -e /dev/sfdsk/gfs01/dsk/volume01 | grep SFCFS_IPV6 <Enter>
SFCFS_IPV6=1
```

If "SFCFS_IPV6=1" is displayed, the communication protocol is IPv6. If "SFCFS_IPV6=0" or nothing is displayed, the communication protocol is IPv4.

📖 See

For notes on registering information in host database, see "11.2.2 Registering information in host database."

11.5.4 Shared node information change

The GFS Shared File System stores information about nodes that share the file system in each partition to restrict access from nodes other than the nodes that share the file system.

The node information mainly contains the following information:

- Host ID

- Host name

To add and delete node information for the file system, use **sfcnode**(8).

📖 See

For details on the **sfcnode**(8) options, see sfcnode(8).

11.5.4.1 Adding node information

Described below is the procedure for using sfcnode(8) to add node information (moony) to an existing file system (/dev/sfdsk/gfs01/dsk/volume01). Execute the command on any one of the nodes where the file system is shared.

1. Check that the target file system is unmounted on all nodes sharing the file system.

   You can do this by checking that "STOP" is displayed in all status fields of the sfcinfo(8) outputs.

```
# sfcinfo -n /dev/sfdsk/gfs01/dsk/volume01 <Enter>
/dev/sfdsk/gfs01/dsk/volume01:
FSID  hostID    status  hostname
   1  80a4f75b  STOP    sunny
```

2. Add the node information.

```
# sfcnode -a moony /dev/sfdsk/gfs01/dsk/volume01 <Enter>
```

3. Confirm that the node information has been added.

   You can do this by checking that the moony field is newly displayed.

```
# sfcinfo -n /dev/sfdsk/gfs01/dsk/volume01 <Enter>
/dev/sfdsk/gfs01/dsk/volume01:
FSID  hostID    status  hostname
   1  80a4f75b  STOP    sunny
   1  80960096  STOP    moony
```

4. Confirm the MDS node.

   You can do this by checking that the moony is not displayed as MDS(S).

```
# sfcrscinfo -m /dev/sfdsk/gfs01/dsk/volume01 <Enter>
/dev/sfdsk/gfs01/dsk/volume01:
FSID  MDS/AC  STATE  S-STATE  RID-1  RID-2  RID-N  hostname
   1  MDS(P)  stop   -            0      0      0  sunny
   1  AC      stop   -            0      0      0  sunny
   1  AC      stop   -            0      0      0  moony
```

5. If it is necessary, set the added node as MDS node.

```
# sfcadm -g ,moony /dev/sfdsk/gfs01/dsk/volume01 <Enter>
```

6. Check if the node to the MDS node is configured has been added.

   If moony is displayed as MDS(S), the node has been added.

```
# sfcrscinfo -m /dev/sfdsk/gfs01/dsk/volume01 <Enter>
/dev/sfdsk/gfs01/dsk/volume01:
FSID  MDS/AC  STATE  S-STATE  RID-1  RID-2  RID-N  hostname
   1  MDS(P)  stop   -            0      0      0  sunny
   1  AC      stop   -            0      0      0  sunny
   1  MDS(S)  stop   -            0      0      0  moony
   1  AC      stop   -            0      0      0  moony
```

11.5.4.2 Deleting node information

Described below is the procedure for using sfcnode(8) to delete node information (moony) from an existing file system (/dev/sfdsk/gfs01/dsk/volume01). If you are deleting node information while the file system is unmounted, execute sfcnode(8) on a node that shares the target file system.

1. Check that the file system is unmounted in all nodes configuring the current file system and all nodes sharing the file system.

   You can do this by checking that "STOP" is displayed in all status fields of the sfcinfo(8) outputs.

```
# sfcinfo -n /dev/sfdsk/gfs01/dsk/volume01 <Enter>
/dev/sfdsk/gfs01/dsk/volume01:
FSID  hostID    status  hostname
   1  80a4f75b  STOP    sunny
   1  80960096  STOP    moony
```

2. Confirm whether the node of the node information to delete is the MDS node.

   You can do this by checking that the moony is displayed as MDS(S).

```
# sfcrscinfo -m /dev/sfdsk/gfs01/dsk/volume01 <Enter>
/dev/sfdsk/gfs01/dsk/volume01:
```

```
FSID  MDS/AC  STATE  S-STATE  RID-1  RID-2  RID-N  hostname
   1  MDS(P)  stop   -            0      0      0  sunny
   1  AC      stop   -            0      0      0  sunny
   1  MDS(S)  stop   -            0      0      0  moony
   1  AC      stop   -            0      0      0  moony
```

3. If the node of the node information to delete is MDS node, delete it from the MDS node.

```
# sfcadm -g moony /dev/sfdsk/gfs01/dsk/volume01 <Enter>
```

4. Check if the node to the MDS node is configured has been deleted.

   If moony is not displayed as MDS(S), the node has been deleted.

```
# sfcrscinfo -m /dev/sfdsk/gfs01/dsk/volume01 <Enter>
/dev/sfdsk/gfs01/dsk/volume01:
FSID  MDS/AC  STATE  S-STATE  RID-1  RID-2  RID-N  hostname
   1  MDS(P)  stop   -            0      0      0  sunny
   1  AC      stop   -            0      0      0  sunny
   1  AC      stop   -            0      0      0  moony
```

5. Delete the node information.

```
# sfcnode -d -h moony /dev/sfdsk/gfs01/dsk/volume01 <Enter>
```

6. Confirm that the node information has been deleted.

   You can do this by checking that the moony field is not displayed in the sfcinfo(8) outputs.

```
# sfcinfo -n /dev/sfdsk/gfs01/dsk/volume01 <Enter>
/dev/sfdsk/gfs01/dsk/volume01:
FSID  hostID    status  hostname
   1  80a4f75b  STOP    sunny
```

11.5.5 Setting fstab

In the GFS Shared File System, it is necessary to add a mount entry to /etc/fstab on all of the nodes that are sharing a file system. In addition, if a mount entry is added to /etc/fstab on the node that is not sharing a file system, the special file is not used as other file systems even if mistaking.

Add the GFS Shared File System to /etc/fstab of nodes that have been added.

If nodes are deleted, change the /fstab settings so that the GFS Shared File System will not be mounted.

See
. . . .
For details on setting /etc/fstab, see "11.2.4 Setting fstab."
. . . .

11.5.6 Mount

After change of the shared node information is completed, mount the file system.

See
. . . .
For details on how to mount the GFS Shared File System, see "11.2.6 Mount."
. . . .

11.6 Change (re-creating a file system)

To recover the GFS Shared File System or restore all the files from the backup file to clear up the fragment, perform the following file system re-creating procedure.

1. Unmount

2. Re-creating the file system

3. Mount

11.6.1 Unmount

To recreate the GFS Shared File System, unmount the file system.

See
....
For details on how to unmount the GFS Shared File System, see "11.3.1 Unmount."
....

11.6.2 Re-creating the file system

To re-create the file system, use **sfcmkfs** (8). You can recreate the file system in the same configuration as before without having to delete it, using **sfcmkfs** (8) with the "-o force" option. Execute the command on any one of the nodes where the file system is shared.

See
....
For details on the **sfcmkfs**(8) options, see sfcmkfs(8).
....

The following is an example of using sfcmkfs(8) to re-create a file system with a single partition.

```
# sfcmkfs -o force,node=mikan,karin /dev/sfdsk/gfs01/dsk/volume01 <Enter>
   Representative partition (meta-data, log, and data): /dev/sfdsk/gfs01/dsk/volume01
   Shared hosts: mikan, karin
   Primary MDS: mikan
   Secondary MDS: karin
```

Note
....
When the file system is re-created, data of the previous file system will be deleted. Back up your data if necessary.
....

11.6.3 Mount

After re-creating of the GFS Shared File System is completed, mount the file system.

See
....
For details on how to mount the GFS Shared File System, see "11.2.6 Mount."
....

11.7 Change (MDS operational information)

To change the MDS operational information of the GFS Shared File System, perform the following tasks:

1. Unmount

2. Changing the MDS operational information

3. Mount

11.7.1 Unmount

To change the MDS operational information of the GFS Shared File System, unmount the file system.

🗂 See
....
For details on how to unmount the GFS Shared File System, see "11.3.1 Unmount."
....

11.7.2 Changing the MDS operational information

In the GFS Shared File System, information about all nodes that operate MDS is retained.

- Priority of nodes on which to place MDS as primary or secondary MDS

Use sfcadm(8) to change information about nodes on which MDS is placed.

🗂 See
....
For details on sfcadm(8) options, see sfcadm(8).
....

The following example shows how to change the priority of the MDS nodes is configured to karin and mikan, using sfcadm(8).

1. Check the current file system configuration.

```
# sfcrscinfo -m /dev/sfdsk/gfs01/dsk/volume01 <Enter>
/dev/sfdsk/gfs01/dsk/volume01:
FSID   MDS/AC   STATE   S-STATE    RID-1    RID-2    RID-N  hostname
   1   MDS(P)   stop    -              0        0        0  mikan
   1   AC       stop    -              0        0        0  mikan
   1   MDS(S)   stop    -              0        0        0  karin
   1   AC       stop    -              0        0        0  karin
```

You can check the priority of the MDS nodes are mikan and karin from the result of sfcrscinfo -m executed.

2. Change the priority of the MDS nodes is configured.

```
# sfcadm -g karin,mikan /dev/sfdsk/gfs01/dsk/volume01 <Enter>
```

3. Check if the priority of the MDS nodes is configured has been changed.

```
# sfcrscinfo -m /dev/sfdsk/gfs01/dsk/volume01 <Enter>
/dev/sfdsk/gfs01/dsk/volume01:
FSID   MDS/AC   STATE   S-STATE    RID-1    RID-2    RID-N  hostname
   1   MDS(P)   stop    -              0        0        0  karin
   1   AC       stop    -              0        0        0  karin
   1   MDS(S)   stop    -              0        0        0  mikan
   1   AC       stop    -              0        0        0  mikan
```

11.7.3 Mount

After change of MDS operational information is completed, mount the file system.

🗂 See
....
For details on how to mount the GFS Shared File System, see "11.2.6 Mount."
....

11.8 Deleting

To delete the GFS Shared File System, perform the following tasks:

1. Unmount

2. Removing the entry in /etc/fstab

3. Deleting the file system

11.8.1 Unmount

To delete the GFS Shared File System, unmount the file system. If there are any applications that are used for the file system, delete them as well.

📑 See

. . . .

For details on how to unmount the GFS Shared File System, see "11.3.1 Unmount."

. . . .

11.8.2 Removing the entry in /etc/fstab

To delete the GFS Shared File System, delete the file system from /etc/fstab of all the nodes.

11.8.3 Deleting the file system

To delete a file system that is not being used in the GFS Shared File System, use **sfcadm**(8) with the **-D** option specified. Execute the command on any one of the nodes where the file system is shared.

📑 See

. . . .

For details on the **sfcadm**(8) options, see sfcadm(8).

. . . .

How to delete a file system using **sfcadm**(8) is shown below.

1. Check the current file system information.

```
# sfcinfo -a <Enter>
/dev/sfdsk/gfs01/dsk/volume01:
FSID special                                  size  Type  mount
   1 /dev/sfdsk/gfs01/dsk/volume01(1e756)     14422  META  -----
   1 /dev/sfdsk/gfs01/dsk/volume01(1e756)      5116  LOG   -----
   1 /dev/sfdsk/gfs01/dsk/volume01(1e756)     95112  DATA  -----
/dev/sfdsk/gfs02/dsk/volume01:
FSID special                                  size  Type  mount
   2 /dev/sfdsk/gfs02/dsk/volume01(1e757)     14422  META  -----
   2 /dev/sfdsk/gfs02/dsk/volume01(1e757)      5116  LOG   -----
   2 /dev/sfdsk/gfs02/dsk/volume01(1e757)     95112  DATA  -----
```

2. Delete the file system whose representative partition is /dev/sfdsk/gfs/dsk/volume01.

```
# sfcadm -D /dev/sfdsk/gfs02/dsk/volume01 <Enter>
```

3. Confirm that the file system has been deleted.

```
# sfcinfo -a <Enter>
/dev/sfdsk/gfs01/dsk/volume01:
FSID special                                  size  Type  mount
   1 /dev/sfdsk/gfs01/dsk/volume01(1e756)     14422  META  -----
```

```
1 /dev/sfdsk/gfs01/dsk/volume01(1e756)          5116   LOG   -----
1 /dev/sfdsk/gfs01/dsk/volume01(1e756)         95112   DATA  -----
```

Chapter 12 File System Management

This chapter describes the procedures for managing the GFS Shared File System using basic commands.

To execute the procedure explained in this Chapter, a file system must have been created in advance.

For details on how to create a file system, see the following:

- "Chapter 10 File System Operations (GUI)"

- "Chapter 11 File System Operations (Command)"

12.1 File System Management Commands

The GFS Shared File System provides exclusive management commands while file system management commands are multipurpose to support various types of file system.

The GFS Shared File System also provides specific management command for its unique features.

The following shows lists of file system management commands provided by the GFS. The lists show description of each command function and whether each command can be used when the target file system is in the mount or unmount state.

Table 12.1 GFS Shared File System exclusive management commands

Command	Description	Used in mount state	Used in unmount state
sfcdf	Displays configuration information and status of use of the GFS Shared File System.	A	A
sfcfsck	Checks a GFS Shared File System for consistency and repairs it.	N/A	A
sfcmkfs	Creates a GFS Shared File System.	A (-m option only)	A
sfcmount	Mounts the GFS Shared File System on a local node.	N/A	A
sfcumount	Unmounts the GFS Shared File System on a local node.	A	N/A

GFS : Global File Services

A: Applicable, N/A: Not applicable

Table 12.2 GFS Shared File System specific management commands

Command	Description	Used in mount state	Used in unmount state
sfcadd	Adds file data partitions.	N/A	A
sfcadm	Changes partition information setting.	N/A	A
sfcfrmstart	Starts sfcfrmd daemon on a local node.	N/A	A
sfcfrmstop	Stops sfcfrmd daemon on a local node.	N/A	A
sfcgetconf	Makes a backup of the management partition.	A	A
sfcinfo	Displays partition information.	A	A
sfcmntgl	Mounts the GFS Shared File System on all nodes.	A (*1)	A
sfcnode	Adds, deletes, and changes node configuration information.	N/A	A

Command	Description	Used in mount state	Used in unmount state
sfcrscinfo	Displays file system information.	A	A
sfcsetup	Initializes the management partition, adds, deletes, and displays node information, displays a path of the management partition, and registers and displays the startup mode of the sfcfrmd daemon.	A (*2)	A
sfcstat	Reports statistical information of a GFS Shared File System.	A	N/A
sfcumntgl	Unmounts the GFS Shared File System from all nodes.	A	A

GFS : Global File Services

A: Applicable, N/A: Not applicable

*1) An error occurs when all nodes are in the mount state.

*2) The following options can be used in the mount state.

```
-p option
-m option without argument
No option
```

12.2 Checking a File System for Consistency and Repairing It

If a file system is damaged and its consistency is lost, for example, due to an automatic recovery error after a primary MDS failure, the file system must be checked and restored using sfcfsck(8).

The GFS Shared File System provides the update log function to ensure high-speed recovery if an error such as a system failure occurs. If the update log function has been enabled, the file system can be repaired quickly regardless of the size of the file system. This is possible because of update log replay, which updates the un-updated portion of the meta-data located in the update log area.

By default, or when "-o log" is specified, sfcfsck(8) repairs the file system by replaying the update log.

If the update log data has been physically damaged, sfcfsck(8) does not execute update log replay, but automatically performs a full check on the file system. To meet the need for immediate system operation resuming, an option "-o elog" that avoids file system restoration while it provides update log replay. If this option is specified, sfcfsck(8) terminates immediately without performing check and recovery when the update log data has been physically damaged. In this event, the file system cannot be mounted unless check and recovery is performed using sfcfsck(8). The mounting on such a file system should be attempted after it is restored through full checking without update log replay by the "-o nolog" option.

The following example repairs a file system with log replay.

```
# sfcfsck /dev/sfdsk/gfs01/dsk/volume01 <Enter>
```

The following example performs a full check on the file system and repairs it without log replay.

```
# sfcfsck -o nolog /dev/sfdsk/gfs01/dsk/volume01 <Enter>
```

See
....
For details about sfcfsck(8) options, see sfcfsck(8) in this manual.
....

12.2.1 Action to be taken when fsck terminates abnormally

The following describes available solutions for sfcfsck(8) abnormal termination.

12.2.1.1 Memory allocation request error

The following message is displayed if sfcfsck(8) fails in requesting for memory allocation for the internal table.

```
Can't allocate memory
```

Add a swap area.

12.2.1.2 File system configuration information acquisition failure

The following message is output if acquisition of the file system configuration information fails.

```
Can not connect to FsRM

other node information get error
```

See
....
Start the file system monitoring facility on all the nodes that provide the GFS Shared File System. See "9.2.4 Starting sfcfrmd daemon."
....

12.2.1.3 File system partition configuration data error

This message indicates that the command terminated abnormally because a mismatch was found in the **partition configuration data** for the GFS Shared File System partitions.

```
Can't open device-name: errno

setup: Volume information error

setup: fsck quit due to unrecoverable error!
```

Use sfcadm(8) to recover the partition information.

The following example shows how to restore partition information on the "/dev/sfdsk/gfs01/dsk/volume01" device.

For a single partition, perform the following operation:

```
# sfcadm -r -a /dev/sfdsk/gfs01/dsk/volume01 <Enter>
```

For a multi-partition configuration, use sfcinfo(8) to check the partitions and then use sfcadm(8) to perform recovery using the **representative partition**.

```
# sfcinfo /dev/sfdsk/gfs01/dsk/volume01 <Enter>
/dev/sfdsk/gfs01/dsk/volume01:
FSID special                             size   Type   mount
   1 /dev/sfdsk/gfs01/dsk/volume01(1e756)    25418   META   -----
   1 /dev/sfdsk/gfs01/dsk/volume01(1e756)     5120   LOG    -----
   1 /dev/sfdsk/gfs01/dsk/volume01(1e756)   232256   DATA   -----
   1 /dev/sfdsk/gfs01/dsk/volume02(1e757)   341778   DATA   -----
   1 /dev/sfdsk/gfs01/dsk/volume03(1e758)   578704   DATA   -----

# sfcadm -r /dev/sfdsk/gfs01/dsk/volume01 <Enter>
```

12.2.1.4 Node that executed sfcfsck(8) is not a shared node of the file system

If the node which executed sfcfsck(8) is not a shared node of the file system, the following message is output.

```
No node volume entry for localhost,file system access forbidden!
```

Execute sfcfsck(8) at the shared node of the file system.

12.2.1.5 Irreparable file system destruction

The following message indicates that the file system is irreparable because of the GFS Shared File System data has been destroyed.

```
BAD SUPER BLOCK

BAD VOLUME INFORMATION

No Meta volume information available!

No Log volume information available!

Only found num data volume, total num in superblock!
```

If a backup of the file system is available, recover the file system from the backup.

12.2.1.6 Operation error

The following message is displayed if there are problems in the parameter or the execution environment of sfcfsck(8).

```
not super user

duplicate option

filesystem lock failed. errno(error-number)

device-name is a mounted file system

device-name is not for sfcfs

Can't check device-name

Can't open /etc/fstab

setmntent failed. errno(error-number)

fs_clean is not FSLOG. cannot log replay!
```

Confirm whether the parameter and the execution user of the command are correct or /etc/fstab and /etc/mtab exist.

12.2.1.7 Repairing of file system is non-completion

The following message is displayed if the file system was not repaired completely.

```
***** PLEASE RERUN FSCK *****
```

```
device-name: UNEXPECTED INCONSISTENCY; RUN sfcfsck MANUALLY.

Log Replay is failed.
```

Execute a full file system check using sfcfsck(8).

The following example shows how to perform check the file system on the "/dev/sfdsk/gfs01/dsk/volume01" device.

```
# sfcfsck -o nolog /dev/sfdsk/gfs01/dsk/volume01 <Enter>
```

12.2.1.8 Other messages

When the message which shows other abnormalities is output and sfcfsck(8) terminates abnormally, please tell field engineers the output message of sfcfsck after collecting the diagnostic data with FJQSS (Information Collection Tool), fjsnap(8), or pclsnap(8).

12.3 Extending a File System

The GFS Shared File System can increase the size of an existing file system without saving or restoring data.

Use the sfcadd(8) command to enlarge the file system. To expand the file system by adding file data partitions, be sure that the file system is being unmounted.

See
. . . .
For the execution procedure using Operation Management View, see "10.1.3 Change (partition addition)." For the execution procedure using line commands, see "11.4 Change (partition addition)."
. . . .

Note
. . . .
If you schedule an expansion of the file system, specify the expanding maximum file data area size (-o maxdsz=n) argument of sfcmkfs(8) when creating the file system. If the maximum data area size of the file system is not specified during file system creation, the resource that is able to manage the file data area's total size of all specified partition is automatically reserved. So, when the file system is expanded, managed resources might be insufficient. If the free file data area is not fragmented, the resources will not be insufficient. If there are too many fragments in the file, part of free area may not be used.
. . . .

Information
. . . .
If managed resources are insufficient, the following message will be output to the console:

NOTICE: sfcfs_mds:3004: *fsid*: too much fragment (extent)
. . . .

See
. . . .
For corrective action of the above message, see "A.2.13 Information messages (MDS (sfcfsmg daemon))."
. . . .

12.4 Displaying File System Information

The usage status and management status of the file system, and the information of partition and node can be displayed in the GFS Shared File System.

12.4.1 Acquiring file system information

Information about a file system, including the usage status of disk blocks, the state of free disk blocks, and the number of files, can be acquired using df(1) or sfcdf(8).

- The following information items is displayed for a **file data area**:

  - Total number of blocks and number of free blocks

The following message specific to the GFS Shared File System is output with sfcdf(8).

- The following information is displayed for a meta-data partition:

  - Use of i-nodes (number of i-nodes, number of free i-nodes)

  - Free **extent** status (maximum and unused leaves)

  - V-data usage status(maximum and unused V-data blocks)

- Only the number of assigned blocks is displayed for an **update log area**.

The following example shows how to display the status of the file system using df(1) and sfcdf(8).

```
# df -k /mnt <Enter>
Filesystem       1K-blocks      Used Available Use% Mounted on
/dev/sfdsk/gfs/dsk/volume01
                   4684576  2543704  1672416   61% /mnt
# sfcdf -v /dev/sfdsk/gfs01/dsk/volume01 <Enter>
Filesystem:/mnt
                   inodes    free    vdata    free   leaf    free    Type
/dev/sfdsk/gfs01/dsk/volume01
                   585524  580417   195174  187973  234208  227772   META
                   kbytes
/dev/sfdsk/gfs01/dsk/volume01
                    51197                                            LOG
                   kbytes    used    free   %used
/dev/sfdsk/gfs01/dsk/volume01
                   4684576  2543704  2140872   54%                   DATA
```

 See

. . . .

For details about df(1), see df(1) in the online manual.
For details about sfcdf(8) options, see sfcdf(8) in this manual.

. . . .

12.4.2 Displaying partition/node information

The partition information about the partition set configuring the GFS Shared File System can be displayed using sfcinfo(8). Node information can also be displayed.

The following example shows how to display the partition and node information for a file system using sfcinfo(8).

```
# sfcinfo /dev/sfdsk/gfs01/dsk/volume01 <Enter>
FSID special                                    size  Type  mount
   1 /dev/sfdsk/gfs01/dsk/volume01(1e756)      14422  META  -----
   1 /dev/sfdsk/gfs01/dsk/volume01(1e756)       5116  LOG   -----
   1 /dev/sfdsk/gfs01/dsk/volume01(1e756)      95112  DATA  -----
# sfcinfo -n /dev/sfdsk/gfs01/dsk/volume01 <Enter>
/dev/sfdsk/gfs01/dsk/volume01:
FSID  hostID    status  hostname
   1  80804266  RUN     mikan
   1  807e7937  RUN     karin
```

12.4.3 Displaying file system management state

The management state of the GFS Shared File System can be checked using sfcrscinfo(8).

Check the information displayed by this command before mounting file systems, unmounting file systems, or changing file system configuration.

The following example shows how to check the management state using sfcrscinfo(8).

```
# sfcrscinfo -am <Enter>
/dev/sfdsk/gfs01/dsk/volume01:
FSID  MDS/AC  STATE  S-STATE  RID-1  RID-2  RID-N  hostname
   2  MDS(P)  run    -             0      0      0  mikan
   2  AC      run    -             0      0      0  mikan
   2  MDS(S)  wait   -             0      0      0  karin
   2  AC      run    -             0      0      0  karin
```

The meaning of sfcrscinfo(8) is as follows:

- MDS/AC

    - MDS(P): The primary MDS on a definition

    - MDS(S): The secondary MDS on a definition

    - AC : The node that shares a file system

- STATE/S-STATE (MDS)

    - run - : The primary MDS is running

    - wait - : The secondary MDS is waiting for failure recovery of the primary MDS

    - stop - : MDS stopped

    - none - : no relation with MDS

    - If S-STATE is other than "-", MDS is in state-transition.

- STATE/S-STATE (AC)

    - run - : file system is in mounted

    - stop - : file system is in unmounted

    - If S-STATE is other than "-", AC is in state-transition.

12.5 How to set GFS Shared File System applications as cluster applications

To set GFS Shared File System applications as cluster applications, switch only applications with the GFS Shared File System mounted on running nodes.

12.5.1 To set GFS Shared File System applications as cluster application

To set GFS Shared File System applications as cluster applications, it is necessary to set an RMS environment first.

- Registering an RMS host name in /etc/hosts on each node

  Define each node where RMS is configured in /etc/hosts so that RMS will be recognized the host name with RMS.

  ```
  192.168.0.1  sunny   sunnyRMS
  192.168.0.2  moony   moonyRMS
  ```

- Setting of cluster.config file or .rhosts file with each node

  In order to distribute RMS composition definition information to each node, an RMS wizard uses CF remote services or rcp (1). For this reason, a setup of the RMS host name to a setup, .rhost, or hosts.equiv of a cluster.config file is needed.

See
. . . .

For details, see "Appendix A Site preparation" of "PRIMECLUSTER Reliant Monitor Services (RMS) with Wizard Tools Configuration and Administration Guide."
. . . .

12.5.2 Notes on cluster application settings

The GFS Shared File System cannot be defined as **Fsystem** resource of the cluster application. Moreover, the disk class to which the GDS volume that is used as management partition or GFS Shared File System belongs cannot be defined as **Gds** resource of the cluster application. If you want to use the application that use the GFS Shared File System as cluster applications, perform the procedure described in "12.5.3 Procedure flow chart for setting GFS Shared File System applications as cluster applications."

Mount information of the GFS Shared File System must be registered in /etc/fstab and "noauto" must be specified in the "mount options field."

Moreover, "noatrc" must not be specified in "mount options" of the file system registered in /etc/fstab to do automatic mount after processing revitalization of the GFS Shared File System.

12.5.3 Procedure flow chart for setting GFS Shared File System applications as cluster applications

You can set GFS Shared File System applications as cluster applications using the following setup flow chart:

Table 12.3 The flow chart for setting GFS Shared File System applications as cluster applications

Shared operation (to be executed on node A)	Node A (active node)	Node B (standby node)
Execution of automatic resource registration (See "Procedure 1" in "12.5.4.")	-	-
-	GDS volume creation (See "Procedure 2" in "12.5.4.") Creating the GFS Shared File System management partition (See "Procedure 3" in "12.5.4.")	-
-	Node registration to the GFS Shared File System (See "Procedure 3" in "12.5.4.")	Node registration to the GFS Shared File System (See "Procedure 3" in "12.5.4.")
-	Creating the GFS Shared File System (See "Procedure 3" in "12.5.4.")	-

Shared operation (to be executed on node A)	Node A (active node)	Node B (standby node)
-	Adding the relevant GFS Shared File System to /etc/fstab (See "Procedure 4" in "12.5.4.")	Adding the relevant GFS Shared File System to /etc/fstab (See "Procedure 4" in "12.5.4.")
Setting cluster applications (See "Procedure 5" in "12.5.4."), (See "PRIMECLUSTER Installation and Administration Guide" and "PRIMECLUSTER Reliant Monitor Services (RMS) with Wizard Tools Configuration and Administration Guide")	-	-
-	Start up RMS and check settings (See "Procedure 6" in "12.5.4.")	Start up RMS and check settings (See "Procedure 6" in "12.5.4.")

GDS : Global Disk Services
GFS : Global File Services

12.5.4 Procedure for setting GFS Shared File System applications as cluster applications

This section describes how to set GFS Shared File System applications as cluster applications when CIP settings are completed.

See

For details on setting, see "PRIMECLUSTER Installation and Administration Guide," "PRIMECLUSTER Reliant Monitor Services (RMS) with Wizard Tools Configuration and Administration Guide," and "PRIMECLUSTER Global Disk Services Configuration and Administration Guide."

Procedure 1. Execution of automatic resource registration (common operation)

When automatic resource registration of disk equipment is not carried out yet, the following command is executed with a node, and a disk device is registered as a cluster resource.

```
# /etc/opt/FJSVcluster/bin/clautoconfig -r -n <Enter>
```

Procedure 2. GDS volume creation (node A (active node), node B (standby node))

Create the GDS volume using the following method.

GFS Shared File Systems require GDS volumes for management partitions. A disk class to which the GDS volume belongs as the management partition is different than that of the GDS volume that is used as a GDS resource of a cluster application.

The following explanation assumes the node A (active node) host name to be sunny and the node B (standby node) host name to be moony.

[When a mirror volume is to be created]

- Register a disk in the disk class on sunny by using the sdxdisk command.
  In the cluster system, specify a node name as the scope of the -a option and then specify shared as the class type.

```
sunny# sdxdisk -M -c gfs -d sda=DISK1,sdb=DISK2 -a scope=sunny:moony,type=shared
<Enter>
sunny# sdxdisk -M -c gfs01 -d sdc=DISK3,sdd=DISK4 -a scope=sunny:moony,type=shared
<Enter>
```

- On sunny, disk connection to the mirror disk group is performed with the sdxdisk command.

```
sunny# sdxdisk -C -c gfs -g grp0001 -d DISK1,DISK2 <Enter>
sunny# sdxdisk -C -c gfs01 -g grp0002 -d DISK3,DISK4 <Enter>
```

- On sunny, GDS volume creation is performed with the sdxvolume command.
  Note that a volume of at least 40 megabytes is necessary for the management partition of the GFS Shared File System.

```
sunny# sdxvolume -M -c gfs -g grp0001 -v control -s 81920 <Enter>
sunny# sdxvolume -M -c gfs01 -g grp0002 -v volume01 -s 17563648 <Enter>
```

- On moony, activation of the GDS volume is performed with the sdxvolume command.

```
moony# sdxvolume -N -c gfs -v control <Enter>
moony# sdxvolume -N -c gfs01 -v volume01 <Enter>
```

[When a single volume is to be created]

- On sunny, disk registration in the disk class is performed with the sdxdisk command.
  In the cluster system, specify a node name as the scope of the **-a** option and then specify shared as the cluster type.

```
sunny# sdxdisk -M -c gfs -d sda=DISK1:single -a scope=sunny:moony,type=shared <Enter>
sunny# sdxdisk -M -c gfs01 -d sdb=DISK2:single -a scope=sunny:moony,type=shared <Enter>
```

- On sunny, GDS volume creation is performed with the sdxvolume command.
  Note that a volume of at least 40 megabytes is necessary for the management partition of the GFS Shared File System.

```
sunny# sdxvolume -M -c gfs -d DISK1 -v control -s 81920 <Enter>
sunny# sdxvolume -M -c gfs01 -d DISK2 -v volume01 -s 17563648 <Enter>
```

- On moony, activation of the GDS volume is performed with the sdxvolume command.

```
moony# sdxvolume -N -c gfs -v control <Enter>
moony# sdxvolume -N -c gfs01 -v volume01 <Enter>
```

Procedure 3. GFS Shared File System creation (node A (active node), node B (standby node))

Create the GFS Shared File System on either one of the nodes.

Use sfcmkfs(8) to create the GFS Shared File System.

When the GFS Shared File System is created for the first time after installation, activation of the GFS Shared File System is necessary before file system creation.

- Create a management partition for the GFS Shared File System on either one of the nodes.

```
# sfcsetup -c /dev/sfdsk/gfs/dsk/control <Enter>
```

- Register node information in the management partition of the GFS Shared File System on each node.

```
sunny# sfcsetup -a /dev/sfdsk/gfs/dsk/control <Enter>
moony# sfcsetup -a /dev/sfdsk/gfs/dsk/control <Enter>
```

- Activate the GFS Shared File System on each node.

```
sunny# sfcfrmstart <Enter>
moony# sfcfrmstart <Enter>
```

- Create the GFS Shared File System on either one of the nodes.

```
# sfcmkfs -o node=sunny,moony /dev/sfdsk/gfs01/dsk/volume01 <Enter>
```

Procedure 4. Adding the relevant GFS Shared File System to /etc/fstab (node A (active node), node B (standby node))

Add the mount information for the relevant GFS Shared File System to /etc/fstab on each node. Be sure to specify "noauto" in the "mount options" field as a parameter of the mount information to be added.

```
/dev/sfdsk/gfs01/dsk/volume01 /sfcfs sfcfs rw,noauto 0 0
```

Procedure 5. Setting the definition of switchover of the RMS application as the RMS configuration (node A (active node), node B (standby node))

Set up the RMS configuration by using the RMS wizard, as follows:

1. Start the RMS wizard by executing the hvw command.

2. Create the RMS application.

   - Select "Application-Create" on the main RMS management menu.

   - Select "STANDBY" on the application type selection menu.

   - Select "ApplicationName" on the STANDBY wizard menu and change the RMS application name.

   - Select "Machines+Basics" on the STANDBY wizard menu, add another node to operate the RMS application by using "AdditionalMachine" on the Machines+Basics menu, and then set the switchover attribute of the RMS application to "HostFailure|ResourceFailure|ShutDown" by executing "AutoSwitchOver."

   - And then add StartCommands and CheckCommands by using "CommandLines" on the STANDBY wizard menu.

3. Create the RMS configuration definition of the RMS application.
   Select "Configuration-Generate" from the main RMS management menu.

4. Activate the RMS configuration definition of the RMS application.
   Select "Configuration-Activate" on the main RMS management menu.

5. End the RMS wizard.

📝 Note

In GFS Shared File System, the following measures are necessary in order to always use GDS volume by active state.

- For cluster application settings, a GDS resource "Gds:Global-Disk-Services" must not be registered in the GDS volume of the GFS Shared File System.

- For cluster application settings, an Fsystem resource "LocalFileSystems" must not be registered in the GFS Shared File System.

- For GDS volume, must not execute /usr/opt/reliant/bin/hvgdsetup.

Procedure 6. Confirming settings (node A (active node), node B (standby node))

After mounting the GFS Shared File System on each node, start RMS and then confirm whether the settings of the /etc/fstab file and the RMS application are proper, paying attention to the following points:

- Whether the GFS Shared File System can be mounted on both nodes according to the /etc/fstab file descriptions.

- Whether the RMS application switches over to a standby node at failover of the active node.

📝 Note

If mounting of the GFS Shared File System fails, the following may be the cause of the failure (apply the indicated solution):

- The type of the file system specified in the "/etc/fstab" file is incorrect

  Solution: Correct the /etc/fstab file.

- The GDS volume is not ACTIVE

  Solution: Activate the GDS volume.

12.5.5 Setup flow chart of adding file data partitions to GFS Shared File Systems of cluster applications

The following setup flow chart shows how to add file data partitions to GFS Shared File Systems of cluster applications.

Table 12.4 The flow chart of adding file data partitions to GFS Shared File Systems of cluster applications

Common operation (to be executed on node A)	Node A (active node)	Node B (standby node)
-	Stop RMS. (See "Procedure 1" in "12.5.6.")	Stop RMS. (See "Procedure 1" in "12.5.6.")
-	GDS volume creation (See "Procedure 2" in "12.5.6.") Add a file data partition to the GFS Shared File System. (See "Procedure 3" in "12.5.6.")	-
-	Start RMS. (See "Procedure 4" in "12.5.6.")	Start RMS. (See "Procedure 4" in "12.5.6.")

GDS : Global Disk Services
GFS : Global File Services

12.5.6 Setup procedure of adding file data partitions to GFS Shared File Systems of cluster applications

The following setup flow chart shows how to add file data partitions to GFS Shared File Systems of cluster applications.

See

For details on setting, see "PRIMECLUSTER Installation and Administration Guide," "PRIMECLUSTER Reliant Monitor Services (RMS) with Wizard Tools Configuration and Administration Guide," and "PRIMECLUSTER Global Disk Services Configuration and Administration Guide."

Procedure 1. Stopping RMS (node A (active node), node B (standby node))

If RMS is activated, stop RMS on each node.

Procedure 2. GDS volume creation (node A (active node), node B (standby node))

See "Procedure 2. GDS volume creation (node A (active node), node B (standby node))."

Procedure 3. Adding a file data partition to a GFS Shared File System (node A (active node))

Add a file data partition to the relevant GFS Shared File System from an active node with the sfcadd(8) command.

```
sunny# sfcadd -D /dev/sfdsk/gfs01/dsk/volume02 /dev/sfdsk/gfs01/dsk/volume01 <Enter>
```

Procedure 4. Starting up RMS (node A (active node), node B (standby node))

Activate RMS on each node.

12.6 How to start up CF by the GUI when a GFS Shared File System is used

This section describes how to start up CF by the GUI after stopping it on all the nodes where a GFS Shared File System is used.

The two start modes of the sfcfrmd daemon of the GFS Shared File System are **wait** (default) and **wait_bg**.

If **wait** is selected, starting the GFS Shared File System will fail because a quorum does not exist until CF is activated on all the nodes. If you need to start CF by the GUI when the CF is stopped on all the nodes, follow the steps below.

🗐 See

For information on how to start up CF, see "Starting CF" of "PRIMECLUSTER Cluster Foundation (CF) Configuration and Administration Guide."

1. Select the "Cluster Admin" menu on the Web-Based Admin View operation menu. Select the node where CF will be started, and click <OK>.
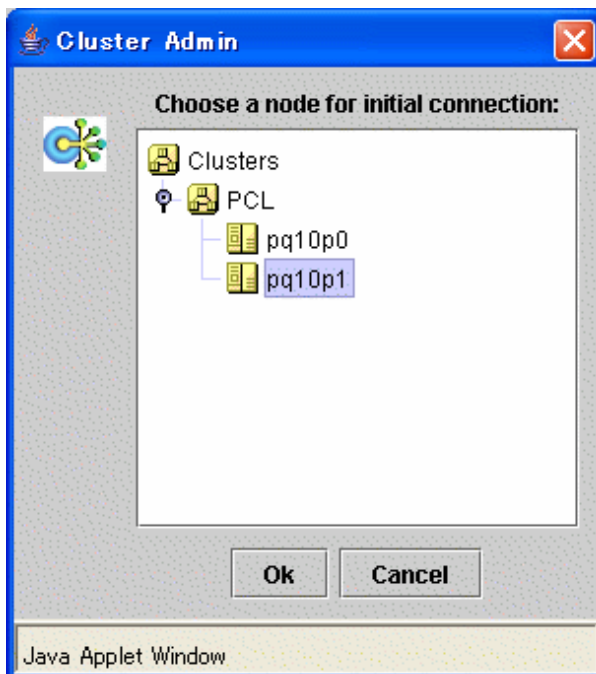
2. Click <Load driver> then activate the CF driver.
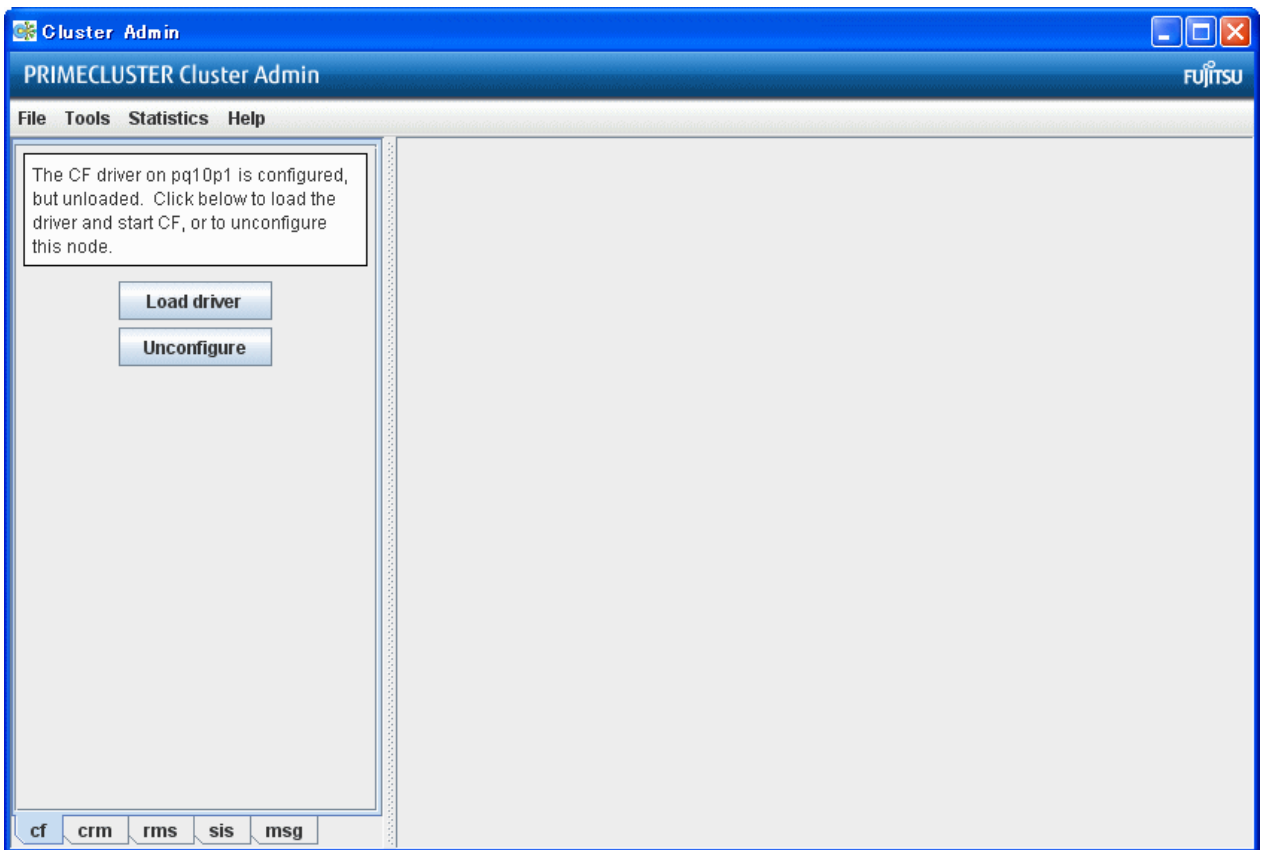
3. The "Start CF" window will pop up. Click <OK> to start up CF.
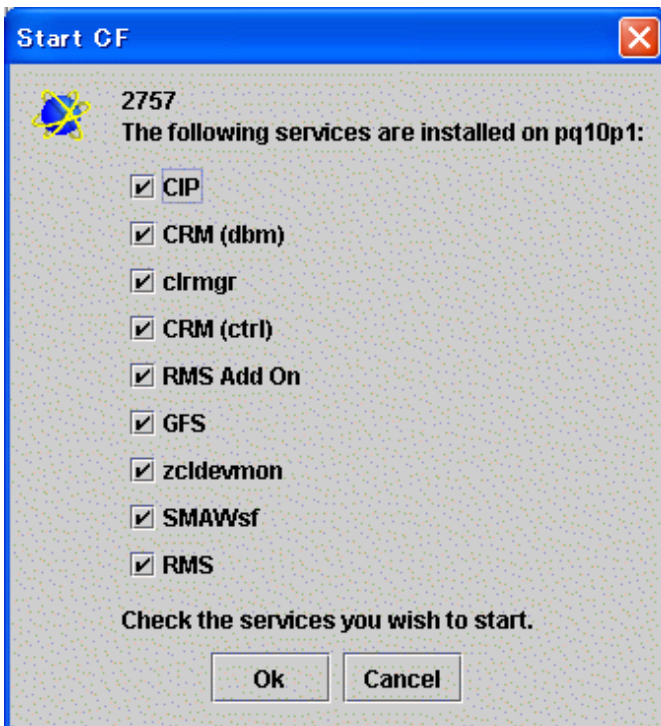
4. Since CF is not activated on all the nodes, GFS startup fails then the "Error" window will pop up. Click <No> then cancel the GFS startup processing.
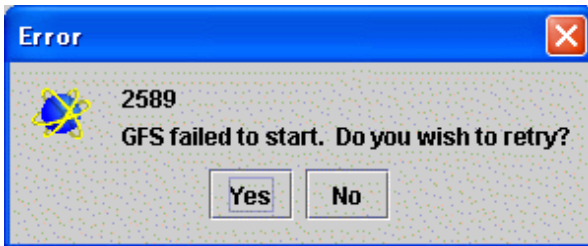
5. If the GFS startup is cancelled, the "Error" window will pop up again. Click <No> then cancel the script processing.
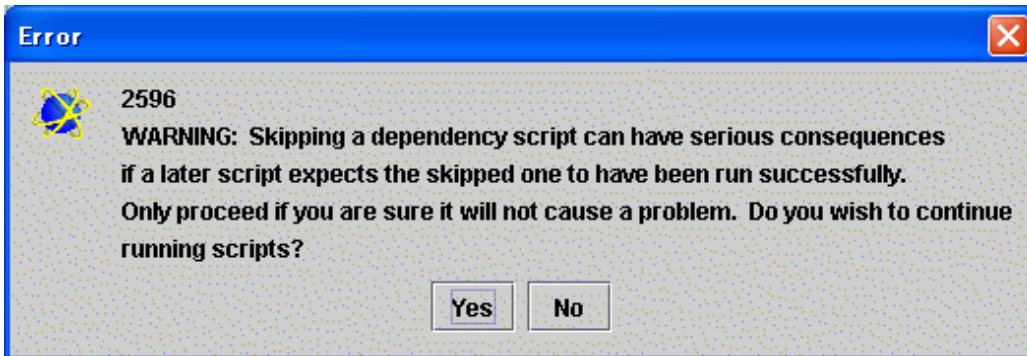
6. Check if all the services before GFS have been activated then click <OK>.
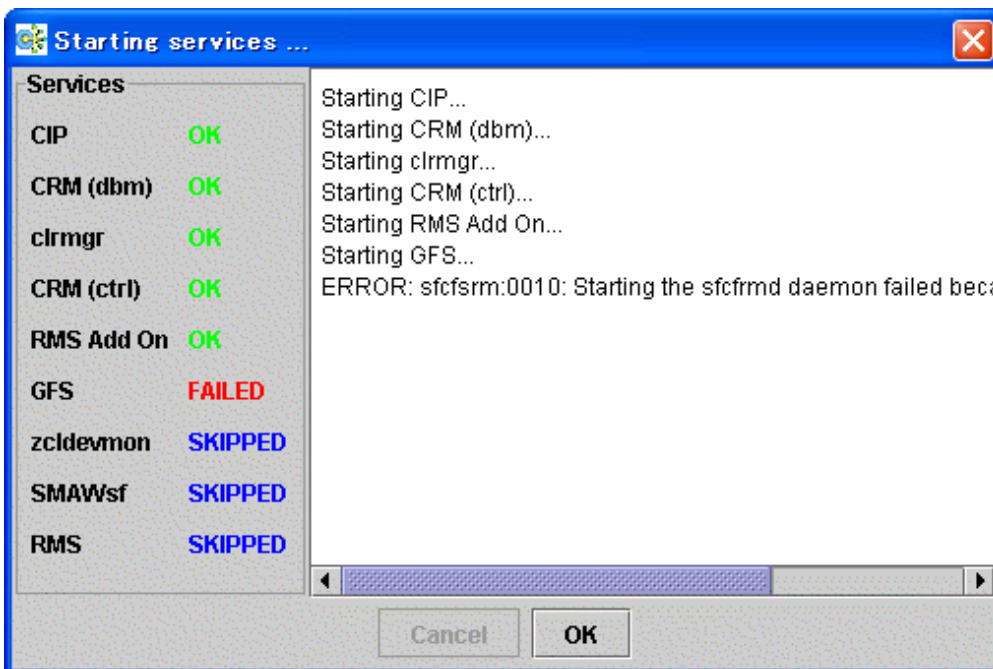
7. Repeat operations from step 1 to 6 on each cluster node then activate all the services before GFS.

📝 **Note**
. . . .
On the node where CF is activated lastly, GFS startup will succeed because CF has been activated on all the nodes.
. . . .

8. On any node that cancelled the GFS startup processing, stop CF and restart up CF.

9. Execute Step 8 then activate all the services after GFS on all the nodes where GFS startup is cancelled.

Chapter 13 File System Backing-up and Restoring

This chapter describes how to backing-up and restoring data in the GFS Shared File System.

13.1 Type of Backing-up and Restoring

This section describes the backing-up and restoring of the GFS Shared File System using.

- Backing-up and restoring file by file

    - General-purpose backing-up tools not dependent on the file system structure (ArcServe/OPEN, etc.)

    - Standard Linux commands cpio(1L) and tar(1)

- Backing-up and restoring entire file system

    - Standard Linux commands dd (1)

Note

dump(8) etc. of ext2 cannot be used to back up the GFS Shared File System since it depends on the file system configuration.

The following describes the backing-up and restoring data of the GFS Shared File System in this Chapter.

- Backing-up and restoring data using Standard Linux commands : cpio(1L), tar(1) and dd(1).

13.2 Backing-up by Standard Linux commands

This section describes the methods for backing-up in units of files and file systems.

13.2.1 Backing-up file by file

To back up a single file or all files under the directory while it is mounted, use cpio(1L) or tar(1).

See

For details about cpio (1L) and tar(1), see the online manual.

The following are examples of backing-up on tape using cpio(1L) and tar(1).

- Backing-up method with cpio(1L)

```
# cd /userdata <Enter>
# find . -depth -print | cpio -oc > /dev/st0 <Enter>
```

- Backing-up method with tar(1)

```
# cd /userdata <Enter>
# tar cvf /dev/st0 . <Enter>
```

Note

Tapes used for data backup should be labeled for identifying the backing-up method.
cron(8) can schedule backing-up startup.

13.2.2 Backing-up entire file system

To back up the GFS Shared File System in file system unit, use dd(1). The following describes the backing-up procedure using dd(1).

🔖 See

....

For details about dd(1), see the online manual.

....

Entire-file-system backing-up allows backing-up of the entire file system more quickly than file-by-file backing-up. However, a file system backed up with this method cannot be restored on a file-by-file basis during restoring.

Before executing backing-up, use the following procedure to check the **partition** information for the desired GFS Shared File System. Estimate the backing-up media capacity required for backing-up and the number of required media and then unmount of the GFS Shared File System.

The following describes examples of tape backup in case of single partition configuration and multi-partition configuration.

For a single partition configuration

Use **sfcinfo**(8) to display partition information and check that the file system has only a single partition (same special file names) and that the capacity of the device, backup destination, is sufficient. The size of the partition can be estimated by adding 1 megabyte to the sum of the values displayed in the size field (in kilobytes).

Example: Backup in case of single partition

```
# sfcinfo /dev/sfdsk/gfs01/dsk/volume01 <Enter>
FSID special                                    size  Type  mount
   1 /dev/sfdsk/gfs01/dsk/volume01(1e756)      14422  META  -----
   1 /dev/sfdsk/gfs01/dsk/volume01(1e756)       5116  LOG   -----
   1 /dev/sfdsk/gfs01/dsk/volume01(1e756)      95112  DATA  -----
# dd if=/dev/sfdsk/gfs01/dsk/volume01 of=/dev/st0 bs=1024k <Enter>
```

For a multi-partition configuration

dd(1) can handle one partition at a time. For a multi-partition configuration, therefore, the partitions must be backed up one at a time. Use sfcinfo(8) to check the partition configuration then back up each partition with the same procedure as used for a single partition configuration.

Example: Backup in case of two partitions

```
# sfcinfo /dev/sfdsk/gfs01/dsk/volume01 <Enter>
FSID special                                    size  Type  mount
   1 /dev/sfdsk/gfs01/dsk/volume01(1e756)      14422  META  -----
   1 /dev/sfdsk/gfs01/dsk/volume01(1e756)       5116  LOG   -----
   1 /dev/sfdsk/gfs01/dsk/volume01(1e756)      95112  DATA  -----
   1 /dev/sfdsk/gfs01/dsk/volume02(1e757)      95112  DATA  -----
# dd if=/dev/sfdsk/gfs01/dsk/volume01 of=/dev/st0 bs=1024k <Enter>
  (Change tape)
# dd if=/dev/sfdsk/gfs01/dsk/volume02 of=/dev/st0 bs=1024k <Enter>
```

📌 Note

....

- All the partitions including the relevant GFS Shared File System must be backed up.

- The tapes used for backups must be labeled so that the backed up partitions can be identified. They must then be stored as a set.

....

dd(1) does not support multi-tape volumes. If the file system is too large to fit on a single tape, data needs to be divided and backed up. In such cases, set bs (block length) and count (number of blocks) and increase skip (input-side offset (number of blocks)) by the value of count.

Example: Backup in case of separating partition as blocks in units of 1 gigabyte

```
# dd if=/dev/sfdsk/gfs01/dsk/volume01 of=/dev/st0 bs=1024k count=1024 <Enter>
  (Change tape)
# dd if=/dev/sfdsk/gfs01/dsk/volume01 of=/dev/st0 bs=1024k count=1024 skip=1024 <Enter>
  (Change tape)
# dd if=/dev/sfdsk/gfs01/dsk/volume01 of=/dev/st0 bs=1024k count=1024 skip=2048 <Enter>
  (Change tape)
# dd if=/dev/sfdsk/gfs01/dsk/volume01 of=/dev/st0 bs=1024k count=1024 skip=3072 <Enter>
  ...
```

13.3 Restoring by Standard Linux commands

This section describes the procedure for restoring a file system from the backup tapes prepared as described earlier.

Two restoring methods are available:

- Restoring file by file using the appropriate Standard Linux commands

- Restoring by file system using dd (1)

13.3.1 File-by-file restoring

Use cpio (1L) or tar (1) to restore files from the backup tape to the disk.

See
....
For details about cpio(1L) and tar(1), see the online manual.
....

The following describes an example of restoring from a tape.

- Method for restoring using cpio(1L)

```
# cd /userdata <Enter>
# cpio -icdm < /dev/st0 <Enter>
```

- Method for restoring using tar(1)

```
# cd /userdata <Enter>
# tar xvf /dev/st0 <Enter>
```

Restoring from a backup tape must be performed using the same command that was used for backing-up.

13.3.2 Entire-file-system restoring

Restore each partition saved on a backup tape onto the disk using dd(1).

Before executing restoring, unmount the desired GFS Shared File System from all the **node**s using the file system, and unmounted of the file system.

Note
....
The size of the partition for the restoring must be the same as the backup partition. Otherwise the partition for the restoring is not available as the GFS Shared File System even if execution of dd(1) has succeeded.
....

The following describes examples of restoring from backup tapes in case of single partition configuration and multi-partition configuration.

- For a single partition configuration

    Example: Restoring in case of single partition

```
# dd if=/dev/st0 of=/dev/sfdsk/gfs01/dsk/volume01 bs=1024k <Enter>
```

- For a multi-partition configuration

Example: Restoring in case of two partitions

```
# dd if=/dev/st0 of=/dev/sfdsk/gfs01/dsk/volume01 bs=1024k <Enter>
  (Change tape)
# dd if=/dev/st0 of=/dev/sfdsk/gfs01/dsk/volume02 bs=1024k <Enter>
```

dd(1) does not support multi-tape volumes. To restore a file system backed up as separate blocks, specify the same values for the bs and count parameters as for where specified for the backup. For seek, specify the same value that was specified for skip.

Example: Restoring the tape backed up as separate blocks in units of 1 gigabyte

```
# dd if=/dev/st0 of=/dev/sfdsk/gfs01/dsk/volume01 bs=1024k count=1024 <Enter>
  (Change tape)
# dd if=/dev/st0 of=/dev/sfdsk/gfs01/dsk/volume01 bs=1024k count=1024 seek=1024 <Enter>
  (Change tape)
# dd if=/dev/st0 of=/dev/sfdsk/gfs01/dsk/volume01 bs=1024k count=1024 seek=2048 <Enter>
  (Change tape)
# dd if=/dev/st0 of=/dev/sfdsk/gfs01/dsk/volume01 bs=1024k count=1024 seek=3072 <Enter>
   ...
```

13.3.3 Set up after Restoration

13.3.3.1 Resetting the partition information

When you restore the GFS Shared File System to a partition different from the backup origin by dd(1) command use, partition information in the inside of the medium does not accord with real partition composition.

Therefore, it is necessary to set partition information again.

Information

When you restore the GFS Shared File System to totally same partition as the backup origin, operations in this section is not necessary because there is no change in partition information in the medium.

Use the sfcadm(8) command to reset the partition information.
If the GFS Shared File System is restored to a partition that is different from a backup, it differs according to the procedure of restoration is whether the partition at the restoration destination is GFS Shared File System. Each procedure is shown as follows.

- When restoring the GFS Shared File System to an unused partition

The following shows how to reset the partition information with sfcadm(8) after you have restored the file system that consists of the three partitions to another partition (/dev/sfdsk/gfs99/dsk/volume01, /dev/sfdsk/gfs99/dsk/volume02, /dev/sfdsk/gfs99/dsk/volume03) that is different from the creation source. Note that the representative partition is /dev/sfdsk/gfs99/dsk/volume01, the port name is sfcfs-1 and that the shared nodes are host01, host02.

```
# sfcadm -m host01,host02 -g host01,host02 -p sfcfs-1,sfcfs-1 /dev/sfdsk/gfs99/dsk/volume01,
/dev/sfdsk/gfs99/dsk/volume02,/dev/sfdsk/gfs99/dsk/volume03 <Enter>
```

- When restoring the GFS Shared File System to a partition in use

You can set the partition information by executing sfcadm(8) after restoring the file system that consists of three partitions to the partitions; /dev/sfdsk/gfs99/dsk/volume01, /dev/sfdsk/gfs99/dsk/volume02, and /dev/sfdsk/gfs99/dsk/volume03. In these partitions, the file system in the same configuration exists. The representative partition is /dev/sfdsk/gfs99/dsk/volume01, and the port name and shared node are the same.

```
# sfcadm /dev/sfdsk/gfs99/dsk/volume01,/dev/sfdsk/gfs99/dsk/volume02,/dev/sfdsk/gfs99/dsk/
volume03 <Enter>
```

Note that the state of the file system creation for the restoration destination can be confirmed with **sfcinfo(8)**.

 See

For details on sfcinfo(8), see sfcinfo(8).
For details on sfcadm(8), see sfcadm(8).

 Note

If nodes where the file system is shared are different before and after backup or restoration, change the sharing node settings.

13.3.3.2 Repairing the file system

If you attempt to back up file systems that are mounted in units of file systems, a file access may occur while backup is being performed, such that an inconsistency occurs in the backup content.

To repair this state, the administrator must execute a file system consistency check and perform a repair with sfcfsck(8). If no inconsistency occurs, sfcfsck(8) will be completed in several seconds to one minute.

An example of file system repair is shown below. The representative partition is /dev/sfdsk/gfs99/dsk/volume01. Before executing the command, be sure to check that the file system is not mounted on all the nodes.

```
# sfcfsck -y /dev/sfdsk/gfs99/dsk/volume01 <Enter>
```

 See

For details on sfcfsck(8), see sfcfsck(8).

Chapter 14 Tuning

This chapter describes how to use a variety of utilities to optimize and make effective use of the GFS Shared File System.

14.1 Tuning Parameters

This section describes the tuning parameter of the **MDS**(Meta Data Server) of the GFS Shared File System.

![Note icon] **Note**
. . . .
The setting value of each tuning parameter is saved in the management partition. If the value is newly added or changed, back up again the management partition by using the sfcgetconf(8) command.
. . . .

![See icon] **See**
. . . .
For how to back up the management partition, refer to "9.6 Backup of the management partition information" in this manual.
. . . .

14.1.1 Amount of Cache

It is possible to tune the following items:

- Amount of free-**extent base**d management information of a file system to be cached in memory
  (**SFCFS_NCACHE_EXTENT**)

- Amount of directory information to be cached in memory
  (**SFCFS_NCACHE_DIRECTORY**)

- Amount of **i-node**s on the disk to be cached in memory
  (**SFCFS_NCACHE_INODE**)

The areas of the following sizes are reserved in the process space of **MDS** for each file system:

- Cache of free-extent based management information of a file system

  Specified value of SFCFS_NCACHE_EXTENT x 1.4 kilobytes

- Cache of directory information

  Specified value of SFCFS_NCACHE_DIRECTORY x 1.4 kilobytes

- Cache of i-node

  Specified value of SFCFS_NCACHE_INODE x 4.4 kilobytes

The following shows the default value and amount of reserved memory for each tuning parameter.

Table 14.1 Default value and amount of reserved memory for tuning parameters

Tuning parameter name	Default value	Amount of reserved memory (megabyte)
SFCFS_NCACHE_EXTENT	2304	3.2
SFCFS_NCACHE_DIRECTORY	10240	14.0
SFCFS_NCACHE_INODE	2048	8.8

Specify the tuning parameters with the -o option of **sfcadm**(8).

The following shows the setting example.

(Example)

```
# sfcadm -o SFCFS_NCACHE_EXTENT=4096 /dev/sfdsk/gfs01/dsk/volume01 <Enter>
# sfcadm -o SFCFS_NCACHE_DIRECTORY=20480 /dev/sfdsk/gfs01/dsk/volume01 <Enter>
# sfcadm -o SFCFS_NCACHE_INODE=5120 /dev/sfdsk/gfs01/dsk/volume01 <Enter>
```

Check the specified values with the -e option of **sfcinfo**(8).

The following shows the example to check the values.

(Example)

```
# sfcinfo -e /dev/sfdsk/gfs01/dsk/volume01 | grep SFCFS_NCACHE <Enter>
SFCFS_NCACHE_EXTENT=4096
SFCFS_NCACHE_DIRECTORY=20480
SFCFS_NCACHE_INODE=5120
```

If the size of a directory exceeds the directory cache's size, performance of file creation in the directory or reference will fall. When a large directory is created, we recommend strongly that users specify a larger value for SFCFS_NCACHE_DIRECTORY.

Also, in the case of tuning, execute **sfcstat**(8) that reports statistics on GFS Shared File System by specifying the -m option, and check the cache-hit ratio of the meta-cache which corresponds to the tuning parameter. The cache-hit ratio is calculated from the number of accesses and the number of cache hits. If the cache-hit ratio is low, consider to specify a larger value of the tuning parameter.

The following shows the correspondence of each tuning parameter and meta-cache type displayed by the -m option of **sfcstat**(8).

Table 14.2 Correspondence of tuning parameter and meta-cache type displayed by -m option of sfcstat(8)

Tuning parameter name	Meta-cache type
SFCFS_NCACHE_EXTENT	NODE, LEAF
SFCFS_NCACHE_DIRECTORY	DIR
SFCFS_NCACHE_INODE	DINODE

14.1.2 Communication timeout value

It is possible to tune the following items:

- Communication timeout value of the GFS Shared File System(**CON_TIMEOUT**)

CON_TIMEOUT set up time in seconds after a communicative response is lost until it estimates a partner node to be an abnormal state in the monitoring of the node by communication. This value can be set for every file system.

The default of a communication timeout value is 180 seconds.

It is not usually necessary to change communication timeout value.

Specify the tuning parameters with the -o option of **sfcadm**(8).

The following shows the setting example.

(Example)

```
# sfcadm -o CON_TIMEOUT=240 /dev/sfdsk/gfs01/dsk/volume01 <Enter>
```

Check the specified values with the -e option of **sfcinfo**(8).

The following shows the example to check the values.

(Example)

```
# sfcinfo -e /dev/sfdsk/gfs01/dsk/volume01 | grep CON_TIMEOUT <Enter>
CON_TIMEOUT=240
```

When it is expected that the loading of system becomes high, please set the value of **CON_TIMEOUT** greatly. Please set up the value of **CON_TIMEOUT** small, when there are few file systems to be used and improve the response of the error return from a system call to the

time of the blockade accompanying the abnormalities in communication etc. When the number of file system use are 10 or less and set a value of **CON_TIMEOUT** small, its value should be 60 seconds or more.

14.1.3 Communication protocol

In the GFS Shared File System, the following item can be tuned for communication between MDS and AC:

- Communication protocol (**SFCFS_IPV6**)

Determine either IPv6 or IPv4 to be used as a communication protocol between MDS and AC. Then, set it for **SFCFS_IPV6**. Set 1 for IPv6 and set 0 for IPv4. This value can be set for each file system.

The default is IPv4.

🛈 **Note**

. . . .

- A communication protocol can be set for each file system; however, we recommend that users set the same communication protocol for all the file systems so that it can be easily managed.

- When changing a communication protocol, change the IP address of the shared node that is registered in host database (for example: /etc/hosts) to the one that matches the changed communication protocol. If the IP address does not match the communication protocol, the GFS Shared File System cannot be mounted.

. . . .

Specify the tuning parameters with the -o option of **sfcadm(8)**.

The following shows the example when using IPv6 as a communication protocol.

(Example)

```
# sfcadm -o SFCFS_IPV6=1 /dev/sfdsk/gfs01/dsk/volume01 <Enter>
```

Check the specified values with the -e option of **sfcinfo(8)**.

The following shows the example to check the values.

(Example)

```
# sfcinfo -e /dev/sfdsk/gfs01/dsk/volume01 | grep SFCFS_IPV6 <Enter>
SFCFS_IPV6=1
```

To restore the communication protocol to IPv4 (default), delete the setting of **SFCFS_IPV6**.

(Example)

```
# sfcadm -o SFCFS_IPV6= /dev/sfdsk/gfs01/dsk/volume01 <Enter>
```

(Example)

```
# sfcinfo -e /dev/sfdsk/gfs01/dsk/volume01 | grep SFCFS_IPV6 <Enter>
```

If no value is displayed here, the setting has been deleted.

Chapter 15 Migration to the GFS Shared File System

This chapter describes how to move the existing files such as ext2 and ext3 that are created on the file system to the GFS Shared File System.

15.1 Moving the existing files

When moving existing files that were created on a file system such as ext2/ext3 to the GFS Shared File System, be aware of the following important information.

📓 Note

. . . .

- The limit value of the quota function in the file system cannot be moved to the GFS Shared File System.

- The ACL settings of the file system cannot be moved to the GFS Shared File System. For archive created with **tar**(1) or **cpio**(1L) and including the ACL settings, ACL related errors will be displayed, but file data can be restored.

- When files with holes are transferred to the GFS Shared File System, some area is reserved on the storage device and such files no longer have holes.

. . . .

File system migration first requires that files and directories be backed up on another location (for example, on a storage device or another file system). Then, a new GFS Shared File System should be configured, and the files and directories can be expanded. Take the following procedures:

1. Check that the file system is not being used.

2. Save all of the files on the file system on a backup device or another file system using a command such as **tar**(1) and **cpio**(1L).

3. Create the GFS Shared File System, as described in "10.2 Creation", "11.2 Creation."

4. Expand the saved files on the created GFS Shared File System.

The following example shows how to move the file of the ext3 file system that is mounted on **/mnt** to the GFS Shared File System on the same device after saving it in **/data** temporarily.

1. Back up files

```
# cd /mnt <Enter>
# tar cvpf - . > /data/backup.tar <Enter>
# cd / <Enter>
# umount /mnt <Enter>
```

2. Configure GFS Shared File System

   Create a GFS Shared File System and mount it according to "10.2 Creation" or "11.2 Creation."

   In this example, the mount point is **/sfcfs**.

3. Restore the file.

```
# cd /sfcfs <Enter>
# tar xvf /data/backup.tar <Enter>
```

Figure 15.1 Migration from ext2, ext3, etc.

GFS : Global File Services
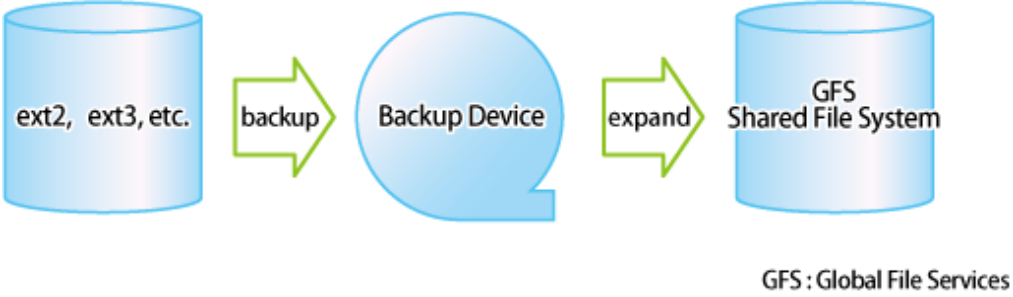
Appendix A List of Messages

This appendix explains GFS Shared File System messages.

If a message has an explanation and Response identical to another message, the description is omitted. In this case, see the explanation and Response for the last message.

A.1 AC Messages of the GFS Shared File System

This section explains AC messages of the GFS Shared File System that appears during operation.

Types and priorities of messages are listed below.

Type	Priority (*facility.level*)
Panic message	kern.emerg
Warning message	kern.warning
Information message	kern.notice

- The priorities of the messages are passed to rsyslogd(8).

- The output destination of the messages is defined in the /etc/rsyslog.conf file. The output destination of the messages is /var/log/ messages by default of the operating system.

See
....
For details about the /etc/rsyslog.conf file, see rsyslog.conf(5) in the online manual.
....

A.1.1 Panic messages

This sub-section describes the AC panic messages of the GFS Shared File System.

PANIC: sfcfs: 0012: Assertion failed: *detail*

Explanation

An internal error is detected. The details of an error are displayed at *detail*.

Response

Collect the following for troubleshooting information and then contact field engineers:

- A crash dump

- A file output with FJQSS, **fjsnap**, or **pclsnap** on all the nodes at the time of rebooting them

PANIC: sfcfs: 0013: *mount_point*: Incomplete blockade

Explanation

As taking time to blockade the file system (*mount_point*), a node panic is carried out.

Response

Collect the following for troubleshooting information and then contact field engineers:

- A crash dump

- A file output with FJQSS, **fjsnap**, or **pclsnap** on all the nodes at the time of rebooting them

A.1.2 Warning messages

This sub-section describes the AC Warning messages of the GFS Shared File System.

WARNING: sfcfs: 1033: Not enough memory: *detail*

Explanation

Memory allocation failed. The details of the failure are displayed at *detail.*

Response

Loads on the memory can be the cause of the failure. Consider to expand the actual memory.

WARNING: sfcfs: 1034: Failed to add module: *detail*

Explanation

Loading of the sfcfs module failed. The details of the failure are displayed at *detail.*

Response

Collect a file output with FJQSS, **fjsnap**, or **pclsnap** for troubleshooting information then contact field engineers.

WARNING: sfcfs: 1035: Invalid mount argument: *detail*

Explanation

Mounting failed, as an error was detected by specifying **mount(2)** in the file system. The details of the failure are displayed at *detail.*

Response

Collect a file output with FJQSS, **fjsnap**, or **pclsnap** for troubleshooting information then contact field engineers.

WARNING: sfcfs: 1036: Mount failed: errcode=*errcode*

Explanation

Mounting failed.

Response

Check following points:

- If any warning message was output immediately before this message, take an action to deal with the warning message.

- Check the network connection state with the node where MDS runs.

- Loads on the memory can be the cause of the failure. Consider to expand the actual memory.

If the problem cannot be solved by the above-mentioned procedures, collect a file output with FJQSS, **fjsnap**, or **pclsnap** for troubleshooting information then contact field engineers.

WARNING: sfcfs: 1037: *mount_point*: Failed to open device=*device* errcode=*errcode*

Explanation

Mounting on the file system (*mount_point*) failed due to the failure of opening a device (*device*) which configures the file system.

Response

Check that devices which configure the file system are connected properly.

WARNING: sfcfs: 1038: *mount_point*: Cannot connect MDS: *detail*

Explanation

Connection with MDS failed.

Mounting failed when the connection with the primary MDS and secondary MDS on a definition cannot be made.

The details of the failure are displayed at *detail.*

Response

Check following points:

- Check the network connection state with the node where MDS runs.

- Check whether the IP address of the node where MDS runs is correctly registered in host database (such as /etc/hosts). For details, see "11.2.2 Registering information in host database."

If the problem cannot be solved after taking the above actions, collect a file output with FJQSS, **fjsnap**, or **pclsnap** for troubleshooting information then contact field engineers.

WARNING: sfcfs: 1039: *mount_point*: MDS reject mount: *detail*

Explanation

Mounting is rejected by MDS. The details of the rejection are displayed at *detail.*

Response

Check following points:

- Check /var/log/messages of the node where MDS runs. If there is any warning message output by MDS immediately before this message has output, take an action to deal with the warning message.

- Check whether the IP address of this node where MDS runs is correctly registered in host database (such as /etc/hosts). For details, see "11.2.2 Registering information in host database."

- When unmounting and mounting are performed in a short time from the blockade of the file system or the node panic, mounting may fail. Spare some time before mounting again.

WARNING: sfcfs: 1040: *mount_point*: Protocol error: *detail*

Explanation

The file system (*mount_point*) is blockaded, as an illegal message was received from MDS. The details of the blockade are displayed at *detail.*

Response

Collect a file output with FJQSS, **fjsnap**, or **pclsnap** for troubleshooting information then contact field engineers.

WARNING: sfcfs: 1041: *mount_point*: Disconnect MDS

Explanation

Disconnect the connection between the file system (*mount_point*) and MDS.

Response

None.

WARNING: sfcfs: 1042: *mount_point*: Blockade: *detail*

Explanation

The file system (*mount_point*) is blockaded. The details of the blockade are displayed at *detail.*

Response

Collect a file output with FJQSS, **fjsnap**, or **pclsnap** for troubleshooting information then contact field engineers.

WARNING: sfcfs: 1043: *mount_point*: File system full

Explanation

There is no more space in the file system (*mount_point*).

As an ordinary user cannot use the space reserved for a superuser, this message may be output even if there is free space.

Response

Perform any of the following procedures to expand the space of the file system (*mount_point*):

- Expand the file system.

- Delete unnecessary files.

- Move files to the file system that has enough space.

Use **sfcadd(8)** to expand the file system.

A.1.3 Information messages

This sub-section describes the AC information messages of the GFS Shared File System.

NOTICE: sfcfs: 2008: *mount_point*: Mount complete

Explanation

Mounting succeeded.

Response

None.

NOTICE: sfcfs: 2009: *mount_point*: Umount complete

Explanation

Unmounting succeeded.

Response

None.

NOTICE: sfcfs: 2010: *mount_point*: Recovery complete

Explanation

Recovery succeeded.

Response

None.

NOTICE: sfcfs: 2011: *mount_point*: Connected with MDS: *detail*

Explanation

Connection was made with MDS. The details of the connection are displayed at *detail.*

Response

None.

NOTICE: sfcfs: 2012: *mount_point*: Open conflict with mmap: ino=*inode_number*

Explanation

The file (*inode_number*) which is to be opened by **open(2)** on the file system (*mount_point*) is being mapped by **mmap(2)** on the other node. To assure data consistency, this file cannot be opened.

Response

Perform the procedure again after completing an application on the other node which is mapping the file with **mmap(2)**.

NOTICE: sfcfs: 2013: *mount_point*: Mmap conflict with open: ino=*inode_number*

Explanation

The file (*inode_number*) which is to be mapped by **mmap(2)** on the file system (*mount_point*) is being opened by **open(2)** on the other node. To assure data consistency, this file cannot be mapped.

Response

Perform the procedure again after completing an application on the other node which is opening the file with **open(2)**.

A.2 GFS Shared File System Daemon messages

This section explains the GFS Shared File System daemon messages.

Types, daemons or commands to be output, and priorities are listed below.

Type	Daemons or commands to be output	Priority (*facility.level*)
Panic message	MDS (sfcfsmg daemon)	local0.emerg
	sfcprmd daemon	daemon.emerg
Error message	sfcfrmd daemon	daemon.err
	MDS (sfcfsmg daemon)	local0.err
	sfcfsd daemon	daemon.err
	sfcfs_mount command	user.err
	sfcpncd daemon	daemon.err
	sfcprmd daemon	daemon.err
Warning message	MDS (sfcfsmg daemon)	local0.warning
	sfcfsd daemon	daemon.warning
	sfchnsd daemon	daemon.warning
Information message	sfcfrmd daemon	daemon.notice
	MDS (sfcfsmg daemon)	local0.notice
	sfcprmd daemon	daemon.notice
	sfchnsd daemon	daemon.notice

- The priorities of the messages are passed to rsyslogd(8).

- The output destination of the messages is defined in the /etc/rsyslog.conf file. The output destination of the messages is /var/log/ messages by default of the operating system.

See
. . . .
For details about the /etc/rsyslog.conf file, see rsyslog.conf(5) in the online manual.
. . . .

A.2.1 Panic messages (MDS (sfcfsmg daemon))

This section explains the MDS (sfcfsmg daemon) panic messages of the GFS Shared File System.

The messages do not cause a node panic.

PANIC: sfcfs_mds:0002: *fsid*: hold count is invalid. type=*number* ino=*ino_num*

Explanation

An error was detected in the reference count of file system data.

Response

Collect a file output with FJQSS, **fjsnap**, or **pclsnap** for troubleshooting information then contact field engineers.

PANIC: sfcfs_mds:0003: *fsid*: Internal Error [*details*]

Explanation

An internal error occurred in the file system (*fsid*). *Details* provide the detailed information.

Response

Collect a file output with FJQSS, **fjsnap**, or **pclsnap** for troubleshooting information then contact field engineers.

A.2.2   Panic messages (sfcprmd daemon)

This section explains the sfcprmd daemon panic messages of the GFS Shared File System.

---

sfcoptdrv:PANIC:Semaphore(semid=*id*) access error, panic for keep data integrity of GFS and continue cluster operation.

Explanation

Operation of a semaphore (*id*) went wrong and it changed into the state where the management continuation of the GFS Shared File System on this node cannot be carried out. Although the panic of this node is carried out for data protection, GFS Shared File System management is continued by other nodes.

Response

Collect a crash dump. Also, collect a file output with FJQSS, **fjsnap**, or **pclsnap** for troubleshooting information when rebooting the system, and then contact field engineers.

---

sfcoptdrv:PANIC:sfcprmd was down, panic for keep data integrity of GFS and continue cluster operation.

Explanation

sfcprmd was downed and it changed into the state where the management continuation of the GFS Shared File System on this node cannot be carried out. Although the panic of this node is carried out for data protection, GFS Shared File System management is continued by other nodes.

Response

Collect a crash dump. Also, collect a file output with FJQSS, **fjsnap**, or **pclsnap** for troubleshooting information when rebooting the system, and then contact field engineers.

---

sfcoptdrv:PANIC:*daemon*(pid=*num*) was down, panic for keep data integrity of GFS and continue cluster operation.

Explanation

The daemon shown by *daemon* was downed and it changed into the state where the management continuation of the GFS Shared File System on this node cannot be carried out. Although the panic of this node is carried out for data protection, GFS Shared File System management is continued by other nodes.

Response

Collect a crash dump. Also, collect a file output with FJQSS, **fjsnap**, or **pclsnap** for troubleshooting information when rebooting the system, and then contact field engineers.

---

sfcoptdrv:PANIC:Panic for keep data integrity of GFS and continue cluster operation.

Explanation

Abnormalities occurred and it changed into the state where the management continuation of the GFS Shared File System on this node cannot be carried out. Although the panic of this node is carried out for data protection, GFS Shared File System management is continued by other nodes.

Response

Collect a crash dump. Also, collect a file output with FJQSS, **fjsnap**, or **pclsnap** for troubleshooting information when rebooting the system, and then contact field engineers.

A.2.3 Error messages (sfcfrmd daemon)

This section explains the sfcfrmd daemon error messages of the GFS Shared File System.

ERROR: sfcfrmd:0001: not superuser

Explanation

sfcfrmd is not performed by superuser.

Response

Please start sfcfrmd by **sfcfrmstart(8)** command as a superuser.

ERROR: sfcfrmd:0002: associated with a terminal device

Explanation

The direct module is performed.

Response

Start the sfcfrmd daemon by executing **sfcfrmstart(8)** with super user access privileges.

ERROR: sfcfrmd:0003: cannot read *pathname* or something wrong

Explanation

Since there was injustice or file reading of *pathname* went wrong, sfcfrmd starting went wrong.

Response

Please check whether the file of *pathname* is set up correctly.

ERROR: sfcfrmd:0004: cannot get node version *sfile*

Explanation

The sfcfrmd daemon failed to start, because a working file *sfile* in **/var/opt/FJSVsfcfs** was broken.

Response

Make sure that /var/opt directory has a space by df(1) command. If there is not enough space, delete unnecessary files or move them to other file systems to increase space. Then re-start the node.

ERROR: sfcfrmd:0005: something wrong in *pathname*

Explanation

The sfcfrmd daemon failed to start, because **/etc/services** has no entry of sfcfsrm or *pathname* is broken.

Response

Check if "**sfcfsrm** *<Port number>***/tcp**" is added to the **/etc/services** file on all nodes in the cluster. If not, add it to the file. For details about the file definition, see **services(5)** of the online manual page.

If the setting above is right, *pathname* seems broken. Collect the diagnostic data with FJQSS, **fjsnap**, or **pclsnap** and contact field engineers.

ERROR: sfcfrmd:0006: cannot allocate memory space

Explanation

Since acquisition of a memory went wrong, sfcfrmd starting went wrong.

Response

Please check the state of a system, and wait for the end of other processes, or increase and reboot a swap domain.

ERROR: sfcfrmd:0007: cannot start service

Explanation

sfcfrmd starting went wrong.

Response

Please confirm whether the port number of sfcfsrm in the /etc/services is the same as the node starting the sfcfrmd daemon.

- When the port number is different

  After changes to the same port number as the node that has started the sfcfrmd daemon, please use sfcfrmstart(8) by the super user authority and start the sfcfrmd daemon.

- When the port number is the same

  Collect the diagnostic data with FJQSS, **fjsnap**, or **pclsnap** in all nodes, and contact field engineers.

ERROR: sfcfrmd:0008: I/O error or data corruption of management partition: *pathname*

Explanation

sfcfrmd failed in reading and writing of *pathname*.

Response

Please check whether a setup as a management partition domain of the GFS Shared File System is right, or a disk operates normally.

ERROR: sfcfrmd:0009: something wrong in cluster settings

Explanation

Since a cluster was not started, sfcfrmd starting went wrong.

Response

Please start a cluster.

ERROR: sfcfrmd:0010: cannot get host name

Explanation

The CIP address registered in the management partition changed.

Response

By remaking and restoring the management partition, back the configuration of the partition to a previous one before a value of the CIP address changed. For details to remake and restore, see "9.6 Backup of the management partition information" and "9.7 Restoring of the management partition information."

ERROR: sfcfrmd:0011: cannot connect to sfcprmd

Explanation

Connection with sfcprmd went wrong.

Response

Make sure that the sfcprmd daemon is running by **ps(1)** command (for example, "**/bin/ps -e | grep sfcprmd**".) If the sfcprmd daemon was stopped by the unload dependency script, restart sfcprmd by the load script ("**/opt/SMAW/SMAWcf/dep/start.d/S81sfcfs load**".) Otherwise, re-start the node.

If this message is output while the sfcprmd daemon is running, collect a core image of the daemon, and a file output with FJQSS, **fjsnap**, or **pclsnap** for troubleshooting information then contact field engineers. Restart the node where this problem occurred.

ERROR: sfcfrmd:0012: cannot connect to sfchnsd

Explanation

Connection with sfchnsd went wrong.

Response

Make sure that the sfchnsd daemon is running by **ps(1)** command (for example, "**/bin/ps -e | grep sfchnsd**".) If the sfchnsd daemon was stopped by the unload dependency script, restart sfchnsd by the load script ("**/opt/SMAW/SMAWcf/dep/start.d/S81sfcfs load**".) Otherwise, re-start the node.

If this message is output while the sfchnsd daemon is running, collect a core image of the daemon, and a file output with FJQSS, **fjsnap**, or **pclsnap** for troubleshooting information then contact field engineers. Restart the node where this problem occurred.

ERROR: sfcfrmd:0013: cannot activate sfchnsd

Explanation

sfchnsd is not activated.

Response

Collect the diagnostic data with FJQSS, **fjsnap**, or **pclsnap** in the node, and contact field engineers.

ERROR: sfcfrmd:0014: cannot get service port

Explanation

A service port is not able to be acquired.

Response

Check if "**sfcfsrm** *<Port number>*/**tcp**" is added to the **/etc/services** file on all nodes in the cluster. If not, add it to the file. For details about the file definition, see **services(5)** of the online manual page.

ERROR: sfcfrmd:0015: cannot get limit of newly created descriptor

Explanation

The maximum value of file descriptor generation was not able to be acquired.

Response

Please check the maximum value of a file descriptor.

ERROR: sfcfrmd:0019: no node information on *special*

Explanation

The node information for this node is not registered into the management partition *special*.

Response

Register the node information using **sfcsetup(8)** command.

A.2.4 Error messages (MDS (sfcfsmg daemon))

This section explains the MDS (sfcfsmg daemon) error messages of the GFS Shared File System.

ERROR: sfcfs_mds:0001: v_count != 0 inode=*ino_pointer* ino=*ino_num*

Explanation

An error was detected in releasing the i-node (*ino_num*) in the file system.

Response

Collect a file output with FJQSS, **fjsnap**, or **pclsnap** for troubleshooting information then contact field engineers.

ERROR: sfcfs_mds:1001: MDS is failed

Explanation

MDS failed in recovery mounting.

Response

Collect a file output with FJQSS, **fjsnap**, or **pclsnap** for troubleshooting information then contact field engineers.

ERROR: sfcfs_mds:1002: primary MDS is terminated

Explanation

Primary MDS was terminated abnormally.

Response

Collect a file output with FJQSS, **fjsnap**, or **pclsnap** for troubleshooting information then contact field engineers.

ERROR: sfcfs_mds:1004: can't execute primary MDS
ERROR: sfcfs_mds:1005: can't execute secondary MDS

Explanation

Startup of the primary MDS or secondary MDS failed.

Response

Check following points.

- Quantity of memory is enough.

    - Check MemFree and SwapFree of /proc/meminfo.

- Connection state of a network is normal.

    - Check whether DNS server is operating normally.

    - Check whether connection of LAN is normal.

    - Check whether network interface card is normal by the following commands:

        - [RHEL6] **netstat(8)**

        - [RHEL7] **ip(8)**

- Setting of network is normal.

    - Check whether all the IP addresses of the nodes that share the file system are correctly registered in host database (such as /etc/hosts). For details, see "11.2.2 Registering information in host database" in this manual.

    - Check whether IP address of a DNS server is described in /etc/resolv.conf.

    - Check whether /etc/nsswitch.conf is described correctly.

    - Check whether fire wall is set up correctly so that access of port used by GFS may not be refused.

- State of volume of GDS is normal.

    - Check state of volume of GDS by **sdxinfo** command. For details on **sdxinfo**, see "Appendix B.1.6 sdxinfo" of "PRIMECLUSTER Global Disk Services Configuration and Administration Guide."

- Isn't the disk breakdown?

    - Confirm whether the error message of GDS or a disk driver is output to /var/log/messages. When the error message is output, please perform disposal corresponding to the message. For details, see "PRIMECLUSTER Global Disk Services Configuration and Administration Guide" or the manual of disk units.

- Is the file system configuration information normal?

    - Please execute the following commands and check that the host name of the shared node which the **sfcinfo(8)** command and the **sfcrscinfo(8)** command display and the host name specified to be **-m** option of the **sfcadm(8)** command execution line obtained by the **sfcgetconf(8)** command are the same.

        - sfcinfo -n <representative partition of the file system>

        - sfcrscinfo -m <representative partition of the file system>

        - sfcgetconf -d <representative partition of the file system> <output destination file name>

    - When not the same, please change the -**m** option to the -u option of the **sfcadm(8)** command execution line obtained by the **sfcgetconf(8)** command, and then execute the command line.

If the problem cannot be solved by the response above, collect the diagnostic data with FJQSS, **fjsnap**, or **pclsnap**, and contact field engineers.

For details on commands, see sfcinfo(8), sfcrscinfo(8) and sfcgetconf(8) in this manual. Moreover for the details of the **sfcadm(8)** command execution line obtained by **sfcgetconf(8)**, see "9.6 Backup of the management partition information" and "9.7 Restoring of the management partition information" in this manual.

ERROR: sfcfs_mds:1006: can't execute MDS

Explanation

Starting up MDS failed.

Response

Collect a file output with FJQSS, **fjsnap**, or **pclsnap** for troubleshooting information then contact field engineers.

ERROR: sfcfs_mds:1007: *fsid*: MDS terminated because logreplay failed. errcode=*errcode*

Explanation

The log replay by MDS of the file system failed.

Response

Collect a file output with FJQSS, **fjsnap**, or **pclsnap** for troubleshooting information then contact field engineers.

ERROR: sfcfs_mds:1008: *fsid*: Filesystem status is FSBAD, so MDS can not run.fs_clean=*fs_clean*

Explanation

Starting up MDS failed, because MDS detected an error in file system.

Response

Unmount the relevant file system from all of the nodes, and execute **sfcfsck(8)** specifying **-o nolog**.

ERROR: sfcfs_mds:1009: *fsid*: MDS terminated abnormally because I/O error occurred continuously on meta-data area of the disk.

Explanation

MDS terminated abnormally because I/O error occurred continuously on meta-data area of the disk.

Response

To resume the system operation, execute **sfcfsck(8)** in the full check mode first.

Then back up the file system and correct the hardware failure by replacing the failed disk with a new one, and restore the backed-up data to recover the file system.

A.2.5 Error messages (sfcfsd daemon)

This section explains the sfcfsd daemon error messages of the GFS Shared File System.

ERROR: sfcfsd: 0001: gfs monitor daemon for mountpoint(*mountpoint*) abort: *detail*
ERROR: sfcfsd: 0002: gfs monitor daemon for fsid(*fsid*) abort: *detail*
ERROR: sfcfsd: 0003: gfs monitor daemon abort: *detail*

Explanation

The GFS Shared File System monitoring daemon ended abnormally.
The error explained in *detail* occurred.

Response

Collect a file output with FJQSS, **fjsnap**, or **pclsnap** for troubleshooting information then contact field engineers.

ERROR: sfcfsd: 0004: can't execute primary MDS : mountpoint(*mountpoint*)

Explanation

Starting up the primary MDS has failed.

Response

Please confirm whether the message of sfcfs_mds is output at the same time as this message. When the message of sfcfs_mds is output, please cope with it according to the message. Otherwise, collect the diagnostic data with FJQSS, **fjsnap**, or **pclsnap**, and contact field engineers.

ERROR: sfcfsd: 0005: can't execute secondary MDS : mountpoint(*mountpoint*)

Explanation

Start up of the secondary MDS has failed.

Response

Please confirm whether the message of sfcfs_mds is output at the same time as this message. When the message of sfcfs_mds is output, please cope with it according to the message. Otherwise, collect the diagnostic data with FJQSS, **fjsnap**, or **pclsnap**, and contact field engineers.

ERROR: sfcfsd: 0006: can't take over MDS : mountpoint(*mountpoint*)

Explanation

Switchover of the MDS failed.

Response

Collect a file output with FJQSS, **fjsnap**, or **pclsnap** for troubleshooting information then contact field engineers.

ERROR: sfcfsd: 0007: primary MDS down : mountpoint(*mountpoint*)

Explanation

The primary MDS has failed.

Response

Please confirm whether the message of sfcfs_mds is output at the same time as this message. When the message of sfcfs_mds is output, please cope with it according to the message. Otherwise, collect the diagnostic data with FJQSS, **fjsnap**, or **pclsnap**, and contact field engineers.

ERROR: sfcfsd: 0008: secondary MDS down : mountpoint(*mountpoint*)

Explanation

The secondary MDS has failed.

Response

Please confirm whether the message of sfcfs_mds is output at the same time as this message. When the message of sfcfs_mds is output, please cope with it according to the message. Otherwise, collect the diagnostic data with FJQSS, **fjsnap**, or **pclsnap**, and contact field engineers.

ERROR: sfcfsd: 0009: can't mount to primary MDS. gfs monitor daemon stop. : mountpoint(*mountpoint*)

Explanation

Mounting on the primary MDS failed. The GFS Shared File System monitoring daemon terminates.

Response

Collect a file output with FJQSS, **fjsnap**, or **pclsnap** for troubleshooting information then contact field engineers.

ERROR: sfcfsd: 0010: can't mount to secondary MDS. : mountpoint(*mountpoint*)

Explanation

Mounting on the secondary MDS failed. There is no problem with the file system operation.

Response

Collect a file output with FJQSS, **fjsnap**, or **pclsnap** for troubleshooting information then contact field engineers.

ERROR: sfcfsd: 0011: blockade : mountpoint(*mountpoint*)

Explanation

The file system (*mountpoint*) has been blockaded. The target mount point is unavailable and only unmounting can be performed.

Response

Collect a file output with FJQSS, **fjsnap**, or **pclsnap** for troubleshooting information then contact field engineers.

A.2.6 Error messages (sfcfs_mount command)

This section explains the sfcfs_mount command messages that is called sfcfsd daemon of GFS Shared File System.

ERROR: sfcfs_mount:0001:*special* : cannot allocate memory

Explanation

The mounting failed due to a memory allocation failure.

Response

Check the system status and wait for the termination of another process or increase the swap space, and re-execute the command.

ERROR: sfcfs_mount:0002:*special* : signal(2) is error: errmsg

Explanation

The mounting failed due to an internal error.

Response

Collect a file output with FJQSS, **fjsnap**, or **pclsnap** for troubleshooting information then contact field engineers.

ERROR: sfcfs_mount:0003:*special* : cannot connect to FsRM

Explanation

The file system monitoring mechanism is not operating.

Response

Check whether the file system monitoring mechanism is running.

ERROR: sfcfs_mount:0004: *special* is not found in file system configuration information

Explanation

The specified partition (*special*) does not exist in the file system configuration information.

Response

Check whether the specified partition is correct.

ERROR: sfcfs_mount:0005: *special* : cannot read file system configuration information

Explanation

Reading the file system configuration information failed.

Response

Collect a file output with FJQSS, **fjsnap**, or **pclsnap** for troubleshooting information then contact field engineers.

ERROR: sfcfs_mount:0006: *special* : Cannot get port number of sfcfs-n for [P-MDS | S-MDS] not found in /etc/services

Explanation

The port number definition used in AC-to-MDS communication is not found.

Response

Review the port number definition in /etc/services.

ERROR: sfcfs_mount:0007: *special* : Cannot get host address of hostname for [P-MDS | S-MDS], not found in /etc/hosts

Explanation

The host name information for AC-to-MDS communication cannot be obtained.

Response

Check following points:

- Setting of network is normal.
  Check whether the IP addresses of the MDS nodes are correctly registered in host database (such as /etc/hosts). For details, see "11.2.2 Registering information in host database."

- The file system configuration information is normal.
  Execute **sfcrscinfo(8)** with the -**m** option to check whether the host names of the MDS nodes are correct.
  If the host names are not correct, change them by using **sfcadm(8)** with the -**u** option.

If the problem cannot be solved after taking the above actions, collect a file output with FJQSS, **fjsnap**, or **pclsnap** for troubleshooting information then contact field engineers.

For details on the commands, see sfcrscinfo(8) and sfcadm(8) in this manual.

ERROR: sfcfs_mount:0009: *special* : can not make directory *path* : *errmsg*

Explanation

Creating a directory failed.

Response

Take action according to **errmsg**.

ERROR: sfcfs_mount:0010: *special* is already mounted, *mount_point* is busy

Explanation

The mounting fails because it has already been mounted.

Response

Check whether the mounting has not been done with **mount(8)**, and take action.

ERROR: sfcfs_mount:0011:*special* : recovery mount is in progress

Explanation

The mounting cannot be performed because recovery processing is in progress.

Response

None.

ERROR: sfcfs_mount:0012:*special* : cannot connect to Primary-MDS
ERROR: sfcfs_mount:0013:*special* : cannot connect to Secondary-MDS

Explanation

AC cannot connect to MDS. Therefore, mount the file system will be failed.

Response

Check MDS down. In this case, try mount again.

ERROR: sfcfs_mount:0014:*special* : sfcfs is not supported by kernel

Explanation

The sfcfs module has not been loaded.

Response

Check whether the sfcfs module has been loaded.

ERROR: sfcfs_mount:0015:*special,mountpoint*: mount(2) is error:*errmsg*

Explanation

An unexpected error was detected on the mounting.
Possible causes:

- The *special* or *mountpoint* specified is invalid.

- The network setting of a shared node is invalid.

- The host name of a shared node specified is invalid.

- The shared node fails during command execution.

Response

Execute **sfcmount(8)** again after correcting the cause of the failure according to the following procedures:

- Check whether a file and a directory specified with *special* and *mountpoint* exist. If the specification is not correct, specify the correct *special* and *mountpoint*. Create the directory specified with *mountpoint*, if it has not been created.

- Check whether the IP address of the node that shares the file system is correctly registered in host database (such as /etc/hosts). For details, see "11.2.2 Registering information in host database."

- Execute **sfcinfo(8)** with the -**n** and -**a** options to check whether the host name of a shared node is correct. If the host name is not correct, change it by using **sfcadm(8)** with -**u** option.

- If any warning message was output immediately before this message, take an action to deal with the warning message.

If the problem cannot be solved by the response above, collect the diagnostic data FJQSS, **fjsnap**, or **pclsnap**, and contact field engineers.

For details on the commands, see sfcinfo(8) and sfcadm(8) in this manual.

ERROR: sfcfs_mount:0016:*special* : cannot add mount information to /etc/mtab

Explanation

The mounting succeeded, but adding the mount information to /etc/mtab failed.

Response

Check the /etc/mtab settings and add the relevant mount information to /etc/mtab.

ERROR: sfcfs_mount:0017:*special* : cannot delete mount information in /etc/mtab

Explanation

Deleting the mount information from /etc/mtab failed.

Response

Check the /etc/mtab settings and delete the relevant mount information from /etc/mtab.

ERROR: sfcfs_mount:0018:*special* : unknown error(*errno*)

Explanation

An unknown error was detected.

Response

Collect a file output with FJQSS, **fjsnap**, or **pclsnap** for troubleshooting information then contact field engineers.

ERROR: sfcfs_mount:0019: cannot open *special*: *errmsg*

Explanation

Opening of the partition (*special*) failed.

Response

Take corrective action according the detailed information *errmsg*.

ERROR: sfcfs_mount:0020: cannot read the {superblock | volume information} of *special*

Explanation

Reading of the super block information or partition configuration information on the partition (*special*) failed.

Response

Collect a file output with FJQSS, **fjsnap**, or **pclsnap** for troubleshooting information then contact field engineers.

ERROR: sfcfs_mount:0021: *special* is not sfcfs file system

Explanation

The specified partition (*special*) is not of the GFS Shared File System.

Response

Specify a special file that is being used for the GFS Shared File System.

ERROR: sfcfs_mount:0023:*special* : invalid parameter(*param*)

Explanation

An invalid value was detected in the tuning parameter *param* for the file system.

Response

Collect a file output with FJQSS, **fjsnap**, or **pclsnap** for troubleshooting information then contact field engineers.

A.2.7 Error messages (sfcpncd daemon)

This section explains the sfcpncd daemon error messages of the GFS Shared File System.

ERROR: sfcpncd:0001: Permission denied. Abort.

Explanation

Execution was attempted with other than super-user permission.

Response

Do not execute the module directly.

ERROR: sfcpncd:0002: Initialize fail. Abort.

Explanation

sfcpncd failed booting.

Response

Collect a file output with FJQSS, **fjsnap**, or **pclsnap** for troubleshooting information then contact field engineers.

ERROR: sfcpncd:0003: Cannot create the semaphore. Abort.

Explanation

Creation of the semaphore failed.

Response

Collect a file output with FJQSS, **fjsnap**, or **pclsnap** for troubleshooting information then contact field engineers.

ERROR: sfcpncd:0004: Fatal error. Abort.

Explanation

sfcpncd terminated abnormally.

Response

Collect a file output with FJQSS, **fjsnap**, or **pclsnap** for troubleshooting information then contact field engineers.

A.2.8 Error messages (sfcprmd daemon)

This section explains the sfcprmd daemon error messages of the GFS Shared File System.

ERROR: sfcprmd:0001: Permission denied. Abort.

Explanation

Execution was attempted with other than super-user permission.

Response

Do not execute the module directly.

ERROR: sfcprmd:0002: Initialize failed. Abort.

Explanation

sfcprmd failed booting.

Response

Collect a file output with FJQSS, **fjsnap**, or **pclsnap** for troubleshooting information then contact field engineers.

ERROR: sfcprmd:0003: Cannot open /dev/null. Abort.

Explanation

Opening **/dev/null** failed.

Response

Collect a file output with FJQSS, **fjsnap**, or **pclsnap** for troubleshooting information then contact field engineers.

ERROR: sfcprmd:0004: Cannot access to the /var/opt/FJSVsfcfs/.sfcprmd_uds. Abort.

Explanation

Accessing /var/opt/FJSVsfcfs/.sfcprmd_uds failed.

Response

Collect a file output with FJQSS, **fjsnap**, or **pclsnap** for troubleshooting information then contact field engineers.

ERROR: sfcprmd:0005: Memory not enough. Abort.

Explanation

There is insufficient memory to allow sfcprmd to boot.

Response

Increase the amount of memory or the size of the swap area.

ERROR: sfcprmd:0007: Cannot access to the semaphore of sfcpncd. Abort.

Explanation

The semaphore created by sfcpncd cannot be accessed.

Response

Collect a file output with FJQSS, **fjsnap**, or **pclsnap** for troubleshooting information then contact field engineers.

A.2.9 Warning messages (MDS (sfcfsmg daemon))

This section explains the MDS (sfcfsmg daemon) warning messages of the GFS Shared File System.

WARNING: sfcfs_mds:2001: *fsid*: daemon create failed(lock_init failed). errcode=*errcode*

Explanation

Generating the daemon failed at mounting the file system (*fsid*) due to a failure in the lock object initialization.

Response

Increase memory or the size of the swap area.

WARNING: sfcfs_mds:2002: daemon create failed(can't create thread).

Explanation

A daemon could not be created when mount file system.

Response

Increase memory or the size of the swap area.

WARNING: sfcfs_mds:2003: *fsid*:daemon create failed(can't creat thread).

Explanation

A daemon could not be created when mount file system (*fsid*).

Response

Increase memory or the size of the swap area.

WARNING: sfcfs_mds:2004: *fsid*: **bad dir: invalid offset. ino=***ino-num* **offset=***offset*
WARNING: sfcfs_mds:2007: *fsid*: **bad dir: size error. ino=***ino-num* **size=***size*

Explanation

An error was detected in a directory entry of *fsid*.

Response

Unmount the relevant file system from all of the nodes, and execute **sfcfsck(8)** specifying **-o nolog**.

WARNING: sfcfs_mds:2008: *fsid*: **meta data read error. type=***number* **dev=***dev_num* **block=***block_num*
size=*size* **errcode=***errcode*
WARNING: sfcfs_mds:2009: *fsid*: **meta data read error. type=***number* **dev=***dev_num* **block=***block_num*
size=*size* **errcode=***errcode*

Explanation

An error occurred during an attempt to read the meta-data area from a device *dev_num* of the file system (*fsid*).

Response

Take action according to the instructions in "C.3.1 Action for I/O errors."

WARNING: sfcfs_mds:2010: *fsid*: **bitmap read error. type=***number* **dev=***dev_num* **block=***block_num*
size=*size* **errcode=***errcode*
WARNING: sfcfs_mds:2011: *fsid*: **bitmap read error. type=***number* **dev=***dev_num* **block=***block_num*
size=*size* **errcode=***errcode*

Explanation

An error occurred during an attempt to read the bitmap area from a device *dev_num* of the file system (*fsid*).

Response

Take action according to the instructions in "C.3.1 Action for I/O errors."

WARNING: sfcfs_mds:2012: bitmap read error. type=*number* **dev=***dev_num* **block=***block_num* **size=***size*
errcode=*errcode*

Explanation

An error occurred during an attempt to read the bitmap area from a device *dev_num* of the file system.

Response

Take action according to the instructions in "C.3.1 Action for I/O errors."

WARNING: sfcfs_mds:2013: *fsid*: **bitmap write error.type=***number* **dev=***dev_num* **block=***block_num*
size=*size* **errcode=***errcode*

Explanation

An error occurred during an attempt to write to the bitmap area on a device *dev_num* of the file system (*fsid*).

Response

Unless message sfcfs_mds:2014 is issued, recovery by retry processing is successful. Therefore, no response is required.

WARNING: sfcfs_mds:2014: *fsid*: **bitmap write error information was lost. type=***number* **dev=***dev_num* **block=***block_num* **size=***size* **errcode=***errcode*

Explanation

An error occurred during an attempt to write to the bitmap area on a device *dev_num* of the file system (*fsid*). Part of the file system information was lost.

Response

Take action according to the instructions in "C.3.1 Action for I/O errors."

WARNING: sfcfs_mds:2016: *fsid*: **meta data write error. type=***number* **dev=***dev_num* **block=***block_num* **size=***size* **errcode=***errcode*

Explanation

An error occurred during an attempt to write to the meta-data area on a device *dev_num* of the file system (*fsid*).

Response

Unless message sfcfs_mds:2017 is issued, recovery by retry processing is successful. Therefore, no response is required.

WARNING: sfcfs_mds:2017: *fsid*: **meta data write error information was lost. type=***number* **dev=***dev_num* **block=***block_num* **size=***size* **errcode=***errcode*

Explanation

An error occurred during an attempt to write to the meta-data area on a device *dev_num* of the file system (*fsid*). Part of the file system information was lost.

Response

Take action according to the instructions in "C.3.1 Action for I/O errors."

WARNING: sfcfs_mds:2018: *fsid*: **super block write error. dev=***dev_num* **block=***block_num* **size=***size* **errcode=***errcode*

Explanation

An error occurred during an attempt to write a super block to a device *dev_num* of the file system (*fsid*).

Response

Unless message sfcfs_mds:2019 is issued, recovery by retry processing is successful. Therefore, no response is required.

WARNING: sfcfs_mds:2019: *fsid*: **super block write error information was lost. dev=***dev_num* **block=***block_num* **size=***size* **errcode=***errcode*

Explanation

An error occurred during an attempt to write a super block to a device *dev_num* of the file system (*fsid*). Part of the file system information was lost.

Response

Take action according to the instructions in "C.3.1 Action for I/O errors."

WARNING: sfcfs_mds:2020: *fsid*: **LOG daemon is dead**

Explanation

The log daemon stopped processing because it could not recover from a failure to write to the update log area on the file system (*fsid*).

Response

Because of an I/O error, the update log daemon was not able to operate and terminated. The file system access performance is degraded because operation is continued without the update log available. Take action according to the instructions in "C.3.1 Action for I/O errors."

WARNING: sfcfs_mds:2021: *fsid*: cannot access, too many I/O errors. dev=*dev_num*

Explanation

The file system (*fsid*) was blockaded due to a writing error to the super block on the device number *dev_num* in the file system. In the future, access to this file system will be an error return.

Response

Take action according to the instructions in "C.3.1 Action for I/O errors."

WARNING: sfcfs_mds:2022: *fsid*: volume information write error. dev=*dev_num* block=*block_num* size=*size* errcode=*errcode*
WARNING: sfcfs_mds:2023: *fsid*: volume information write error. dev=*dev_num* block=*block_num* size=*size* errcode=*errcode*

Explanation

An error occurred during an attempt to write a partition configuration data to a device *dev_num* in the file system (*fsid*).

Response

Take action according to the instructions in "C.3.1 Action for I/O errors."

WARNING: sfcfs_mds:2024: *fsid*: cannot delete extents. ino=*number* offset=*offset* block=*block* errcode=*errcode*

Explanation

An error in extent-based management of a file system (*fsid*) was detected.

Response

Unmount the relevant file system from all of the nodes, and execute **sfcfsck(8)** specifying **-o nolog**.

WARNING: sfcfs_mds:2026: *fsid*: super block read error. dev=*dev_num* errcode=*errcode*

Explanation

The super block on the device of device number (*dev_num*) cannot be read, when mounting the file system (*fsid*).

Response

Unmount the relevant file system, and execute **sfcfsck(8)** specifying **-o nolog**.

WARNING: sfcfs_mds:2027: super block read error. errcode=*errcode*

Explanation

In mounting the file system, reading from a super block on the device failed.

Response

Unmount the relevant file system, and execute **sfcfsck(8)** specifying **-o nolog**.

WARNING: sfcfs_mds:2028: *fsid*: partition configuration read error. dev=*dev_num* errcode=*errcode*

Explanation

Partition configuration data could not be read when mount the file system (*fsid*) started.

Response

Take action according to the instructions in "C.3.1 Action for I/O errors."

WARNING: sfcfs_mds:2029: partition configuration read error. errcode=*errcode*

Explanation

Partition configuration data could not be read when mount the file system started.

Response

Take action according to the instructions in "C.3.1 Action for I/O errors."

WARNING: sfcfs_mds:2030: *fsid*: node information not found for node

Explanation

Information for *node* was not found in the file system (*fsid*).

Response

Register correct node information using **sfcsetup(8)**.

WARNING: sfcfs_mds:2031: *fsid*: primary and secondary super blocks are different. dev=*dev_num*

Explanation

An inconsistency was detected between the contents of the primary super block and the contents of the secondary super block on the indicated device.

Response

Specify the partition of device number *dev_num* using **sfcinfo(8)**, and perform the procedure of "C.3.2 Corrective action in the event of data inconsistency."

```
# sfcinfo -d dev_num <Enter>
/dev/sfdsk/gfs01/dsk/volume01:
FSID special                                    size  Type  mount
   1 /dev/sfdsk/gfs01/dsk/volume01(dev_num)    14422  META  /mnt
   1 /dev/sfdsk/gfs01/dsk/volume01(dev_num)     5116  LOG   /mnt
   1 /dev/sfdsk/gfs01/dsk/volume01(dev_num)    95112  DATA  /mnt
```

WARNING: sfcfs_mds:2032: *fsid*: primary and secondary volume informations are different. dev=*dev_num*

Explanation

An inconsistency was detected between the contents of the primary partition configuration data and the contents of the secondary partition configuration data on the device.

Response

Specify the partition of device number *dev_num* using **sfcinfo(8)**, and perform the procedure of "C.3.2 Corrective action in the event of data inconsistency."

```
# sfcinfo -d dev_num <Enter>
/dev/sfdsk/gfs01/dsk/volume01:
FSID special                                    size  Type  mount
   1 /dev/sfdsk/gfs01/dsk/volume01(dev_num)    14422  META  /mnt
   1 /dev/sfdsk/gfs01/dsk/volume01(dev_num)     5116  LOG   /mnt
   1 /dev/sfdsk/gfs01/dsk/volume01(dev_num)    95112  DATA  /mnt
```

WARNING: sfcfs_mds:2033: *fsid*: super blocks differ. dev *dev_num1* to *dev_num2*

Explanation

An inconsistency was found in the super block information in the indicated file system (*fsid*) in a multiple partition configuration.

Response

Specify the partition of device number *dev_num2* using **sfcinfo(8)**, and perform the procedure of "C.3.2 Corrective action in the event of data inconsistency."

```
# sfcinfo -d dev_num2 <Enter>
/dev/sfdsk/gfs01/dsk/volume01:
FSID special                                     size   Type   mount
   1 /dev/sfdsk/gfs01/dsk/volume01(dev_num2)    14422   META   /mnt
   1 /dev/sfdsk/gfs01/dsk/volume01(dev_num2)     5116   LOG    /mnt
   1 /dev/sfdsk/gfs01/dsk/volume01(dev_num2)    95112   DATA   /mnt
```

WARNING: sfcfs_mds:2034: mmap can't access between other nodes fsid=*fsid* ino=*i-number* nodename(*node1*,*node2*) mode(*mode1*,*mode2*)

Explanation

open(2) and **mmap(2)** that cannot be maintain the consistency of the data were issued to the same file.

Response

Please wait for the termination of the process running on the "*node1*" indicated this message, and execute again.
To determine the process using the *i-number*, act according to the following sequence:

1. determine the filename of the *i-number*, using **find(1)** with **-inum** option.

2. determine the process ID that uses the filename, using **fuser(1)**.

3. determine the command name of the process ID, using **ps(1)**.

WARNING: sfcfs_mds:2035: MDS(recover) is timed out

Explanation

MDS detected a timeout at recovery mounting.

Response

Reboot the system and perform the mounting again.

WARNING: sfcfs_mds:2036: can't allocate memory in primary

Explanation

Memory could not be allocated on the file system.

Response

Increase memory or the size of the swap area.

WARNING: sfcfs_mds:2037: read error on log partition(errno=*errno* errcode=*errcode*): *mount_point*

Explanation

An error occurred, in reading from the update log area in the file system (*mount_point*).

Response

Take action according to the instructions in "C.3.1 Action for I/O errors."

WARNING: sfcfs_mds:2038: read error on meta partition(errno=*errno* errcode=*error_code*): *mount_point*

Explanation

An error occurred, in reading from the meta-data area in the file system (*mount_point*).

Response

Take action according to the instructions in "C.3.1 Action for I/O errors."

WARNING: sfcfs_mds:2039: write error on log partition(errno=*errno* errcode=*error_code*): *mount_point*

Explanation

An error occurred, in writing to the update log area in the file system (*mount_point*).

Response

Take action according to the instructions in "C.3.1 Action for I/O errors."

WARNING: sfcfs_mds:2040: write error on meta partition(errno=*errno* errcode=*error_code*): *mount_point*

Explanation

An error occurred during an attempt to write the meta partition.

Response

Take action according to the instructions in "C.3.1 Action for I/O errors."

WARNING: sfcfs_mds:2042: *fsid*: The size of the partition might have been changed after the file system was made.

Explanation

The size of the partition which constitutes the file system (*fsid*) might have been changed.

Response

Perform the procedure (from step 6 on) described on "C.3.2 Corrective action in the event of data inconsistency."

GFS Shared File System can not be expanded by size expansion of the partition which shares the GFS Shared File System. Perform the procedure above promptly.

WARNING: sfcfs_mds:2043: *fsid*: volume informations differ. dev *dev_num1* to *dev_num2*

Explanation

Partition configuration information is inconsistent in the multipartition file system (*fsid*).

Response

Specify the partition of device number *dev_num2* using **sfcinfo(8)**, and perform the procedure of "C.3.2 Corrective action in the event of data inconsistency."

```
# sfcinfo -d dev_num2 <Enter>
/dev/sfdsk/gfs01/dsk/volume01:
FSID special                                 size  Type  mount
   1 /dev/sfdsk/gfs01/dsk/volume01(dev_num2)  14422  META  /mnt
   1 /dev/sfdsk/gfs01/dsk/volume01(dev_num2)   5116  LOG   /mnt
   1 /dev/sfdsk/gfs01/dsk/volume01(dev_num2)  95112  DATA  /mnt
```

WARNING: sfcfs_mds:2046: *fsid*: invalid parameter(*param*)

Explanation

An invalid value was detected in the tuning parameter *param* for the file system.

Response

Check whether the tuning parameter is correctly set. For details, see "14.1 Tuning Parameters" in this manual.

A.2.10 Warning messages (sfcfsd daemon)

This section explains the sfcfsd daemon warning messages of the GFS Shared File System.

WARNING: sfcfsd: 0001: invalid parameter(*param*), so it is ignored

Explanation

An invalid value was detected in tunable parameters for the file system, and it is ignored.

Response

Review the file system's tunable parameters.

A.2.11 Warning messages (sfchnsd daemon)

When the warning message of the sfchnsd daemon was output, please collect a file output with FJQSS, **fjsnap**, or **pclsnap**, contact field engineers.

A.2.12 Information messages (sfcfrmd daemon)

This section explains the sfcfrmd daemon information messages of the GFS Shared File System.

NOTICE: sfcfrmd:0001: serve as master

Explanation

sfcfrmd was started as a master.

Response

None.

NOTICE: sfcfrmd:0002: serve as slave

Explanation

sfcfrmd was started as a slave.

Response

None.

NOTICE: sfcfrmd:0003: service start

Explanation

Service of sfcfrmd was started.

Response

None.

NOTICE: sfcfrmd:0004: service stop

Explanation

Service of sfcfrmd was completed.

Response

None.

A.2.13 Information messages (MDS (sfcfsmg daemon))

This section explains the MDS (sfcfsmg daemon) information messages of the GFS Shared File System.

NOTICE: sfcfs_mds:3001: *fsid*: file system full

Explanation

There is no more space in file data area of the file system (*fsid*).
(For an ordinary user, this message is also output when space drops below the **minfree** limit.)

Response

Perform any of the following procedures to expand the space of the file system:

- Expand the file system.

- Delete unnecessary files.

- Move files to the file system that has enough space.

Use **sfcadd(8)** to expand the file system.

NOTICE: sfcfs_mds:3002: *fsid*: out of inodes

Explanation

There is no more space in i-node area of the file system (*fsid*).

Response

Delete unnecessary files or move files to a file system that has enough area.

NOTICE: sfcfs_mds:3003: *fsid*: file system meta data full(Vdata) type=*number*

Explanation

There is no more space in V-data area of the file system (*fsid*).

Response

Delete unnecessary files or move files to a file system that has enough area.

NOTICE: sfcfs_mds:3004: *fsid*: too much fragment(extent)

Explanation

There is no more area for managing unused space in the file system (*fsid*) because of area fragmentation. An area released by subsequent area release processing may become an invalid area.

Response

To sort out fragmented area in the file system, back up each file in the file system, and recreate the file system then restore the backed up data.

NOTICE: sfcfs_mds:3006: *fsid*: I/O error on log. dev=*dev_num* block=*block_num,* size=*size* errcode=*errcode*

Explanation

The system could not write an update log on device *dev_num* of the file system (*fsid*).

Response

Unless message sfcfs_mds:2020 is issued, recovery by retry processing was successful. So no response is required.

NOTICE: sfcfs_mds:3007: *fsid*: primary super block read failed, but secondary read. dev=*dev_num* block=*block_num* size=*size* errcode=*errcode*

Explanation

The primary super block on the device *dev_num* of the file system (*fsid*) could not be read. Processing was continued because the secondary super block was read correctly.

Response

File system operation can be continued. However, because a block error occurred, promptly take action according to the instructions in "C.3.1 Action for I/O errors."

NOTICE: sfcfs_mds:3008: primary super block read failed, but secondary read. errcode=*errcode*

Explanation

The primary super block on the device of the file system could not be read. Processing was continued because the secondary super block was read correctly.

Response

File system operation can be continued. However, because a block error occurred, promptly take action according to the instructions in "C.3.1 Action for I/O errors."

NOTICE: sfcfs_mds:3009: *fsid*: secondary super block read failed, but primary read. dev=*dev_num* block=*block_num* size=*size* errcode=*errcode*

Explanation

The secondary super block on the device *dev_num* of the file system (*fsid*) could not be read. Processing was continued because the primary super block was read correctly.

Response

File system operation can be continued. However, because a block error occurred, promptly take action according to the instructions in "C.3.1 Action for I/O errors."

NOTICE: sfcfs_mds:3010: secondary super block read failed, but primary read. errcode=*errcode*

Explanation

The secondary super block on the device of the file system could not be read. Processing was continued because the primary super block was read correctly.

Response

File system operation can be continued. However, because a block error occurred, promptly take action according to the instructions in "C.3.1 Action for I/O errors."

NOTICE: sfcfs_mds:3011: *fsid*: primary partition configuration read failed, but secondary read. dev=*dev_num* block=*block_num* size=*size* errcode=*errcode*

Explanation

Primary partition configuration data on the device *dev_num* of the file system (*fsid*) could not be read. Processing was continued because the secondary partition configuration data was read correctly.

Response

File system operation can be continued. However, because a block error occurred, promptly take action according to the instructions in "C.3.1 Action for I/O errors."

NOTICE: sfcfs_mds:3012: primary partition configuration read failed, but secondary read. errcode=*errcode*

Explanation

Primary partition configuration data on the device of the file system could not be read. Processing was continued because the secondary partition configuration data was read correctly.

Response

File system operation can be continued. However, because a block error occurred, promptly take action according to the instructions in "C.3.1 Action for I/O errors."

NOTICE: sfcfs_mds:3013: *fsid*: secondary partition configuration read failed, but primary read. dev=*dev_num* block=*block_num* size=*size* errcode=*errcode*

Explanation

Secondary partition configuration data on the device *dev_num* of the file system (*fsid*) could not be read. Processing was continued because the primary partition configuration data was read correctly.

Response

File system operation can be continued. However, because a block error occurred, promptly take action according to the instructions in "C.3.1 Action for I/O errors."

NOTICE: sfcfs_mds:3014: secondary partition configuration read failed, but primary read. errcode=*errcode*

Explanation

Secondary partition configuration data on the device of the file system could not be read. Processing was continued because the primary partition configuration data was read correctly.

Response

File system operation can be continued. However, because a block error occurred, promptly take action according to the instructions in "C.3.1 Action for I/O errors."

NOTICE: sfcfs_mds:3015: *fsid*: can't get partition size. dev=*dev_num*: error=*errcode*

Explanation

The size of the specified partition on the device *dev_num* of the file system (*fsid*) could not be obtained. Processing was continued because the primary super block was read correctly.

Response

Unmount the file system, and make the super blocks consistent using **sfcadm(8)**.

NOTICE: sfcfs_mds:3016: can't get partition size

Explanation

The size of the specified partition on the device of the file system could not be obtained. Processing was continued because the primary super block was read correctly.

Response

Unmount the file system, and make the super blocks consistent using **sfcadm(8)**.

NOTICE: sfcfs_mds:3017: nodename mismatch. *node<->node*

Explanation

A node with an existing hostid and nonexistent node name was registered in the file system.

Response

Unless sfcfs_mds:2030 is issued, no response is required.

NOTICE: sfcfs_mds:3020: *mountpoint:* The total number of vnodes used reached the upper bound. vnodes used=*total_num*: upper bound=*high_num*: lower bound=*low_num*

Explanation

i-node's cache memory (vnode) is released until *total_num* falls below the *low_num* because *total_num* of i-node's cache memory amounted to *high_num*.

Response

The file system operation can be continued. But limit operations to the bare minimum until message (sfcfs_mds:3021) is output, because i-node's cache memory is being released.

NOTICE: sfcfs_mds:3021: *mountpoint:* The total number of vnodes used(=*total_num*) is the lower bound(=*low_num*).

Explanation

Releasing i-node's cache memory (vnode) is stopped because i-node's cache memory reached the *low_num*.

Response

None.

A.2.14 Information messages (sfcprmd daemon)

This section explains the sfcprmd daemon information messages of the GFS Shared File System.

NOTICE: sfcprmd:0006: *daemon* was respawn fail.

Explanation

The daemon indicated by *daemon* failed to reboot.

Response

None.

A.2.15 Information messages (sfchnsd daemon)

This section explains the sfchnsd daemon information messages of the GFS Shared File System.

---

already connected from sfcfrmd

Explanation

The UNIX domain that is used for communication with the "sfcfrmd" daemon is requested to establish communication from non-sfcfrmd daemon process.

Response

Check privileges to access /var/opt/FJSVsfcfs. Also, check if there is a process to access files in the above directory as a super user.

Cannot open libcf.so: *errmsg*

Explanation

Loading of libcf.so failed.

Response

To use GFS, the PRIMECLUSTER Clustering Base must be installed. Check if it is installed correctly.

---

libcf.so loaded

Explanation

Loading of libcf.so succeeded.

Response

None.

---

receive *signal*

Explanation

The "sfchnsd" daemon received *signal*.

Response

None.

---

start

Explanation

The "sfchnsd" daemon has been activated.

Response

None.

A.3 GFS Shared File System Script Messages

This section describes the script messages of the GFS Shared File System.

Types and priorities of messages are listed below.

Type	Priority (*facility.level*)
ERROR	daemon.err
WARNING	daemon.warning
NOTICE	daemon.notice

- The priorities of the messages are passed to rsyslogd(8).

- The output destination of the messages is defined in the /etc/rsyslog.conf file. The output destination of the messages is /var/log/ messages by default of the operating system.

See

For details about the /etc/rsyslog.conf file, see rsyslog.conf(5) in the online manual.

A.3.1 sfcfsrm script

ERROR: sfcfsrm:0001: can not execute *file_name*

Explanation

file_name is not executable.

Response

See "Installation Guide" of PRIMECLUSTER and reinstall the package.

ERROR: sfcfsrm:0002: *management_partition* is not available

Explanation

The status of the GDS volume that was specified as the management partition is not ACTIVE. Or, there is a possibility that the GDS volume is deleted without deleting node configuration information from the management partition by **sfcsetup -d** on the target node when the use of GFS is stopped.

Response

Check whether the volume exists by executing the following procedures:

```
# ls -l management_partition <Enter>
```

- If GFS is used

  - If the volume does not exist

    Create a volume according to the instructions in "Operation" "GDS Management View" in "PRIMECLUSTER Global Disk Services Configuration and Administration Guide."

    - If there is a backup of the management partition information

      Restore the management partition according to the instructions in "9.7 Restoring of the management partition information" in this manual.

    - If there is no backup of the management partition information

      See "9.2 Creating of the management partition" and set up the management partition.

  - If the volume exists

    If the volume status is not ACTIVE, change the status to ACTIVE. For instructions on how to change the volume status, see "Operation" "GDS Management View" in "PRIMECLUSTER Global Disk Services Configuration and Administration Guide."

- If GFS is not used

  - If the volume does not exist

    Delete information on the deleted GDS volume by executing the following procedures:

    1. Create a temporary GDS volume of 40 megabytes or more. (GDS volume name is arbitrary)

    2. sfcsetup -c temporary_GDS_volume (GDS volume name made at 1.)

    3. sfcsetup -a temporary_GDS_volume (GDS volume name made at 1.)

    4. sfcsetup -d

  - If the volume exists

    Delete the node configuration information from the management partition by executing the following procedures:

    1. Set the volume status to ACTIVE. For instructions on how to change the volume status, see "Operation" "GDS Management View" in "PRIMECLUSTER Global Disk Services Configuration and Administration Guide."

    2. See "9.4 Deleting node configuration information from the management partition"and delete the node configuration information from the management partition.

ERROR: sfcfsrm:0003: sfcfrmstart failed

Explanation

sfcfrmd daemon failed booting.

Response

Check the error message of **sfcfrmstart(8)** displayed previously to this message, and take corrective action about that message according to "A.5.4 sfcfrmstart command."

ERROR: sfcfsrm:0004: can not terminate *daemon*

Explanation

The *daemon* could not be stopped.

Response

Collect a file output FJQSS, **fjsnap**, or **pclsnap** for troubleshooting information then contact field engineers.

ERROR: sfcfsrm:0005: can not load *module*

Explanation

Loading of *module* failed.

Response

Check if an urgent fix of GFS corresponding to the kernel version of the system is applied.

If the error still exists after applying the fix above, collect a file output with FJQSS, **fjsnap**, or **pclsnap** for troubleshooting information then contact field engineers.

ERROR: sfcfsrm:0006: can not unload *module*

Explanation

Unloading of *module* failed.

Response

Collect a crash dump and file output with FJQSS, **fjsnap**, or **pclsnap** for troubleshooting information then contact field engineers.

ERROR: sfcfsrm:0007: cluster is not available

Explanation

Cluster resource management has not been activated.

Response

Please follow the procedure in an error message of the Cluster resource management output to /var/log/messages. For details, see "PRIMECLUSTER Messages."

ERROR: sfcfsrm:0008: An error occurred while accessing a file *path* : *errmsg*

Explanation

File access to cancel the suspension of the sfcfrmd daemon's startup failed.

Response

Check that the **/var/opt/FJSVsfcfs** directory exists. If it does, restart the system. If it doesn't, install the package according to the installation guide.

ERROR: sfcfsrm:0009: Querying the state of quorum failed

Explanation

Checking a quorum exists failed.

Response

Check the message of **rcqquery(1M)** displayed previously to this message, and after taking corrective action about that message according to "rcqquery messages" of "PRIMECLUSTER Messages," restart the system.

ERROR: sfcfsrm:0010: Starting the sfcfrmd daemon failed because quorum does not exist

Explanation

Activation of the GFS service is disrupted because a quorum does not exist when CF is started by the GUI. If the sfcfrmd daemon's start mode is set to **wait**, and a quorum does not exist, the GFS service will be discontinued during CF startup.

A quorum does not exist, and the GFS service is discontinued under the following conditions:

- CF is activated by the GUI after CF is stopped on all the nodes, then GFS is operated.

- A cluster partition error exists.

Response

If you activate CF by the GUI after stopping it on all the nodes, see the procedure described in "12.6 How to start up CF by the GUI when a GFS Shared File System is used."

If you did not stop CF on all the nodes, or the GFS service activation fails after performing the procedure detailed above, a cluster partition error might have occurred. Take corrective action according to "C.3.4 Corrective action when the sfcfrmd daemon is not started."

WARNING: sfcfsrm:5001: Starting the sfcfrmd daemon was suspended because quorum does not exist

Explanation

Activation of the sfcfrmd daemon is suspended because a quorum does not exist when the node is started up, the node enters multi-user mode, or CF is started by the GUI. Normally, a quorum be not existing is not maintained until all the nodes are activated. As soon as it is exist, the sfcfrmd daemon will be activated.

Accordingly, it is necessary to take corrective actions under the following conditions otherwise data remains inconsistent.

- A cluster partition error exists.

- When all the nodes are stopped, or CF is stopped on all the nodes, part of nodes or only CF is activated and then GFS is operated.

Response

Take corrective action according to "C.3.4 Corrective action when the sfcfrmd daemon is not started."

WARNING: sfcfsrm:5002: Considering the start mode of the sfcfrmd daemon to be wait because getting the start mode failed

Explanation

Since acquisition of the startup mode of the sfcfrmd daemon fails, the daemon is activated with the **wait** setting.

Response

If GFS is activated properly, no corrective action needs to be taken. If an error message of the sfcfsrm script is output, it is necessary to take corrective action according to the message.

NOTICE: sfcfsrm:2001: The suspended sfcfrmd daemon is resumed because quorum exist

Explanation

Since a quorum exists, the suspended sfcfrmd daemon was started.

Response

None.

NOTICE: sfcfsrm:2002: The suspended sfcfrmd daemon is resumed by ordering of sfcfrmstart

Explanation

Since **sfcfrmstart -f** was executed, the suspended sfcfrmd daemon was started.

Response

None.

A.4 GFS Shared File System Implementations of Standard File System Management Commands' Messages

This section describes messages that can be output when GFS Shared File System implementations of standard File System management commands are used.

A.4.1 sfcdf command

sfcdf: ERROR: cannot open *special*: *errmsg*

Explanation

Opening the partition (special) failed.

Response

Take action according to **errmsg**.

sfcdf: ERROR: cannot allocate memory

Explanation

Memory allocation failed.

Response

Wait for the termination of another process or increase the swap space, and re-execute the command.

sfcdf: ERROR: cannot connect to FsRM

Explanation

The file system monitoring mechanism is not operating.

Response

Check whether the file system monitoring mechanism is running.

sfcdf: ERROR: cannot read file system configuration information

Explanation

Reading the file system configuration information failed.

Response

Collect a file output with FJQSS, **fjsnap**, or **pclsnap** for troubleshooting information then contact field engineers.

sfcdf: ERROR: *special* file system configuration information is full of entry

Explanation

The file system configuration information has reached the limit.

Response

Collect a file output with FJQSS, **fjsnap**, or **pclsnap** for troubleshooting information then contact field engineers.

sfcdf: ERROR: *special* is not a sfcfs file system

Explanation

The specified partition (*special*) is not a sfcfs file system partition.

Response

Check whether the specified partition is a sfcfs file system partition.

sfcdf: ERROR: unknown error = *errid*

Explanation

An unexpected error was detected.

Response

Collect the diagnostic data with FJQSS, **fjsnap**, or **pclsnap**, and contact field engineers submitting it along with text of the error message displayed.

sfcdf: ERROR: filesystem(*special*) is already locked

Explanation

The partition (*special*) is unavailable because another command is using it.

Response

Upon termination of the other command, re-execute the command.

sfcdf: ERROR: cannot read the {superblock | volume information} of *special*

Explanation

Reading the super block or volume information on the partition (*special*) failed.

Response

Delete the target file system where it is unnecessary. For details, see "11.8 Deleting" in this manual. When it is a necessary file system, restore it running a full scan with **sfcfsck(8)**. For details, see "C.3.1 Action for I/O errors" in this manual.

sfcdf: ERROR: cannot stat *special*: *errmsg*

Explanation

Obtaining the attribute of the specified partition (*special*) failed.

Response

Take action according to **errmsg**.

sfcdf: ERROR: filesystem(*special*) is not shared for this node.

Explanation

The command was executed on a node other than the nodes sharing the file system.

Response

Execute the command on a file system sharing node.

sfcdf: ERROR: meta volume is not found

Explanation

An error was detected in volume information.

Response

Collect a file output with FJQSS, **fjsnap**, or **pclsnap** for troubleshooting information then contact field engineers.

sfcdf: ERROR: invalid magic word 0x*magic*

Explanation

The partition configuration data of the specified partition is invalid.

Response

Check whether the specified partition is a sfcfs file system partition.

---

sfcdf: ERROR: invalid device: *special*

Explanation

The specified partition (*special*) is invalid.

Response

Check whether the specified partition is correct.

---

sfcdf: ERROR: cannot ioctl(BLKGETSIZE) *special*: *errmsg*

Explanation

Obtaining the size of the device (*special*) failed.

Response

Take action according to *errmsg*.

---

sfcdf: ERROR: lseek64 error *retval*: *errmsg*

Explanation

Special file I/O processing failed.

Response

Take action according to *errmsg*.

---

sfcdf: ERROR: {primary | secondary} {superblock | volinfo} read error: *errno*

Explanation

A read error occurred in the {primary | secondary} super block.

Response

Check the disk status.

---

sfcdf: ERROR: fs structure is wrong

Explanation

The specified file system is not a sfcfs file system.

Response

Check whether the specified file system is a sfcfs file system.

---

sfcdf: ERROR: volume information for this node is not found

Explanation

The command was executed on a node other than the nodes sharing the file system.

Response

Execute this command on a file system sharing node.

---

sfcdf: ERROR: *special* is not a block device

Explanation

special is not a block special file.

Response

Check whether the specified partition is correct.

---

sfcdf: ERROR: *special* is not representative partition

Explanation

The specified partition *special* is not a representative partition.

Response

Specify a representative partition then execute the command.

A.4.2　sfcfsck command

For details about **sfcfsck(8)** error messages, see "12.2.1 Action to be taken when fsck terminates abnormally."

A.4.3　sfcmkfs command

---

sfcmkfs:ERROR: {Node | Partition} is not same

Explanation

The specified configuration is not the same as the configuration of the previously created file system.

Response

When specifying the **-o force** option, specify the same configuration as that of the previously created file system and re-execute the command.

---

sfcmkfs:ERROR: Cannot find {any | free} port of sfcfs in file

Explanation

A failure occurred because all communication port numbers have been used.

Response

Add a new communication port number for sfcfs to /etc/services.

---

sfcmkfs:ERROR: Cannot connect to cluster
sfcmkfs:ERROR: can not connect to CLUSTER SYSTEM

Explanation

Connecting to the cluster control facility failed.

Response

Make sure that the cluster control facility is operating using the ps command, and re-execute the command.

---

sfcmkfs:ERROR: Cannot look up *special*

Explanation

Specified *special* was not found.

Response

Check whether the specified special file name is correct, and re-execute the command.

---

sfcmkfs:ERROR: Cannot open file : *filename*

Explanation

Opening the file *filename* failed.

Response

Check whether the file *filename* exists, and re-execute the command.

---

sfcmkfs:ERROR: Data area is too small, can't mkfs

Explanation

The file data area is not sufficient in size, and therefore, a file system cannot be created.

Response

Review parameters so that a sufficiently large file data area can be reserved.

---

sfcmkfs:ERROR: can not open special : *filename*(*errmsg*)

Explanation

Opening the special file *filename* failed.

Response

Take action according to the detailed information *errmsg*.

---

sfcmkfs:ERROR: sfcfs_synctab : Cannot write to filesystem configuration information : *errmsg*

Explanation

Updating the file system configuration information failed.

Response

Collect the diagnostic data with FJQSS, **fjsnap**, or **pclsnap**, and contact field engineers submitting it along with text of *errmsg*.

---

sfcmkfs:ERROR: Illegal option : *option*

Explanation

An illegal option was specified.

Response

Specify a correct option and argument.

---

sfcmkfs:ERROR: Master partition is small, log area space not enough

Explanation

The update log area cannot be created because the representative partition is too small.

Response

Check the specified parameter, allocate enough space for the representative partition, and re-execute the command.

---

sfcmkfs:ERROR: Master partition is small, meta area not enough

Explanation

The meta-data area cannot be created because the representative partition is too small.

Response

Review the specified parameter, secure sufficient space on the representative partition, and re-execute the command.

---

sfcmkfs:ERROR: Not super user

Explanation

The user executing this command is not a super-user.

Response

Re-execute the command as a super-user.

---

sfcmkfs:ERROR: Number of hostname is over *num*

Explanation

The number of specified sharing hosts exceeds the maximum **num** for GFS Shared File System.

Response

Specify no more than **num** hosts that will share the file system.

---

sfcmkfs:ERROR: not entry file system

Explanation

No corresponding file system exists while the **-o force** option was specified.

Response

Specify the same configuration as the file system configuration previously created the **-o force** option should be set, and re-execute the command.

---

sfcmkfs:ERROR: No more add filesystem

Explanation

The number limit of registrable GFS Shared File Systems was exceeded, and any more file system cannot be added.

Response

Delete an unnecessary GFS Shared File System where it is possible and re-execute the command. For details about deleting a file system, see sfcadm(8) and sfcinfo(8).

---

sfcmkfs:ERROR: Same hostname *string* is specified more than once

Explanation

The specified host name indicated by **string** has already been specified.

Response

Specify a unique host name for the **-o node** option.

---

sfcmkfs:ERROR: fstat(2) failed *special* : *errmsg*

Explanation

Obtaining the status of the special file **special** failed.

Response

Take action according to detailed information indicated by **errmsg**. For details, see the description of **fstat(2)**.

---

sfcmkfs:ERROR: ioctl(2) { BLKGETSIZE } error : *errmsg*

Explanation

Obtaining the device size failed.

Response

Take action according to the detailed information **errmsg**.

---

sfcmkfs:ERROR: lseek(2) failed *special* : *errmsg*

Explanation

lseek(2) could not be executed for **special**.

Response

Take action according to detailed information indicated by **errmsg**. For details, see the description of **lseek(2)**.

sfcmkfs:ERROR: malloc(3C) failed : *errmsg*

Explanation

Memory could not be allocated.

Response

Wait for the termination of another process or increase the swap space according to the detailed information **errmsg**.

sfcmkfs:ERROR: *special* is already used

Explanation

The specified special device *special* has been used in another GFS Shared File System.

Response

Re-execute the command after taking one of the following actions:

Use an unused special device.

Delete the file system if it is not being used.

Specify **-o force** to re-create a file system with exactly the same configuration.

sfcmkfs:ERROR: Cannot add data, file *filetable*

Explanation

The number limit of registrable GFS Shared File Systems was exceeded, and any other file system cannot be added in *filetable*.

Response

Delete an unnecessary GFS Shared File System where it is possible and re-execute the command. For details about deleting a file system, see sfcadm(8) and sfcinfo(8).

sfcmkfs:ERROR: *special* is mounted, can't mkfs

Explanation

The specified partition indicated by **special** cannot be changed because the file system is mounted.

Response

Unmount the file system and re-execute the command specifying **-o force**, or delete the file system with **sfcadm(8)** and re-execute the command.

sfcmkfs:ERROR: *special* is not a sfcfs file system

Explanation

The special device **special** is not a GFS Shared File System device while the **-m** option was specified.

Response

Specify a special file created already as a GFS Shared File System, and re-execute the command.

sfcmkfs:ERROR: *special* is not special device, can't mkfs

Explanation

Specified **special** is not a special device file.

Response

Specify a special file.

---

sfcmkfs:ERROR: read(2) error at *special* : *errmsg*

Explanation

read(2) for *special* failed.

Response

Take action according to detailed information indicated by *errmsg*. For details, see the description of **read(2)**.

---

sfcmkfs:ERROR: write(2) error at *special* : *errmsg*

Explanation

The **write(2)** operation for *special* failed.

Response

Take action according to detailed information indicated by *errmsg*. For details, see the description of **write(2)**.

---

sfcmkfs:ERROR: Fsid out of range

Explanation

An error was detected in the internal processing.

Response

Collect the diagnostic data with FJQSS, **fjsnap**, or **pclsnap**, and contact field engineers submitting it along with text of the error message displayed.

---

sfcmkfs:ERROR: *special* partition size is over 1TB

Explanation

The size of the partition of *special* has exceeded 1 terabyte.

Response

Reduce the size of *special* or specify another partition when it is exceeding the maximum partition size, and re-execute the command.

---

sfcmkfs:ERROR: fsid bad entry

Explanation

An internal error was detected in the process of newly adding a file system.

Response

Collect the diagnostic data with FJQSS, **fjsnap**, or **pclsnap**, and contact field engineers submitting it along with text of the error message displayed.

---

sfcmkfs:ERROR: Not enough memory

Explanation

Memory allocation failed.

Response

Wait for the termination of another process or increase the swap space.

---

sfcmkfs:ERROR: Not read file system configuration information

Explanation

Obtaining the file system configuration information failed.

Response

Collect a file output with FJQSS, **fjsnap**, or **pclsnap** for troubleshooting information then contact field engineers.

sfcmkfs:ERROR: Fsid already in use

Explanation

An internal error was detected in the process of newly adding a file system.

Response

Collect the diagnostic data with FJQSS, **fjsnap**, or **pclsnap**, and contact field engineers submitting it along with text of the error message displayed.

sfcmkfs:ERROR: can not sfcfs_lock

Explanation

Obtaining the lock in the file system failed.

Response

Re-execute the command after some interval because another process may currently retain the lock.

sfcmkfs:ERROR: can not unlock

Explanation

Releasing the lock in the file system failed.

Response

Re-execute the command after the **sfcmkfs** termination releases the lock.

sfcmkfs:ERROR: can not get fsid

Explanation

Obtaining Fsid for newly adding a file system failed.

Response

Collect the diagnostic data with FJQSS, **fjsnap**, or **pclsnap**, and contact field engineers submitting it along with text of the error message displayed.

sfcmkfs:ERROR: can not configlock

Explanation

Obtaining the lock in the domain failed.

Response

Re-execute the command after some interval because another process may currently retain the lock.

sfcmkfs:ERROR: *special* is already used for the management partition.

Explanation

Specified *special* has been used for the file system management.

Response

Specified *special* is not available because it is in use for the file system management.

Specify another special file to execute the command.

---

sfcmkfs:ERROR: can not acquire fsid

Explanation

An internal error was detected in the process of newly adding a file system.

Response

Collect the diagnostic data with FJQSS, **fjsnap**, or **pclsnap**, and contact field engineers submitting it along with text of the error message displayed.

---

sfcmkfs:ERROR: no entry fsid = *fsidnumber*

Explanation

Obtaining the file system configuration information of *fsidnumber* failed.

Response

Specify the same configuration as the file system configuration previously created where the **-o force** option is set, or specify the representative partition of the file system configured already where the **-m** option is set.

---

sfcmkfs:ERROR: *special* is not Master Partition

Explanation

special is not the representative partition.

Response

Specify the representative partition of the file system created already.

---

sfcmkfs:ERROR: Global Memory Broken

Explanation

An error was detected in the internal processing.

Response

Collect the diagnostic data with FJQSS, **fjsnap**, or **pclsnap**, and contact field engineers submitting it along with text of the error message displayed.

---

sfcmkfs:ERROR: cluster error

Explanation

An error was detected in the internal processing with the cluster control facility.

Response

Collect the diagnostic data with FJQSS, **fjsnap**, or **pclsnap**, and contact field engineers submitting it along with text of the error message displayed.

---

sfcmkfs:ERROR: The specified device does not exist.

Explanation

The specified special file is not found.

Response

Check whether the special file exists on any node of the file system.

---

sfcmkfs:ERROR: partition [*special*] is small, can't acquire data area

Explanation

Securing the file data area failed because the size of the partition *special* is small.

Response

Specify the bigger partition for *special* or adjust such parameters as **metasz** and **logsz** for smaller size in order to secure the file data area, and re-execute the command.

---

sfcmkfs:ERROR: cannot read the FsRM configuration file: errno = *value*

Explanation

Obtaining the FsRM management information failed.

Response

Collect the diagnostic data with FJQSS, **fjsnap**, or **pclsnap**, and contact field engineers submitting it along with text of the error message displayed.

---

sfcmkfs:ERROR: Can not get host informations

Explanation

Acquisition of host information of the specified node failed.
Possible causes:

- The node specified is invalid.

- The network setting of the specified node is invalid.

- The specified node is stopped.

- sfcfrmd daemon is not running on the specified nodes.

- A cluster partition error exists.

Response

Execute **sfcmkfs(8)** again after correcting the cause of the failure according to the following procedures.

1. Check whether the node specified is correct. If a mistake is found, specify the correct host name.

2. On the specified node, check whether the IP address of the specified node is correctly registered in host database (such as /etc/ hosts). For details, see "11.2.2 Registering information in host database" in this manual.

3. Confirm that the specified node is not stopped. If the node was stopped, start it.

4. Execute the following operations on all the specified nodes.

    a. Check if the sfcfrmd daemon is operating by **ps(1)** command.
       An example is shown below.

```
# /bin/ps -e | /bin/grep sfcfrmd <Enter>
27198 ?        00:00:00 sfcfrmd
# echo $? <Enter>
0
```

    If **grep(1)**'s return code is 0, sfcfrmd daemon is running.

    ### 🗃 See
    ....
    For details of **ps(1)** and **grep(1)**, see the online manual page.
    ....

    b. If the sfcfrmd daemon is not running, confirm that the following message is not output to **/var/log/messages**.

```
WARNING: sfcfsrm:5001: Starting the sfcfrmd daemon was suspended
because quorum does not exist
```

If the message is not output, execute **sfcfrmstart(8)** command.

5. When a problem can not be resolved using operations from 1 to 4, a cluster partition error may have occurred. Correct the cluster partition error according to "C.3.4 Corrective action when the sfcfrmd daemon is not started."

sfcmkfs:ERROR: cannot update the management partition: errno = *value*

Explanation

Registration to the file system management information failed.

Response

Collect the diagnostic data with FJQSS, **fjsnap**, or **pclsnap**, and contact field engineers submitting it along with text of the error message displayed.

sfcmkfs:ERROR: Unknown error : *value*

Explanation

An unexpected error was detected in the internal processing.

Response

Collect the diagnostic data with FJQSS, **fjsnap**, or **pclsnap**, and contact field engineers submitting it along with text of the error message displayed.

sfcmkfs:ERROR: Number of partition is over *count* in maxvol : *special*

Explanation

The number of partitions specified exceeds the upper limit **count** specified with the maxvol option.

Response

Increase the value of the **maxvol** option or decrease the number of partitions configuring the file system by grouping the partitions.

sfcmkfs:ERROR: Same device is specified more than once

Explanation

A partition same as the file data partition specified with **-o data** or same as a representative partition specified with an operand is being used.

Response

Specify a file data partition with **-o data** only once, or specify a representative partition with an operand different than that specified with **-o data**.

sfcmkfs:ERROR: total file system size is too large

Explanation

The total of specified partition sizes exceeds the maximum file system size, so a file system cannot be created.

Response

Make sure that the total of partition sizes of the file data partition specified with **-o data** and the representative partition specified with operand is less than the maximum file system size.

sfcmkfs:ERROR: fsid no match. *special1* fsid *value1*, *special2* fsid *value2*

Explanation

The configuration of the specified partition is different than that of the current partition, so a file system cannot be recreated.

Response

Check the configuration of the current partition using the "**sfcmkfs -m**" command. To make the configuration of the specified partition same as that of the displayed current partition by executing **sfcmkfs(8)** with the "**-o data**" option and correct operand.

---

sfcmkfs:WARNING: *value* : bogus logsz reset to system set value

Explanation

The parameter value, **value**, specified by the **logsz** option lies outside the valid range, and the value is automatically calculated to create the file system.

Response

If there is a problem, specify a correct parameter value, and re-execute the command.

---

sfcmkfs:WARNING: *value* : bogus maxnode reset to 16

Explanation

The **value** of the **maxnode** option is out of the valid range, so the default value 16 will be used to create a file system.

Response

If there is a problem, specify a correct parameter value, and re-execute the command.

---

sfcmkfs:WARNING: *value* : bogus metasz reset

Explanation

The value **value** specified for the **metasz** option is outside the valid range. The file system is created using a calculated size.

Response

If there is a problem, specify a correct parameter value, and re-execute the command.

---

sfcmkfs:WARNING: *value* : bogus mminfree reset to 10%

Explanation

The value **value** specified for the **mfree** option is outside the valid range. The file system is created using the default value (10%).

Response

If there is a problem, specify a correct parameter value, and re-execute the command.

---

sfcmkfs:WARNING: *value* : bogus nblkpext reset to 1

Explanation

The value **value** specified for the **nblkpext** option is outside the valid range. The file system is created using the default value (1).

Response

If there is a problem, specify a correct parameter value, and re-execute the command.

---

sfcmkfs:WARNING: *value* : bogus nbpi reset to 8192

Explanation

The value **value** specified for the **nbpi** option is outside the valid range. The file system is created using the default value (8192).

Response

If there is a problem, specify a correct parameter value, and re-execute the command.

---

sfcmkfs:WARNING: *value* : logsize of partition is over : Retry Auto size

Explanation

The log area of the size (*value*) specified by the **logsz** option cannot be secured because the capacity of the partition is insufficient. The log area size is automatically calculated to create the file system.

Response

When this action comes to an issue, specify a proper parameter value and re-execute the command.

---

sfcmkfs:WARNING: *value* : metasize of partition is over : Retry Auto size

Explanation

The meta-data area of the size (*value*) specified by the **metasz** option cannot be secured because the capacity of the partition is insufficient. The meta-data area size is automatically calculated to create the file system.

Response

When this action comes to an issue, specify a proper parameter value and re-execute the command.

---

sfcmkfs:WARNING: *value* : bogus free reset to 10%

Explanation

The parameter value, **value**, specified by the **free** option lies outside the range, and the default value 10% is applied to create the file system.

Response

When this action comes to an issue, specify a proper parameter value and re-execute the command.

---

sfcmkfs:WARNING: *value* : bogus maxdsz reset

Explanation

As the specified value **value** of the **maxdsz** option is out of the valid range or is small for the specified file system configuration, the file system is created, assuming that the maxdsz option is omitted.

Response

The meta-data area size and update log area size of the created file system is calculated from the file system size of the specified partition. No corrective action should be taken if this configuration is used.

If it is assumed to add file data partitions to the file system in the future, delete the file system using **sfcadm(8)**.

Then, specify the maximum total size of the file data area in case of expanding area by adding file data partitions in the "**-o maxdsz**" of **sfcmkfs(8)** and execute the command again.

---

sfcmkfs:WARNING: *value* : bogus maxvol reset to 16

Explanation

As the specified value **value** of the **maxvol** option is out of the valid range, the default value 16 will be used to create a file system.

Response

The maximum number of partitions of the file system is 16. No corrective action should be taken if this configuration is used.

If it is assumed that the number of partitions of the file system exceeds 16 in the future, delete the file system using **sfcadm(8)**.

Then, specify the maximum total number of partitions in case of expanding area by adding file data partitions in the "**-o maxvol**" of **sfcmkfs(8)** and execute the command again.

A.4.4 sfcmount command

ERROR: sfcmount: 0001: Permission denied

Explanation

Execution was attempted with other than super-user permission.

Response

Retry execution with super-user permission.

ERROR: sfcmount: 0002: Cannot open fstab file

Explanation

/etc/fstab could not be opened.

Response

Check if /etc/fstab exists. If it does not, restore the file from backed up data, or define the file again according to "11.2.4 Setting fstab."

ERROR: sfcmount: 0003: Both *special* and *mount_point* need to be specified in fstab

Explanation

Special file *special*, or mount point *mount_point* is not described in /etc/fstab.

Response

Check the status of /etc/fstab.

ERROR: sfcmount: 0004: { *special* | *mount_point* }: Cannot connect to cluster foundation

Explanation

Connection to the cluster control facility failed.

Response

Confirm that the cluster control facility has started.

ERROR: sfcmount: 0005: { *special* | *mount_point* }: Cannot allocate memory

Explanation

Memory acquisition failed. System memory is insufficient.

Response

Check the system status then wait for the other process to be completed or expand the swap area or actual memory.

ERROR: sfcmount: 0006: { *special* | *mount_point* }: signal(2) is error: *errmsg*

Explanation

Signal handling failed.

Response

Collect a file output with FJQSS, **fjsnap**, or **pclsnap** for troubleshooting information then contact field engineers.

ERROR: sfcmount: 0007: { *special* | *mount_point* }: Cannot get lock: config

Explanation

The lock to change the file system (*mount_point*) configuration information cannot be obtained. In an optional process within the cluster configuration, the lock to change the file system configuration information is being obtained.

Response

Execution must be performed after waiting for the lock to change the domain configuration information to be released.

ERROR: sfcmount: 0008: { *special* | *mount_point* }: Read error on management partition to get fsid

Explanation

Reading of the management partition to obtain the file system (*mount_point*) id has failed.

Response

After confirming the state of sfcfrmd by using the like of the ps(1) command (for example, "**/bin/ps -e | grep sfcfrmd**".), contact field engineers.

ERROR: sfcmount: 0010: { *special* | *mount_point* }: Read error on management partition to get node information

Explanation

Reading of the management partition to obtain the cluster configuration node information has failed.

Response

After confirming the state of sfcfrmd by using the like of the ps(1) command (for example, "**/bin/ps -e | grep sfcfrmd**".), contact field engineers.

ERROR: sfcmount: 0011: { *special* | *mount_point* }: Node information for this node was not found

Explanation

The node is not set as a member of sharing the GFS Shared File System.

Response

Check node information by executing the **sfcinfo(8)** with the "-an" option. If the host name is missing, add it by executing **sfcnode(8)**. If the host name is incorrect, correct it by executing **sfcadm(8)**.

ERROR: sfcmount: 0012: Cannot mount *special*

Explanation

Mount processing for *special* failed.

Response

Check the status of the specified special file and mount point, and check whether the mount option described by /etc/fstab is correct. Then retry mount processing.

ERROR: sfcmount: 0013: { *special* | *mount_point* }: sfcfrmd may not exist

Explanation

A communications channel with sfcfrmd could not be established.

Response

After confirming the state of sfcfrmd by using the like of the ps(1) command (for example, "**/bin/ps -e | grep sfcfrmd**".), contact field engineers.

ERROR: sfcmount: 0015: { *special* | *mount_point* }: No such file or directory

Explanation

Special file *special*, or mount point *mount_point* does not exist.

Response

Check whether specified *special* file and *mount_point* exists on any node.

ERROR: sfcmount: 0016: { *special* | *mount_point* }: stat(2) is error: *errmsg*

Explanation

Obtaining the status of the special file *special*, or mount point *mount_point* failed.

Response

Take action according to detailed information indicated by *errmsg*. For details, see the description of stat(2).

ERROR: sfcmount: 0017: *special*: Character device can not be specified as mount point. Please specify block device

Explanation

The character special file was specified.

Response

Specify the block special file.

ERROR: sfcmount: 0018: *special* is already mounted, *mount_point* is busy, or the allowable number or mount points has been exceeded

Explanation

Special file *special* may have already been mounted, or the mount point *mount_point* may be in use.

Response

Check if some process is using *mount_point* by executing fuser(1), and if it is, stop the process to use the mount point. If the special file is already mounted, unmount it.

ERROR: sfcmount: 0019: { *special* | *mount_point* }: Mount demand processing to sfcfrmd went wrong

Explanation

Mount processing request to sfcfrmd has failed.

Response

After confirming the state of sfcfrmd by using the like of the ps(1) command (for example, "**/bin/ps -e | grep sfcfrmd**".), contact field engineers.

ERROR: sfcmount: 0020: { *special* | *mount_point* }: Time out occurred in mount demand processing to sfcfrmd

Explanation

Timeout occurred for a mount processing request to sfcfrmd.

Response

After confirming the state of sfcfrmd by using the like of the ps(1) command (for example, "**/bin/ps -e | grep sfcfrmd**".), contact field engineers.

ERROR: sfcmount: 0021: *special* is not block device

Explanation

The specified *special* is not a block special file.

Response

Check whether specified special file and described special file in /etc/fstab is a block special file.

ERROR: sfcmount: 0022: *mount_point* is not directory

Explanation

The specified *mount_point* is not a directory.

Response

Check whether specified mount point and described mount point in /etc/fstab is a directory.

ERROR: sfcmount: 0023: fstype is described as an invalid value in fstab

Explanation

An unexpected error was detected in the fstype field in /etc/fstab.

Response

Check the contents of /etc/fstab.

ERROR: sfcmount: 0024: fstype is described not as sfcfs but as *fstype* in fstab

Explanation

The specified file system is not the GFS Shared File System.

Response

Check the contents of /etc/fstab and specify the GFS Shared File System.

ERROR: sfcmount: 0025: The description form of fstab is wrong

Explanation

The description format of /etc/fstab is incorrect.

Response

Confirm the contents of /etc/fstab.

ERROR: sfcmount: 0026: { *special* | *mount_point* }: filesystem may not be created

Explanation

The specified file system was not registered in the management partition.

Response

After confirming that the file system has already been created, re-execute the mount command.

ERROR: sfcmount: 0029: { *special* | *mount_point* }: self-node is already processing mount or umount

Explanation

In the self-node, mount or umount processing has already been performed.

Response

When mounting or unmounting of the file system is completed, execute the command again.

ERROR: sfcmount: 0030: { *special* | *mount_point* }: mount_point need to be representative partition

Explanation

The block special file specified as a mount point was not a representation partition.

Response

Please specify a representation partition name and execute a command again.

ERROR: sfcmount: 0031: *special* not in fstab

Explanation

Special file *special* is not described in /etc/fstab.

Response

Define /etc/fstab or specify a correct argument.

WARNING: sfcmount: 0027: { *special | mount_point* **}: mount_option** *mntopts* **is ignored, used** *vfs_mntopts* **in fstab**

Explanation

Although mount option *mntopts* has been specified as the mount command parameter, mount processing was performed with mount option *vfs_mntopts*, described by /etc/fstab.

Response

If you want to mount the file system using a mount option, set the mount point in the file system definition of /etc/fstab.

WARNING: sfcmount: 0028: { *special | mount_point* **}: mount_option** *mntopts* **is ignored**

Explanation

Although mount option *mntopts* has been adopted as the mount command parameter, it was ignored and mount processing was performed with no parameters.

Response

To specify the mount option and perform mount processing, describe the mount option for the mount point in question in /etc/fstab.

A.4.5 sfcumount command

ERROR: sfcumount: 0001: Permission denied

Explanation

Execution was attempted with other than super-user permission.

Response

Retry execution with super-user permission.

ERROR: sfcumount: 0002: Cannot open mtab file

Explanation

Opening of the /etc/mtab file failed. The file may be lost.

Response

Recover /etc/mtab then execute the command again.

ERROR: sfcumount: 0003: Both *special* **and** *mount_point* **need to be specified in mtab**

Explanation

A special file *special* and mount point *mount_point* were not registered in /etc/mtab.

Response

Check if the values specified are correct.

ERROR: sfcumount: 0004: { *special | mount_point* **} not in mtab**

Explanation

A special file *special* and mount *point mount_point* were not registered in /etc/mtab.

Response

Check the status of /etc/mtab.

ERROR: sfcumount: 0005: { *special | mount_point* **}: Cannot allocate memory**

Explanation

Memory acquisition failed. System memory is insufficient.

Response

Check the system status then wait for the other process to be completed or expand the swap area or actual memory.

ERROR: sfcumount: 0006: { *special* | *mount_point* }: Cannot connect to cluster foundation

Explanation

Connection to the cluster control facility failed.

Response

Confirm that the cluster control facility has already been started.

ERROR: sfcumount: 0007: { *special* | *mount_point* }: Read error on management partition to get fsid

Explanation

Reading of the management partition to obtain the file system ID failed.

Response

After confirming the state of sfcfrmd by using the like of the ps(1) command (for example, "**/bin/ps -e | grep sfcfrmd**".), contact field engineers.

ERROR: sfcumount: 0009: { *special* | *mount_point* }: Read error on management partition to get node information

Explanation

Reading of the management partition to obtain the cluster configuration node information has failed.

Response

After confirming the state of sfcfrmd by using the like of the ps(1) command (for example, "**/bin/ps -e | grep sfcfrmd**".), contact field engineers.

ERROR: sfcumount: 0010: { *special* | *mount_point* }: Node information for this node was not found

Explanation

Although the management partition containing the cluster node information was read successfully, no self-node information was found to exist.

Response

Check the cluster configuration, and then contact field engineers.

ERROR: sfcumount: 0011: { *special* | *mount_point* }: sfcfrmd may not exist

Explanation

A communication path to sfcfrmd could not be established.

Response

After confirming the state of sfcfrmd by using the like of the ps(1) command (for example, "**/bin/ps -e | grep sfcfrmd**".), contact field engineers.

ERROR: sfcumount: 0012: *special* not mounted

Explanation

Special had already been unmounted.

Response

Specify a file system that is being mounted.

ERROR: sfcumount: 0013: *special*: mount is in progress

Explanation

Mount processing had already been performed to *special.*

Response

When the mount process is completed, unmount the file system again.

ERROR: sfcumount: 0015: *special*: umount is in progress

Explanation

Umount processing had already been performed to *special.*

Response

After the **sfcumount(8)** is completed, check if the file system is unmounted.

ERROR: sfcumount: 0016: { *special* | *mount_point* } busy

Explanation

Unmount processing for special file *special* or mountpoint *mount_point* failed.

Response

In the command execution node, try to find a process that is using the file system by fuser(1) command (for example, "fuser -m *mount_point*"). If such a process exists, kill the process and unmount the file system again. Otherwise, collect the diagnostic data with FJQSS, **fjsnap**, or **pclsnap**, and collect the kernel dump by push the NMI button and contact field engineers.

ERROR: sfcumount: 0017: *special*: Umount demand processing to sfcfrmd went wrong

Explanation

The unmount processing request for sfcfrmd failed.

Response

After confirming the state of sfcfrmd by using the like of the ps(1) command (for example, "**/bin/ps -e | grep sfcfrmd**".), contact field engineers.

ERROR: sfcumount: 0018: *special*: Time out occurred in umount demand processing to sfcfrmd

Explanation

A timeout occurred for the sfcfrmd unmount processing request.

Response

After confirming the state of sfcfrmd by using the like of the ps(1) command (for example, "**/bin/ps -e | grep sfcfrmd**".), contact field engineers.

ERROR: sfcumount: 0019: { *special* | *mount_point* }: No such file or directory

Explanation

Special file *special*, or mount point *mount_point* does not exist.

Response

Check whether specified *special* file and *mount_point* exists on any node.

ERROR: sfcumount: 0020: { *special* | *mount_point* }: stat(2) is error: *errmsg*

Explanation

Obtaining the status of the special file *special*, or mount point *mount_point* failed.

Response

Take action according to detailed information indicated by *errmsg*. For details, see the description of stat(2).

ERROR: sfcumount: 0021: *special*: Character device can not be specified as mount point. Please specify block device

Explanation

The character special file was specified.

Response

Specify the block special file.

ERROR: sfcumount: 0022: *special* is not block device

Explanation

As the argument of the umount command, a block special file was not specified.

Response

Specify a block special file.

ERROR: sfcumount: 0023: *mount_point* is not directory

Explanation

The specified **mount_point** is not a directory.

Response

Change the mount point that is defined in /etc/fstab to the directory path name.

ERROR: sfcumount: 0024: fstype is described as an invalid value in { fstab | mtab }

Explanation

An unexpected error was detected in the fstype field in /etc/fstab or /etc/mtab.

Response

Recover /etc/fstab or /etc/mtab then execute the command again.

ERROR: sfcumount: 0025: fstype is described not as sfcfs but as *fstype* in { fstab | mtab }

Explanation

The specified file system is not the GFS Shared File System.

Response

Check /etc/fstab and /etc/mtab then specify the correct mount point or block special file of the GFS Shared File System for the argument then execute the command again.

ERROR: sfcumount: 0027: { *special* | *mount_point* }: self-node is already processing mount or umount

Explanation

In the self-node, mount or umount processing has already been performed.

Response

Please set time a little and execute a command again.

ERROR: sfcumount: 0028: { *special* **|** *mount_point* **}: mount_point need to be representative partition**

Explanation

The block special file specified as a mount point was not a representation partition.

Response

Please specify a representation partition name and execute a command again.

A.5 GFS Shared File System Specific Management Commands' Messages

This section describes messages that can be output when GFS Shared File System specific management commands are used.

A.5.1 libsfcfs library

libsfcfs: ERROR: can not get nodeid by the clgetnode command
ERROR:libsfcfs:0001: can not get nodeid by the clgetnode command

Explanation

It failed in node ID acquisition by the cluster resource management facility.

Response

There is a possibility that the cluster resource management facility operated abnormally. Please execute clgettree(1), and confirm that cluster resource management facility works normally. Please follow the action of the message of clgettree(1) when clgettree(1) terminates abnormally. Please re-execute the command that outputs this message.

 See
. . . .
For a corrective action of the message of clgettree(1), see "PRIMECLUSTER Messages."
. . . .

A.5.2 sfcadd command

sfcadd: ERROR: 0001: not super user

Explanation

Execution was attempted with other than super-user permission.

Response

Retry execution with super-user permission.

sfcadd: ERROR: 0002: cannot allocate memory

Explanation

Memory acquisition failed. System memory is insufficient.

Response

Check the system status then wait for the other process to be completed or expand the swap area or actual memory.

sfcadd: ERROR: 0003: the number of partitions exceeded the maximum number

Explanation

Adding a partition failed because the number of partitions specified exceeds the upper limit. The upper limit of the number of partitions that can configure the file system is the value specified with the maxvol parameter of **sfcmkfs(8)**. The upper limit cannot be changed.

Response

Specify the upper limit or less and execute the command again.

---

sfcadd: ERROR: 0004: same device is specified more than once

Explanation

The same partitions exists in the partition specified.

Response

Check if the specified partition is correct.

---

sfcadd: ERROR: 0005: cannot connect to FsRM

Explanation

The file system monitoring mechanism is not operating.

Response

Confirming the state of sfcfrmd by using the likes of the ps(1) command. If it does not, activate the file system monitoring facility by executing the **sfcfrmstart(8)**.

---

sfcadd: ERROR: 0006: unknown error(*errnum*)

Explanation

An error is detected.

Response

Check if the DNS settings are correct. If they are, collect troubleshooting information with FJQSS, **fjsnap**, or **pclsnap** the contact field engineers.

---

sfcadd: ERROR: 0007: cannot acquire lock: *name*(*errnum*)

Explanation

Acquisition of the file system lock failed. The other command has acquired the lock.

Response

Wait until the management command of the GFS Shared File System to be completed on all the cluster nodes then execute the command again.

---

sfcadd: ERROR: 0008: cannot read management partition

Explanation

Reading of the management partition failed.

Response

Check if the management partition is set using the **sfcsetup(8)** command.
If it is not, set the management partition using **sfcsetup(8)**. If it is, collect troubleshooting information with FJQSS, **fjsnap**, or **pclsnap** the contact field engineers.

---

sfcadd: ERROR: 0009: *special* is already used for the management partition

Explanation

The special file *special* cannot be used because it is being used as a management partition.

Response

Specify the other partition and execute the command again.

sfcadd: ERROR: 0010: file system configuration information not found

Explanation

Acquisition of the file configuration failed.

Response

Collect a file output with FJQSS, **fjsnap**, or **pclsnap** for troubleshooting information then contact field engineers.

sfcadd: ERROR: 0011: *special* is not found in file system configuration table

Explanation

The *special* file special does not exist in the management partition.

Response

Check whether the specified partition is correct.

sfcadd: ERROR: 0012: fsid(*num*) is invalid

Explanation

The identifier number *num* of the specified file system is not registered.

Response

Collect a file output with FJQSS, **fjsnap**, or **pclsnap** for troubleshooting information then contact field engineers.

sfcadd: ERROR: 0013: *special* is illegal device

Explanation

The special file *special* specified is nonexistent.

Response

Check if the special file *special* exists.

sfcadd: ERROR: 0014: *special* is not block special device

Explanation

The specified *special* is not a block special file.

Response

Specify the block special file.

sfcadd: ERROR: 0015: *special* already used as sfcfs file system

Explanation

The partition *special* is already being used for the GFS Shared File System.

Response

Specify a partition that is not used for the GFS Shared File System.

sfcadd: ERROR: 0016: shared nodes of *special* is inconsistent

Explanation

A shared node of the specified file system is not included in the scope of the disk class to which GDS logical volume of *special* belongs.

Response

Specify GDS logical volume that belongs to the disk class that includes all shared nodes of the specified file system in the scope to *special*.

sfcadd: ERROR: 0017: *special* is mounted, can't sfcadd

Explanation

The special file *special* is being mounted, so the **sfcadd(8)** cannot be executed.

Response

Unmount the special file special then execute the command again.

sfcadd: ERROR: 0018: *special* is mounted from other node

Explanation

The special file *special* is being mounted on the other node, so the **sfcadd(8)** cannot be executed.

Response

Unmount the special file special then execute the command again.

sfcadd: ERROR: 0019: *special* cannot open

Explanation

Opening of the special file *special* failed.

Response

Check if the special file special is correct.

sfcadd: EROOR: 0020: *special* fstat64 error

Explanation

Acquisition of the special file *special* failed.

Response

Check if the special file special is correct.

sfcadd: ERROR: 0022: *special* is already used for other file system

Explanation

The additional partition *special* is described in /etc/fstab.

Response

Check the "/etc/fstab" file. If the partition *special* is added, delete *special* from /etc/fstab.

sfcadd: ERROR: 0023: /etc/mtab cannot open

Explanation

Opening of /etc/mtab failed.

Response

Recover /etc/mtab then execute the command again.

sfcadd: ERROR: 0024: /etc/fstab cannot open

Explanation

Opening of /etc/mtab failed.

Response

Check if /etc/fstab exists. If it does not, restore the file from backed up data, or define the file again according to "11.2.4 Setting fstab."

sfcadd: ERROR: 0025: cannot read /etc/fstab

Explanation

There is an invalid line in /etc/fstab.

Response

Modify the "/etc/fstab" file.

sfcadd: ERROR: 0026: *special* partition size is over the maximum file system capacity

Explanation

The size of the partition *special* exceeds the maximum file system size.

Response

Set the size of the partition *special* not to exceed the maximum file system size, or specify the other partition.

sfcadd: ERROR: 0027: *special* ioctl{(BLKGETSIZE)|(DKIOCGVTOC)|(DKIOCINFO)} error

Explanation

Acquisition of the device *special* failed.

Response

Check if a disk I/O error is output then take corrective action according to the error message.

sfcadd: ERROR: 0028: *special* is not representative partition

Explanation

Special is not a representative partition.

Response

Specify a representative partition for operand.

sfcadd: ERROR: 0029: *special* cannot read super block

Explanation

Reading of the *special* super block failed.

Response

Recover the file system using **sfcfsck(8)**. If the message "*special* is clean" is output, recover the file system by executing **sfcfsck(8)** with the options -o, f, and nolog. If the file system cannot be recovered due to a disk device failure, recover the failed part then recreate the file system and restore the backed up data.

sfcadd: ERROR: 0030: *special* is not sfcfs file system

Explanation

Special is not of the GFS Shared File System.

Response

Recreate the file system then restore the backed up data.

sfcadd: ERROR: 0031: *special* cannot write super block

Explanation

Writing of the *special* super block failed.

Response

If *special* is the specified additional partition, recover failed parts, or prepare the other partition then specify it to execute the command.

If *special* is a partition that configures the file system, recover the file system by executing **sfcfsck(8)** with the options -o, f, and nolog. If the file system cannot be recovered due to a disk device failure, recover the failed part then recreate the file system and restore the backed up data.

sfcadd: ERROR: 0032: *special* is invalid disk device

Explanation

Special is not of the GFS Shared File System.

Response

Specify a GDS volume.

sfcadd: ERROR: 0033: total file system size is too large

Explanation

If the specified partition is added, the file system size will exceed the maximum file system size, so the partition cannot be added.

Response

Specify a partition where the file system size will not exceed the maximum file system size.

sfcadd: ERROR: 0034: *special* cannot write volume information

Explanation

Writing of the *special* partition configuration failed.

Response

If *special* is the specified additional partition, recover failed parts, or prepare the other partition then specify it to execute the command.

If *special* is a partition that configures the file system, recover the file system by executing **sfcfsck(8)** with the options -o, f, and nolog. If the file system cannot be recovered due to a disk device failure, recover the failed part then recreate the file system and restore the backed up data.

sfcadd: ERROR: 0035: *special* super block restore failed

Explanation

Recovering of the *special* super block failed.

Response

Copy the super block of the representative partition to the partition that configures the file system by executing **sfcadm(8)**. Then, recover the file system by executing **sfcfsck(8)** with the options **-o**, **f**, and **nolog**. If the file system cannot be recovered due to a disk device failure, recover the failed part then recreate the file system and restore the backed up data.

sfcadd: ERROR: 0036: device number of *special* and *special* is the same

Explanation

special and the device number of *special* are the same.

Response

Check if the specified partition is correct.

sfcadd: ERROR: 0037: cannot update meta area

Explanation

Updating the meta-data area failed.

Response

Recover the file system using **sfcfsck(8)**. If the message "*special is* clean" is output, recover the file system by executing **sfcfsck(8)** with the options -o, f, and nolog. If the file system cannot be recovered due to a disk device failure, recover the failed part then recreate the file system and restore the backed up data.

---

sfcadd: ERROR: 0038: meta area restore failed

Explanation

Recovering the meta-data area failed.

Response

Recreate the file system then restore the backed up data.

---

sfcadd: ERROR: 0039: file system (*special*) is corrupted. To Fix: sfcfsck *special*

Explanation

sfcadd(8) cannot be executed because the state of the file system of the partition *special* was not stopped properly.

Response

Recover the file system using **sfcfsck(8)**. If the file system cannot be recovered due to a disk device failure, recover the failed part then recreate the file system and restore the backed up data.

---

sfcadd: ERROR: 0040: *special* volume information restore failed

Explanation

Recovering of the special partition configuration failed.

Response

Copy configuration information of the representative partition to the partition that configures the file system by executing **sfcadm(8)**. Then, recover the file system by executing **sfcfsck(8)** with the options -o, f, and nolog. If the file system cannot be recovered due to a disk device failure, recover the failed part then recreate the file system and restore the backed up data.

---

sfcadd: ERROR: 0041: *special* cannot read volume information

Explanation

Reading of *special* partition configuration information failed.

Response

Recover the file system by executing **sfcfsck(8)** with the options -o, f, and nolog. If the file system cannot be recovered due to a disk device failure, recover the failed part then recreate the file system and restore the backed up data.

---

sfcadd: ERROR: 0042: cannot read /etc/mtab

Explanation

Reading of the /etc/mtab file failed.

Response

Recover the /etc/mtab file then execute the command again.

---

sfcadd: ERROR: 0043: cannot get host informations

Explanation

The sfcfrmd daemon of the node that shares the file system has not been activated.

Response

Check if the sfcfrmd daemon exists by executing the **ps(1)** command on all the nodes that share the file system. The nodes can be checked with **sfcinfo(8)**.

If the nodes are stopped, start them. If the sfcfrmd daemon does not exist, activate the daemon using **sfcfrmstart(8)** on the node.

If it does, collect troubleshooting information using FJQSS, **fjsnap**, or **pclsnap** then contact field engineers.

---

sfcadd: ERROR: 0044: cannot read sfcfsrm.conf

Explanation

Reading of the setup file /var/opt/FJSVsfcfs/sfcfsrm.conf failed.

Response

Recover the system in the event of a disk failure. If a disk failure does not occur, collect troubleshooting information using FJQSS, **fjsnap**, or **pclsnap** then contact field engineers.

---

sfcadd: ERROR: 0045: file system (*special*) is not shared for this node

Explanation

The command is executed on the node that does not share the file system.

Response

Execute the command on the node that shares the file system.

A.5.3   sfcadm and sfcnode commands

---

sfcnode: all node entry delete, can't sfcnode

Explanation

The command cannot be executed because information about all nodes sharing the file system will be deleted.

Response

Reduce the specified host information, and re-execute the command.

{sfcadm | sfcnode}: too many host count

Explanation

Too many host names were specified.

Response

Check whether the number of specified host names is correct, and re-execute the command.

{sfcadm | sfcnode}: too many partition count

Explanation

Too many special files were specified.

Response

Check whether the number of specified partitions is correct, and re-execute the command.

{sfcadm | sfcnode}: argument error

Explanation

The specified command format is wrong.

Response

Check the command specification format, and re-execute the command.

{sfcadm | sfcnode}: cannot allocate memory

Explanation

Memory could not be allocated.

Response

Wait for the termination of another process or increase the size of the swap area, and re-execute the command.

{sfcadm | sfcnode}: cannot get host information

Explanation

Acquisition of the host information of node which is specified or which shares file system failed.
Possible causes:

- The specified node or nodes sharing the file system is stopped.

- sfcfrmd daemon is not running on the specified node or nodes sharing the file system.

- A cluster partition error exists.

Response

Execute **sfcadm(8)** or **sfcnode(8)** again after correcting the cause of the failure according to the following procedures.

1. Confirm that the specified node or nodes sharing the file system is not stopped. If the specified node or nodes sharing the file system was stopped, start it.

2. Execute the following operations on the specified node or nodes sharing the file system.

    a. Check if the sfcfrmd daemon is operating by **ps(1)** command.
    An example is shown below.

    ```
    # /bin/ps -e | /bin/grep sfcfrmd <Enter>
    27198 ?        00:00:00 sfcfrmd
    # echo $? <Enter>
    0
    ```

    If **grep(1)**'s return code is 0, sfcfrmd daemon is running.

    ### 🗊 See
    ●
    For details of **ps(1)** and **grep(1)**, see the online manual page.
    ●

    b. If the sfcfrmd daemon is not running, confirm that the following message is not output to **/var/log/messages**.

    ```
    WARNING: sfcfsrm:5001: Starting the sfcfrmd daemon was suspended
    because quorum does not exist
    ```

    If the message is not output, execute **sfcfrmstart(8)** command.

3. When a problem can not be resolved using operations 1 and 2, a cluster partition error may have occurred. Correct the cluster partition error according to "C.3.4 Corrective action when the sfcfrmd daemon is not started."

{sfcadm | sfcnode}: Cannot connect to FsRM

Explanation

FsRM is not running.

Response

Confirm that FsRM is running.

{sfcadm | sfcnode}: *filename* cannot open

Explanation

Relevant file (*filename*) does not exist.

Response

Confirm if *filename* is available.

Note that, if the management cannot be found, restore it from a backup or create a new one.

{sfcadm | sfcnode}: FSID(*fsid_num*) is invalid

Explanation

File system identification numbers (*fsid_num*) are not registered in the management partition.

Response

Check whether the relevant FSID number has been registered with **sfcinfo(8)**.

{sfcadm | sfcnode}: hostid(*hostid*) not found in volume information

Explanation

The specified host ID (*hostid*) is not a host ID of the node which shares the GFS Shared File System. Otherwise, the node which executed command is not a node which shares the GFS Shared File System.

Response

Specify the host ID of the node which shares the GFS Shared File System. Otherwise, execute the command again in the node which shares the GFS Shared File System.

{sfcadm | sfcnode}: invalid hostname

Explanation

An invalid host name was specified.

Response

Check the specified host name.

sfcnode: hostname(*hostname*) is not found

Explanation

The specified host name (*hostname*) was not found in the node information.

Response

Check whether the specified host name exists in the node information.

{sfcadm | sfcnode}: *hostname* is not shared

Explanation

The specified node (*hostname*) does not share the device.
Possible causes:

- The *hostname* specified is invalid.

- The network setting of the *hostname* is invalid.

- The *hostname* is stopped.

- The sfcfrmd daemon is not running on the *hostname*.

- A cluster partition error exists.

Response

Execute **sfcadm(8)** or **sfcnode(8)** again after correcting the cause of the failure according to the following procedures.

1. Check whether the *hostname* specified is correct. If a mistake is found, specify the correct host name.

2. On the node specified with the *hostname*, check whether the IP address of the *hostname* is correctly registered in host database (such as/etc/hosts). For details, see "11.2.2 Registering information in host database" in this manual.

3. Confirm that the *hostname* is not stopped. If the node was stopped, start it.

4. Execute the following operations on the *hostname*.

    a. Check if the sfcfrmd daemon is operating by **ps(1)** command.
       An example is shown below.

    ```
    # /bin/ps -e | /bin/grep sfcfrmd <Enter>
    27198 ?        00:00:00 sfcfrmd
    # echo $? <Enter>
    0
    ```

    If **grep(1)**'s return code is 0, sfcfrmd daemon is running.

    ![See icon] See
    ・・
    For details of **ps(1)** and **grep(1)**, see the online manual page.
    ・・

    b. If the sfcfrmd daemon is not running, confirm that the following message is not output to **/var/log/messages**.

    ```
    WARNING: sfcfsrm:5001: Starting the sfcfrmd daemon was suspended
    because quorum does not exist
    ```

    If the message is not output, execute **sfcfrmstart(8)** command.

5. When a problem can not be resolved using operations from 1 to 4, a cluster partition error may have occurred. Correct the cluster partition error according to "C.3.4 Corrective action when the sfcfrmd daemon is not started."

sfcnode: *hostname* registered as MDS, can't sfcnode -d

Explanation

The specified host (*hostname*) could not delete, because it is registered as the MDS.

Response

Change the MDS host, and re-execute the command.

{sfcadm | sfcnode}: invalid devicename

Explanation

An invalid partition name was specified.

Response

Check whether the specified partition name is correct.

{sfcadm | sfcnode}: invalid magic word *magic_no*

Explanation

The partition configuration data of the specified partition is invalid.

Response

Check whether the specified partition belongs to a sfcfs file system.

{sfcadm | sfcnode}: ioctl(BLKGETSIZE) error

Explanation

Device information could not be obtained.

Response

Check the system environment to see whether device information is correct.

{sfcadm | sfcnode}: read error (*block_no*)
{sfcadm | sfcnode}: seek error (*block_no*)
{sfcadm | sfcnode}: write error (*block_no*)

Explanation

Special file I/O processing failed.

Response

Check whether the specified special file is correct.

{sfcadm | sfcnode}: log partition size too small

Explanation

The specified log size exceeds the size of the entire log partition.

Response

Check the size of the log partition.

{sfcadm | sfcnode}: major number get error

Explanation

The major number could not be obtained.

Response

Check whether the system environment is correct.

sfcadm: invalid MDS hostname

Explanation

The specification of MDS contains an error.

Response

Confirm the management partition and specified MDS information.

{sfcadm | sfcnode}: node full, can't {sfcadm | sfcnode}

Explanation

The number of nodes sharing the file system exceeded the maximum. When the file system is created by sfcmkfs(8), the maximum value is set by the maxnode option.

Response

If you want to change the maximum values of the number of nodes in which the file system is shared, please execute sfcmkfs(8) again, and create the file system again.

{sfcadm | sfcnode}: file system configuration information is full of entry

Explanation

No more file system information could be added.

Response

Delete unnecessary data with **sfcadm(8)** and re-execute the command.

sfcadm: not enough element of sfcfs partition

Explanation

Not all the partitions belonging to the file system were specified.

Response

Specify all partitions belonging to the file system.

sfcnode: this node entry(hostID:*hostid*) not found

Explanation

Self-node information was not found in the information for nodes sharing the file system.

Response

Check whether self-node information exists.

{sfcadm | sfcnode}: not super user

Explanation

The command cannot be executed without super-user authority.

Response

Execute the command with super-user authority.

{sfcadm | sfcnode}: not support command

Explanation

The specified command is not supported.

Response

Check the specified command name.

{sfcadm | sfcnode}: now locked, can't {sfcadm | sfcnode}

Explanation

The command could not be executed because another command is being executed.

Response

Upon termination of the other command, re-execute the command.

sfcadm: please specify MDS setting

Explanation

A MDS name must be specified.

Response

Specify a MDS name, and re-execute the command.

sfcadm: please specify PORTNAME setting

Explanation

A service port name must be specified.

Response

Specify a service port name, and re-execute the command.

{sfcadm | sfcnode}: same device is specified more than once

Explanation

A partition that is the same as the one specified was found.

Response

Check whether the specified partition name is correct.

---

sfcadm: invalid service portname

Explanation

An invalid service port name was specified.

Response

Check the specified service port name.

{sfcadm | sfcnode}: file system configuration information sync error

Explanation

Distribution to another node sharing the file system failed.

Response

Check that a file system watch mechanism works normally, and re-execute the command.

{sfcadm | sfcnode}: *special* cannot open

Explanation

A special file (*special*) could not be opened.

Response

Check whether the specified special file is correct.

{sfcadm | sfcnode}: *special* fstat64 error

Explanation

The status of a special file (*special*) could not be obtained.

Response

Check whether the specified special file is correct.

{sfcadm | sfcnode}: *special* is illegal device

Explanation

An invalid partition (*special*) that is not part of the file system configuration was specified.

Response

Specify all partitions configuring the file system.

{sfcadm | sfcnode}: *special* is mounted, can't {sfcadm | sfcnode}

Explanation

The command could not be executed because the specified partition (*special*) has already been mounted.

Response

Unmount the file system, and re-execute the command.

{sfcadm | sfcnode}: *special* is not block special device

Explanation

The specified file (*special*) is not a character special file.

Response

Specify a block special file.

{sfcadm | sfcnode}: *special* is not found in file system configuration information

Explanation

The specified partition (*special*) is not included in the file system configuration information.

Response

Reexamine whether the specified partition is correct.

{sfcadm | sfcnode}: *special* is not shared device

Explanation

The specified partition (*special*) is not shared by nodes.

Response

Check whether the system environment is correct.

{sfcadm | sfcnode}: system error(*function_name*)

Explanation

A system error occurred.

Response

Contact field engineers.

{sfcadm | sfcnode}: unknown error(*error_code*)

Explanation

An invalid error was detected.

Response

Contact field engineers.

---

sfcadm: FSID(*fsid_num*) is already exist

Explanation

The file system with the file system ID *fsid_num* has already been created.

Response

Check whether the relevant FSID number has been registered with **sfcinfo(8)**, and re-execute the command.

---

sfcadm: too many option

Explanation

Overfull options were specified.

Response

Make sure that the number of specified options is proper, and re-execute the command.

{sfcadm | sfcnode}: *hostid* is illegal hostid

Explanation

The specified host ID is illegal.

Response

Check whether the specified host ID is correct executing **sfcsetup(8)**.

sfcadm: *option* is illegal option

Explanation

The specified tunable parameter for the file system cannot be registered to the file system configuration information.

Response

Check whether the specified file system's tunable parameter is correct.

{sfcadm | sfcnode}: *filename*: stat(2) is error: *errmsg*

Explanation

The stat(2) operation for filename failed.

Response

Take action according to **errmsg**.

sfcadm: *special* is already used in file system configuration information

Explanation

The specified partition already exists in the file system configuration information.

Response

Check whether the specified partition is proper.

{sfcadm | sfcnode}: file system configuration information not found

Explanation

The file system configuration information cannot be obtained.

Response

See "Chapter 5 Management Partition" and restore the file system configuration information.

{sfcadm | sfcnode}: *special* is not sfcfs file system

Explanation

The specified partition (*special*) is not the GFS Shared File System.

Response

Specify a partition of the GFS Shared File System.

{sfcadm | sfcnode}: ERROR: *special* is not representative partition

Explanation

The partition *special* is not a representative partition.

Response

Specify a representative partition then execute the command.

---

sfcadm: The size of the partition(*special*) might have been changed after the file system was made

Explanation

The size of the partition (*special*) which shares the GFS Shared File System might have been changed.

Response

Perform the procedure (from step 2 on) described on "C.3.2 Corrective action in the event of data inconsistency."

GFS Shared File System can not be expanded by size expansion of the partition which shares the GFS Shared File System.

Perform the procedure above promptly.

---

sfcadm: The size of the partition(*special*) might have been reduced

Explanation

The size of the partition (*special*) currently used for a GFS Shared File System may have been reduced.

Response

The file system cannot be repaired. If a backup of the file system is available, recover the file system from the backup.

A.5.4   sfcfrmstart command

ERROR: sfcfrmstart: 0001: Permission denied

Explanation

Execution was attempted with other than super-user permission.

Response

Retry execution with super-user permission.

ERROR: sfcfrmstart: 0002: Cannot open file *path*: *errmsg*

Explanation

File open processing for command lock failed.

Response

Apply countermeasures as described in *errmsg*.

ERROR: sfcfrmstart: 0003: sfcfrmstart or sfcfrmstop is already running

Explanation

Execution of this command failed as another **sfcfrmstart(8)** or **sfcfrmstop(8)** command was being executed.

Response

Once execution of the other command has terminated, re-execute this command.

ERROR: sfcfrmstart: 0004: Cannot connect to cluster faundation

Explanation

Connection with the cluster control facility failed.

Response

Confirm that the cluster control facility has already been started.

ERROR: sfcfrmstart: 0005: Operation of management partition may be performed

Explanation

The management partition may be being used on either the self-node or another node.

Response

Wait for the operation of the management partition to end, and then retry.

ERROR: sfcfrmstart: 0006: Cannot allocate memory

Explanation

Memory acquisition failed.

Response

Either wait for another process to end, or increase the size of the swap area.

ERROR: sfcfrmstart: 0007: Read error on sfcfsrm.conf: *errmsg*

Explanation

Setting file sfcfsrm.conf could not be read.

Response

Perform management according to *errmsg*.

ERROR: sfcfrmstart: 0008: Read error on management partition

Explanation

Reading of the management partition failed.

Response

If a disk of the management partition fails, restore the management partition by referring to "9.6 Backup of the management partition information" and "9.7 Restoring of the management partition information."
If a disk failure does not occur, field engineers.

ERROR: sfcfrmstart: 0009: Self-node does not belong to host group

Explanation

The self-node does not belong to a host group.

Response

Using the **sfcsetup(8)** command, add the self-node to a host group.

ERROR: sfcfrmstart: 0010: Starting processing of sfcfrmd went wrong

Explanation

Start processing for sfcfrmd failed.

Response

After confirming the state of the self-node and that sfcprmd is available, contact field engineers.

ERROR: sfcfrmstart: 0011: sfcfrmd has already existed

Explanation

The sfcfrmd has already been started.

Response

None.

ERROR: sfcfrmstart: 0012: sfcsetup may not be performed

Explanation

As no setting is made to the management partition, the management partition cannot be accessed.

Response

Use the **sfcsetup(8)** command to set up the management partition. For details, see "9.2 Creating of the management partition."

ERROR: sfcfrmstart: 0013: sfcprmd may not exist

Explanation

The communication path to sfcprmd cannot be established.

Response

After confirming the state of the self-node and that sfcprmd is available, contact field engineers.

ERROR: sfcfrmstart: 0014: State check processing to sfcprmd went wrong

Explanation

State confirmation processing for sfcprmd using several different daemons failed.

Response

After confirming the state of self-node and that sfcprmd is available, contact field engineers.

ERROR: sfcfrmstart: 0015: Write error on management partition

Explanation

Write processing for the management partition failed.

Response

After confirming that no disk error occurred, contact field engineers.

ERROR: sfcfrmstart: 0016: *filename*: open(2) error: *errmsg*

Explanation

File open processing for editing the sfcfsrm.vers file failed.

Response

Apply countermeasures as described in *errmsg*.

ERROR: sfcfrmstart: 0017: *filename*: write(2) error: *errmsg*

Explanation

File write processing for editing the sfcfsrm.vers file failed.

Response

Apply countermeasures as described in *errmsg*.

ERROR: sfcfrmstart: 0018: *filename*: rename(2) error: *errmsg*

Explanation

File rename processing for editing the sfcfsrm.vers file failed.

Response

Apply countermeasures as described in *errmsg*.

ERROR: sfcfrmstart: 0020: An error occurred while accessing a file *path* : *errmsg*

Explanation

File access to cancel the suspension of the sfcfrmd daemon's startup failed.

Response

Check if the sfcfrmd daemon is running by **ps(1)** command.
An example is shown below.

```
# /bin/ps -e | /bin/grep sfcfrmd <Enter>
27198 ?        00:00:00 sfcfrmd
# echo $? <Enter>
0
```

If **grep(1)**'s return code is 0, sfcfrmd daemon is running.

🛈 See

For details of **ps(1)** and **grep(1)**, see the online manual page.

If the sfcfrmd daemon is not running, take corrective action according to *errmsg*. If the error message indicates "broken file", the file to cancel the suspension of the sfcfrmd daemon's startup might be broken. Reboot the system because this file is remade by booting the system.

If sfcfrmd daemon is running, no corrective action needs to be taken because it is thought the sfcfrmd daemon had been started before this file is broken.

ERROR: sfcfrmstart: 0021: The suspended sfcfrmd daemon does not exist

Explanation

Startup of the sfcfrmd daemon has not been suspended.

Response

If the sfcfrmd daemon that is not being suspended is activated, execute the **sfcfrmstart(8)** command without the **-f** option.

A.5.5   sfcfrmstop command

ERROR: sfcfrmstop: 0001: Permission denied

Explanation

Execution was attempted with other than super-user permission.

Response

Retry execution with super-user permission.

ERROR: sfcfrmstop: 0002: Cannot open file *path*: *errmsg*

Explanation

File open processing for command lock failed.

Response

Apply countermeasures as described in *errmsg*.

ERROR: sfcfrmstop: 0003: sfcfrmstart or sfcfrmstop is already running

Explanation

Execution of this command failed as another **sfcfrmstart(8)** or **sfcfrmstop(8)** command was being executed.

Response

Once execution of the other command has terminated, re-execute this command.

ERROR: sfcfrmstop: 0004: Cannot allocate memory

Explanation

Memory acquisition failed.

Response

Either wait for another process to end, or increase the size of the swap area.

ERROR: sfcfrmstop: 0005: sfcprmd may not exist

Explanation

The communication path to the process monitoring daemon sfcprmd cannot be established.

Response

After confirming the state of the self-node and that sfcprmd is available, contact field engineers.

ERROR: sfcfrmstop: 0006: stop processing of sfcfrmd went wrong

Explanation

Stop processing for sfcfrmd failed.

Response

The GFS may be mounted on the self-node. Perform umount processing, and then re-execute the sfcfrmstop command.

ERROR: sfcfrmstop: 0007: Self-node is still mounted

Explanation

The GFS Shared File System is mounted by the self-node.

Response

Please execute a command again after unmounting all the GFS Shared File Systems of a self-node.

ERROR: sfcfrmstop: 0011: Cannot open sfcfsrm.conf

Explanation

File open processing for the sfcfsrm.conf file failed.

Response

Check the status for the /var/opt/FJSVsfcfs/sfcfsrm.conf file.

ERROR: sfcfrmstop: 0012: No entry of management partition path in sfcfsrm.conf

Explanation

The sfcfsrm.conf file contains no entry that indicates the path to the management partition.

Response

Confirm the contents of the setup file /var/opt/FJSVsfcfs/sfcfsrm.conf.

ERROR: sfcfrmstop: 0013: Read error occured on sfcfsrm.conf

Explanation

Error occurred when reading of the sfcfsrm.conf file.

Response

Check the status of the /var/opt/FJSVsfcfs/sfcfsrm.conf file.

ERROR: sfcfrmstop: 0014: Cannot connect to cluster foundation

Explanation

Connection to the cluster control facility failed.

Response

Confirm that the cluster control facility has started.

INFO: sfcfrmstop: 0004: wait for other node to stop

Explanation

The completion of the stop of sfcfrmd in another node is waited.

Response

None.

INFO: sfcfrmstop: 0005: wait for LEFTCLUSTR elimination

Explanation

A recovery from a LEFTCLUSTER state is waited because there is a node that is in a LEFTCLUSTER state in the cluster system.

Response

None.

A.5.6   sfcgetconf command

ERROR: sfcgetconf: 0001: Permission denied

Explanation

Execution was attempted with other than super-user permission.

Response

Perform execution with super-user permission.

ERROR: sfcgetconf: 0002: Cannot allocate memory

Explanation

Memory allocation failed to execute the command.

Response

Wait for another application to end, and then retry.

ERROR: sfcgetconf: 0003: Invalid `-i' parameter: "*iparam*"

Explanation

The parameter specified for the "**-i**" option is invalid.

Response

Re-execute the command with a valid (numerical) parameter in the "**-i**" option.

ERROR: sfcgetconf: 0004: Invalid parameter

Explanation

The format of the command parameter is invalid.

Response

Re-execute the command with a valid parameter according to the format.

ERROR: sfcgetconf: 0005: *backupfile* is already exist

Explanation

The specified file as a backup (backupfile) has already existed.

Response

Re-execute the command with a file that does not exist.

ERROR: sfcgetconf: 0006: Cannot open *backupfile*: open(2) returned errcode=*errno*

Explanation

The open(2) processing for the backup file (backupfile) failed.

Response

Re-execute the command with the correct name of the backup file after checking the location where the backup file will be created.

ERROR: sfcgetconf: 0007: Cannot fdopen *backupfile*: fdopen(3C) returned errcode=*errno*

Explanation

The fdopen(3) processing for the backup file (backupfile) failed.

Response

Re-execute the command with the correct name of the backup file after checking the location where the backup file will be created.

ERROR: sfcgetconf: 0008: Cannot access to cluster foundation: frm_cl_open() returned errcode=*errno*

Explanation

Connection to the cluster control facility failed.

Response

Confirm that the cluster control facility is operating normally.

ERROR: sfcgetconf: 0009: Cannot get domain wide lock: config

Explanation

Execution of this command failed as another command is currently being executed.

Response

Re-execute this command after the execution of the other command.

ERROR: sfcgetconf: 0010: Cannot open *path_to_sfcfsrm.conf*: fopen(3C) returned errcode=*errno*

Explanation

The attempt to get a backup of the sfcfsrm.conf file according to the **-a** option setting, but sfcfsrm.conf failed in but fopen(3).

Response

Check whether the path indicated by a message is destined for sfcfsrm.conf. If the file does not exist, initialize the management partition.

ERROR: sfcgetconf: 0011: file system configuration information parse error: frm_tabent() returned errcode=*errno*

Explanation

There is invalid data in file system configuration information of the management partition.

Response

Recover the file system configuration by referring to "9.7 Restoring of the management partition information."

ERROR: sfcgetconf: 0014: Cannot open sfcfsrm.conf

Explanation

File open processing for the sfcfsrm.conf file failed.

Response

Check the status for the /var/opt/FJSVsfcfs/sfcfsrm.conf file.

ERROR: sfcgetconf: 0015: No entry of management partition path in sfcfsrm.conf

Explanation

The sfcfsrm.conf file contains no entry that indicates the path to the management partition.

Response

Confirm the contents of the setup file /var/opt/FJSVsfcfs/sfcfsrm.conf.

ERROR: sfcgetconf: 0016: Read error occured on sfcfsrm.conf

Explanation

Error occurred when reading of the sfcfsrm.conf file.

Response

Check the status of the /var/opt/FJSVsfcfs/sfcfsrm.conf file.

NOTICE: sfcgetconf: 0012: fsid=*fsid* is not found

Explanation

The file system configuration information that corresponds to *fsid*, which was specified as the target of the search with the **-i** option, could not be found.

Backup processing of the file system configuration information, that corresponds to the specified *fsid*, was ignored by the command.

Response

By eliminating the target of the *fsid* specification from the command argument specification, message output can be disabled.

NOTICE: sfcgetconf: 0013: dev=*special* is not found

Explanation

The file system configuration information that corresponds to *special*, which was specified as the target of the search with the **-d** option, could not be found.

Backup processing of the file system configuration information that corresponds to the specified *special* was ignored by the command.

Response

By eliminating the target of the *spceial* specification from the command argument specification, message output can be disabled.

A.5.7 sfcinfo command

sfcinfo: ERROR: Cannot allocate memory

Explanation

Failed to allocate memory. Processing cannot be continued.

Response

Wait until other processes are finished, or increase a swap area, then re-execute the command.

sfcinfo: ERROR: Cannot get major number of *drivename*

Explanation

Failed to get major number of **drivename**. Processing cannot be continued.

Response

Check the system environment of device driver.

sfcinfo: ERROR: Cannot open {/etc/mtab|/etc/fstab}: *errmsg*

Explanation

Failed to open /etc/mtab or /etc/fstab file because of **errmsg**. Processing cannot be continued.

Response

Check the system environment file.

sfcinfo: ERROR: cannot read file system configuration information

Explanation

A read error occurred in the management partition.

Response

Collect a file output with FJQSS, **fjsnap**, or **pclsnap** for troubleshooting information then contact field engineers.

sfcinfo: ERROR: file system configuration information not found: *special*

Explanation

Failed to get the file system configuration information in the management partition.

Response

Restore the file system configuration information. For detail, see "Chapter 5 Management Partition."

sfcinfo: ERROR: Invalid device number *device_no*

Explanation

User specified an invalid device number.

Response

Correct device number and re-execute this command.

sfcinfo: ERROR: Unknown error(*error_code*)

Explanation

Failed work with unknown error.

Response

Contact field engineers.

---

sfcinfo: ERROR: Cannot connect to FsRM

Explanation

The file system monitoring mechanism is not operating.

Response

Make sure that the file system monitoring mechanism is running.

---

sfcinfo: ERROR: cannot read file system configuration information: *special*

Explanation

A read error occurred in the management partition.

Response

Collect a file output with FJQSS, **fjsnap**, or **pclsnap** for troubleshooting information then contact field engineers.

---

sfcinfo: ERROR: system error(gethostid)

Explanation

The obtained host ID is illegal.

Response

Collect a file output with FJQSS, **fjsnap**, or **pclsnap** for troubleshooting information then contact field engineers.

---

sfcinfo: WARNING: Cannot open {*special* | *mntpoint*}: *errmsg*

Explanation

Failed to open *special* or *mntpoint* because of *errmsg*. Other file systems are processed continuously.

Response

Check the arguments of command or the system environment file.

If a GDS logical volume is being used and is ACTIVE, the volume information will be displayed. Bring the volume into ACTIVE then execute the command again.

If you are operating the GFS Management View, bring the volume into ACTIVE then execute the command again.

---

sfcinfo: WARNING: Invalid device number *device_no*

Explanation

User specified an invalid device number (*device_no*).

Response

Check if the specified device number (*device_no*) is correct. If it is not, correct it and execute the command again.

---

sfcinfo: WARNING: Invalid magic word 0x*magic* in *special*

Explanation

Configuration data of the specified partition is invalid. Other file systems are processed continuously.

Response

Check the specified partition to determine if it is an sfcfs file system.

---

sfcinfo: WARNING: ioctl(BLKGETSIZE) failed in *special*: *errmsg*

Explanation

Failed to obtain the device data of partition *special*. Other file systems are processed continuously.

Response

Check the system environment to determine if partition *special* is correct.

---

sfcinfo: WARNING: Not a block device(*special*)

Explanation

The specified partition must be a block special file but it is a character special file. Other file systems are processed continuously.

Response

Specify a block special file as a partition.

---

sfcinfo: WARNING: Not a directory(*file*)

Explanation

The specified *file* must be a block special file or a directory. Other file systems are processed continuously.

Response

Specify a block special file or mount point directory.

---

sfcinfo: WARNING: lseek error at *special*: *errmsg*

Explanation

The I/O operation to the special file *special* failed.

Response

Check if the special file is correct.

---

sfcinfo: WARNING: read error in {super block | volume information} of *special*: *errmsg* So try to read alternative block

Explanation

Failed to read normal super block data. Therefore, redundant data is read out.

Response

Delete the target file system where it is unnecessary. For details, see "11.8 Deleting" in this manual. When it is a necessary file system, restore it running a full scan with **sfcfsck(8)**. For details, see "C.3.1 Action for I/O errors" in this manual.

---

sfcinfo: WARNING: read error in alternative {super block | volume information} of *special*: *errmsg*

Explanation

Reading the redundant volume information on *special* failed.

Response

Delete the target file system where it is unnecessary. For details, see "11.8 Deleting" in this manual. When it is a necessary file system, restore it running a full scan with **sfcfsck(8)**. For details, see "C.3.1 Action for I/O errors" in this manual.

A.5.8   sfcmntgl command

ERROR: sfcmntgl: 0001: Permission denied

Explanation

Execution was attempted with other than super-user permission.

Response

Retry execution with super-user permission.

ERROR: sfcmntgl: 0002: Cannot open fstab file

Explanation

/etc/fstab could not be opened.

Response

Check if /etc/fstab exists. If it does not, restore the file from backed up data, or define the file again according to "11.2.4 Setting fstab."

ERROR: sfcmntgl: 0003: Both *special* and *mount_point* need to be specified in fstab

Explanation

Special file *special*, or mount point *mount_point* is not described in /etc/fstab.

Response

Check the status of /etc/fstab.

ERROR: sfcmntgl: 0004: {*special* | *mount_point*} not in fstab

Explanation

The specified *special* or *mount_point* cannot be found in fstab.

Response

Check the status of /etc/fstab.

ERROR: sfcmntgl: 0005: { *special* | *mount_point* }: Cannot allocate memory

Explanation

Memory acquisition failed.

Response

Either wait for another process to end, or increase the size of the swap area.

ERROR: sfcmntgl: 0006: { *special* | *mount_point* }: signal(2) is error: *errmsg*

Explanation

Signal handling failed.

Response

Collect a file output with FJQSS, **fjsnap**, or **pclsnap** for troubleshooting information then contact field engineers.

ERROR: sfcmntgl: 0007: { *special* | *mount_point* }: Cannot connect to cluster foundation

Explanation

Connection to the cluster control facility failed.

Response

Confirm that the cluster control facility has already been started.

ERROR: sfcmntgl: 0008: { *special* | *mount_point* }: Cannot get lock: config

Explanation

The lock to change the file system configuration information cannot be obtained. The lock to change the file system configuration information is now being obtained in an optional process of the node within the cluster configuration.

Response

Release the lock for changing the domain configuration information, and then re-execute.

ERROR: sfcmntgl: 0009: { *special* | *mount_point* }: Read error on management partition to get fsid

Explanation

Reading of the management partition to obtain the file system ID failed.

Response

Confirm, using the likes of the ps command, that sfcfrmd is available.

ERROR: sfcmntgl: 0011: { *special* | *mount_point* }: Read error on management partition to get node information

Explanation

Reading of the management partition to obtain the cluster configuration node information failed.

Response

Confirm, using the likes of the ps command, that sfcfrmd is available.

ERROR: sfcmntgl: 0012: *hostname*: *special* was not able to mount. Other process performed umount or some error occurred. Refer console information on each node for detailed error information.

Explanation

The mount processing for special file *special* failed.

Response

Check the status of special file *special*, and then contact field engineers.

ERROR: sfcmntgl: 0013: { *special* | *mount_point* }: sfcfrmd may not exist

Explanation

A communications channel with sfcfrmd could not be established.

Response

Check whether FsRM is operating, and then contact field engineers.

ERROR: sfcmntgl: 0015: { *special* | *mount_point* }: No such file or directory

Explanation

Special file *special*, or mount point *mount_point* does not exist.

Response

Check whether specified *special* file and *mount_point* exists on any node.

ERROR: sfcmntgl: 0016: { *special* | *mount_point* }: stat(2) is error: *errmsg*

Explanation

Obtaining the status of the special file *special*, or mount point *mount_point* failed.

Response

Take action according to detailed information indicated by *errmsg*. For details, see the description of stat(2).

ERROR: sfcmntgl: 0017: *special*: Character device can not be specified as mount point. Please specify block device

Explanation

The character device was specified.

Response

Specify the block special file.

ERROR: sfcmntgl: 0018: The description form of fstab is wrong

Explanation

The description format of /etc/fstab is incorrect.

Response

Confirm the contents of /etc/fstab.

ERROR: sfcmntgl: 0019: { *special* | *mount_point* }: filesystem may not be created

Explanation

The specified file system { *special* | *mount_point* } was not registered in the management partition.

Response

After confirming that the file system has been created, re-execute the **sfcmntgl(8)** command.

ERROR: sfcmntgl: 0021: *special* is not block device

Explanation

The specified **special** is not a block special file.

Response

Check whether specified special file and described special file in /etc/fstab is a block special file.

ERROR: sfcmntgl: 0022: *mount_point* is not directory

Explanation

The specified *mount_point* is not a directory.

Response

Check whether specified mount point and described mount point in /etc/fstab is a directory.

ERROR: sfcmntgl: 0023: fstype is described as an invalid value in fstab

Explanation

An unexpected error was detected in the fstype field in /etc/fstab.

Response

Check the contents of /etc/fstab.

ERROR: sfcmntgl: 0024: fstype is described not as sfcfs but as *fstype* in fstab

Explanation

The specified file system is not the GFS Shared File System.

Response

Check the contents of /etc/fstab and specify the GFS Shared File System.

ERROR: sfcmntgl: 0028: { *special | mount_point* }: *hostname* is already processing mount or umount

Explanation

In the *hostname*, mount or umount processing has already been performed.

Response

Please set time a little and execute a command again.

ERROR: sfcmntgl: 0029: { *special | mount_point* }: mount_point need to be representative partition

Explanation

The block special file specified as a mount point was not a representation partition.

Response

Please specify a representation partition name and execute a command again.

WARNING: sfcmntgl: 0020: *hostname*: *special* is already mounted, *mount_point* is busy, or the allowable number or mount points has been exceeded

Explanation

Either special file *special* has already been mounted in *hostname*, or mount point *mount_point* is already in use. Mount processing is not performed for *hostname*.

Response

None.

WARNING: sfcmntgl: 0025: *hostname*: { *special | mount_point* }: Send mount demand processing to sfcfrmd went wrong

Explanation

Transmission processing for sfcfrmd of the mount processing request made to *hostname* failed.

Response

After confirming the state of sfcfrmd with the likes of the ps(1) command, contact field engineers.

WARNING: sfcmntgl: 0026: { *special | mount_point* }: Waiting processing for a response mount demand went wrong

Explanation

Reception processing of the mount processing request made to sfcfrmd failed.

Response

After confirming the state of sfcfrmd with the likes of the ps command, contact field engineers.

WARNING: sfcmntgl: 0027: { *special | mount_point* }: Time out occurred in mount demand processing to sfcfrmd

Explanation

A timeout occurred in the reception processing of the mount processing request made to sfcfrmd.

Response

After confirming the state of sfcfrmd with the likes of the ps command, contact field engineers.

A.5.9 sfcsetup command

ERROR: sfcsetup: 0001: Permission denied

Explanation

Execution was attempted with other than super-user permission.

Response

Perform execution with super-user permission.

ERROR: sfcsetup: 0002: Cannot connect to cluster foundation

Explanation

Connection to the cluster control facility failed.

Response

Confirm that the cluster control facility is operating normally.

ERROR: sfcsetup: 0003: sfcfrmd is active within cluster foundation

Explanation

FsRM is operating within a cluster.

Response

Stop FsRM and then re-execute the command.

ERROR: sfcsetup: 0004: Cannot get lock: config

Explanation

Execution of this command failed as another command was being executed.

Response

Once execution of the other command has terminated, re-execute this command.

ERROR: sfcsetup: 0005: Too many arguments

Explanation

Too many arguments have been specified.

Response

Check the command format, and then re-execute.

ERROR: sfcsetup: 0006: A few option

Explanation

Too few arguments have been specified.

Response

Check the command format, and then re-execute.

ERROR: sfcsetup: 0007: Wrong command format

Explanation

The command format is invalid.

Response

Check the command format, and then re-execute.

ERROR: sfcsetup: 0008: Cannot get cipname

Explanation

The CIP host name for the node executing the command could not be acquired.

Response

Check the cluster system settings.

ERROR: sfcsetup: 0009: Bad data exists in management partition

Explanation

The management partition contain invalid data.

Response

Contact field engineers.

ERROR: sfcsetup: 0010: Cannot write to sfcfsrm.conf

Explanation

Writing to the sfcfsrm.conf file failed.

Response

Check the status of the /var/opt/FJSVsfcfs/sfcfsrm.conf file.

ERROR: sfcsetup: 0011: Cannot open management partition

Explanation

The management partition could not be opened.

Response

Check whether the path to the management partition, specified in the /var/opt/FJSVsfcfs/sfcfsrm.conf file, is correct.

ERROR: sfcsetup: 0012: Node information area is full

Explanation

The maximum possible amount of node information has already been stored.

Response

Delete any node information that is no longer necessary, and then re-execute the command.

ERROR: sfcsetup: 0013: Cannot write node information

Explanation

An error occurred when trying to write new node information.

Response

Contact field engineers.

ERROR: sfcsetup: 0015: Invalid argument

Explanation

An argument specified with the command contains an invalid character string.

Response

Check the argument, and then re-execute the command.

ERROR: sfcsetup: 0016: Cannot get lock: master. sfcfrmd may be working now

Explanation

The command cannot be executed, either because another command is being executed, or because sfcfrmd is operating.

Response

Wait for execution of the other command to terminate or, after executing sfcfrmstop on all the nodes, re-execute the command.

ERROR: sfcsetup: 0017: Node(hostid=0x*hostid*) is busy

Explanation

Node (hostid=0x*hostid*) has still file system information in management partition. It is configured as sharing node of a GFS Shared File System. **sfcsetup -d** did not do anything.

Response

Delete node information by following procedures:

1. Execute **sfcfrmstart(8)**.

2. Check sharing node information of all GFS Shared File System by **sfcinfo -an**.

3. Reconfigure sharing node information of GFS Shared File System by **sfcadm -u**, or delete file system information by **sfcadm -D**.

4. Execute **sfcfrmstop(8)**.

5. Retry **sfcsetup -d**.

ERROR: sfcsetup: 0021: Error occurred. errcode=*errcode*

Explanation

An internal error occurred.

Response

Contact field engineers.

ERROR: sfcsetup: 0022: Cannot allocate memory

Explanation

Memory acquisition failed due to a shortage of memory.

Response

Terminate any applications that are not required, and then re-execute the command.

Alternatively, review the system environment.

ERROR: sfcsetup: 0023: Read error of sfcfsrm.conf

Explanation

Reading of sfcfsrm.conf failed.

Response

Check the status of the /var/opt/FJSVsfcfs/sfcfsrm.conf file.

ERROR: sfcsetup: 0026: Cluster function may not be working

Explanation

The cluster system may not be operating normally.

Response

Contact field engineers.

ERROR: sfcsetup: 0027: Input error

Explanation

An internal error occurred.

Response

Collect a file output with FJQSS, **fjsnap**, or **pclsnap** for troubleshooting information then contact field engineers.

ERROR: sfcsetup: 0028: *special* is already used as management partition

Explanation

The block special file *special* is already being used as a management partition.

Response

If you want to initialize the management partition again, execute the command again with the "**-f**" option.

ERROR: sfcsetup: 0029: Management partition read error

Explanation

Reading of the management partition failed.

Response

Check the status of the disk containing the management partition.

ERROR: sfcsetup: 0030: Management partition write error

Explanation

Writing to the management partition failed.

Response

Check the status of the disk containing the management partition.

ERROR: sfcsetup: 0031: Cannot initialize management partition. errcode=*errno*

Explanation

Initialization of the management partition failed.

Response

Check the status of the disk containing the management partition.

ERROR: sfcsetup: 0032: *special* is invalid disk device

Explanation

Either the specified file *(special)* does not exist, or is not a GDS volume.

Response

Check that the specified device is a GDS volume and whether it does, in fact, exist.

ERROR: sfcsetup: 0033: Cannot open fstab: *errmsg*

Explanation

The /etc/fstab file could not be opened.

Response

Check the status of /etc/fstab, and whether the file actually exists.

ERROR: sfcsetup: 0034: *special* is already exists in fstab

Explanation

The file specified *(special)* in the management partition may already be being used by the system.

Response

Either specify a different device in the management partition or, if an entry in /etc/fstab is not needed, delete it and then re-execute the command.

WARNING: sfcsetup: 0035: Illegal line exists in fstab

Explanation

An invalid row is included in /etc/fstab.

Response

Confirm the contents of /etc/fstab and correct it if necessary.

ERROR: sfcsetup: 0036: Cannot open mtab: *errmsg*

Explanation

The /etc/mtab file could not be opened.

Response

Check the status of the /etc/mtab file and whether it does, in fact, exist.

ERROR: sfcsetup: 0037: *special* is already exists in mtab

Explanation

The device specified *(special)* in the management partition may already be being used by the system.

Response

Specify the other special file, or unmount and delete the file then execute the command again.

WARNING: sfcsetup: 0038: Illegal line exists in mtab

Explanation

There is an illegal line in /etc/mtab.

Response

Recover the contents of /etc/mtab correctly then execute the command again.

ERROR: sfcsetup: 0039: Cannot open *special*: *errmsg*

Explanation

The specified special file *(special)* could not be opened.

Response

Check whether the physical connection to the disk is satisfactory.

ERROR: sfcsetup: 0040: Cannot get the disk size for the management partition

Explanation

The size of the disk on which the management partition resides could not be obtained.

Response

Check the status of the disk.

ERROR: sfcsetup: 0041: Cannot get major number

Explanation

The major number of disk containing the management partition could not be obtained.

Response

Check whether the system environment is satisfactory, and whether the specified special file is correct.

ERROR: sfcsetup: 0042: Cannot get driver name

Explanation

The diver name for the disk containing the management partition could not be obtained.

Response

Check whether the system environment is satisfactory, and whether the specified special file is correct.

ERROR: sfcsetup: 0043: Cannot get node information

Explanation

Node information could not be acquired from the management partition.

Response

Check the status and contents of /var/opt/FJSVsfcfs/sfcfsrm.conf.

ERROR: sfcsetup: 0044: Node information does not exist

Explanation

The information for the node for which the command was executed has already been deleted.

Response

Execute the command on the node which node information will be deleted from config information.

ERROR: sfcsetup: 0046: Illegal entry exists ,or no entry of path to management partition in sfcfsrm.conf

Explanation

The sfcfsrm.conf file contains no entry that indicates the path to the management partition. Either node information has not yet been added to the management partition, or it was on a node that has already been deleted.

Response

Check whether node information has not yet been added to the management partition, or whether the node containing the entry has already been deleted.

ERROR: sfcsetup: 0048: Too small disk size, need more than 40MB

Explanation

The size of the device specified in the management partition is too small.

Response

Specify a shared disk with a volume size of 40 megabytes or more.

ERROR: sfcsetup: 0049: Illegal entry exists ,or no entry of path to management partition in sfcfsrm.conf

Explanation

An invalid row is included in setup file sfcfsrm.conf.

Response

An invalid row is included in setup file /var/opt/FJSVsfcfs/sfcfsrm.conf.

ERROR: sfcsetup: 0050: Cannot open sfcfsrm.conf: *message*

Explanation

Due to the reason displayed in the *message*, opening of the FsRM setup file sfcfsrm.conf failed.

Response

Apply the countermeasures described in the message.

ERROR: sfcsetup: 0051: *special* is not management partition

Explanation

The volume of the specified shared disk *(special)* is not initialized for the management partition.

Response

Initialize the volume of the shared disk and then re-execute.

ERROR: sfcsetup: 0052: sfcfsrm.conf does not exist. sfcsetup may not be executed

Explanation

The setup file sfcfsrm.conf cannot be found. Node information may not be registered in the node.

Response

Register node information in the management partition by using the **sfcsetup(8)** command.

ERROR: sfcsetup: 0053: Volume(*special*) may not be active. Please activate volume

Explanation

The shared disk volume *special* specified for a management partition may not be ACTIVE.

Response

Check if the shared disk volume is ACTIVE by executing the "**sdxinfo(8)**" command then execute the command again.

ERROR: sfcsetup: 0054: *special* is not block_device_file

Explanation

The device special is not a block special file.

Response

Check the device name then execute the command again.

INFO: sfcsetup: 0002: Waiting to be able to get lock:master

Explanation

The acquisition of domain-wide lock is in progress.

Response

Wait for a short while.

INFO: sfcsetup: 0003: Waiting to be able to get lock:config

Explanation

The acquisition of domain-wide lock is in progress.

Response

Wait for a short while.

A.5.10 sfcstat command

sfcstat: *special* is not a sfcfs file system

Explanation

The specified partition (special) is not an sfcfs file system partition.

Response

Check whether the specified partition is an sfcfs file system partition.

sfcstat: Can't get information about mounted file systems.

Explanation

Reading /etc/mtab failed.

Response

Check the /etc/mtab settings and the access right to the file.

sfcstat: can't open *mnt_pnt*

Explanation

Opening the mount point mnt_pnt failed.

Response

Check the access right to the mount point. When you have the access right or an I/O error occurs in accessing the mount point, collect the diagnostic data with FJQSS, **fjsnap**, or **pclsnap** and contact field engineers.

sfcstat: missing entry of sfcfs fstotal

Explanation

Statistical information could not be obtained.

Response

Check whether the file system is normal.

sfcstat: no active sfcfs

Explanation

The file system is not mounted.

Response

After mount the file system, re-execute the command.

sfcstat: *special* not mounted

Explanation

special is not mounted.

Response

After mount the file system, re-execute the command.

A.5.11 sfcumntgl command

ERROR: sfcumntgl: 0001: Permission denied

Explanation

Execution was attempted with other than super-user permission.

Response

Retry execution with super-user permission.

ERROR: sfcumntgl: 0002: Cannot open fstab file

Explanation

/etc/fstab could not be opened.

Response

Check the status of /etc/fstab.

ERROR: sfcumntgl: 0003: Both *special* and *mount_point* need to be specified in fstab

Explanation

It needs to be described by /etc/fstab when both the special file *special* and mount point *mount_point* are specified as an operand.

Response

Correct the definition of the "etc/fstab" file system.

ERROR: sfcumntgl: 0004: { *special | mount_point* } not in fstab

Explanation

It needs to be described by /etc/fstab when both the special file *special* and mount point *mount_point* are specified as an operand.

Response

Check the status of /etc/fstab.

ERROR: sfcumntgl: 0005: { *special | mount_point* }: Cannot allocate memory

Explanation

Memory acquisition failed.

Response

Either wait for another process to end, or increase the size of the swap area.

ERROR: sfcumntgl: 0006: { *special | mount_point* }: Cannot connect to cluster foundation

Explanation

Connection with the cluster control facility failed.

Response

Confirm that the cluster control facility has already been started.

ERROR: sfcumntgl: 0007: { *special | mount_point* }: Cannot get lock: config

Explanation

The lock to change the file system configuration information cannot be obtained. The lock to change the file system configuration information is now being obtained in an optional process of the node within the cluster configuration.

Response

Re-execute after releasing the lock for changing the domain configuration information.

ERROR: sfcumntgl: 0008: { *special* | *mount_point* }: Read error on management partition to get fsid

Explanation

Reading of the management partition to obtain the file system ID failed.

Response

Confirm, using the likes of the ps(1) command, that sfcfrmd is available.

ERROR: sfcumntgl: 0010: { *special* | *mount_point* }: Read error on management partition to get node information

Explanation

Reading of the management partition to obtain the cluster configuration node information failed.

Response

Confirm, using the likes of the ps(1) command, that sfcfrmd is available.

ERROR: sfcumntgl: 0011: { *special* | *mount_point* }: sfcfrmd may not exist

Explanation

A communication path to sfcfrmd could not be established.

Response

Confirm, using the likes of the ps(1) command, that sfcfrmd is available.

ERROR: sfcumntgl: 0012: *hostname*: { *special* | *mount_point* } busy

Explanation

Unmount processing for special file *special* or mountpoint *mount_point* failed.

Response

In the command execution node, try to find a process that is using the file system by fuser(1) command (for example, "fuser -mu *special*"). If such a process exists, kill the process and unmount the file system again. Otherwise, collect the diagnostic data with FJQSS, **fjsnap**, or **pclsnap**, and collect the kernel dump by push the NMI button and contact field engineers.

ERROR: sfcumntgl: 0013: The description form of fstab is wrong

Explanation

The description format of /etc/fstab is incorrect.

Response

Confirm the contents of /etc/fstab.

ERROR: sfcumntgl: 0019: { *special* | *mount_point* }: No such file or directory

Explanation

Special file *special*, or mount point *mount_point* does not exist.

Response

Check whether specified *special* file and *mount_point* exists on any node.

ERROR: sfcumntgl: 0020: { *special* | *mount_point* }: stat(2) is error: *errmsg*

Explanation

Obtaining the status of the special file *special*, or mount point *mount_point* failed.

Response

Take action according to detailed information indicated by *errmsg*. For details, see the description of stat(2).

ERROR: sfcumntgl: 0021: *filename*: Character device can not be specified as mount point. Please specify block device

Explanation

Character-type special device name *filename* was specified as the mount point.

Response

Specifying a block special file name and then re-execute.

ERROR: sfcumntgl: 0022: *special* is not block device

Explanation

A block special file was not specified as the argument of the **sfcumntgl(8)** command.

Response

Specify a block special file.

ERROR: sfcumntgl: 0023: *mount_point* is not directory

Explanation

The specified *mount_point* is not a directory.

Response

Check whether specified mount point and described mount point in /etc/fstab is a directory.

ERROR: sfcumntgl: 0024: fstype is described as an invalid value in { fstab | mtab }

Explanation

An unexpected error was detected in the fstype field in /etc/fstab or /etc/mtab.

Response

Check the contents of /etc/fstab or /etc/mtab.

ERROR: sfcumntgl: 0025: fstype is described not as sfcfs but as *fstype* in { fstab | mtab }

Explanation

The specified file system is not the GFS Shared File System.

Response

Check **/etc/fstab** and **/etc/mtab** then specify a correct mountpoint or block special file of the GFS Shared File System.

ERROR: sfcumntgl: 0029: { *special* | *mount_point* }: *self-node* is already processing mount or umount

Explanation

In the self-node, mount or umount processing has already been performed.

Response

Please set time a little and execute a command again.

ERROR: sfcumntgl: 0030: { *special* | *mount_point* }: mount_point need to be representative partition

Explanation

The block special file specified as a mount point was not a representation partition.

Response

Please specify a representation partition name and execute a command again.

WARNING: sfcumntgl: 0015: *hostname*: *special* not mounted

Explanation

special had already been unmounted at *hostname*.

Response

None.

WARNING: sfcumntgl: 0016: *hostname*: *special*: mount is in progress

Explanation

Mount processing had already been performed to *special* at *hostname*.

Response

Either wait a short while and retry the unmount processing.

WARNING: sfcumntgl: 0017: *hostname*: *special*: fail over is in progress

Explanation

Failover processing had already been performed to *special* at *hostname*.

Response

Either wait a short while and retry the unmount processing.

WARNING: sfcumntgl: 0018: *hostname*: *special*: umount is in progress

Explanation

Umount processing had already been performed to *special* at *hostname*.

Response

None.

WARNING: sfcumntgl: 0026: *hostname*: { *special* | *mount_point* }: Send umount demand processing to sfcfrmd went wrong

Explanation

Transmission of the unmount processing request to sfcfrmd failed with *hostname*.

Response

After confirming the state of sfcfrmd with the likes of the ps(1) command, contact field engineers.

WARNING: sfcumntgl: 0027: Waiting processing for a response umount demand went wrong

Explanation

Reception processing of the unmount processing request to sfcfrmd failed.

Response

After confirming the state of sfcfrmd with the likes of the ps(1) command, contact field engineers.

WARNING: sfcumntgl: 0028: Time out occurred in umount demand processing to sfcfrmd

Explanation

A timeout occurred in the reception processing of the unmount processing request made to sfcfrmd.

Response

After confirming the state of sfcfrmd with the likes of the ps command, contact field engineers.

A.5.12 sfcrscinfo command

sfcrscinfo: Cannot allocate communication buffer

Explanation

The initialization processing of communication failed due to insufficient memory or swap space.

Response

Check the system environment whether memory or swap space insufficiency has not occurred.

---

sfcrscinfo: Cannot connect to fsrm

Explanation

The file system monitoring mechanism is not operating.

Response

Check whether the sfcfrmd process exists using the ps(1) command.

When it does not exist, activate the file system monitoring mechanism using the **sfcfrmstart(8)** command.

---

sfcrscinfo: Cannot get device name of fsid(*FSID*)

Explanation

Illegal data exists in the file system configuration information for the file system whose fsid is FSID.

Response

Restore the management partition using its backup previously saved.

---

sfcrscinfo: Cannot open /etc/mtab: *errmsg*

Explanation

Opening /etc/mtab failed.

Response

Check the system environment using errormsg.

---

sfcrscinfo: File system config data of fsid(*FSID*) is broken

Explanation

Illegal data exists in the file system configuration information for the file system whose fsid is FSID.

Response

Correct the file system configuration according to "9.7 Restoring of the management partition information."

---

sfcrscinfo: Fs(*FSID*) config db broken

Explanation

Illegal data exists in the file system configuration information for the file system whose fsid is FSID.

Response

Recover the file system configuration by referring to "9.7 Restoring of the management partition information."

sfcrscinfo: fsid(*FSID*) is out of range

Explanation

The specified fsid value lies outside the valid range.

Response

Specify a correct file system ID then execute the command again.

sfcrscinfo: Fsrm is not active

Explanation

The file system monitoring mechanism is not operating.

Response

Check whether the file system monitoring mechanism is running.

sfcrscinfo: *special* is not a representative partition of GFS shared file system

Explanation

special is not the GFS Shared File System's representative partition.

Response

Specify the representative partition and re-execute the command.

sfcrscinfo: There is not GFS shared file system with fsid(*FSID*)

Explanation

A GFS Shared File System identified by the specified fsid does not exist.

Response

Check specified FSID is correct.

A.6 GFS Management View Messages

This appendix provides information on message dialog associated with GFS Management View, and wherever possible, a recovery procedure is provided to locate and correct the problem.

A.6.1 Error messages

6000: Executing the command was unsuccessful.

Explanation

Executing the command was unsuccessful.

Response

To proceed with operation, click <Yes> button. To cancel click <No> button.

6001: Command line output: (detailed information)

Explanation

The error given in (detailed information) occurred when executing the command.

Response

Identify the cause referring to the command line output. When you cannot identify the cause, contact field engineers.

6002: Abnormality occurred in the log monitoring mechanism on node *node*. Monitoring will stop. Restart Global File Services.

Explanation

Abnormality occurred in the log monitoring mechanism on node *node*.

Response

Restart GFS management View.

6003: Specified meta partition is not found at server side.

Explanation

The specified meta partition could not be found.

Response

Specify the name of a meta partition that can be shared by the specified hosts, and re-execute the command.

6004: Specified data partition is not found at server side.

Explanation

The specified data partition could not be found.

Response

Specify the name of a data partition that can be shared by the specified hosts, and re-execute the command.

6005: Specified log partition is not found at server side.

Explanation

The specified log partition could not be found.

Response

Specify the name of a log partition that can be shared by the specified hosts, and re-execute the command.

6006: Specified File system is not found at server side.

Explanation

The specified file system has been registered in /etc/fstab.

Response

Specify the name of a file system that can be shared by the specified hosts, and re-execute the command.

6007: Specified file system still exists in /etc/fstab.

Explanation

The specified file system was registered in the **/etc/fstab**.

Response

Check whether the specified file system is correct.

6008: The path of the management partition is inconsistent. Redo the setting from the command line.

Explanation

The management partition is inconsistent.

Response

Re-execute the operation after deleting the management partition with the **sfcsetup(8)** command and then re-creating the partition by selecting [Operation]:[Operate Management Partition]:[Create] on the Management View menu.

A.6.2 Warning messages

6200: The file system is active. Cannot change the status.

Explanation

The active file system cannot delete or change.

Response

Unmount the file system, and do again.

6201: Specified file system has error status, so can not modify status.

Explanation

Specified file system has error status.

Response

Investigate the cause of abnormalities by reference of a Log.

6202: There is no enough update log area in the specified configuration. Please do reconfiguration.

Explanation

The update log area is too small.

Response

Change partition configuration to get enough update log area, and click <Next> or <Create> button.

6203: The input mount point is not available. Change the mount point.

Explanation

The input mount point is not available.

Response

Change the mount point, and click <Next> button.

6204: There is no disk shared by the specified node. Please change the selection of the node.

Explanation

There is no disk shared by the specified node.

Response

Please change the selection of the node.

6205: There are no unused disks. so You can not create new file system.

Explanation

There are no unused disks.

Response

None.

6206: A management partition cannot be created because there is no unused volume.

Explanation

You cannot create a management partition because all the disk volumes are being used.

Response

Re-execute the operation after you add a new shared disk and then set up the GDS logical volume. For information on making additions and performing setup, see "Operation" "GDS Management View" in "PRIMECLUSTER Global Disk Services Configuration and Administration Guide."

6207: Since *node* is out of the range of the scope, it cannot be added to the management partition.

Explanation

A node cannot be added because it is out of the scope of the volumes that are set to the management partition.

Response

Register the node in the scope of the management partition volume then execute the command again. For information on how to register a node with the scope, see "Operation" "GDS Management View" in "PRIMECLUSTER Global Disk Services Configuration and Administration Guide."

6209: There is a file system that is being started. You cannot add the management partition nodes.

Explanation

If a GFS Shared File System is being started, you cannot change the configuration of the management partition.

Response

Unmount all GFS Shared File Systems that are being started.

A.6.3 Information messages

6300: Exit GFS management?

Explanation

Exits GFS Management View.

Response

Click <Yes> to exit, <No> to continue operation.

6301: A file system has been created.

Explanation

A file system was created.

Response

Click the <OK> button to close the message.

Start the cluster control facility to make available the file system.

6302: Changing of attribute has been completed. (*detail | mount | scope*)

Explanation

Changing a file system attribute of the details, mount, or scope is completed.

Response

Click the <OK> button to close the message.

6303: Changing configuration was completed.

Explanation

Changing configuration of the file system was completed.

Response

Click the <OK> button to close the message.

6304: Specified file system will be deleted. Data in the file system is lost. Are you sure?

Explanation

When the file system is deleted, data in the file system is lost.

Response

To delete the file system, click the <Yes> button. To cancel the deletion of the file system, click the <No> button.

6305: A file system has been deleted.

Explanation

A file system was deleted.

Response

Click the <OK> button to close the message.

6306: The management partition is not set up. Set it up from the Operate management partition.

Explanation

The management partition is not set up on this node.

Response

Click the <OK> button to close the message.
If necessary, execute either of the following countermeasures:

- If the management partition has not been created:

  Re-execute the operation after you create the management partition by selecting [Operation]:[Operate management partition]: [Create].

- If the management partition has been created on another node:

  Re-execute the operation after adding a node to the management partition by selecting [Operation]:[Operate management partition]: [Create].

6307: sfcfrmd is not operating. Start up sfcfrmd.

Explanation

sfcfrmd is not activated on this node.

Response

Click the <OK> button to close the message.

Re-execute the operation after starting up the sfcfrmd daemon by selecting [Operation]: [Operate sfcfrmd]: [Start].

6308: The management partition is created with the following settings. Do you want to continue?

Explanation

The management partition is being created.

Response

To create the management partition, click the <Yes> button. To abort the creation, click the <No> button.

6309: *node* is going to be added to the following management partition. Do you want to continue?

Explanation

This node is going to be added to the management partition.

Response

To add the node, click the <Yes> button. To abandon the addition, click the <No> button.

6310: The selected volume is offline on *node*. Switch it to online?

Explanation

The selected volume is offline on that node.

Response

To place the node online, click the <Yes> button. To leave the node offline, click the <No> button.

6311: The processing of *node* could not be completed. Continue with the processing?

Explanation

The processing for the stated node could not be completed.

Response

To continue with the processing, click the <Yes> button. To abort the processing, click the <No> button.

6312: The processing is suspended.

Explanation

The processing has been suspended.

Response

Response is unnecessary.
Click the <OK> button to clear the message.

6313: Management partition operation has been completed.{(Create)|(Add Node)}

Explanation

Management partition operation has been completed.

Response

Response is unnecessary.
Click the <OK> button to clear the message.

6314: sfcfrmd start or stop has been completed.{(Start)|(Stop)}

Explanation

sfcfrmd has been started or stopped.

Response

Response is unnecessary.
Click the <OK> button to clear the message.

6315: The selected nodes are added to a management partition. Do you want to continue?

Explanation

The selected nodes are added to a management partition.

Response

Processing will be interrupted, if you click the <No> button. The selected nodes are added to a management partition, if a <Yes> button is clicked.

6316: sfcfrmd is started by the selected nodes. Do you want to continue?

Explanation

sfcfrmd is started by the selected nodes.

Response

Processing will be interrupted, if you click the <No> button. Starting sfcfrmd daemon is started, if you click <Yes> button.

6317: sfcfrmd is stopped by the selected nodes. Do you want to continue?

Explanation

sfcfrmd is stopped by the selected nodes.

Response

Processing will be interrupted, if you click the <No> button. Stopping sfcfrmd daemon is started, if you click <Yes> button.

A.7 Web-Based Admin View Messages

For information about the Web-Based Admin View error messages displayed when GFS Shared File System is active, see "PRIMECLUSTER Web-Based Admin View Operation Guide."

A.8 Messages Of Functions For GFS To Use

For information about the Messages of functions for GFS to use, see "PRIMECLUSTER Installation and Administration Guide," "PRIMECLUSTER Cluster Foundation (CF) Configuration and Administration Guide," "PRIMECLUSTER Global Disk Services Configuration and Administration Guide," and "PRIMECLUSTER Web-Based Admin View Operation Guide."

A.9 Installation Error Messages

For information about the messages output during installation, see "Installation Guide" of PRIMECLUSTER.

Appendix B Reference Manual

This appendix describes management commands of the GFS Shared File System.

B.1 GFS Shared File System Exclusive Management Commands

This section describes the commands which are same function as operation commands of Linux standard file system to operate the GFS Shared File System.

B.1.1 sfcdf(8) Display the GFS Shared File System use status and configuration information

<SYNOPSIS>

sfcdf -v [*device* ...]

sfcdf -m *device*

<DESCRIPTION>

The **sfcdf** command displays detailed information about the use status and the configuration of a GFS Shared File System.

A path name of the block special file can be specified for a *device* to indicate the partition of a file system. When no value is specified for *device*, information on the mounted file system is displayed.

In GFS Shared File System, a part of an empty area in the data area has been reserved for each AC to allocate the file data on the AC side. This reservation area is treated as occupied. Please note to what is reported as occupied that areas more than the file data which were really assigned.

Moreover, so that the allocation of the file data may give priority and use this reservation area, please necessarily note that the area reported while using it doesn't increase by new writing.

The following information can be displayed for file system use statuses.

For the meta-data area, the following information is displayed:

- inodes: Total number of i-nodes

- free: Total number of available i-nodes

- vdata: Total number of V-data blocks

- free: Total number of available V-data blocks

- leaf: Total number of leaf for empty extent management

- free: Total number of available leaf for empty extent management

- Type: META

For the update log area, the following information is displayed:

- kbytes: Size of the update log area (kilobyte)

- Type: LOG

For the file data area, the following information is displayed:

- kbytes: Size of the file data area (kilobyte)

- used: Size of the file data area in use (kilobyte)

- free: Size of the available file data area in use (kilobyte)

- %used: Utilization of the file data area (including the reserved area)

- Type: DATA

<OPTIONS>

The following options can be specified:

-v

Outputs if i-nodes, V-data, and file data area are used, how much the extent is empty, and what size the update log area is.

In case of multi-partitions, the configuration will be displayed by partition.

-m

Displays the super block and the partition configuration data on the partition.

<OPERANDS>

The following operands can be specified:

device

Specifies the path to a path name of the block special file on the partition of the GFS Shared File System. When **-v** option is specified and no value is specified for *device*, information on the mounted file system is displayed.

<EXAMPLES>

Single-partition configuration

```
# sfcdf -v /dev/sfdsk/gfs01/dsk/volume01 <Enter>
Filesystem:/mnt
                 inodes     free     vdata      free      leaf      free    Type
/dev/sfdsk/gfs01/dsk/volume01
                 45156     45151     15052     15034     22584     22582    META
             kbytes
/dev/sfdsk/gfs01/dsk/volume01
                  5113                                                      LOG
             kbytes      used      free     %used

/dev/sfdsk/gfs01/dsk/volume01
                 361240        8    361232        0%                        DATA
```

Multi-partition configuration

```
# sfcdf -v /dev/sfdsk/gfs01/dsk/volume01 <Enter>
Filesystem:/mnt
                 inodes     free     vdata      free      leaf      free    Type
/dev/sfdsk/gfs01/dsk/volume01

                 45156     45151     15052     15034     22584     22582    META
             kbytes
/dev/sfdsk/gfs01/dsk/volume01
                  5113                                                      LOG
             kbytes      used      free     %used
/dev/sfdsk/gfs01/dsk/volume01
                 361240        8    361232        0%                        DATA
             kbytes      used      free     %used
/dev/sfdsk/gfs01/dsk/volume02
                 65504        0     65504        0%                         DATA
```

<FILE>

/etc/mtab

The table of a mounted system file

sfcmkfs(8)

df(1), getmntent(3) and mount(8) in the online manual

B.1.2   sfcfsck(8) Verify and restore GFS Shared File System consistency

<SYNOPSIS>

sfcfsck [*generic_options*] [-o *specific_options*] *device*

<DESCRIPTION>

The **sfcfsck** command checks the GFS Shared File System for consistency and restores the file system.

The GFS Shared File System can perform high-speed recovery by using the update log function. The command replays update logging during recovery of data integrity then determines if quick recovery is enabled. If an update log replay cannot be executed because of a damage or for some other reason, **sfcfsck** executes a full file system check.

For *device*, specify a path name of the block special file where the file system has been configured, or specify the mount point. **sfcfsck** is not executed when a path name of the block special file has been mounted.

If there is a serious inconsistency, the operator is prompted to reply yes or no to repair before **sfcfsck** executes recovery. If the operator does not have write permission for the file system, **sfcfsck** executes the operation with the **-n** option (no corrections are made).

sfcfsck executes automatic recovery for minor inconsistencies such as:

- i-node not referenced

- Too many links in an i-node

- An incorrect block included in a free list

- A block appearing in both a free list and a file

- An invalid value detected in a super block

After the inconsistency is corrected, a message is displayed. When recovery of the file system is successful, **sfcfsck** displays the number of files on the file system, the counts for used blocks and free blocks, and the fragmentation rate.

A full check is executed for the following inconsistencies:

- A block acquired by multiple i-nodes or free lists.

- A block requested by an i-node or free list outside the range of the file system.

- Invalid link count

- Invalid directory size

- i-node with invalid syntax

- A block not referenced from anywhere

- Directory check, a file pointing to an unallocated i-node, i-node number falling outside the range, the first two entries (. and ..) in a directory

- Super block check: i-node blocks exceeding the number of blocks in the file system.

- A free block list with invalid syntax

- An invalid total for the free block count and free i-node count

Independent files and directories (allocated, but not referenced) are placed in a lost+found directory and reunited with the operator's consent. They are assigned i-node numbers as names. If a lost+found directory does not exist, create it. If area for the lost+found directory is not sufficient, expand the directory.

<OPTIONS>

generic_options

The following options can be specified:

-V

Displays a command line, but does not execute the command.

-m

Executes a check, but does not make corrections.

One of the following messages is output at the time of execution.

- **sfcfsck: sanity check:** *device* **okay**

    device is consistent.

- **sfcfsck: sanity check:** *device* **need log replay**

    device needs to perform update log replay.

- **sfcfsck: sanity check:** *device* **needs checking**

    device needs to perform full check of file system.

Update log replay is automatically performed if needed at the time of mount.

The performing method of the update log replay or full check by **sfcfsck** is described in **-o** option.

-n|N

Automatically replies "no" to all inquiries. Does not open the file system for writing.

-y|Y

Automatically replies **yes** to all inquiries.

-o *specific_options*

Allows options specific to the GFS Shared File System to be specified. Insert a comma delimiter between options.

f

Executes a forcible check, regardless of the super block status flag.

p

Checks for non-conversational synopsis and makes corrections.

w

Checks only for write-enabled file systems.

log|elog|nolog

Specifies whether to perform an update log replay. The default is **log**.

log

Executes an update log replay. Executes a full file system check if the log data is physically destroyed.

elog

Executes an update log replay. The command terminates immediately if the log data is physically destroyed without a full file system check being executed.

nolog

Executes a full file system check without executing an update log replay.

<OPERAND>

The following operand can be specified:

device

Specifies the path to a path name of the block special file on the representative partition of the GFS Shared File System or the path to the mount point.

<EXAMPLES>

```
# sfcfsck /dev/sfdsk/gfs01/dsk/volume01 <Enter>
# sfcfsck -o nolog /dev/sfdsk/gfs01/dsk/volume01 <Enter>
```

<EXIT STATUS>

0

File system recovery has been completed normally, or no recovery was necessary.

1

File system recovery has been completed normally.

2

The system must be rebooted.

4

The file system error has left unrepaired.

8

The processing has terminated because of an abnormal event occurring during command execution.

16

An incorrect parameter was specified.

<SEE ALSO>

sfcmkfs(8)

reboot(8), **stat(2)**, and **fstab(5)** in the online manual

<NOTE>

A file system can be checked only when all hosts that share the file system have started operation. The reason is that the information for these hosts is required.

B.1.3 sfcmkfs(8) Construct the GFS Shared File System

<SYNOPSIS>

sfcmkfs [**-m**] [**-o** *specific_options*] *device*

<DESCRIPTION>

The **sfcmkfs** command builds the GFS Shared File System on specified disks.

Once the GFS Shared File System has been created, its entry is added to file system configure information in a management partition. Thereafter, file system configure information in a management partition is used to manage the GFS Shared File System.

<OPTIONS>

The following options can be specified.

-m

Outputs command lines of file system creation time. The file system needs to be present.

-o *specific_options*

Allows options specific to GFS Shared File System to be specified. Insert a comma delimiter between options.

N

Displays parameters used for creating a file system.
Does not actually create a file system.

data=device

Specifies a path name of the block special file of a file data partition.
If you use multiple file data partitions, specify the parameter for each file data partition.

dataopt=y|n

Determines whether or not the file data area is detached from a representative partition. If y is specified, the file data area will not be included in the representative partition. If n is specified, the file data area will be included in the representative partition.
If this parameter is omitted, n will be selected as default.
If a file data partition is not specified for the data option, n will be selected as default.

free=*n*

Specifies the minimum percentage of free area in the file system.

The default is **10**%.
A value can be specified in the range from 0 to 99.

force

Specifies the **force** option to re-create a file system previously registered in file system configure information in a management partition when its configuration is unchanged. If, however, the target file system is being mounted, specifying this option results in an error.

logsz=*n*

Specifies the size of the update log area. The unit is megabytes.
The default is about 1% of the file system size (if maxdsz is specified, it is taken maxdsz). If the specified value is smaller than 5 megabytes, 5 megabytes is set. If the specified value is greater than 50 megabytes, 50 megabytes is set.
A value can be specified in the range from 5 to 100.

maxdsz=*n*

Specifies the maximum size of the file data area of a file system that includes additional partitions. The unit is megabytes. Based on the value specified for this parameter, the size of meta-data and the update log area size are set to proper values after taking into account the additional partitions.

If this option is not specified, the sizes of meta-data area and the update log area are determined based on the file system size of the specified partition.

A minimum value is the same as default. A maximum value is **1048576** (1 terabyte).

maxnode=*n*

Specifies the maximum number of nodes, including expanded nodes by sfcnode(8) in the future, that are shared in the file system configuration on disks.
The default is 16.
A value can be specified in the range from 1 to 32.

maxvol=*n*

Specifies the maximum number of partitions that configure a file system on a disk including additional partitions that will be expanded with sfcadd(8).
When the value is omitted, the default value 16 will be selected.
The minimum value is 1, and the maximum value is 32.

metasz=*n*

Specifies the size of the meta-data area. The unit is megabytes.
The default is about 10% of the file system size (if maxdsz is specified, it is taken maxdsz). However, when the file system size is larger, the ratio of the meta-data area becomes smaller.

The minimum value that can be specified is same as default.

The larger value than the size of the representative partition cannot be specified because the meta-data area is allocated in the representative partition. Moreover, if the file system size is smaller than 1-terabyte, the larger value than the minimum requirement size for managing 1-mega V-data cannot be specified. When the specified value doesn't meet these requirements, either of following messages is displayed and the file system is built by the default value.

```
sfcmkfs:WARNING: value : bogus metasz reset

sfcmkfs:WARNING: value : metasize of partition is over : Retry Auto size
```

mfree=*n*

Specifies the minimum percentage of free V-data area.
The default is 10%.
A value can be specified in the range from 0 to 99.

nblkpext=*n*

Specifies an average number of blocks to be managed by one extent.
This parameter is used as a fragmentation factor for the file data area.
The default is 1.
A value can be specified in the range from 1 to 100.

nbpi=*n*

Specifies the byte count of data per i-node. Number of total i-nodes is decided by dividing the size of the file system's file data area by this parameter. Specify a large value for this parameter if the number of total i-nodes is to be reduced. And specify a larger value for this parameter if the meta-data area is to be saved, too.
The default is 8192 or a value adjusted as the number of total i-nodes should not exceed the maximum value (16 mega).
A value can be specified by the N power of 2. (N is a range from 9 to 20.)

node=*host*,...

Registers the name of a *host* that will share the file system. If multiple nodes should share the file system, specify multiple entries for this parameter using commas.
Hosts whose names are to be specified for *host* must be able to communicate with each other. MDS (meta-data server) placing nodes are selected in order of descriptions for this option.
Unless this option is specified, only the host name of the local node is registered.

<OPERAND>

The following operand can be specified:

device

Specifies a path name of the block special file on the representative partition of the GFS Shared File System.

<EXAMPLES>

```
# sfcmkfs -o node=host1,host2 /dev/sfdsk/gfs01/dsk/volume01 <Enter>
# sfcmkfs -o dataopt=y,data=/dev/sfdsk/gfs01/dsk/volume02,data=/dev/sfdsk/gfs01/dsk/
volume03,node=host1,host2 /dev/sfdsk/gfs01/dsk/volume01 <Enter>
```

<FILE>

/etc/fstab

Environment file for the file system.

<SEE ALSO>

sfcadd(8), sfcadm(8), sfcdf(8), sfcnode(8), sfcsetup(8)

fstab(5) in the online manual

<NOTES>

- *host* must be the host name of the node configuring the GFS Shared File System registered on the management partition by sfcsetup(8) in advance.

- A file system can be created only when all hosts that share the file system have started operation and GFS Shared File System has been installed. The reason is that the information for these hosts is required.

- If another option is specified after a comma (,) that follows a **node** option, the option will have no effect because it will be interpreted as a host name.

- In the GFS Shared File System, the data block size is fixed at **8192** bytes and cannot be altered.

- To re-create a GFS Shared File System by modifying its configuration, first delete the file system with sfcadm(8) .

- The parameter value of the **metasz** option that is included in the command line displayed by **-m** option is a value in which the size of an actual meta-data area was rounded up every a megabyte. Therefore, if the size of an actual meta-data area is same as upper limit, the following warning message might be displayed when the file system is built by using the command line displayed by **-m** option. However, the action is unnecessary.

```
sfcmkfs:WARNING: value : bogus metasz reset
```

B.1.4   sfcmount(8) Mount the GFS Shared File System on the local node

<SYNOPSIS>

sfcmount *device* | *mount_point*
sfcmount *device mount_point*

<DESCRIPTION>

The **sfcmount** command mounts the GFS Shared File System on the local node only. This command can be executed from any nodes that can access the GFS Shared File System.

During the mount processing, this command activates sfcfsd and, if necessary, MDS. Pre-defining /etc/fstab is necessary.

Only a super user can perform mount processing.

<OPTIONS>

The following mount options can be specified for /etc/fstab. Multiple options must be delimited with commas.

atime | noatime | relatime | strictatime

Sets the method to update the last access time of a file.
If **atime** or **relatime** is specified, the last access time of a file is updated when the previous last access time is the same as or older than the last modify time or the last change time. Also, when the previous last access time of a file is old for over a day, it is updated.
If **noatime** is specified, the last access time of a file is not updated.
If **strictatime** is specified, the last access time of a file is updated for each file access.
The default is **relatime**.

atrc | noatrc

Specifies whether to mount (**atrc**) or not to mount (**noatrc**) at the node startup.
The default is **atrc**.

async

Mounts in the manner of processing any I/O for the file system asynchronously.
The default is **async**.

defaults

Synonymous with **rw**, **suid**, **dev**, **exec** and **async**.

dev | nodev

Mounts making block special files available(**dev**) or unavailable(**nodev**).
The default is **dev**.

exec | noexec

Mounts giving permission(**exec**) or no permission(**noexec**) to execute binary files on the file system.
The default is **exec**.

rw | ro

Mounts the file system as read/write enabled(**rw**) or read only(**ro**).

The default is **rw**.

suid | nosuid

Disallows a **setuid** execution. By default, allows **setuid** execution.
The default is **suid**.

<OPERANDS>

The following operands can be specified:

device

Specifies path name of the block or character special file of a representative partition of the file system to be mounted.

mount_point

Specifies the name of the directory to be mounted.

<EXAMPLES>

```
# sfcmount /dev/sfdsk/gfs01/dsk/volume01 /mount_point <Enter>
# sfcmount /dev/sfdsk/gfs01/dsk/volume01 <Enter>
# sfcmount /mount_point <Enter>
```

<FILES>

/etc/mtab

Table for the currently mounted file system.

/etc/fstab

Environment file for the file system.

<SEE ALSO>

sfcumount(8), sfcmntgl(8), sfcumntgl(8)

mount(8) and **fstab(5)** in the online manual

<NOTES>

- sfcfrmd daemon must be running to enable the execution of this command.

- Nodes outside the MDS positioning range alone cannot be mounted.

- The same set of the descriptions for *device* and *mount_point* must be contained in /etc/fstab.

- Mount processing is always performed using the mount options for the mount point described in /etc/fstab. When no mount options are described in /etc/fstab, mount processing without any options is performed.

B.1.5 sfcumount(8) Unmount the GFS Shared File System from the local node

<SYNOPSIS>

sfcumount *device* | *mount_point*

<DESCRIPTION>

The **sfcumount** command unmounts the GFS Shared File System from a local node.

Besides, it stops a MDS if the sfcfsd daemon and node are running.

Only a superuser can perform unmount processing.

<OPERANDS>

The following operands can be specified:

device

Specifies the name of path name of the block special file of a representative partition of the file system to be unmounted.

mount_point

Specifies the directory of the mount point of the file system that will be unmounted.

<EXAMPLES>

```
# sfcumount /dev/sfdsk/gfs01/dsk/volume01 <Enter>
# sfcumount /mount_point <Enter>
```

<FILES>

/etc/mtab

Table for the currently mounted file system.

/etc/fstab

Environment file for the file system.

<SEE ALSO>

sfcmount(8), sfcmntgl(8), sfcumntgl(8)

umount(8) and **fstab(5)** in the online manual

B.2 GFS Shared File System Specific Management Commands

B.2.1 sfcadd(8) Adds file data partitions

<SYNOPSIS>

sfcadd -D *device*[,*device*,...] *t_device*

<DESCRIPTION>

The **sfcadd** command adds file data partitions to the GFS Shared File System that includes a representative partition using the *t_device* operand. The file data partition of the file system can be added only if the file system is unmounted status. No data has to be saved or restored.

<OPTIONS>

The following options can be specified:

-D *device*[,*device*,...]

    Specifies a partition that will be added as the file data area using a path name of the block special file. If multiple partitions are added, separate path name of the block special file by a comma.

<EXAMPLES>

When a representative partition adds /dev/sfdsk/gfs01/dsk/volume02 and /dev/sfdsk/gfs01/dsk/volume03 to the existing GFS Shared File System (/dev/sfdsk/gfs01/dsk/volume01).

```
# sfcadd –D /dev/sfdsk/gfs01/dsk/volume02,/dev/sfdsk/gfs01/dsk/volume03 /dev/sfdsk/gfs01/dsk/
volume01 <Enter>
```

<OPERANDS>

The following operands can be specified:

t_device

    Specifies a path of a path name of the block special file of a representative partition of the GFS Shared File System.

<SEE ALSO>

sfcfsck(8), sfcmkfs(8), sfcadm(8), sfcinfo(8), sfcnode(8)

<NOTES>

- To expand a file system, all the nodes that share the file system must be active because information of each node is required.

- **sfcadd** cannot add a partition if the total number of partitions including the additional partition exceeds the maximum number of partitions (maxvol) defined with sfcmkfs(8). When file data region size after adding partition exceeds the maximum file data region size specified by sfcmkfs(8) (-o maxdsz=n), a part of free area in file data region might not be able to be used because of management region shortage.

- To execute **sfcadd**, the file system must be normal. If it is not, the message "file system (*partition_name*) is corrupted. To Fix: **sfcfsck** *partition_name*" will be output. Execute **sfcadd** again after recovering the file system using sfcfsck(8).

B.2.2   sfcadm(8) Change partition information setting

<SYNOPSIS>

sfcadm -r [-a] *device*
sfcadm -N *device*,...
sfcadm -u *host*,... [-g *p_mds*[,*s_mds*,*mds*,...]] [-p *p_port*[,*s_port*]] [-o *keyword=value*,...] *device*,...
sfcadm -m *host*,... -g *p_mds* [,*s_mds*,*mds*,...] -p *p_port* [,*s_port*] [-o *keyword=value*,...] *device*,...
sfcadm -g *p_mds* [,*s_mds*,*mds*,...] *device*
sfcadm -o *keyword=value*,... *device*
sfcadm -p *p_port* [,*s_port*] *device*
sfcadm -h *ohost nhost*
sfcadm -D *device*,...
sfcadm *device*,...

<DESCRIPTION>

The **sfcadm** command changes the partition information settings for the GFS Shared File System.

In the GFS Shared File System, the information about the group of partitions that makes up the file system is retained in each partition.

Partition information settings must be changed in the following cases:

- Device information (path name, device number) for the file system is changed.

- Information of nodes that share the file system is changed.

- The node name is changed

- The file system is migrated to another system.

Use the command in the following cases:

- Recovery for a super block and partition configuration data.

- For the partition configuration data and copying of a super block

- Recovery if a copy of the partition configuration data no longer matches the original information

- Changing the information in the file system configure information in a management partition and deleting the file system that is specified

- Changing the number of placing MDS (meta-data server) nodes.

- Changing the priority of nodes that places MDS (meta-data server) as primary MDS or secondary MDS.

The **sfcadm** updates the entries in the file system configure information in a management partition, which changes the partition information for the GFS Shared File System and/or the file system environment information. If changes are made to the partition information settings, the information retained in each of the partitions is updated.

This command can only be used if the file system is unmounted in all nodes beforehand.

<OPTIONS>

The following options can be specified:

-D *device,...*

Deletes the GFS Shared File System in which the representative partition provided by file system configure information in a management partition has been registered. Use a comma delimiter to specify multiple files.

-N **or no options specified**

Updates the node and device information for all nodes in the partition configuration data. Use **-N** option if the minor numbers of the specified devices on each nodes are different after the path name of the block or character special file of the file system in use has been changed or copying between partitions is performed. If all hosts have the same minor number, specify only the device information, without specifying the **-N** option. Then updates the file system configure information in a management partition.

-a

When the **-r** option is specified, performs the recovery using a copy of the super block in the partition of the recovery source and a copy of the partition configuration data.

-g *p_mds,s_mds,mds,...*

Specifies the node name that MDS can place delimited with a comma. Primary MDS and secondary MDS can be specified in order of the descriptions of node names. To change either primary MDS or secondary MDS, use the "*p_mds,*" or ",*s_mds*" and specify the new node name. In order to change node except primary and secondary, all nodes must be specified.
To change a configuration with only a primary MDS, specify only "***p_mds***" (a comma delimiter is not required). The file system configure information in a management partition is also updated.
The node name to be specified needs to be a node which is sharing the file system. When specified with the option **-u** or **-m** at the same time, it can only specify the node name specified as either **-u** or **-m** option.

-h *o_host n_host*

Changes a node name to be changed. Use the "*o_host n_host*" syntax and specify the node names. In the file system configure information in a management partition, then, all occurrences of *o_host* are changed to *n_host*. The node names registered in the partition configuration data also change.

-m *host,...*

Changes the node and device information in the partition configuration data to the information specified. Use a comma delimiter to specify multiple node names.
The **-m** option with the option **-g** and **-p** must be specified at the same time.
This option also updates the file system configuration information on the management partition. Use this option to migrate from the current file system to another system.

-o *keyword=value,...*

Specifies a variable setting that can be tuned in the GFS Shared File System using the syntax "*variable-name=value*" syntax. Use a comma delimiter to specify multiple variables.

When deleting a registered variable, use the specification "*variable-name=*". You can specify the timeout of communication between nodes such as follow.

In the following example, the timeout of communication between nodes is set to 180 seconds.

```
-o CON_TIMEOUT=180
```

Changes the time after a communicative response is lost until it estimates a partner node to be an abnormal state in the monitoring of the node by communication.

Please set the large value in the above-mentioned timeout of communication when it is forecast that the load of the system rises.

-p *p_port,s_port*

Specifies the primary port name and the secondary port name delimited with a comma. To change either port name, use the *p_port,* or *,s_port* syntax and specify the new port name.

When primary, secondary register the same port name, specify only p_port (a comma delimiter is not required).

The file system configure information in a management partition is also updated.

-r *device*

Specifies path name of the block special file that will be the recovery source when a super block and partition configuration data are restored. When the **-a** option is specified, redundant blocks in the specified source partition are used for restoring the data into other partitions.

-u *host,...*

Changes the node and device information in the partition configuration data to the information specified. Use a comma delimiter to specify multiple node names.

The file system configure information in a management partition is also updated. Use this option when the node and device information must be rewritten after the path name of the block or character special file of the file system in use has been changed or copying between partitions is performed.

<OPERANDS>

The following operands can be specified:

device,...

Specify path name of the block special files of all partitions that configure the GFS Shared File System by separating by a comma.

device

Specifies a path name of the block special file on the representative partition of the GFS Shared File System. If the -r option is used, non-representative partitions can also be specified.

<EXAMPLES>

Perform recovery of another partition using a partition information of the recovery source.

```
# sfcadm -r /dev/sfdsk/gfs01/dsk/volume01 <Enter>
```

Perform recovery of another partition using a copy of a partition information of the recovery source.

```
# sfcadm -r -a /dev/sfdsk/gfs01/dsk/volume01 <Enter>
```

Update the node and device information for all nodes in the partition information.

```
# sfcadm /dev/sfdsk/gfs01/dsk/volume01,/dev/sfdsk/gfs01/dsk/volume02 <Enter>
```

Rewrite the node and device information in the partition information

```
# sfcadm -u host1,host2 -g host1,host2 -p sfcfs-1,sfcfs-1 /dev/sfdsk/gfs01/dsk/
volume01 <Enter>
```

Migrate a specific file system to another system.

```
# sfcadm -m host1,host2 -g host1,host2 -p sfcfs-1,sfcfs-1 /dev/sfdsk/gfs01/dsk/
volume01 <Enter>
```

Change the priority ranking of nodes placing MDS.

    - Change primary MDS to host1.

```
# sfcadm -g host1, /dev/sfdsk/gfs01/dsk/volume01 <Enter>
```

    - Change secondary MDS to host2.

```
# sfcadm -g ,host2 /dev/sfdsk/gfs01/dsk/volume01 <Enter>
```

Change the priority ranking of nodes placing MDS.

```
# sfcadm -g host1,host2 /dev/sfdsk/gfs01/dsk/volume01 <Enter>
```

Assign the same port name to the primary and secondary ports.

```
# sfcadm -p sfcfs-1 /dev/sfdsk/gfs01/dsk/volume01 <Enter>
```

Register the variable XXX=60.

```
# sfcadm -o XXX=60 /dev/sfdsk/gfs01/dsk/volume01 <Enter>
```

Change the node name host1 to host2.

```
# sfcadm -h host1 host2 <Enter>
```

Delete a specific file system.

```
# sfcadm -D /dev/sfdsk/gfs01/dsk/volume01 <Enter>
```

<SEE ALSO>

sfcmkfs(8), sfcinfo(8), sfcmntgl(8), sfcnode(8), sfcsetup(8)

<NOTES>

- A file system can be changed only when all nodes that share the file system have started operation and GFS Shared File System has been installed. The reason is that the information for these nodes is required.

- *host* must be the node name of the node configuring the GFS Shared File System registered on the management partition by sfcsetup(8) in advance.

- Using **-u** option with specified new node name or device name, **-g** options must be specified.

- When the name of the representative partition of a GFS Shared File System is changed or migration from the file system to another file system is performed with **sfcadm**, /etc/fstab needs to be set.

- When only one node is specified as MDS node by the **-g** option, the file system cannot be mounted on the other node individually. In this case, the file system must be mounted on all nodes by using sfcmntgl(8). Note that the file system will be blockaded on the node which is not MDS node, when unmount, shutting down or panic occurs on the MDS node.

B.2.3 sfcfrmstart(8) Start sfcfrmd daemon on the local node

<SYNOPSIS>

sfcfrmstart [-f]

<DESCRIPTION>

The **sfcfrmstart** command is used to perform sfcfrmd daemon start processing on the local node.

Only a super user can perform sfcfrmd daemon start processing.

<OPTION>

The following option can be specified:

-f

> A sfcfrmd daemon will be suspended until it can confirm that a cluster quorum exists. If the **sfcfrmstart -f** command is executed, the sfcfrmd daemon will start without waiting for a quorum.
> You can confirm if startup of the sfcfrmd daemon is suspended by checking if the message "WARNING: sfcfsrm:5001: Starting the sfcfrmd daemon was suspended because quorum does not exist" is output.

<EXAMPLES>

Starting the sfcfrmd daemon when the sfcfrmd is stopped.

```
# sfcfrmstart <Enter>
```

Starting the suspended sfcfrmd daemon without waiting for a quorum.

```
# sfcfrmstart -f <Enter>
```

<FILE>

/var/opt/FJSVsfcfs/sfcfsrm.conf

Environment file of sfcfrmd daemon

<SEE ALSO>

sfcfrmstop(8), sfcsetup(8)

cftool(1M) in the online manual of Cluster Foundation

<NOTES>

- To enable the use of this command, the local node must be multi user mode. This command cannot be used by single user mode.

- Within a given node, this command cannot be executed more than once.

- Within a given node, this command cannot be executed concurrently with sfcfrmstop(8).

- When you execute this command with the **-f** option specified, confirm that there is no cluster partition by execute **cftool -n** at all nodes. If a cluster partition occurred and this command with the **-f** option specified is executed, inconsistency of the GFS Shared File System might occur.

B.2.4 sfcfrmstop(8) Stop sfcfrmd daemon on the local node

<SYNOPSIS>

sfcfrmstop

<DESCRIPTION>

The **sfcfrmstop** command is used to stop sfcfrmd daemon on the local node.

The GFS Shared File System can perform only in the state where it is not mounted by the applicable node.

Only a super user can perform sfcfrmd daemon stop processing.

<EXAMPLE>

```
# sfcfrmstop <Enter>
```

<SEE ALSO>

sfcfrmstart(8)

<NOTES>

- To enable the use of this command, the local node must be multi user mode. This command cannot be used by single user mode.

- Within a given node, this command cannot be executed more than once.

- Within a given node, this command cannot be executed concurrently with sfcfrmstart(8).

B.2.5   sfcgetconf(8) Make a backup of the management partition

<SYNOPSIS>

sfcgetconf *backupfile*

sfcgetconf -a *backupfile*

sfcgetconf -i *fsid* [-i *fsid*]... *backupfile*

sfcgetconf -d *device* [-d *device*]... *backupfile*

<DESCRIPTION>

The **sfcgetconf** command outputs next items to **backupfile**:

- The file system configuration information in management partition.

- The procedure for using the sfcadm(8) command to reconfigure the file system.

The contents of the sfcfsrm.conf file when the **-a** option is specified.

The output backup file contains the procedure for executing the sfcadm(8) command in the sh(1) shell script format.

Management partition recovery can be done using the script output by **sfcgetconf** after it is initialized with sfcsetup(8).

<OPTIONS>

The following options can be specified:

-a

Outputs the contents of sfcfsrm.conf to the first part of the backup file in the form of comments. (The output for no **-a** option will follow.)

-i *fsid*

Outputs only information about the file system corresponding to a specified fsid value and the procedure for executing the sfcadm(8) command that re-creates information on the management partition for that file system when this option is specified (not for all the file systems).

-d *device*

Specifies a path name of the block special file on the partition configuring the file system. When this option is specified (not for all the file systems), only information about the file system corresponding to a specified device, and the procedure for executing the sfcadm(8) command that re-creates information on the management partition for that file system, are output.

<OPERANDS>

The following operands can be specified:

backupfile

> Specifies the name of the output destination file. When the output destination file already exists, backup is not performed.

<EXAMPLES>

When a backup of the entire file system is to be output to file sfccfg.bkup.1.

```
# sfcgetconf sfccfg.bkup.1 <Enter>
```

When the contents of the sfcfsrm.conf file are to be recorded at the head of the backup file, and the backup for the entire file system is to be output to sfccfg.bkup.2.

```
# sfcgetconf -a sfccfg.bkup.2 <Enter>
```

When backups of the file systems corresponding to fsid values 1 and 2 are to be output to sfccfg.bkup.3

```
# sfcgetconf -i 1 -i 2  sfccfg.bkup.3 <Enter>
```

When backups of the file system corresponding to /dev/sfdsk/gfs/dsk/volume1 are to be output to sfccfg.bkup.4.

```
# sfcgetconf -d /dev/sfdsk/gfs01/dsk/volume01 sfccfg.bkup.4 <Enter>
```

<FILE>

/var/opt/FJSVsfcfs/sfcfsrm.conf

Environment file of sfcfrmd daemon

<SEE ALSO>

sfcadm(8), sfcsetup(8)

<NOTE>

To enable the use of this command, the local node must be multi user mode.

B.2.6   sfcinfo(8) Display partition information

<SYNOPSIS>

sfcinfo [-V] [-n] { *device* | *mount_point* } [{ *device* | *mount_point* }...]
sfcinfo [-V] [-n] -d *dev*[,*dev*,...]
sfcinfo -a [-V] [-n]
sfcinfo -e { *device* | *mount_point* }
sfcinfo -e [-a]

<DESCRIPTION>

sfcinfo displays the partition information of the GFS Shared File System specified by *device* or *mount_point*.

device indicates the partition configuring the file system by a path name of the block special file.

mount_point indicates the path to the mount point of the file system. To use the path to the mount point, the file system must already be described in /etc/fstab.

If neither *device* nor *mount_point* is specified and the **-a** option is omitted, **sfcinfo** displays the information on all GFS Shared File Systems being mounted in the command execution node.

sfcinfo outputs the following detailed information:

- FSID

  FS identification number

- special

  Special device name (device number)

- size

  Partition size (in kilobytes)

- Type

  Usage of partitions

  - META: meta-data area

  - DATA: file data area

  - LOG: update log area

- mount

  Mount point in a command execution node
  "-----" when no file system is mounted

- hostID

  Host ID

- status

  Start status

  - RUN: The node is operating

  - STOP: The node is stopped

- hostname

  Node name

<OPTIONS>

The following options can be specified:

-V

  Displays the version of the GFS Shared File System.

-a

  Displays the information for all file systems described in the file system configure information in a management partition.

-d *dev*[,*dev*,...]

  Specifies a device number in hexadecimal for the partition to be displayed.
  Use a comma delimiter to specify multiple device numbers.

-e

  Displays file system variables set by sfcadm(8).

-n

  Displays node information.
  If no parameters are specified, the information for all mounted file systems is displayed.

<OPERAND>

The following operand can be specified:

device | mount_point

Specifies a path name of the block special file on the partition configuring the file system or the path to the mount point of the file system.

<EXAMPLES>

```
# sfcinfo /dev/sfdsk/gfs01/dsk/volume01 <Enter>
/dev/sfdsk/gfs01/dsk/volume01:
FSID special                                  size  Type  mount
   1 /dev/sfdsk/gfs01/dsk/volume01(1e756)     14422  META  /mnt
   1 /dev/sfdsk/gfs01/dsk/volume01(1e756)      5116  LOG   /mnt
   1 /dev/sfdsk/gfs01/dsk/volume01(1e756)     95112  DATA  /mnt

# sfcinfo -V /dev/sfdsk/gfs01/dsk/volume01 <Enter>
sfcfs v1.0 file system
/dev/sfdsk/gfs01/dsk/volume01:
FSID special                                  size  Type  mount
   1 /dev/sfdsk/gfs01/dsk/volume01(1e756)     14422  META  /mnt
   1 /dev/sfdsk/gfs01/dsk/volume01(1e756)      5116  LOG   /mnt
   1 /dev/sfdsk/gfs01/dsk/volume01(1e756)     95112  DATA  /mnt

# sfcinfo /mnt <Enter>
/dev/sfdsk/gfs01/dsk/volume01:
FSID special                                  size  Type  mount
   1 /dev/sfdsk/gfs01/dsk/volume01(1e756)     14422  META  /mnt
   1 /dev/sfdsk/gfs01/dsk/volume01(1e756)      5116  LOG   /mnt
   1 /dev/sfdsk/gfs01/dsk/volume01(1e756)     95112  DATA  /mnt

# sfcinfo -d 1e756 <Enter>
/dev/sfdsk/gfs01/dsk/volume01:
FSID special                                  size  Type  mount
   1 /dev/sfdsk/gfs01/dsk/volume01(1e756)     14422  META  /mnt
   1 /dev/sfdsk/gfs01/dsk/volume01(1e756)      5116  LOG   /mnt
   1 /dev/sfdsk/gfs01/dsk/volume01(1e756)     95112  DATA  /mnt

# sfcinfo -n /dev/sfdsk/gfs01/dsk/volume01 <Enter>
/dev/sfdsk/gfs01/dsk/volume01:
FSID  hostID     status  hostname
   1  80000000   RUN     prime01
   1  80000001   STOP    prime02

# sfcinfo -a <Enter>
/dev/sfdsk/gfs01/dsk/volume01:
FSID special                                  size  Type  mount
   1 /dev/sfdsk/gfs01/dsk/volume01(1e756)     14422  META  /mnt
   1 /dev/sfdsk/gfs01/dsk/volume01(1e756)      5116  LOG   /mnt
   1 /dev/sfdsk/gfs01/dsk/volume01(1e756)     95112  DATA  /mnt

/dev/sfdsk/gfs01/dsk/volume02:
FSID special                                  size  Type  mount
   2 /dev/sfdsk/gfs01/dsk/volume02(1e757)     25418  META  -----
   2 /dev/sfdsk/gfs01/dsk/volume02(1e757)      5120  LOG   -----
   2 /dev/sfdsk/gfs01/dsk/volume02(1e757)    232256  DATA  -----
```

<FILES>

/etc/mtab

Mounted file system table

/etc/fstab

List of default parameters for each file system

<SEE ALSO>

sfcadm(8)

fstab(5) in the online manual

B.2.7 sfcmntgl(8) Mount the GFS Shared File System on all nodes

<SYNOPSIS>

sfcmntgl *device* | *mount_point*
sfcmntgl *device mount_point*

<DESCRIPTION>

The **sfcmntgl** mounts all nodes sharing the same file system according to the specified operands.

sfcmntgl can be executed from any nodes as the GFS configuring nodes that are specified with sfcsetup(8). The pre-definitions in /etc/fstab are necessary.

Only a super user can perform mount processing.

<OPERAND>

The following operand can be specified:

device

Specifies path name of the block special file of a representative partition of the file system to be mounted.

mount_point

Specifies the name of the directory to be mounted.

<EXAMPLE>

```
# sfcmntgl /dev/sfdsk/gfs01/dsk/volume01 /mount_point <Enter>
# sfcmntgl /dev/sfdsk/gfs01/dsk/volume01 <Enter>
# sfcmntgl /mount_point <Enter>
```

<FILES>

/etc/mtab

Table for the currently mounted file system.

/etc/fstab

Environment file for the file system.

<SEE ALSO>

sfcmount(8), sfcsetup(8), sfcumntgl(8), sfcumount(8)

mount(8) and **fstab(5)** in the online manual

<NOTES>

- sfcfrmd daemon must be running to enable the execution of this command.

- When both *device* and *mount_point* are specified, the same set of the descriptions must be contained in /etc/fstab for all nodes sharing the file system.

- Mount processing is always performed using the mount options for the mount point described in /etc/fstab. When no mount options are described in /etc/fstab, mount processing without any options is performed.

B.2.8 sfcnode(8) Add, delete, and change node configuration information

<SYNOPSIS>

sfcnode -a *host*,... *device*

sfcnode -d [-h *host*,... | -i *hostid*,...] *device*

<DESCRIPTION>

The **sfcnode** command adds, deletes, and changes GFS Shared File System node information.

To add or update node information, specify the **-a** option. If a specified host name already exists, **sfcnode** updates node and device information for the host. If the host does not exist, **sfcnode** adds the node and device information. Use a comma delimiter to specify multiple host names.

To delete node information, specify the **-d** option. For the **-h** or **-i** option, specify the host name (*host*) or host ID (*hostid*) of the host to be deleted. **sfcnode** deletes the node information of the host. If neither the **-h** nor **-i** option is specified, **sfcnode** deletes the node information for the host on which the **sfcnode** is executed. Use a comma delimiter to specify multiple host names or IDs. The **-h** and **-i** options cannot be specified together.

After the **sfcnode** command changes the GFS Shared File System node information, the entries in the file system configure information in a management partition are updated. The information retained in the partitions is also changed.

<OPTIONS>

The following options can be specified:

-a *host*,...

   Specifies the name of a host whose node information is to be added.

-d

   Deletes node information.

-h *host*,...

   Specifies the name of a host whose node information is to be deleted. The host must be one that can be used by the GFS Shared File System.

-i *hostid*,...

   Specifies the ID in hexadecimal of a host to be deleted from node information. The host must be one that can be used by the GFS Shared File System. Note that a host ID specified for this option differs from a host ID obtained with **hostid(1)**.

<OPERAND>

The following operand can be specified:

device

   Specify a path name of the block special file on the representative partition of the GFS Shared File System whose node information is to be updated.

<EXAMPLES>

Add new node that was added by sfcsetup(8) to the sharing nodes of GFS Shared File System.

```
# sfcnode -a host2 /dev/sfdsk/gfs01/dsk/volume01 <Enter>
```

Add or update node/device information for hosts by specifying host names.

```
# sfcnode -a host1,host2 /dev/sfdsk/gfs01/dsk/volume01 <Enter>
```

Delete node/device information for hosts by specifying host names.

```
# sfcnode -d -h host2 /dev/sfdsk/gfs01/dsk/volume01 <Enter>
```

Delete node/device information for a host by specifying its ID.

```
# sfcnode -d -i 80000000 /dev/sfdsk/gfs01/dsk/volume01 <Enter>
```

Delete node/device information for the local host.

```
# sfcnode -d /dev/sfdsk/gfs01/dsk/volume01 <Enter>
```

<SEE ALSO>

sfcmkfs(8), sfcadm(8), sfcinfo(8), sfcsetup(8)

hostid(1) in the online manual

<NOTES>

- A file system can be changed only when all hosts that share the file system have started operation and GFS Shared File System has been installed. The reason is that the information for these hosts is required.

- *host* must be the host name of the node configuring the GFS Shared File System registered on the management partition by sfcsetup(8) in advance.

- If **sfcnode** fails outputting a message as "node full, can't sfcnode," the number of nodes has exceeded the maximum number of sharing nodes set when the GFS Shared File System was created with sfcmkfs(8). Check the registered node information with sfcinfo(8), and then add nodes after deleting node information no longer necessary using **sfcnode** with the **-d** option.

- Cannot delete node information of host that placing primary MDS, and secondary MDS.

B.2.9   sfcrscinfo(8) Display file system information

<SYNOPSIS>

sfcrscinfo -m [-a | -i *fsid* | *device*]

<DESCRIPTION>

The **sfcrscinfo** command outputs file system information for GFS Shared File Systems.

sfcrscinfo outputs the following detailed information:

- FSID (File system ID)

- MDS/AC

    - MDS(P): The primary MDS on a definition

    - MDS(S): The secondary MDS on a definition

    - AC : The node that mount the file system

- STATE/S-STATE (MDS)

    - run - : The primary MDS is running

    - wait - : The secondary MDS is waiting for failure recovery of the primary MDS

    - stop - : MDS stopped

- none - : no relation with MDS

- If S-STATE is other than "-", MDS is in state-transition.

- STATE/S-STATE (AC)

- run - : file system is in mounted

- stop - : file system is in unmounted

- If S-STATE is other than "-", AC is in state-transition.

- RID-1/ RID-2/ RID-N (Resource No.: Always 0)

- hostname (MDS/AC host name)

<OPTIONS>

The following options can be specified:

-a

Outputs information for all GFS Shared File Systems that have been created.

-m

Outputs the MDS and AC information for GFS Shared File Systems.
If **-m** specified with **-a**, outputs information for all GFS Shared File Systems that have been created.
If only specified **-m**, outputs information for GFS Shared File System which status is mounted.

-i *fsid*

Displays the information for the file system specified by *fsid*.

<OPERAND>

The following operand can be specified:

device

Specifies path name of the block special file of a representative partition of the GFS Shared File System.

<EXAMPLES>

```
# sfcrscinfo -m /dev/sfdsk/gfs01/dsk/volume01 <Enter>
/dev/sfdsk/gfs01/dsk/volume01:
FSID  MDS/AC  STATE  S-STATE  RID-1  RID-2  RID-N  hostname
   1  MDS(P)  run    -            0      0      0  prime01
   1  AC      run    -            0      0      0  prime01
   1  MDS(S)  wait   -            0      0      0  prime02
   1  AC      run    -            0      0      0  prime02
```

B.2.10 sfcsetup(8) Initialize the management partition, add, delete and display node information, display path of management partition, and register and display the startup mode of the sfcfrmd daemon

<SYNOPSIS>

sfcsetup -c [-f] *device*

sfcsetup -a *device*

sfcsetup -d

sfcsetup -m [*startup_mode*]

sfcsetup -p

sfcsetup

<DESCRIPTION>

The **sfcsetup** command initializes the management partition, adds node information to the management partition, deletes, displays the node information, displays a path of the management partition, and registers and displays the startup mode of sfcfrmd daemon.

To initialize the management partition, specify the **-c** option and path name of the block special file. When the specified device has already been used for the management partition, the **sfcsetup** command returns the error in order to prevent incorrect operation. In this case, perform initialization with the **-f** option specified. The size of a management partition needs to be **40** megabytes or more.

To add node information, specify the **-a** option. When the specified block special file has already been initialized for the management partition, node information is added to the management partition for which the command was executed.

To delete node information, specify the **-d** option. Note, however, that sfcfrmd daemon must be stopped in all node of the cluster foundation.

To register the startup mode of the sfcfrmd daemon, specify the **-m** option and the startup mode. When the command is executed specify the **-m** option without the startup mode, display the registered startup mode of the sfcfrmd daemon.

When the command is executed with **-p** options specified, display the path name of a management partition.

When the command is executed with no options specified, a listing of the node information recorded in the management partition is displayed.

When addition of node information and deletion are performed by the **sfcsetup** command, add/delete entries of management partition to/from the sfcfsrm.conf file.

<OPTIONS>

The following options can be specified:

-c *device*

Initializes the partition specified in *device* for use as a management partition.

-f

Initializes a block special file that has been used as a management partition.

-a *device*

Adds node information.

-d

Deletes node information.

-m [*startup_mode*]

Registers the startup mode of the sfcfrmd daemon specified in **startup_mode**. When the **startup_mode** is not specified, displays the registered startup mode of the sfcfrmd daemon.
The following startup mode can be specified in **startup_mode**. When initializing as management partition, **wait** is registered as startup mode of sfcfrmd daemon.

wait

When a node is started up, the startup of the sfcfrmd daemon will be suspended until it can confirm that a cluster quorum exists. Node startup also will be suspended.
You can confirm if startup of the sfcfrmd daemon is suspended by checking if the message "WARNING: sfcfsrm:5001: Starting the sfcfrmd daemon was suspended because quorum does not exist" is output.
And, when the CF is started up by the GUI, startup of the sfcfrmd daemon is given up if a quorum does not exist.

wait_bg

When a node is started up or the CF is started up by the GUI, the startup of the sfcfrmd daemon will be suspended until it can confirm that a cluster quorum exists.
This differs from **wait** in that node startup or CF startup will continue.
You can confirm if startup of the sfcfrmd daemon is suspended by checking if the message "WARNING: sfcfsrm:5001: Starting the sfcfrmd daemon was suspended because quorum does not exist" is output.

nowait

> When a node is started up or the CF is started up by the GUI, the sfcfrmd daemon will be started even if it cannot confirm that a cluster quorum exists.
> This startup mode exists for backward compatibility only.

-p

> Displays the path name of a management partition.

<OPERAND>

The following operand can be specified:

device

> Specifies path name of the block special file to be used for the management partition.

<EXAMPLES>

Initializing a management partition

```
# sfcsetup -c /dev/sfdsk/gfs/dsk/control <Enter>
```

Initializing a parturition that has previously been used as a management partition.

```
# sfcsetup -c -f /dev/sfdsk/gfs/dsk/control <Enter>
```

When sfcfrmd is not running in all node of the cluster and adding information for a node.

```
# sfcsetup -a /dev/sfdsk/gfs/dsk/control <Enter>
```

When sfcfrmd is not running in all node of the cluster and deleting information for a node.

```
# sfcsetup -d <Enter>
```

Registering the startup mode as **wait**. When a startup mode is **wait**, sfcfrmd daemon will be suspended until it can confirm that a cluster quorum exists, and node startup also will be suspended.

```
# sfcsetup -m wait <Enter>
```

Displaying the startup mode of the sfcfrmd daemon registered.

```
# sfcsetup -m <Enter>
wait
```

<FILE>

/var/opt/FJSVsfcfs/sfcfsrm.conf

Environment file of sfcfrmd daemon

<SEE ALSO>

sfcadm(8), sfcfrmstop(8)

<NOTES>

- Initialization of a management partition, addition and deletion of node information, and registration the startup mode of the sfcfrmd daemon cannot be performed during the sfcfrmd daemon activation.

- Node information used by a file system cannot be deleted. Please delete the file system configuration information related the node in a management partition by sfcadm(8).

- If a cluster partition condition occurs during node startup or when the CF is started by the GUI, and **nowait** is set as the startup mode of the sfcfrmd daemon, a data inconsistent may be caused in the GFS Shared File System.

B.2.11 sfcstat(8) Report statistics on GFS Shared File Systems

<SYNOPSIS>

sfcstat [-abcdefghijklmnopqrstA] [*interval* [*count*]] [*mount_point*...]

<DESCRIPTION>

The **sfcstat** command displays statistics about GFS Shared File Systems.

If **sfcstat** is used without parameters, it reports statistical information recorded after the file system is mounted. The next time, it reports the information recorded since the last report.

If *interval* is specified, the command reports statistical information each number of seconds defined by *interval*.

If *count* is specified, the command reports statistical information repeatedly for the value defined by *count*.

If *mount_point* is specified, the command reports the information for the specified file system.

<OPTIONS>

The following options can be specified:

-A

Reports all information.

-a

Reports **AC** information about read/write operations. Reports the number of RPC messages for the following purposes:

- To change file atime or mtime

- To change file attributes

- To access file attributes

- To access file size

- To acquire tokens when changing file size

- To change file size

- To purge tokens when changing file size

- To access information about extent allocations

- To change information about extent allocations

-b

Reports AC information about meta-data management buffers. Reports the number of buffer allocations, the number of dynamic buffer allocations, and the number of sleep occurred during buffer requests.

-c

Reports MDS information about meta-data management buffers. Reports the number of buffer allocations, the number of dynamic buffer allocations, and the number of sleep occurred during buffer requests.

-d

Reports AC information about direct I/O. Reports the following information:

- Number of direct I/O read operations

- Number of direct I/O write operations

- Number of blocks read

- Number of blocks written

- Number of read requests changed to normal I/O

- Number of write requests changed to normal I/O

- Number of page flush after reading

- Number of page flush after writing

-e

Reports MDS information about purging tokens. Reports the following information:

- Number of purging tokens

- Number of AC that purged tokens

- Number of purging tokens (by token type)

- Number of AC that purged tokens (by token type)

-f

Reports AC information about VFS operations. Reports the number of RPC messages for the following purposes:

- To access file handle

- To access file attributes

- To purge tokens

-g

Reports the number of times MDS was received requests for tokens (by token type).

-h

Reports MDS information about the meta-daemon. Reports the number of times the meta-daemon was activated, the number of buffers in the meta-daemon I/O queues, the number of times the meta-daemon was activated because the disk was full, and the number of times the meta-daemon was activated because the meta-cache was full.

-i

Reports AC information about actual I/O. Reports the number of actual I/O reading, the number of actual I/O writing, the number of I/O blocks actually read, and the number of I/O blocks actually written.

-j

Reports MDS information about the log daemon. Reports the following information:

- Number of times log daemon was activated

- Number of sync requests

- Number of fsync requests

- Number of times log buffer full

- Number of requests issued from meta-write daemon

- Number of times sleep occurs because transaction limit exceeded

- Number of times sleep occurs because log disk full

- Number of times sleep occurs because of a preceding large transaction

- Number of times sleep occurs because a preceding transaction enlarged

- Number of requests for I/O operation issued from log daemon

- Time(msec) required for writing

- Number of blocks written

- Number of valid blocks

- Number of requests to log transactions

- Number of parallel operation transactions to be logged

- Average number of parallel operation transactions

- Number of times huge buffer was used

- Number of sub-transactions generated

-k

Reports MDS information about operations for free extents. Reports the number of search requests, the number of times blocks acquired, the number of acquired blocks, the number of blocks acquisition failed, the number of times blocks released, the number of released blocks, the number of times released blocks successfully linked, the number of division of free extent management, and the number of joining free extent management.

-l

Reports AC information about file and record locks. Reports the number of RPC messages for the following purposes:

- To open file

- To close file

- To map file to address space

- To read file

- To write file

- To set lock, clear lock, or get lock information

- To set mandatory lock, or clear mandatory lock

-m

Reports MDS information about meta cache (by meta type). Reports the number of accesses, the number of cache hits, the number of updates, the number of read operations, the number of write operations, writing time, and average time(msec).

-n

Reports AC information about RNODE operations. Reports the number of RPC messages for the following purposes:

- To be requested to purge tokens

- To be requested to purge tokens of freed RNODE

- To purge tokens to free RNODE

-o

Reports AC information about VNODE operations. Reports the following information:

- Number of RPC messages to open file

- Number of RPC messages to close file

- Number of access to file attributes

- Number of RPC messages to access file attributes

- Number of re-sent RPC messages to access file attributes

- Number of RPC messages to access the number of blocks of regular file

- Number of access to symbolic link

- Number of RPC messages to access symbolic link

- Number of RPC messages to move modified attributes to storage devices

- Number of access to directory

- Number of RPC messages to access directory

- Number of re-sent RPC messages to access directory

-p

Reports MDS information about file operation. Reports the following information:

- Number of requests to operate files (by operation type)

- Size of RPC messages to operate files (by operation type)

- Number of successful replies (by operation type)

- Size of successful replies (by operation type)

- Number of failed replies (by operation type)

- Time(msec) required for operating files (by operation type)

-q

Reports MDS information about operations for allocated extents. Reports the number of search requests, the number of times blocks acquired, the number of acquired blocks, the number of blocks acquisition failed, the number of times blocks released, the number of released blocks, the number of times released blocks successfully linked, the number of division of free extent management, and the number of joining free extent management.

-r

Reports AC information about file operation. Reports the following information:

- Number of sending requests to operate files

- Number of receiving replies

- Number of timeouts while receiving replies

- Number of requests waiting to be sent (by operation type)

- Time for requests to wait to be sent (by operation type)

- Number of sending requests (by operation type)

- Time(msec) to wait to receive replies (by operation type)

- Number of successful replies (by operation type)

- Number of replies failed because no file existed (by operation type)

- Number of replies failed because file attributes changed (by operation type)

- Number of replies failed because file existed (by operation type)

- Number of replies failed because of other error (by operation type)

- Number of failure to receive replies (by operation type)

-s

Reports MDS information about operations for reserving extents. Reports the number of requests to reserve extents, the number of reserved extents, the number of reserved blocks, the number of requests to reserve extents contiguously, and the number of extents failed to reserve.

-t

Reports the number of times MDS is request to purging tokens. Reports the following information:

- Number of changing information about extent allocations

- Number of extents whose allocation information were changed

- Number of blocks whose allocation information were changed

- Number of changing information about extent allocations partially

- Number of extents whose allocation information were changed partially

- Number of blocks whose allocation information were changed partially

- Number of purging tokens (by token type)

If the options are omitted, the command reports all information.

<OPERAND>

The following operand can be specified:

mount_point

Specifies the path name of the mount point where the file system has been mounted.

<EXAMPLE>

```
# sfcstat 5 /mnt <Enter>
```

<SEE ALSO>

sar(1) in the online manual

<NOTE>

The AC does not get meta-data management buffers dynamically at present. So, AC information about meta-data management buffers such as the numbers of buffer allocations, dynamic buffer allocations, and sleep that occurred during buffer requests will be 0 at all times.

B.2.12 sfcumntgl(8) Unmount the GFS Shared File System from all nodes

<SYNOPSIS>

sfcumntgl *device* | *mount_point*
sfcumntgl *device mount_point*

<DESCRIPTION>

The **sfcumntgl** unmounts all nodes sharing the same file system according to the specified operands. sfcfsd and MDS will be stopped.

sfcumntgl can be executed from any nodes as the GFS configuring nodes that are set with sfcsetup(8).

Only a super user can perform unmount processing.

<OPERANDS>

The following operands can be specified:

device

Specifies path name of the block special file for a representative partition of the file system to be unmounted.

mount_point

Specifies the name of the directory to be unmounted.

<EXAMPLES>

```
# sfcumntgl /dev/sfdsk/gfs01/dsk/volume01 /mount_point <Enter>
# sfcumntgl /dev/sfdsk/gfs01/dsk/volume01 <Enter>
# sfcumntgl /mount_point <Enter>
```

<FILES>

/etc/mtab

Table for the currently mounted file system.

/etc/fstab

Environment file for the file system.

<SEE ALSO>

sfcmntgl(8), sfcsetup(8), sfcumount(8)

umount(8) and **fstab(5)** in the online manual

<NOTES>

- sfcfrmd daemon must be running to enable the execution of this command.

- When both *device* and *mount_point* are specified, the same set of the descriptions must be contained in /etc/fstab for all nodes sharing the file system.

Appendix C Troubleshooting

This appendix describes how to collect troubleshooting information and take corrective action in the event of a GFS Shared File System failure.

Please cope with a GFS Shared File System failure in the following procedures:

- "C.1 Corrective actions of messages"

- "C.2 Collecting information for troubleshooting"

C.1 Corrective actions of messages

If a command of the GFS Shared File System fails, take corrective action of the message. See "Appendix A List of Messages."

If the problem could not be resolved, take corrective action according to the message output to the console or the message stored in /var/log/messages.

See
. . . .

For corrective actions of messages related to cluster systems, see "PRIMECLUSTER Messages."

For corrective actions of messages related to GDS, see "Appendix C GDS Messages" of "PRIMECLUSTER Global Disk Services Configuration and Administration Guide."

For corrective actions of messages related to GLS, see "Appendix A List of Error Messages" of "PRIMECLUSTER Global Link Services Configuration and Administration Guide: Redundant Line Control Function."

For corrective actions of messages related to GFS Management View, see "Appendix A Messages" of "PRIMECLUSTER Web-Based Admin View Operation Guide."

For corrective actions of messages related to GFS Shared File System, see "Appendix A List of Messages"of this document.
. . . .

C.2 Collecting information for troubleshooting

Collect following material for investigation using super user access privileges when you request troubleshooting to field engineers and if your manual instructs you to do so.

1. Material collected by FJQSS, fjsnap, or pclsnap

2. Crash dump file (/var/crash/*/*)

3. Output of the following command: "sfcfsck -N -o nolog *device*"

   For *device*, specify a block special file of the file system where the problem occurs.

If the investigation material must urgently be sent, collect the following materials a and b for the initial investigation:

a. Material collected by sfcsnap

b. /var/log/messages*

Note
. . . .

- The above materials 1, 2, a, and b should be collected from all nodes where the GFS Shared File System is shared.

- The above material 3 should be collected when a file system needs to be recovered.

- These materials should be collected as soon as the symptom occurred. Some material might not be collected after rebooting the system.

- The materials a and b are of initial investigation. Collect materials a and b first, and then make sure to collect materials 1 to 3.
. . . .

ⓅPoint

••

Use FJQSS if you want to collect the above material 1 with other middleware products that support FJQSS.

••

The following indicates how to collect materials by FJQSS, fjsnap, pclsnap, and sfcsnap.sh.

Collecting material by FJQSS (Information Collection Tool)

    1. Execute the following command.

```
# /opt/FJSVqstl/fjqss_collect
```

       The product selection menu appears.

    2. Input the number of the product of which you want to collect the investigation material.
       Select either of the following products:

        - PRIMECLUSTER Enterprise Edition

        - PRIMECLUSTER HA Server

       If you select one of the above, you can collect investigation material for PRIMECLUSTER including GFS in a batch.

    3. Press the [Y] key according to the instruction in the prompt.

       After the FJQSS has completed the collection, the name of the output directory of the collected material appears.

    4. Verify that the material have been collected in the directory.

       The following file is created in the output directory of the collected material.

       result*YYYYMMDDHHMMSS*.tar.gz
       (*YYYYMMDDHHMMSS*: time (year, month, day, hour, minute, and second) that the collection started)

    5. Send it to field engineers.

Collecting material by fjsnap

    1. Execute the following command.

```
# /usr/sbin/fjsnap -a filename
```

       Specify a non-existent file name to *filename*. The material is collected with this file name.

    2. Send this file to field engineers.

Collecting material by pclsnap

    1. Execute the following command.

```
# /opt/FJSVpclsnap/bin/pclsnap -a filename
```

       Specify a non-existent file name to *filename*. The material is collected with this file name.

    2. Send this file to field engineers.

Collecting material by sfcsnap

    1. Execute the following command.

```
# /etc/opt/FJSVsfcfs/bin/sfcsnap.sh -a -d dirname
```

       Specify an existent directory to *dirname*. The material is collected under this directory.

    2. Compress this directory and send it to field engineers.

See
. . . .

- For information on how to collect investigation material on PRIMECLUSTER systems, see "PRIMECLUSTER Installation and Administration Guide."

- For FJQSS, see "FJQSS (Information Collection Tool) User's Guide" included in the product's DVD. For how to see "FJQSS (Information Collection Tool) User's Guide", open the following file in the product's DVD from your browser.

  documents/fjqss-manual_sollnx/index_en.html

. . . .

C.2.1 Collecting a crash dump file

Collect a crash dump file as troubleshooting information of the GFS Shared File System in the event of a node panic.

A crash dump will be saved as a file during node startup after the panic. The storage directory will be /var/crash/<*Time when a panic occurred*> as default.

C.2.2 Collecting a core image of the daemon

It may be necessary to collect a core image of each daemon according to the corrective action of the GFS Shared File System.

The following procedure shows how to collect a core image of the sfcfrmd daemon on all the nodes that share the GFS Shared File System using super user access privileges.

1. Determine the process ID

   Determine the process ID using ps(1). For non sfcfrmd daemons, change argument of grep(1).

   ```
   # /bin/ps -e | /bin/grep sfcfrmd <Enter>
     5639 ?        00:00:25 sfcfrmd
   ```

   The beginning of the above output shows a process ID of the sfcfrmd daemon. If the sfcfrmd daemon is not running, the process number will not be output. In such a case, collect troubleshooting information on the other node.

   See
   . . . .
   For details about ps(1) and grep(1), see the online manual page.
   . . . .

2. Acquire a core image

   Acquire the core image of sfcfrmd in the "/var/tmp/sfcfrmd_node1.5639" file using gcore(1). After that, compress the file using tar(1).

   ```
   # /usr/bin/gcore -o /var/tmp/sfcfrmd_node1 5639 <Enter>
     :
   Saved corefile /var/tmp/sfcfrmd_node1.5639
   # /bin/tar czvf /var/tmp/sfcfrmd_node1.5639.tar.gz /var/tmp/sfcfrmd_node1.5639
   <Enter>
     :
   /var/tmp/sfcfrmd_node1.5639
   # /bin/ls -l /var/tmp/sfcfrmd_node1.5639.tar.gz <Enter>
   -rw-rw-r--   1 root     other     1075577  Jun 12  16:30 /var/tmp/
   sfcfrmd_node1.5639.tar.gz
   #
   ```

   See
   . . . .
   For details about tar(1), see the online manual page.
   . . . .

C.3 Common corrective actions against failures

This section describes most common corrective actions against system failures.

C.3.1 Action for I/O errors

The "sfcfs...:read error" message indicates that an I/O error occurred while file system data was being read from a disk. The "sfcfs...:write error" message indicates that an I/O error occurred while file system data was being written to a disk.

If the messages are displayed, recovery processing is required because an I/O path or device is faulty.

Perform the recovery processing using the following procedures:

1. Locate the file system and partition in which the error occurred

   Pinpoint it with **sfcinfo**(8), a GFS Shared File System management command, using the device number in the error message as a key.

   ```
   # sfcinfo -d device-number <Enter>
   ```

2. If an sfcfs...:write error occurs

   This error indicates that file system information updated in memory cannot be reflected on the disk. If this happens, files in the file system should be backed up before unmounting to prevent the updated information from being lost by unmounting.

3. Unmount the file system service in which the error occurred

   Unmount the file system service.

4. Correct the file system in which the error occurred

   Restore data integrity of the file system using **sfcfsck**(8) after creating a copy of the file system data of the failed partition. After the file system is restored, replay update log and perform full check.

   ```
   # sfcfsck -o log special-device-name <Enter>
   # sfcfsck -o nolog special-device-name <Enter>
   ```

5. If the file system in which the error occurred cannot be corrected

   When step 4 cannot repair the file system, swap the faulty part, re-create the file system, and restore the backed up data.

🔖 See
....

For corrective action against an I/O error of a GDS volume, see "Troubleshooting" of "PRIMECLUSTER Global Disk Services Configuration and Administration Guide." For details of backing-up and restoring, see "Chapter 13 File System Backing-up and Restoring."
....

C.3.2 Corrective action in the event of data inconsistency

This section explains procedure when it becomes impossible to mount the GFS Shared File System because super block or partition configuration data differ from each copy data.

Perform the recovery in accordance with the following procedures.

1. Recover super block and partition configuration information in the partition using sfcadm(8).

   ```
   # sfcadm -r /dev/sfdsk/gfs01/dsk/volume01 <Enter>
   ```

2. Recover the consistency of the file system.

   ```
   # sfcfsck /dev/sfdsk/gfs01/dsk/volume01 <Enter>
     :
   # sfcfsck -o f,nolog /dev/sfdsk/gfs01/dsk/volume01 <Enter>
   ```

```
    :
#
```

3. Mount the recovered file system.

```
# sfcmntgl /dev/sfdsk/gfs01/dsk/volume01 <Enter>
```

4. Confirm that the recovered file system can be mounted on all nodes sharing GFS Shared File System.

```
# mount <Enter>
```

5. Confirm on all nodes that the following messages are not output in /var/log/messages when sfcmntgl(8) was performed in step 3.

```
WARNING: sfcfs_mds:2042: fsid: The size of the partition might have been
changed after the file system was made.
```

No need the following operations if the message is not output.

If the message is output, there is a possibility that the GDS volume used for the GFS Shared File System was expanded. The file system can keep on using, but please perform the following operations promptly because GFS does not deal with expanded volume of GDS.

6. Back up the files in the recovered file system. For the way to back up, refer to "13.2.1 Backing-up file by file."

📒 Note

. . . .

Please be sure to perform the backup with each file.

. . . .

7. Unmount the recovered file system.

```
# sfcumntgl /dev/sfdsk/gfs01/dsk/volume01 <Enter>
```

8. Confirm that the recovered file system can be unmounted on all nodes sharing GFS Shared File System.

```
# mount <Enter>
```

9. The parameter information on a GFS Shared File System is evacuated.

```
# sfcmkfs -m /dev/sfdsk/gfs01/dsk/volume01 > /tmp/mkfs_cmd <Enter>
```

The command line used when the file system was created is stored in /tmp/mkfs_cmd by the above procedure.

10. Delete the recovered file system.

```
# sfcadm -D /dev/sfdsk/gfs01/dsk/volume01 <Enter>
```

11. Confirm the file system is deleted.

```
# sfcinfo -a <Enter>
```

12. Depending on the situation, re-create the partition.

Please re-create GDS volume only when you will return to the volume size before expanding GDS volume and resume manipulation. It is not necessary to re-create if resume manipulation in the volume size after expanding GDS volume.

Please refer to "Operation" "GDS Management View" in "PRIMECLUSTER Global Disk Services Configuration and Administration Guide" about the operation method of GDS volume.

13. Create the GFS Shared File System using the same partition or the partition created in step 12.

```
# sfcmkfs -o node=sunny,moony /dev/sfdsk/gfs01/dsk/volume01 <Enter>
```

When creating the GFS Shared File System with the same parameters that used before volume expansion of GDS is performed, it is created using the information which evacuated in step 9.

```
# /bin/sh /tmp/mkfs_cmd <Enter>
```

14. Confirm that the GFS Shared File System is created.

```
# sfcinfo -a <Enter>
```

15. Mount the created file system.

```
# sfcmntgl /dev/sfdsk/gfs01/dsk/volume01 <Enter>
```

16. Restore the file backed up in step 6. For the way to restore, refer to "13.3.1 File-by-file restoring."

C.3.3 Action to Be Taken When Unmounted by umount(8)

This section describes the action to be taken when a GFS Shared File System is unmounted by umount(8).

Perform the recovery processing taking the following procedures.

1. Unmount all of the file systems on the node where the file system was unmounted using umount(8).

```
# mount | /bin/grep sfcfs <Enter>
 :
/dev/sfdsk/gfs/dsk/volume1 on /mnt/fs1 type sfcfs (rw)
/dev/sfdsk/gfs/dsk/volume2 on /mnt/fs2 type sfcfs (rw)
# sfcumount /mnt/fs1 <Enter>
# sfcumount /mnt/fs2 <Enter>
```

2. Stop the operation of the GFS Shared File System.

[RHEL6]

```
# /etc/init.d/sfcfsrm stop <Enter>
```

[RHEL7]

```
# systemctl stop fjsvgfsfsrm2.service <Enter>
# systemctl stop fjsvgfsfsrm.service <Enter>
```

3. Start the operation of the GFS Shared File System.

[RHEL6]

```
# /etc/init.d/sfcfsrm start <Enter>
```

[RHEL7]

```
# systemctl start fjsvgfsfsrm.service <Enter>
# systemctl start fjsvgfsfsrm2.service <Enter>
```

C.3.4 Corrective action when the sfcfrmd daemon is not started

This section describes corrective actions when the sfcfrmd daemon is not started during node is started up, the node enters multi-user mode, or CF is started by the GUI.

For integrated file system access being maintained, startup of the sfcfrmd daemon will be suspended until a quorum exists.

If activation of the daemon is suspended, the following message will be output;

```
WARNING: sfcfsrm:5001: Starting the sfcfrmd daemon was suspended because quorum
dose not exist
```

Normally, no corrective action should be taken to activate the sfcfrmd daemon because it will be started as soon as a quorum exists.

In the following cases, a quorum does not exist, so corrective action should be taken to start operation of the GFS Shared File System.

- A cluster partition error exists.

- When all the cluster nodes or CF are stopped then part of nodes or CF are started and operated due to a failure.

If GFS cannot be operated because the sfcfrmd daemon is not activated, using the following steps:

Procedure 1. Check the state of all the cluster nodes.

Connect all the operating nodes and check if the same state is displayed using the **cftool(1M)** command or Cluster Admin GUI.

```
# cftool -n <Enter>
Node  Number State      Os     Cpu
sunny 1       UP        Linux  Pentium
monny 2       UP        Linux  Pentium
```

If the node state is different on all the operating nodes, a cluster partition error exists.

📑 See
....
For details about **cftool(1M)**, see "Node details" of "PRIMECLUSTER Cluster Foundation (CF) Configuration and Administration Guide."
....

Procedure 2. If a cluster partition error occurs, take the following corrective action:

- When a LEFTCLUSTER node exists

  If Shutdown Facility (SF) is running properly on all the cluster nodes, it will solve the cluster partition error, so no corrective action is necessary. If SF is not running properly, or forced shutdown of the node through SF fails, you need to recover the node manually. Take corrective action according to "Caused by a cluster partition" of "PRIMECLUSTER Cluster Foundation (CF) Configuration and Administration Guide."

- When a LEFTCLUSTER node does not exist

  - Perform a CF or node restart

    Take corrective action according to "Join-related problems" of "PRIMECLUSTER Cluster Foundation (CF) Configuration and Administration Guide."

  - Activate all the CF's

    After stopping all the CFs, solve the cluster partition error then restart CF according to "12.6 How to start up CF by the GUI when a GFS Shared File System is used."

  - Activate all nodes

    After stopping all nodes by **shutdown(8)**, solve the cluster partition error then restart all nodes.

If startup of a node is suspended, complete the startup of a node using operations below, before executing **shutdown(8)** on each node.

1. Confirm the process ID of the sfcfsrm script by **ps(1)**.

   [RHEL6]

   The beginning column of the output below shows a process ID of the sfcfsrm script.

   ```
   # /bin/ps -e | /bin/grep sfcfsrm <Enter>
   18550 ?        00:00:00 S57sfcfsrm
   ```

   [RHEL7]

   The second column of the output below shows a process ID of the sfcfsrm script.

   ```
   # /bin/ps -ef | /bin/grep sfcfsrm | /bin/grep -v grep <Enter>
   root     18550     1  0 02:35 ?        00:00:00 /bin/bash -c /opt/FJSVsfcfs/
   lib/systemd/sfcfsrm  start
   ```

2. Stop the sfcfsrm script by **kill(1)**.

   ```
   # /bin/kill -9 18550 <Enter>
   ```

If you want to resume operation immediately with a part of the nodes, take the following action:

1. Decide which nodes will be used in the cluster after performing "Procedure 1" (above).
   Typically, you will choose the nodes that have the largest number of nodes in it or the one where the most important hardware is connected or the most important application runs. In addition, the nodes that will be used need to include at least one MDS node for using the GFS Shared File System.

2. Stop all nodes that will not be used with **shutdown(8)**.
   If startup of a node is suspended, complete the startup of a node before executing **shutdown(8)**. Procedure is shown [Note] (above).

3. Check the state of all the nodes that will be used using "Procedure 1" again. The state of all the nodes must be the same.

4. Forcibly restart the sfcfrmd daemon that has been suspended by executing the sfcfrmstart(8) command on all the nodes that will be used.
   An example of **sfcfrmstart(8)** is shown "Procedure 3" (below).

For details about **shutdown(8)**, **ps(1)**, **grep(1)** and **kill(1)**, see the online manual page.

Procedure 3. If a cluster partition error does not exist, forcibly restart the sfcfrmd daemon that has been suspended by executing the sfcfrmstart(8) command on all the nodes where operation will be resumed.

```
# sfcfrmstart -f <Enter>
```

C.3.5 Corrective action when the mount processing fails

This section describes common corrective actions, if the mount processing fails due to the blockade on mounting the file system, or if any of the following messages is output.

```
ERROR: sfcfs_mds:1004: can't execute primary MDS
ERROR: sfcfs_mds:1005: can't execute secondary MDS
```

Check following points:

- Setting of network is normal.

  Check whether all the IP addresses of the nodes that share the file system are correctly registered in host database (such as /etc/hosts). For details, see "11.2.2 Registering information in host database" in this manual.

- State of the file system configuration information is normal.

  Execute the following commands and check that the host name of the shared node which the **sfcinfo(8)** command and the **sfcrscinfo(8)** command display and the host name specified by the "**sfcadm -m**" command line in the backup file created by the execution of the **sfcgetconf(8)** command are the same.

    - sfcinfo -n <*representative partition of the file system*>

    - sfcrscinfo -m <*representative partition of the file system*>

    - sfcgetconf -d <*representative partition of the file system*> <*output destination file name*>

  When not the same, please change the -**m** option to the -**u** option of the **sfcadm(8)** command execution line obtained by the **sfcgetconf(8)** command, and then execute the command line.

If no problems are found in the checks above, see also corrective actions described in "Appendix A List of Messages."

See
··
For details on above commands, see sfcinfo(8), sfcrscinfo(8), sfcgetconf(8) in this manual.
··

C.3.6 Corrective action when write fails by ENOSPC

This section describes corrective actions when the write processing fails by ENOSPC (no free areas) even if data space of the file system contains free space.

When V-data area usage exceeds 80%, the write processing may fail by ENOSPC.

Execute the following command to the file system with a write failure, and then check the number of V-data blocks (vdata) and the total number of available V-data blocks (free).

```
sfcdf –v <representative partition of the file system>
```

Deleting unused files or directories, or moving them into other file systems helps you to reduce V-data area usage.

See
··
For details when the write processing fails by ENOSPC, see "1.8.19 Notes on writing file data on the GFS Shared File System."
··

See
··
For details on sfcdf, see sfcdf(8) in this manual.
··

Appendix D Startup Scripts and Startup Daemons, and Port Numbers in GFS [RHEL6]

This appendix provides the explanations on scripts and daemons that are started by GFS, and the port numbers being used.

D.1 Explanation Formats

Scripts and daemons, and port numbers are explained with the following formats:

Startup script

Name of startup script.

Function

Function of startup script and daemon.

Effect if stopped

Effect if startup script and daemon are stopped.

Startup daemon

Daemon started by startup script.

Utilized port

Port

Port number.

Protocol

Protocol - TCP or UDP.

Send/Receive

"s" if port sends data, "r" if it receives data, "s, r" for both.

Network

Utilized network - any of Cluster interconnect, administrative LAN, or public LAN.

Target

Node that uses the port.

Communication target

Port

Port number of communication target.

Target

Node or device that uses the port of the communication target.

Remarks

Remarks.

D.2 Startup Script Lists

S57sfcfsrm

Function

Startup control for monitoring facility of GFS shared file system, mount control for GFS shared file system.

Effect if stopped

Functions of GFS shared file system cannot be used.

Startup daemon

/usr/lib/fs/sfcfs/sfcpncd
/usr/lib/fs/sfcfs/sfcprmd
/usr/lib/fs/sfcfs/sfchnsd
/usr/lib/fs/sfcfs/sfcfrmd
/usr/lib/fs/sfcfs/sfcfsd
/usr/lib/fs/sfcfs/sfcfsmg

Utilized port

Port	Protocol	Send/ Receive	Network	Target	Communication target	
					Port	Target
9300 (*1)	TCP	s, r	Interconnect	Cluster node	ANY	Remote cluster node
9200-9263 (*2)	TCP	s, r	Interconnect Administrative LAN	Cluster node	ANY	Local and remote cluster nodes

Remarks

(*1) No. 9300 is set to support the service name "sfcfsrm."

(*2) From No. 9200 to No. 9263 are set to support the service names from sfcfs-1 to sfcfs-64.

Appendix E systemd Services and Startup Daemons, and Port Numbers in GFS [RHEL7]

This appendix provides the explanations on the systemd services and the startup daemons in GFS, and the port numbers being used.

E.1 Explanation Formats

systemd services and startup daemons, and port numbers are explained with the following formats:

Unit name

Name of Unit.

Function

Function of Unit.

Effect if stopped

Effect if Unit is stopped.

Dependency with another Unit

Requires

Prior Unit that the local Unit needs. If the specified Unit failed to start, the local Unit does not start either.

Wants

Prior Unit that the local Unit needs. Even if the specified Unit failed to start, the local Unit starts.

Before

Another Unit that starts after the local Unit.

After

Another Unit that starts before the local Unit.

Startup daemon

Daemon started by Unit.

Utilized port

Port

Port number.

Protocol

Protocol - TCP or UDP.

Send/ Receive

"s" if port sends data, "r" if it receives data, "s, r" for both.

Network

Utilized network - any of Cluster interconnect, administrative LAN, or public LAN.

Target

Node that uses the port.

Communication target

Port

Port number of communication target.

Target

    Node or device that uses the port of the communication target.

Remarks

    Remarks.

E.2　systemd Service Lists

WantedBy: multi-user.target

fjsvgfsfsrm.service

Function

    Startup control for monitoring facility of GFS shared file system, mount control for GFS shared file system.

Effect if stopped

    Functions of GFS shared file system cannot be used.

Dependency with another Unit

    Requires

        None

    Wants

        None

    Before

        smawrrms.service

    After

        fjsvclctrl.service
        fjsvclrmgr2.service

Startup daemon

    /usr/lib/fs/sfcfs/sfcpncd
    /usr/lib/fs/sfcfs/sfcprmd
    /usr/lib/fs/sfcfs/sfchnsd
    /usr/lib/fs/sfcfs/sfcfrmd
    /usr/lib/fs/sfcfs/sfcfsd
    /usr/lib/fs/sfcfs/sfcfsmg

Utilized port

Port	Protocol	Send/ Receive	Network	Target	Communication target	
					Port	Target
9300 (*1)	TCP	s, r	Interconnect	Cluster node	ANY	Remote cluster node
9200-9263 (*2)	TCP	s, r	Interconnect Administrative LAN	Cluster node	ANY	Local and remote cluster nodes

Remarks

(*1) No. 9300 is set to support the service name "sfcfsrm."

(*2) From No. 9200 to No. 9263 are set to support the service names from sfcfs-1 to sfcfs-64.

fjsvgfsfsrm2.service

Function

Stop control for monitoring facility of GFS shared file system, unmount control for GFS shared file system.

Effect if stopped

GFS shared file system cannot be stopped properly at the system stop.

Dependency with another Unit

Requires

None

Wants

None

Before

smawrrms.service

After

fjsvclctrl.service
fjsvclrmgr2.service
fjsvgfsfsrm.service

Startup daemon

None

Utilized port

None

Remarks

None

Appendix F Release information

This appendix describes the main contents for change of this manual.

F.1 Change of a function

This section explains outline, content, and details (the pertinence point in the manual) of function added or changed from a former version and level.

Number	VL	Function name	Contents	Details
1	4.3A00	Changing the maximum number of sharable nodes	The maximum number of sharable nodes is four.	- 1.1.2 File system sharing with PRIMEQUEST - 1.3.1 Uninterrupted operation in case of node failure - 1.5 Extendibility - 1.5.2 Addition of Shared node - 1.7 Upper Limits of the File System - Maximum number of sharable nodes: 4 nodes - 1.8.16 Notes when starting and stopping - 2.2 Component configuration - 3.1 MDS Failure Recovery - 3.1.1 Automatic recovery for primary MDS failures - 4.3.2 Node layout - 5.1 Management Partition - 6.3 Starting GFS Management View - 7.1 Screen Configuration - 7.2.1 General - 7.2.2 Operation - 8.2.2 Creating the management partition, registering node configuration information and starting sfcfrmd daemon - 8.3.2 Stopping sfcfrmd daemon - 8.3.3 Adding node configuration information to the management partition - 8.3.4 Starting sfcfrmd daemon - 9.6.1 Backup of the management partition information - 10.2.1 Creating a file system - 10.3.3 Changing the shared node information - 11.2.2.2 Examples of creating a representative file system

Number	VL	Function name	Contents	Details
				- 11.2.2.4 Setting MDS operational information
				- 11.2.4.2 Mount
				- 11.2.5 Checking file system status
				- 11.3.1.2 Unmount
				- 11.5.3.1 Adding node information
				- 11.5.3.2 Deleting node information
				- 11.6.2 Re-creating the file system
				- 11.7.2 Changing the MDS operational information
				- 12.4.3 Displaying file system management state
				- 13.3.3.1 Resetting the partition information
				- A.4.4 sfcmount command
				- B.1.4 sfcmount(8) Mount the GFS Shared File System on the local node
				- B.1.5 sfcumount(8) Unmount the GFS Shared File System from the local node
				- B.2.2 sfcadm(8) Change partition information setting
				- B.2.9 sfcrscinfo(8) Display file system information
				- D.2.2.10 Addition of the i-node's cache memory release function at the time of accessing many files
				- Glossary
				- "shared node"
2	4.3A00	Changing the maximum size of the file system	The maximum size of the file system is 32 terabytes.	- 1.7 Upper Limits of the File System
				- Maximum file system capacity: 32 terabytes - 1 kilobyte
				- Maximum partitions size: 1 terabyte
				- Maximum number of directories per file system: 16 mega
				- 11.2.2.3 Customizing a GFS Shared File System
				- Meta-data area size
				- Maximum data area size
				- A.4.3 sfcmkfs command
				- sfcmkfs:ERROR: special exceeds the maximum size of a partition
				- sfcmkfs:ERROR: total file system size is too large

Number	VL	Function name	Contents	Details
				- A.5.2 sfcadd command
				- sfcadd: ERROR: 0026:
				- B.1.3 sfcmkfs(8) Construct the GFS Shared File System
				- maxdsz
				- metasz
				- nbpi
				- D.2.3 Incompatibilities from 4.2A00
				- D.2.4 Incompatibilities from 4.2A30
3	4.3A00	Addition of "Advisory file lock"	Addition of the Advisory file lock in the "application interface is commonly available:"	- 1.1 File System Overview
4	4.3A00	Delete of kernel of Linux2.4	Deletes it because it is not supported in the GFS 4.3A00 Linux version	- Preface
				- 1.8.4 Notes on use of logical volumes of GDS
				- 1.8.9 Deleting a file being created from the other node
				- 1.8.15 Notes on file lock function
				- 4.3.5 For future expansion
				- 10.1.3 Change (partition addition)
				- 10.2.1 Creating a file system
				- 10.3.2 Changing the file system configuration (partition addition)
				- 11.1.3 Change (partition addition)
				- 11.2.2.2 Examples of creating a representative file system
				- 11.2.2.3 Customizing a GFS Shared File System
				- 11.4 Change (partition addition)
				- 12.1 File System Management Commands
				- 12.3 Extending a File System
				- 12.5.6 Setup procedure of adding file data partitions to GFS Shared File Systems of cluster applications
				- B.1.3 sfcmkfs(8) Construct the GFS Shared File System
				- B.2.1 sfcadd (8) Adds file data partitions
				Deleted contents:
				- Unavailable functions on Linux kernel version 2.4

Number	VL	Function name	Contents	Details
				- Open processing from other nodes to a file mapped by mmap(2) [Linux2.4]
				- Creating a regular file by mknod(2) [Linux2.4]
5	4.3A20	IPv6 support	IPv6 is supported on networks (cluster interconnect) used by GFS.	- 1.8.2.1 Functions not provided by the GFS Shared File System
				- 1.8.22 Notes in IPv6 environments
6	4.3A20	Changing the maximum number of sharable nodes	The maximum number of sharable nodes is two.	- 1.1.2 File system sharing with PRIMEQUEST
				- 1.3.1 Uninterrupted operation in case of node failure
				- 1.5 Extendibility
				- 1.7 Upper Limits of the File System
				- Maximum number of sharable nodes: 2 nodes
				- 1.8.16 Notes on stopping a node
				- 1.8.17 Notes when CF is started and stopped by the GUI
				- 2.2 Component configuration
				- 3.1 MDS Failure Recovery
				- 3.1.1 Automatic recovery for primary MDS failures
				- 4.3.2 Node layout
				- 5.1 Management Partition
				- 6.3 Starting GFS Management View
				- 7.1 Screen Configuration
				- 7.2.1 General
				- 7.2.2 Operation
				- 8.2.2 Creating the management partition, registering node configuration information and starting sfcfrmd daemon
				- 8.3.2 Stopping sfcfrmd daemon
				- 8.3.3 Adding node configuration information to the management partition
				- 8.3.4 Starting sfcfrmd daemon
				- 9.6.1 Backup of the management partition information
				- 10.2.1 Creating a file system
				- 10.3.3 Changing the shared node information
				- 11.2.3.2 Examples of creating a representative file system

Number	VL	Function name	Contents	Details
				- 11.2.3.4 Setting MDS operational information
				- 11.2.5.2 Mount
				- 11.2.6 Checking file system status
				- 11.3.1.2 Unmount
				- 11.5.4.1 Adding node information
				- 11.5.4.2 Deleting node information
				- 11.6.2 Re-creating the file system
				- 11.7.2 Changing the MDS operational information
				- 12.4.3 Displaying file system management state
				- 13.3.3.1 Resetting the partition information
				- A.4.4 sfcmount command
				- B.1.4 sfcmount(8) Mount the GFS Shared File System on the local node
				- B.1.5 sfcumount(8) Unmount the GFS Shared File System from the local node
				- B.2.2 sfcadm(8) Change partition information setting
				- B.2.9 sfcrscinfo(8) Display file system information
				- Glossary
				- "shared node"
				Deleted contents:
				- 1.5.2 Addition of Shared node
7	4.3A20	Changing the maximum size of the file system	The maximum size of the file system is 1 terabyte.	- 1.7 Upper Limits of the File System
				- Maximum file system capacity: 1 terabyte - 1 kilobyte
				- Maximum number of directories per file system: 1 mega
				- 11.2.3.3 Customizing a GFS Shared File System
				- Meta-data area size
				- Maximum data area size
				- A.5.2 sfcadd command
				- sfcadd: ERROR: 0026:
				- B.1.3 sfcmkfs(8) Construct the GFS Shared File System
				- maxdsz

Number	VL	Function name	Contents	Details
				- metasz
				- nbpi
				- E.2.3 Incompatibilities from 4.2A00
				- E.2.4 Incompatibilities from 4.2A30
				Deleted contents:
				- E.2.3.1 Change in error message of sfcadd(8)
				- E.2.4.1 Change in error message of sfcadd(8)
8	4.3A30	IPv6 support	IPv6 is available for communication between MDS and AC.	- 1.8.2.1 Functions not provided by the GFS Shared File System
				- 10.1.1 Creation
				- 10.1.4 Change (shared node information)
				- 10.2.1 Creating a file system
				- 10.3.1 Changing the file system attributes
				- 11.1.1 Creation
				- 11.1.4 Change (shared node information)
				- 11.2 Creation
				- 11.2.2 Registering information in host database
				- 11.2.5 Selecting a communication protocol
				- 11.3.2.2 Tuning file system
				- 11.5 Change (shared node information)
				- 11.5.3 Registering information in host database (When adding node information)
				- 14.1.3 Communication protocol
				- A.1.2 Warning messages
				- A.1.3 Information messages
				- A.2.4 Error messages (MDS (sfcfsmg daemon))
				- A.2.6 Error messages (sfcfs_mount command)
				- A.2.9 Warning messages (MDS (sfcfsmg daemon))
				- A.4.3 sfcmkfs command
				- A.5.3 sfcadm and sfcnode commands
				- C.3.5 Corrective action when the mount processing fails
				- E.2.1 Incompatibilities from 4.1A20
				- E.2.1.14 Change in AC warning messages

Number	VL	Function name	Contents	Details
				- E.2.2 Incompatibilities from 4.1A30
				- E.2.2.12 Change in AC warning messages
				- E.2.3 Incompatibilities from 4.2A00
				- E.2.3.3 Change in AC warning messages
				- E.2.4 Incompatibilities from 4.2A30
				- E.2.4.3 Change in AC warning messages
				- E.2.5 Incompatibilities from 4.3A00
				- E.2.5.3 Change in AC warning messages
				- E.2.6 Incompatibilities from 4.3A10
				- E.2.6.1 Change in AC warning messages
				- E.2.7 Incompatibilities from 4.3A20
				- E.2.7.1 Change in AC warning messages
				Deleted contents:
				- 1.8.22 Notes in IPv6 environments
9	4.3A40	FJQSS support	FJQSS is available as the investigation material collection tool for GFS.	- 12.2.1.8 Other messages - Appendix A List of Messages - C.2 Collecting information for troubleshooting
10	4.3A40	RHEL7 environment support	GFS is available in the RHEL7 environment.	- 1.8.8 Notes on operation stop - 1.8.23 Notes on lseek(2) with SEEK_DATA or SEEK_HOLE - 4.1 Mount and unmount opportunity of the file system at the time of system starting and ending - 4.3.4 LAN selection - A.2.4 Error messages (MDS (sfcfsmg daemon) - C.3.3 Action to Be Taken When Unmounted by umount(8) - C.3.4 Corrective action when the sfcfrmd daemon is not started - Appendix E systemd Services and Startup Daemons, and Port Numbers in GFS [RHEL7]
11	4.3A40	Support of the Migration function of KVM	The Migration function of KVM is available in the environment where GFS is used.	- 1.8.22 Notes on using the Migration function in a KVM environment
12	4.3A40	RHEL7.1 environment support	GFS is available in the RHEL7.1 environment.	- 1.8.16 Notes on starting and stopping a node - 1.8.17 Notes on starting and stopping CF by GUI

Number	VL	Function name	Contents	Details
				- 1.8.24 Notes on updating OS to RHEL7.1 or later version
13	4.5A00	Support of the function to start Web-Based Admin View by using Java Web Start	Web-Based Admin View can be started by using Java Web Start.	- Chapter 6 Starting and Exiting the Management View - 6.1 Starting Web-Based Admin View - 6.5 Exiting Web-Based Admin View
14	4.5A00	RHOSP environment support	GFS is available in RHOSP environment.	- 1.8.25 Notes on RHOSP environment

GFS : Global File Services

GDS : Global Disk Services

Note

....

The number one and two are unsupported functions that are described as restrictions in PRIMECLUSTER Software Release Guide and Installation Guide for 4.3A10 or earlier. For the number six and seven, descriptions of the unsupported functions are deleted from the manual; however, the available scope of the function is not changed.

....

F.2   Information of Incompatibility

This section explains incompatibility of the GFS Shared File System.

F.2.1   Incompatibilities from 4.1A20

There are some incompatibilities with the following functions compared to 4.1A20:

Table F.1 Incompatibilities from 4.1A20

Number	Item
1	Changes operation of truncate(2), ftruncate(2), creat(2), and open(2) specified with O_TRUNC from the other node to the file where mmap(2) is being executed
2	Changes port numbers reserved by GFS
3	Changes the type of a special file specified for the GFS command argument
4	Change in error message of sfcdf(8)
5	Change in error message of sfcfsck(8)
6	Check strengthening of parameter specified for -o *specific_options* option of sfcmkfs(8)
7	Changes allocated reserved area size by writing operations
8	Add the upper limit of concurrently open files
9	Addition of the i-node's cache memory release function at the time of accessing many files
10	Omitted value of the communication timeout between nodes
11	Deletion of the sfcfrmd information message
12	Change in operation at the time of specifying the mount option atime
13	Cannot create sockets
14	Cannot create named pipes

GFS : Global File Services

F.2.1.1 Changes operation of truncate(2), ftruncate(2), creat(2), and open(2) specified with O_TRUNC from the other node to the file where mmap(2) is being executed

Incompatibility description

When the following system calls are issued on another node for a file within mmap(2) on the GFS Shared File System, action of them is changed.

- truncate(2)

- ftruncate(2)

- creat(2)

- open(2) specified with O_TRUNC

Before change: Normal exit

After change: Abnormal exit (return code: -1, errno: EBUSY)

Corrective action

None.

F.2.1.2 Changes port numbers reserved by GFS

Incompatibility description

Port numbers reserved by GFS in the /etc/services file have been changed.

Before change: 9200 - 9300

After change: 9200 - 9263, 9300

Corrective action

No corrective action is required.

F.2.1.3 Changes the type of a special file specified for the GFS command argument

Incompatibility description

The type of special files has been changed in the following commands:

- sfcdf(8)

- sfcfsck(8)

- sfcmkfs(8)

- sfcadm(8)

- sfcgetconf(8)

- sfcnode(8)

- sfcsetup(8)

Before change: Specify a character special file

After change: Specify a block special file. A character special file cannot be specified.

Corrective action

If a special file is specified for the argument of the above command, specify a block special file. If a character special file is specified, the commands will output error messages and abnormal exit.

F.2.1.4 Change in error message of sfcdf(8)

Incompatibility description

In the sfcdf(8) command, the error message output at the following conditions has been changed.

1) -m option is specified. And,

2-1) A special file name which does not exist is specified for the argument. Or,

2-2) A special file name other than the GFS Shared File System is specified for the argument.

Before change:

For 2-1)

sfcdf: ERROR: cannot open *special*: *errmsg*

For 2-2)

sfcdf: ERROR: invalid device: *special*

After change:

sfcdf: ERROR: *special* is not a sfcfs file system

Corrective action

None.

F.2.1.5 Change in error message of sfcfsck(8)

Incompatibility description

In the sfcfsck(8) command, the error message output at the following conditions has been changed.

A special file name other than the GFS Shared File System is specified for the argument.

Before change:

device-name is not for sfcfs

After change:

Can't check *device-name*

Corrective action

None.

F.2.1.6 Check strengthening of parameter specified for -o specific_options option of sfcmkfs(8)

Incompatibility description

The operation of sfcmkfs(8) for the following conditions was changed.

1) Any of the following -o *specific_options* options is specified.

free=*n*
logsz=*n*
maxnode=*n*
metasz=*n*
mfree=*n*
nbpi=*n*
nblkpext=*n*

2) The first character of *n* is a digit. And,

3) When characters other than the figure are included since the second character of *n*.

Before change:

The string only of a digit consecutive from the first character in *n* is recognized as a parameter.

Example) "-o free=9a0" is considered to be "-o free=9".

After change:

The following error messages are output, and it terminates abnormally.

sfcmkfs:ERROR: Illegal option : *value*

Corrective action

Please specify the string only of the digit for the parameter.

F.2.1.7 Changes allocated reserved area size by writing operations

Incompatibility description

When executing write processing for a file on the GFS Shared File System, reserve area size by a node basis was increased. Within this change, used data area that is obtained by **df(1)**, **sfcdf(8)**, or **statfs(2)** is increased, after executing write processing.

As the example, an output result of **df(1)** is shown below in case an 1 kilobyte file is created on a GFS Shared File System just after making.

Before change:

```
Filesystem            1K-blocks      Used Available Use% Mounted on
/dev/sfdsk/c1/dsk/vol1  56569128      5200  50907016   1% /mnt/gfs1
```

After change:

```
Filesystem            1K-blocks      Used Available Use% Mounted on
/dev/sfdsk/c1/dsk/vol1  56569128     20480  50891736   1% /mnt/gfs1
```

Reserve area is increased because it is interpreted as used date area (column of "Used"), and available data area (column of "Available") is decreased.

Corrective action

None.

F.2.1.8 Add the upper limit of concurrently open files

Incompatibility description

When 5000 different files are being concurrently opened on a GFS Shared File System, it was changed that operation of **open(2)** or **creat(2)**'s execution for 5001th file.

Before change: Normal exit

After change: Abnormal exit (return code: -1, errno: ENFILE)

Corrective action

Please check the following message in **/var/log/messages**.

```
kernel: VFS: file-max limit number reached
```

If this message is not output, 5000 different files are already being opened on the GFS Shared File System where the file tried to open exists.

Please execute **close(2)** to files that are being opened and can be closed (have already been transacted).

If you need to open 5000 and over files on the GFS Shared File System, please disperse the target files to multiple GFS Shared File Systems.

F.2.1.9 Addition of the i-node's cache memory release function at the time of accessing many files

Incompatibility description

When the amount of i-node's cache memory used by MDS exceeded 30000, processing that decreases the amount of i-node's cache memory from 30000 to 22000 was added.

Before change:

The function of releasing i-node's cache memory used by MDS doesn't work.

After change:

1. When the amount of i-node's cache memory used by MDS exceeds 30000, the following message is output and release processing of i-node's cache memory will be started.

   ```
   NOTICE: sfcfs_mds:3020: mountpoint: The total number of vnodes used reached the
   upper bound. vnodes used=totalnum: upper bound=30000: lower bound=22000
   ```

2. When the i-node's cache memory used by MDS has been released until 22000, release processing will stop with the following indication.

   ```
   NOTICE: sfcfs_mds:3021: mountpoint: The total number of vnodes used(=totalnum) is
   the lower bound(=22000).
   ```

Corrective action

None.
But the i-node's cache memory release is in progress in the period of between the two message indications, "NOTICE : sfcfs_mds:3020" and "NOTICE : sfcfs_mds:3021". During release processing, file accessibility on the file system will be affected. Please minimize file access.

F.2.1.10 Communication timeout value between nodes

Incompatible contents

In the monitoring of the node by communication, the default of timeout value is changed after a communicative response is lost until it estimates a partner node to be an abnormal state.

Before change: 60 seconds

After change: 180 seconds

Corrective action

No corrective action is required

Information

You do not have to change the communication timeout between nodes. The value can be changed.

See

Please refer to "14.1.2 Communication timeout value", "10.1.2 Change (in file system attributes)" and "11.1.2 Change (in file system attributes)" about the change method of the communication timeout value between nodes.

F.2.1.11 Deletion of information message of sfcfrmd daemon

Incompatibility description

The information message output at the time of starting sfcfrmd is deleted.

Before change: The following messages are output.

```
sfcfrmd[pid]: INFO: sfcfrmd:0001: set number of threads 16
sfcfrmd[pid]: INFO: sfcfrmd:0002: set service port number 9300
sfcfrmd[pid]: INFO: sfcfrmd:0003: set connection timeout value 30 [sec]
```

After change: The above message is not output.

Corrective action

No corrective action is required.

F.2.1.12 Change in operation at the time of specifying the mount option atime

Incompatibility description

Operation at the time of specifying the mount option **atime** has been changed.

Before change:

The last access time of a file is updated for each file access.

After change:

The last access time of a file is updated when the previous last access time is the same as or older than the last modify time or the last change time. Also, when the previous last access time of a file is old for over a day, it is updated.

Corrective action

To execute the operation performed before change, change the mount option which is described in **/etc/fstab** from **atime** to **strictatime**.

F.2.1.13 Cannot create sockets

Incompatibility description

Operation when creating sockets on the GFS Shared File System by specifying S_IFSOCK in the second argument of mknod(2) has been changed.

Before change: Normal exit

After change: Abnormal exit (return code: -1, errno: EPERM)

Corrective action

None.

F.2.1.14 Cannot create named pipes

Incompatibility description

Operation when creating named pipes on the GFS Shared File System by specifying S_IFIFO in the second argument of mknod(2) has been changed.

Before change: Normal exit

After change: Abnormal exit (return code: -1, errno: EPERM)

Corrective action

None.

F.2.2 Incompatibilities from 4.2A00

There are some incompatibilities with the following functions compared to 4.1A30:

Table F.2 Incompatibilities from 4.1A30

Number	Item
1	Changes operation of truncate(2), ftruncate(2), creat(2), and open(2) specified with O_TRUNC from the other node to the file where mmap(2) is being executed
2	Changes port numbers reserved by GFS
3	Changes the type of a special file specified for the GFS command argument
4	Change in error message of sfcdf(8)
5	Change in error message of sfcfsck(8)
6	Check strengthening of parameter specified for -o *specific_options* option of sfcmkfs(8)
7	Changes allocated reserved area size by writing operations
8	Add the upper limit of concurrently open files
9	Addition of the i-node's cache memory release function at the time of accessing many files
10	Change in operation at the time of specifying the mount option atime
11	Cannot create sockets
12	Cannot create named pipes

GFS : Global File Services

F.2.2.1 Changes operation of truncate(2), ftruncate(2), creat(2), and open(2) specified with O_TRUNC from the other node to the file where mmap(2) is being executed

Incompatibility description

When the following system calls are issued on another node for a file within mmap(2) on the GFS Shared File System, action of them is changed.

- truncate(2)

- ftruncate(2)

- creat(2)

- open(2) specified with O_TRUNC

Before change: Normal exit

After change: Abnormal exit (return code: -1, errno: EBUSY)

Corrective action

None.

F.2.2.2 Changes port numbers reserved by GFS

Incompatibility description

Port numbers reserved by GFS in the /etc/services file have been changed.

Before change: 9200 - 9300

After change: 9200 - 9263, 9300

Corrective action

No corrective action is required.

F.2.2.3 Changes the type of a special file specified for the GFS command argument

Incompatibility description

The type of special files has been changed in the following commands:

- sfcdf(8)

- sfcfsck(8)

- sfcmkfs(8)

- sfcadm(8)

- sfcgetconf(8)

- sfcnode(8)

- sfcsetup(8)

Before change: Specify a character special file

After change: Specify a block special file. A character special file cannot be specified.

Corrective action

If a special file is specified for the argument of the above command, specify a block special file. If a character special file is specified, the commands will output error messages and abnormal exit.

F.2.2.4 Change in error message of sfcdf(8)

Incompatibility description

In the sfcdf(8) command, the error message output at the following conditions has been changed.

1) -m option is specified. And,

2-1) A special file name which does not exist is specified for the argument. Or,

2-2) A special file name other than the GFS Shared File System is specified for the argument.

Before change:

For 2-1)

sfcdf: ERROR: cannot open *special*: *errmsg*

For 2-2)

sfcdf: ERROR: invalid device: *special*

After change:

sfcdf: ERROR: *special* is not a sfcfs file system

Corrective action

None.

F.2.2.5 Change in error message of sfcfsck(8)

Incompatibility description

In the sfcfsck(8) command, the error message output at the following conditions has been changed.

A special file name other than the GFS Shared File System is specified for the argument.

Before change:

device-name is not for sfcfs

After change:

Can't check *device-name*

Corrective action

None.

F.2.2.6 Check strengthening of parameter specified for -o specific_options option of sfcmkfs(8)

Incompatibility description

The operation of sfcmkfs(8) for the following conditions was changed.

1) Any of the following -o *specific_options* options is specified.

free=*n*
logsz=*n*
maxnode=*n*
metasz=*n*
mfree=*n*
nbpi=*n*
nblkpext=*n*

2) The first character of *n* is a digit. And,

3) When characters other than the figure are included since the second character of *n*.

Before change:

The string only of a digit consecutive from the first character in *n* is recognized as a parameter.

Example) "-o free=9a0" is considered to be "-o free=9".

After change:

The following error messages are output, and it terminates abnormally.

sfcmkfs:ERROR: Illegal option : *value*

Corrective action

Please specify the string only of the digit for the parameter.

F.2.2.7 Changes allocated reserved area size by writing operations

Incompatibility description

When executing write processing for a file on the GFS Shared File System, reserve area size by a node basis was increased. Within this change, used data area that is obtained by **df(1)**, **sfcdf(8)**, or **statfs(2)** is increased, after executing write processing.

As the example, an output result of **df(1)** is shown below in case an 1 kilobyte file is created on a GFS Shared File System just after making.

Before change:

```
Filesystem              1K-blocks      Used Available Use% Mounted on
/dev/sfdsk/c1/dsk/vol1   56569128      5200  50907016   1% /mnt/gfs1
```

After change:

```
Filesystem               1K-blocks      Used Available Use% Mounted on
/dev/sfdsk/c1/dsk/vol1   56569128      20480 50891736   1% /mnt/gfs1
```

Reserve area is increased because it is interpreted as used date area (column of "Used"), and available data area (column of "Available") is decreased.

Corrective action

None.

F.2.2.8 Add the upper limit of concurrently open files

Incompatibility description

When 5000 different files are being concurrently opened on a GFS Shared File System, it was changed that operation of **open(2)** or **creat(2)**'s execution for 5001th file.

Before change: Normal exit

After change: Abnormal exit (return code: -1, errno: ENFILE)

Corrective action

Please check the following message in **/var/log/messages**.

```
kernel: VFS: file-max limit number reached
```

If this message is not output, 5000 different files are already being opened on the GFS Shared File System where the file tried to open exists.

Please execute **close(2)** to files that are being opened and can be closed (have already been transacted).

If you need to open 5000 and over files on the GFS Shared File System, please disperse the target files to multiple GFS Shared File Systems.

F.2.2.9 Addition of the i-node's cache memory release function at the time of accessing many files

Incompatibility description

When the amount of i-node's cache memory used by MDS exceeded 30000, processing that decreases the amount of i-node's cache memory from 30000 to 22000 was added.

Before change:

The function of releasing i-node's cache memory used by MDS doesn't work.

After change:

1. When the amount of i-node's cache memory used by MDS exceeds 30000, the following message is output and release processing of i-node's cache memory will be started.

```
NOTICE: sfcfs_mds:3020: mountpoint: The total number of vnodes used reached the
upper bound. vnodes used=totalnum: upper bound=30000: lower bound=22000
```

2. When the i-node's cache memory used by MDS has been released until 22000, release processing will stop with the following indication.

```
NOTICE: sfcfs_mds:3021: mountpoint: The total number of vnodes used(=totalnum) is
the lower bound(=22000).
```

Corrective action

None.

But the i-node's cache memory release is in progress in the period of between the two message indications, "NOTICE : sfcfs_mds:3020" and "NOTICE : sfcfs_mds:3021". During release processing, file accessibility on the file system will be affected. Please minimize file access.

F.2.2.10 Change in operation at the time of specifying the mount option atime

Incompatibility description

Operation at the time of specifying the mount option **atime** has been changed.

Before change:

The last access time of a file is updated for each file access.

After change:

The last access time of a file is updated when the previous last access time is the same as or older than the last modify time or the last change time. Also, when the previous last access time of a file is old for over a day, it is updated.

Corrective action

To execute the operation performed before change, change the mount option which is described in **/etc/fstab** from **atime** to **strictatime**.

F.2.2.11 Cannot create sockets

Incompatibility description

Operation when creating sockets on the GFS Shared File System by specifying S_IFSOCK in the second argument of mknod(2) has been changed.

Before change: Normal exit

After change: Abnormal exit (return code: -1, errno: EPERM)

Corrective action

None.

F.2.2.12 Cannot create named pipes

Incompatibility description

Operation when creating named pipes on the GFS Shared File System by specifying S_IFIFO in the second argument of mknod(2) has been changed.

Before change: Normal exit

After change: Abnormal exit (return code: -1, errno: EPERM)

Corrective action

None.

F.2.3 Incompatibilities from 4.2A00

There are some incompatibilities with the following functions compared to 4.2A00:

Table F.3 Incompatibilities from 4.2A00

Number	Item
1	Support of sendfile(2)
2	Change in operation at the time of specifying the mount option atime

Number	Item
3	Cannot create sockets
4	Cannot create named pipes

📖 **Note**

. . . .

When the OS before upgrading is RHEL4, an incompatibility of the number 1 occurs.

. . . .

F.2.3.1 Support of sendfile(2)

Incompatibility description

Action of sendfile(2) to the GFS Shared File System has been changed.

Before change: Abnormal exit (return code: -1, errno: EINVAL)

After change: Normal exit

Corrective action

None.

F.2.3.2 Change in operation at the time of specifying the mount option atime

Incompatibility description

Operation at the time of specifying the mount option **atime** has been changed.

Before change:

The last access time of a file is updated for each file access.

After change:

The last access time of a file is updated when the previous last access time is the same as or older than the last modify time or the last change time. Also, when the previous last access time of a file is old for over a day, it is updated.

Corrective action

To execute the operation performed before change, change the mount option which is described in **/etc/fstab** from **atime** to **strictatime**.

F.2.3.3 Cannot create sockets

Incompatibility description

Operation when creating sockets on the GFS Shared File System by specifying S_IFSOCK in the second argument of mknod(2) has been changed.

Before change: Normal exit

After change: Abnormal exit (return code: -1, errno: EPERM)

Corrective action

None.

F.2.3.4 Cannot create named pipes

Incompatibility description

Operation when creating named pipes on the GFS Shared File System by specifying S_IFIFO in the second argument of mknod(2) has been changed.

Before change: Normal exit

After change: Abnormal exit (return code: -1, errno: EPERM)

Corrective action

None.

F.2.4 Incompatibilities from 4.2A30

There are some incompatibilities with the following functions compared to 4.2A30:

Table F.4 Incompatibilities from 4.2A30

Number	Item
1	Support of sendfile(2)
2	Change in operation at the time of specifying the mount option atime
3	Cannot create sockets
4	Cannot create named pipes

F.2.4.1 Support of sendfile(2)

Incompatibility description

Action of sendfile(2) to the GFS Shared File System has been changed.

Before change: Abnormal exit (return code: -1, errno: EINVAL)

After change: Normal exit

Corrective action

None.

F.2.4.2 Change in operation at the time of specifying the mount option atime

Incompatibility description

Operation at the time of specifying the mount option **atime** has been changed.

Before change:

The last access time of a file is updated for each file access.

After change:

The last access time of a file is updated when the previous last access time is the same as or older than the last modify time or the last change time. Also, when the previous last access time of a file is old for over a day, it is updated.

Corrective action

To execute the operation performed before change, change the mount option which is described in **/etc/fstab** from **atime** to **strictatime**.

F.2.4.3 Cannot create sockets

Incompatibility description

Operation when creating sockets on the GFS Shared File System by specifying S_IFSOCK in the second argument of mknod(2) has been changed.

Before change: Normal exit

After change: Abnormal exit (return code: -1, errno: EPERM)

Corrective action

None.

F.2.4.4 Cannot create named pipes

Incompatibility description

Operation when creating named pipes on the GFS Shared File System by specifying S_IFIFO in the second argument of mknod(2) has been changed.

Before change: Normal exit

After change: Abnormal exit (return code: -1, errno: EPERM)

Corrective action

None.

F.2.5　Incompatibilities from 4.3A00

There are some incompatibilities with the following functions compared to 4.3A00:

Table F.5 Incompatibilities from 4.3A00

Number	Item
1	Support of sendfile(2)
2	Change in operation at the time of specifying the mount option atime
3	Cannot create sockets
4	Cannot create named pipes

F.2.5.1 Support of sendfile(2)

Incompatibility description

Action of sendfile(2) to the GFS Shared File System has been changed.

Before change: Abnormal exit (return code: -1, errno: EINVAL)

After change: Normal exit

Corrective action

None.

F.2.5.2 Change in operation at the time of specifying the mount option atime

Incompatibility description

Operation at the time of specifying the mount option **atime** has been changed.

Before change:

The last access time of a file is updated for each file access.

After change:

The last access time of a file is updated when the previous last access time is the same as or older than the last modify time or the last change time. Also, when the previous last access time of a file is old for over a day, it is updated.

Corrective action

To execute the operation performed before change, change the mount option which is described in **/etc/fstab** from **atime** to **strictatime**.

F.2.5.3 Cannot create sockets

Incompatibility description

Operation when creating sockets on the GFS Shared File System by specifying S_IFSOCK in the second argument of mknod(2) has been changed.

Before change: Normal exit

After change: Abnormal exit (return code: -1, errno: EPERM)

Corrective action

None.

F.2.5.4 Cannot create named pipes

Incompatibility description

Operation when creating named pipes on the GFS Shared File System by specifying S_IFIFO in the second argument of mknod(2) has been changed.

Before change: Normal exit

After change: Abnormal exit (return code: -1, errno: EPERM)

Corrective action

None.

F.2.6　Incompatibilities from 4.3A10

There are some incompatibilities with the following function compared to 4.3A10:

Table F.6 Incompatibilities from 4.3A10

Number	Item
1	Support of sendfile(2)
2	Change in operation at the time of specifying the mount option atime
3	Cannot create sockets
4	Cannot create named pipes

 Note
. . . .
When the OS before upgrading is RHEL5, incompatibilities of the number 1 to 3 occur.
. . . .

F.2.6.1 Support of sendfile(2)

Incompatibility description

Action of sendfile(2) to the GFS Shared File System has been changed.

Before change: Abnormal exit (return code: -1, errno: EINVAL)

After change: Normal exit

Corrective action

None.

F.2.6.2 Change in operation at the time of specifying the mount option atime

Incompatibility description

Operation at the time of specifying the mount option **atime** has been changed.

Before change:

The last access time of a file is updated for each file access.

After change:

The last access time of a file is updated when the previous last access time is the same as or older than the last modify time or the last change time. Also, when the previous last access time of a file is old for over a day, it is updated.

Corrective action

To execute the operation performed before change, change the mount option which is described in **/etc/fstab** from **atime** to **strictatime**.

F.2.6.3 Cannot create sockets

Incompatibility description

Operation when creating sockets on the GFS Shared File System by specifying S_IFSOCK in the second argument of mknod(2) has been changed.

Before change: Normal exit

After change: Abnormal exit (return code: -1, errno: EPERM)

Corrective action

None.

F.2.6.4 Cannot create named pipes

Incompatibility description

Operation when creating named pipes on the GFS Shared File System by specifying S_IFIFO in the second argument of mknod(2) has been changed.

Before change: Normal exit

After change: Abnormal exit (return code: -1, errno: EPERM)

Corrective action

None.

F.2.7 Incompatibilities from 4.3A20

There are some incompatibilities with the following function compared to 4.3A20:

Table F.7 Incompatibilities from 4.3A20

Number	Item
1	Support of sendfile(2)
2	Change in operation at the time of specifying the mount option atime
3	Cannot create sockets
4	Cannot create named pipes

F.2.7.1 Support of sendfile(2)

Incompatibility description

Action of sendfile(2) to the GFS Shared File System has been changed.

Before change: Abnormal exit (return code: -1, errno: EINVAL)

After change: Normal exit

Corrective action

None.

F.2.7.2 Change in operation at the time of specifying the mount option atime

Incompatibility description

Operation at the time of specifying the mount option **atime** has been changed.

Before change:

The last access time of a file is updated for each file access.

After change:

The last access time of a file is updated when the previous last access time is the same as or older than the last modify time or the last change time. Also, when the previous last access time of a file is old for over a day, it is updated.

Corrective action

To execute the operation performed before change, change the mount option which is described in **/etc/fstab** from **atime** to **strictatime**.

F.2.7.3 Cannot create sockets

Incompatibility description

Operation when creating sockets on the GFS Shared File System by specifying S_IFSOCK in the second argument of mknod(2) has been changed.

Before change: Normal exit

After change: Abnormal exit (return code: -1, errno: EPERM)

Corrective action

None.

F.2.7.4 Cannot create named pipes

Incompatibility description

Operation when creating named pipes on the GFS Shared File System by specifying S_IFIFO in the second argument of mknod(2) has been changed.

Before change: Normal exit

After change: Abnormal exit (return code: -1, errno: EPERM)

Corrective action

None.

F.2.8 Incompatibilities from 4.3A30

There are some incompatibilities with the following function compared to 4.3A30:

Table F.8 Incompatibilities from 4.3A30

Number	Item
1	Support of sendfile(2)
2	Change in operation at the time of specifying the mount option atime
3	Cannot create sockets
4	Cannot create named pipes

 Note

. . . .

When the OS before upgrading is RHEL5, incompatibilities of the number 1 to 3 occur.

. . . .

F.2.8.1 Support of sendfile(2)

Incompatibility description

Action of sendfile(2) to the GFS Shared File System has been changed.

Before change: Abnormal exit (return code: -1, errno: EINVAL)

After change: Normal exit

Corrective action

None.

F.2.8.2 Change in operation at the time of specifying the mount option atime

Incompatibility description

Operation at the time of specifying the mount option **atime** has been changed.

Before change:

The last access time of a file is updated for each file access.

After change:

The last access time of a file is updated when the previous last access time is the same as or older than the last modify time or the last change time. Also, when the previous last access time of a file is old for over a day, it is updated.

Corrective action

To execute the operation performed before change, change the mount option which is described in **/etc/fstab** from **atime** to **strictatime**.

F.2.8.3 Cannot create sockets

Incompatibility description

Operation when creating sockets on the GFS Shared File System by specifying S_IFSOCK in the second argument of mknod(2) has been changed.

Before change: Normal exit

After change: Abnormal exit (return code: -1, errno: EPERM)

Corrective action

None.

F.2.8.4 Cannot create named pipes

Incompatibility description

Operation when creating named pipes on the GFS Shared File System by specifying S_IFIFO in the second argument of mknod(2) has been changed.

Before change: Normal exit

After change: Abnormal exit (return code: -1, errno: EPERM)

Corrective action

None.

F.2.9   Incompatibilities from 4.3A40

There are no incompatibilities from 4.3A40.

F.2.10   Incompatibilities from 4.4A00

There are no incompatibilities from 4.4A00.

F.2.11   Incompatibilities from 4.5A00

There are no incompatibilities from 4.5A00.

Glossary

64-bit file system interface

This interface can handle files that exceed 2 gigabytes.

AC

See Access client.

Access client

Access client. Module which operates in the kernel layer (file system layer) on each node sharing GFS Shared File System. An AC communicates with the MDS if a user application operating on the node sends an access request in a system call. File operations are performed using meta-data sent from the MDS and direct access to the real.

ACL

ACL is a file security system that specifies file access authority for specific users or groups in addition to the owner and owner group. For details, see **setfacl(1)** in the online manual.

alternate file system

The file system that made on the shared disk device is exclusive used in active node and the standing node in RMS Cluster.

area management data

Area management data includes information on i-nodes, V-data, and file data allocation information. It is stored in the meta-data area.

CF

See Cluster Foundation.

Cluster Foundation

The set of PRIMECLUSTER modules that provides basic clustering communication services. For details, see "PRIMECLUSTER Cluster Foundation (CF) Configuration and Administration Guide."

cluster interconnect

The set of private network connections used exclusively for PRIMECLUSTER communications. For details, see "PRIMECLUSTER Cluster Foundation (CF) Configuration and Administration Guide."

cluster partition

A cluster state in which communications with some of the nodes that constitute the cluster have been interrupted. For details, see "PRIMECLUSTER Cluster Foundation (CF) Configuration and Administration Guide."

copy of partition structure information

If a block failure on a disk device destroys partition structure information, a copy is available for continued processing of the file system.

data

The term data refers to file data.

data area

The term data area refers to a file data area.

data block

A data block is the fundamental unit of file allocation.

direct I/O

Direct I/O refers to a mode of data input/output that does not use kernel buffering.

Domain

A single cluster system that is configured through PRIMECLUSTER, or a single server (computing machine) that is not belonging to a cluster system

DOWN

A node state that indicates that the node is unavailable (marked as down). A LEFTCLUSTER node must be marked as DOWN before it can rejoin a cluster. For details, see "PRIMECLUSTER Cluster Foundation (CF) Configuration and Administration Guide."

extent

An extent is a group of contiguous data blocks of the file system that are handled as a single unit. Each extent is defined by its start block and length.

extent-base

The extent base is a method of managing file system data based on extents.

fibre channel

One of the interfaces using optical fibre cables as transmission media. In a cluster system, a fibre channel is used to connect an array disk and the computer.

file data

File data is the data part of a file.

file data area

The file data area in disk partitions contains file data. It is also called data area. The file data area is allocated in the representative partition and the file data partitions.

file data partition

The file data partition is a partition that consists of a file data area (data area).

The file data partitions are added to an existing file system in order to extend it. The file data partition will be specified in file data allocation attribute information.

file lock

File lock is a means for synchronization among processes that concurrently access data in the same file. There are two types of file locks: mandatory locks and advisory locks.

GFS Management View

Software which allows various operations including configuration settings and operation management of GFS using the Web screen.

Global Disk Services

Volume management software for protecting data by managing a disk storage device totally within a SAN environment. For details, refer to "PRIMECLUSTER Global Disk Services Configuration and Administration Guide."

Global Link Services

GLS(Global Link Services) software provides High Reliability of network communications by duplexing network communication paths using redundant NIC(Network Interface Card)s. For details, refer to "PRIMECLUSTER Global Link Services Configuration and Administration Guide: Redundant Line Control Function."

i-node

The i-node contains information on the file locations on the disk, on file sizes, and on the time when the respective files were last. It also includes other information for identification.

i-node area

The i-node area contains an i-node. It is located in the meta-data area.

indirect block

Same as indirect extent management information.

indirect extent management information

V-data to manage the extent that cannot be stored in i-node.

LEFTCLUSTER

A node state that indicates that the node cannot communicate with other nodes in the cluster. That is, the node has left the cluster. The reason for the intermediate LEFTCLUSTER state is to avoid the network partition problem. For details, see "PRIMECLUSTER Cluster Foundation (CF) Configuration and Administration Guide."

logical volume

A logical disk device that provides the user with a block special file interface as well as a physical slice (disk partition). The user can access data on a physical disk through accessing the logical volume. Logical volumes are classified into single volumes, mirror volumes, stripe volumes, the volumes in concatenation groups ,and switch volumes. Logical volumes may be called "volumes." For details, see "PRIMECLUSTER Global Disk Services Configuration and Administration Guide."

management partition

Required information is managed when employing the GFS Shared File System. Moreover, this is used in order for sfcfrmd daemon to maintain informational adjustment. The following information is included in the management partition.

- Information on each node that shares the GFS Shared File System.

- Correspondence information on shared node of each GFS Shared File System, and shared device.

- MDS placement information on each GFS Shared File System.

MDS

See meta-data server.

MDS node

The node with the possibility that MDS performs. MDS performs in either node that it is defined as MDS node in nodes that share the file system. The default MDS node is 2 for the primary and secondary MDS.

meta area

Same as the meta-data area.

meta-data

Various types of management data for the file system are collectively referred to as meta-data. Meta-data includes information on free areas, i-nodes, indirect extent management information, directories and symbolic links.

meta-data area

The meta-data area contains meta-data. It is allocated in the representative partition.

meta-data server

Program that provides integrated management of file system management information (Meta data). The program is running on the primary server or secondary server. Each AC communicates with MDS then operates files based on the Meta data from MDS.

Mount point

Point that is used to connect a file system on the directory tree.

network partition

This condition exists when two or more nodes in a cluster cannot communicate over the interconnect; however, with applications still running, the nodes can continue to read and write to a shared device, compromising data integrity.

node

A node is a computer system that runs under one operating system.

node identifier

The node identifier is a universal name that is assigned to each node in the PRIMECLUSTER system. For each node, the node identifier is normally the same as the node name. However, if the node name was assigned to the standby node such as in hot standby mode, the correspondence between physical name and node name is temporarily disrupted. In this event, the node can be managed via the node identifier, which is the name for the physical device. The node identifier is displayed as each node name using PRIMECLUSTER Management View.

node name

The node name is the name of a node that is displayed with the "uname -n" command and identified in the communication network. For details, see **uname(1L)** in the online manual.

node state

Every node in a cluster maintains a local state for every other node in that cluster. The node state of every node in the cluster must be either UP, DOWN, or LEFTCLUSTER. For details, see "PRIMECLUSTER Cluster Foundation (CF) Configuration and Administration Guide."

partition

The partition is a physical or logical disk slice.

partition configuration data

This information refers to the partitions that shares the GFS Shared File System. It includes information about the size of the meta-data area, update log area, file data area (data area) and on their layout in the configuration partition.

primary management server

Node on which the management server daemon of the Web-Based Admin View operates during normal operation. The primary management server uses the menus of the GFS Management View to accept operation and display requests from the client and to transfer the requests to the monitored nodes.

primary server

Node on which the MDS operates during normal operation in a cluster system. This server should be set during file system registration.

public LAN

The local area network (LAN) by which normal users access a machine.

quorum

The state in which integrity is maintained among the nodes that configure the cluster system. Specifically, the CF state in all nodes that configure the cluster system is either UP or DOWN (there is no LEFTCLUSTER node). For details, see "PRIMECLUSTER Cluster Foundation (CF) Configuration and Administration Guide."

quota function

quota is a function for limiting disk space or the number of i-nodes that can be used by the user. For details, see edquota(8) in the online manual.

Reliant Monitor Services (RMS)

The package that maintains high availability of user-specified resources by providing monitoring and switchover capabilities. For details, see "PRIMECLUSTER Reliant Monitor Services (RMS) with Wizard Tools Configuration and Administration Guide."

representative partition

The representative partition is a GFS Shared File System partition that consists of meta-data, update log, and file data (data) areas. It may also consist of meta-data and file data (data) areas.

RMS

See Reliant Monitor Services.

Rolling update

Update method used to fix an application or maintenance within the cluster system. Fix application is enabled by applying fixes to each node sequentially without stopping jobs.

round robin allocation

Round robin allocation is a file data allocation mode for allocating separate file data to multiple file data areas.

scope

A range of node set that can share an object when a disk class is shared.

secondary management server

Node on which the Web-Based Admin View management server daemon operates if an error occurs on the primary management server. A secondary management server is set when the Web-Based Admin View operating environment is set. If an error occurs on the primary management server, the management server function of the Web-Based Admin View passes to the secondary management server.

secondary server

Server that takes over the MDS function if an error occurs on the primary server in a cluster system. The secondary server must be set during file system registration before it can assume the MDS function when an error occurs on the primary server. The file system status is restored to assure continuous access from each AC.

sfcfrmd

The demand from the commands or sfcfsd daemon is received and the communication way for realizing commands sfcfsd, commands sfcprmd, and communication between nodes is offered. Management of a composition database and a management partition I/O function are offered.

sfcfsd

Employment start and employment stop of the GFS Shared File System, taking over employment (failover) of MDS, an on-line node addition and deletion, file system composition cut return in the definition state (my gray SHON), etc. offer the function which controls MDS and AC.

sfcprmd

Starting and stop, down surveillance of the process which constitutes a GFS Shared File System is performed. When a composition process starts and executes other daemons and commands, a demand is transmitted to sfcprmd and sfcprmd executes a daemon and a command according to a demand.

The process which transmits a demand is called client process and the process started from sfcprmd according to a demand is called target process.

A target process notifies the situation to a client process, when it is supervised by sfcprmd and a process carries out a normal end and unusual end.

shared objects

Object set which belongs to shared disk class. Each object may also be referred to as shared class, shared group and shared disk.

shared disk

One of operation modes of a shared disk device. Applications in each node can simultaneously access a shared disk device on their own. (Note that the terms shared disk device and shared disk have different meanings. To distinguish between the two terms, a "shared disk" may be referred to as a "shared disk for simultaneous access.")

shared disk device

Disk device connected to multiple nodes. FDU/FUs and disk arrays are available as PRIMECLUSTER shared disk devices. Shared disk devices can be designated as shared disks or switched disks according to whether multiple nodes are allowed to access them simultaneously. (Note that the terms shared disk device and shared disk has different meanings.)

shared node

The nodes shared the data of GFS Shared File System. Up to 2 nodes are allowed.

slice

A slice is an alias of a partition.

super block

The super block contains information on the type of the file system, dates of creation and updating, label information, information on size and layout of the file system, and summary information on free areas.

super block copy

A copy of the super block is created so that file system processing can continue even if the original super block is damaged by a block fault in the disk device.

system disk

Disk on which the operating Linux OS is installed. The entire disk containing a slice currently operating as the file system (or swap area) of one of the following:

- /

- /usr

- /var

- swap area

UP

A node state that indicates that the node can communicate with other nodes in the cluster. For details, see "PRIMECLUSTER Cluster Foundation (CF) Configuration and Administration Guide."

update log

The update log contains the history of updates made to the file system structure. The history information is used to make a high-speed recovery of a file system possible. The area that contains the update log is called the update log area. Data in this area is overwritten cyclically.

update log area

The update log area contains update log.

update log buffer

Update log data that was loaded into memory.

update log replay

Update log replay updates meta-data in the update log area for which an update had been interrupted. This feature is used to recover the file system in the case of failures.

V-data

Directory blocks, symbolic link path names, and indirect extent management information are collectively referred to as V-data. This data is stored in the V-data area that was allocated in the meta-data area.

V-data area

The V-data area contains V-data. It is allocated in the meta-data area.

Volume

See logical volume.

Web-Based Admin View

Software for system operation and management using the Web screen. For details, see "PRIMECLUSTER Web-Based Admin View Operation Guide."

Index