

IBM General Parallel File System for AIX: Problem Determination Guide

Before using this information and the product it supports, be sure to read the general information under "Notices" on page v.

| **Second Edition (November 1998)**

| This edition applies to Release 2 of the IBM General Parallel File System for AIX, and to all subsequent releases and modifications
| until otherwise indicated in new editions or technical newsletters.

© **Copyright International Business Machines Corporation 1998. All rights reserved.**

Note to U.S. Government Users — Documentation related to restricted rights — Use, duplication or disclosure is subject to restrictions set forth in GSA ADP Schedule Contract with IBM Corp.

Contents

Notices	v
Trademarks and Service Marks	v
About This Document	vii
Who Should Use This Information	vii
What You Need to Know First	vii
How to Use This Information	viii
Typography and Terminology	viii
Related Information	ix
Obtaining Documentation	ix
A Technical Overview of GPFS	1
The GPFS Operating Environment	2
Basic GPFS Structure	2
GPFS Daemon	3
Special Managers	4
GPFS and Memory	5
Component Interfaces	6
Administrative Commands	6
Key Data Structures	7
The GPFS Configuration	7
The GPFS File System	7
Logging	8
Token Management	8
Application Calls	9
1–Initialization	9
2–Mount	9
3–Open	10
4–Read	10
5–Write	11
Command Processing	13
The mmchdisk Command	13
The mmfsck Command	14
Recovery	14
System Data Repository	15
Common GPFS Problems	17
Network Problems	17
Unable to Access Disks	18
Disk Media Failures	20
Incorrect VSD States After Failures in mmcrfs and mmdelfs	21
Multiple Stripe Group Manager Failures	22
I/O Operations Cannot Complete	22
Problems Reported by Group Services	22
GPFS Join Problems	23
Quorum Problems	23
Installation and Administration Problems	24
Possible Migration Problems	24
Unable to Start GPFS after the Installation of GPFS 1.2	25
Unable to Access a File System	25

New GPFS 1.2 Functions Will Not Operate	25
Unable to Access Data Outside the Configuration	25
Mount Failures	26
Restricted Mode Mount	27
Quota Problems	27
Quota Problems that Interfere with File System Mount	27
Quota is Denied	28
Cannot Write to Quota File	28
GPFS Not Responding to Commands	28
Data Corruption Errors	29
Loss of File System Access	29
sysctl Problems	30
The SDR is Locked	30
GPFS Problem Determination Tools	33
Non-Standard Errors	33
GPFS Error Logs	34
MMFS_LONGDISKIO	35
MMFS_DISKFAIL	35
MMFS_SGPANIC	35
MMFS_ABNORMAL_SHUTDOWN	35
MMFS_GENERIC	36
MMFS_ENVIRON	36
MMFS_PHOENIX	36
MMFS_FSSTRUCT	36
MMFS_QUOTA	37
MMFS_RECOVERY	37
Error Log Entry Example	37
The Console Log	38
Command and Application Returns	39
Returns from User Applications	39
Return from mfsadm Command	39
Return from mmcrfs Command	39
Return from errpt Command	39
Return from df Command	40
GPFS Trace Facility	40
The mmfsadm Command	40
A Script for Generating GPFS Trace Reports	41
A Script for Collecting GPFS Service Information	42
Enabling the SPLM Error Log for Use by GPFS	42
mmsnap2 Utility	42
Hints and Tips	43
Cannot Unmount File System? (Device Busy)	43
What Disk is Your IBM Virtual Shared Disk On?	44
Reporting a Problem to IBM	44
Messages	47

Notices

References in this publication to IBM products, programs, or services do not imply that IBM intends to make these available in all countries in which IBM operates. Any reference to an IBM product, program, or service is not intended to state or imply that only IBM's product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any of IBM's intellectual property rights may be used instead of the IBM product, program, or service. Evaluation and verification of operation in conjunction with other products, except those expressly designated by IBM, are the user's responsibility.

IBM may have patents or pending patent applications covering subject matter in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to

IBM Director of Licensing
IBM Corporation
500 Columbus Avenue
Thornwood, NY 10594
USA

Licensees of these programs who wish to have information about it for the purpose of enabling: (i) the exchange of information between independently created programs and other programs (including this one) and (ii) the mutual use of the information which has been exchanged, should contact:

IBM Corporation
Mail Station P131
522 South Road
Poughkeepsie, NY 12601-5400
USA
Attention: Information Request

Such information may be available, subject to appropriate terms and conditions, including in some cases, payment or a fee.

Trademarks and Service Marks

The following terms are trademarks of the IBM Corporation in the United States or other countries or both:

AIX

IBM

RS/6000

RS/6000 Scalable POWERParallel Systems

Scalable POWERParallel Systems

SP

UNIX is a registered trademark in the United States and other countries licensed exclusively through X/Open Company Limited.

Other company, product, and service names may be the trademarks or service marks of others.

About This Document

This document is provided to help you understand how the General Parallel File System for AIX (GPFS) works and to equip you to deal with any problems you may encounter during its normal usage.

Throughout this publication you will see various command and component names beginning with the prefix **mmfs**. This is not an error. GPFS shares many components with the IBM Multi-Media File Server, a related product.

Who Should Use This Information

Use this information only if you are knowledgeable in AIX administration concepts and familiar with AIX internals.

What You Need to Know First

Before using this information, you should be familiar with, and have documentation for, the following:

- The AIX operating system, including the Logical Volume Manager subsystem as described in *IBM AIX Version 4.3 System Management Guide: Operating Systems and Devices*, SC23–4126.
- The IBM Virtual Shared Disk information in *IBM Parallel Systems Support Programs: Managing Shared Disks*, SA22–7349.
- The IBM Virtual Shared Disk and switch information in *IBM Parallel Systems Support Programs: Administration Guide*, SA22–7348.
- The problem determination information in *IBM Parallel Systems Support Programs: Diagnosis Guide*, GA22–7350.
- The problem determination information in *IBM Parallel Systems Support Programs: Messages Reference*, GA22–7352.
- The command information, particularly IBM Virtual Shared Disk commands, in *IBM Parallel Systems Support Programs: Command and Technical Reference*, SA22–7351.
- The Group Services information in *IBM Parallel Systems Support Programs: Group Services Programming Guide and Reference*, SA22–7355.
- The High Availability information in *IBM Parallel Systems Support Programs: Administration Guide*, SA22–7348.
- Configuration and tuning information in *IBM RS/6000 SP Management: Easy, Lean, and Mean*, GG24–2563

Most of the RS/6000 SP hardware and software books are available from the IBM RS/6000 web site at <http://www.rs6000.ibm.com>.

How to Use This Information

This information is organized in the following topics:

1. “A Technical Overview of GPFS” on page 1, describing its behavior under normal operating conditions
2. “Common GPFS Problems” on page 17, a discussion of errors that might occur in the use of GPFS.
3. “GPFS Problem Determination Tools” on page 33, provided to assist you in isolating and correcting error conditions.
4. “Messages” on page 47, with explanations and suggested responses to error messages.

Typography and Terminology

This book uses the following typographical conventions:

Convention	Usage
Bold	Bold words or characters represent system elements that you must use literally, such as commands, subcommands, flags, path names, directories, file names, values, and selected menu options.
<u>Bold Underlined</u>	<u>Bold Underlined</u> keywords are defaults. These take effect if you fail to specify a different keyword.
<i>Italic</i>	<ul style="list-style-type: none">• <i>Italic</i> words or characters represent variable values that you must supply• <i>Italics</i> are used for book titles• <i>Italics</i> are used for general emphasis.
Monospace	All of the following are displayed in monospace type: <ul style="list-style-type: none">• Displayed information• Message text• Example text• Specified text typed by the user• Field names as displayed on the screen• Prompts from the system• References to example text
[]	Brackets enclose optional items in format and syntax descriptions.
{ }	Braces enclose a list from which you must choose an item in format and syntax descriptions.
< >	Angle brackets (less-than and greater-than) enclose the name of a key on the keyboard. For example, <Enter> refers to the key on your terminal or workstation that is labeled with the word Enter.
...	An ellipsis indicates that you can repeat the preceding item one or more times.
<Ctrl-x>	The notation <Ctrl-x> indicates a control character sequence. For example, <Ctrl-c> means that you hold down the control key while pressing <c>.

Related Information

The *IBM General Parallel File System for AIX: Installation and Administration Guide*, SA22-7278, shipped on the product media as a softcopy document and installed in the **gpfsdocs** file set, contains procedures for installing, configuring, and maintaining GPFS. Performance tuning, command reference, and programming interface material is also included.

The *International Technical Support Organization GPFS: A Parallel File System*, SG24-5165 contains additional information on GPFS.

If you have questions about the SP, PSSP, or a related product, the following online information resources make it easy to find the information you are looking for:

- Access the new SP Resource Center by issuing the command:

`/usr/lpp/ssp/bin/resource_center`

If you have the Resource Center on CD ROM, see the readme.txt file for information on how to run it.

- Access the RS/6000 Web Site at:

`http://www.rs6000.ibm.com`

Obtaining Documentation

The **gpfsdocs** file set includes HTML files that contain online versions of the GPFS publications. Once you have installed the **gpfsdocs** file set, the GPFS publications will be located at **`/usr/lpp/mmfs/gpfsdocs`**. Since other parts of GPFS link to the HTML publications, these files should not be removed from the **`/usr/lpp/mmfs/gpfsdocs`** directory.

You can also view the latest update to this document by visiting the RS/6000 SP Product Documentation Library site on the World Wide Web at:

`http://www.rs6000.ibm.com/resource/aix_resource/sp_books`

A Technical Overview of GPFS

General Parallel File System for AIX (GPFS) is a shared disk file system that provides all file system services at the application node. GPFS uses the IBM Virtual Shared Disk component of the Parallel System Support Programs for AIX (PSSP) to emulate disks that are connected to every node. This model provides increased performance because, unlike higher semantic file system protocols, the node-to-node connection is at the simple disk read or write level. In this way GPFS allows applications on any node to access any byte of any file.

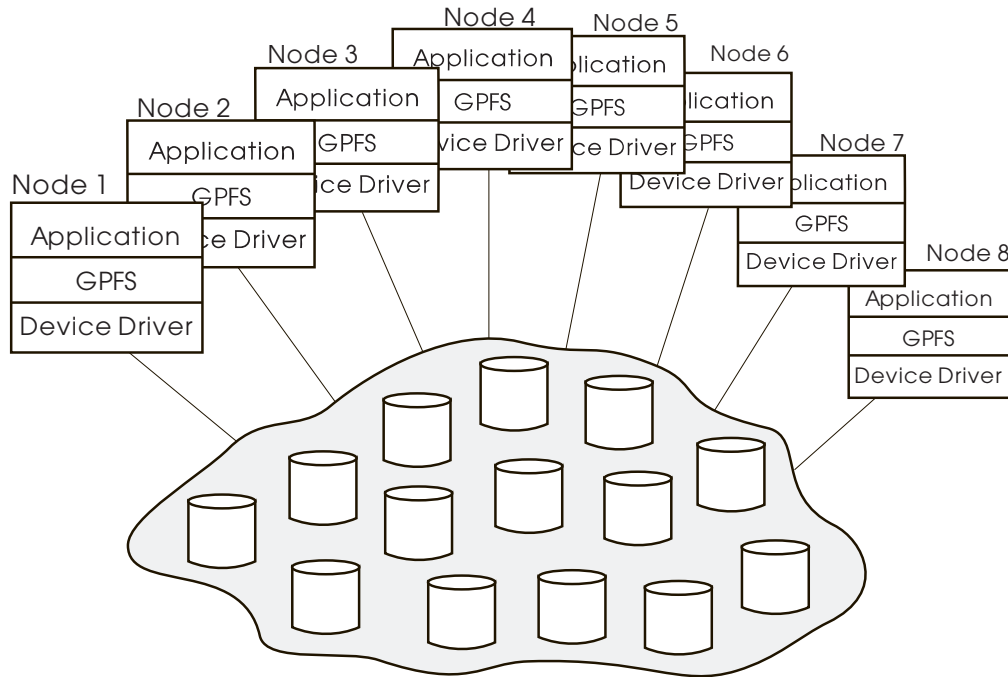


Figure 1. GPFS Provides File System Access from Every Node

Interaction between nodes at the file system level is limited to the locks and control flows required to maintain data and metadata integrity in the parallel environment.

This overview is organized in the following topics:

- “The GPFS Operating Environment” on page 2
- “Basic GPFS Structure” on page 2
- “GPFS and Memory” on page 5
- “Component Interfaces” on page 6
- “Key Data Structures” on page 7
- “Logging” on page 8
- “Application Calls” on page 9
- “Command Processing” on page 13
- “Recovery” on page 14
- “System Data Repository” on page 15

The GPFS Operating Environment

GPFS runs on the IBM RS/6000 SP equipped with switch connectivity and the following software products:

AIX release 4.3.2 or later, provides basic operating system services and the routing of file system calls requiring GPFS data.

PSSP 3.1 or later, provides:

- The IBM Virtual Shared Disk component for disk driver level support of universally attached disks.
- The Group Services component for notification when nodes or processes fail. It also provides sequencing of recovery on multiple nodes, and initialization information on the IBM Virtual Shared Disk.
- The IBM Recoverable Virtual Shared Disk component for the capability to fence a node from accessing certain disks, which is a prerequisite for successful recovery of that node. It also provides for transparent failover of disk access in the event of the failure of a disk server.

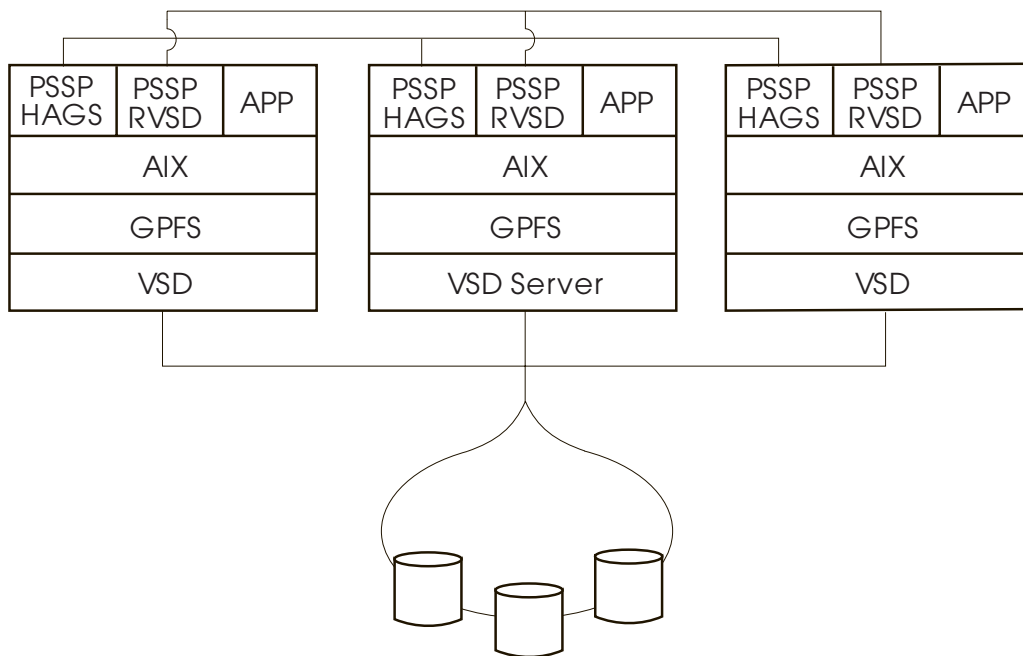


Figure 2. GPFS Provides Simultaneous Access to Files From All Nodes

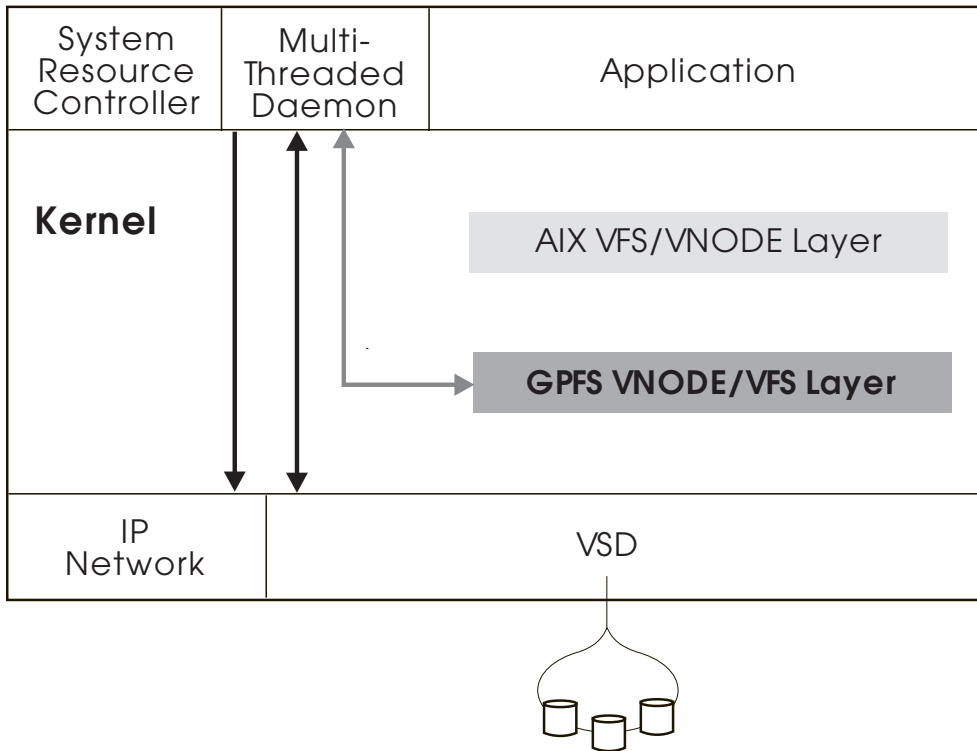
Basic GPFS Structure

GPFS resides on each node as a kernel extension, a multi-threaded daemon, and a number of commands. The kernel extension provides the interfaces to the AIX VNODE and VFS interfaces for adding a file system.

Structurally, applications make file system calls to AIX, which presents them to the GPFS file system kernel extension. In this way GPFS appears to applications as just another file system. The GPFS kernel extension will either satisfy these

requests using resources which are already available in the system, or send a message to the daemon to complete the request.

The messaging facility is a special purpose one where threads of the daemon wait for work. The daemon performs all I/O and buffer management for GPFS. This includes read-ahead for sequential reads and write-behind for all writes not specified as synchronous. All I/O is protected by a token manager, which insures that the file system on multiple nodes honors the atomicity and data consistency guarantees of a file system.



| Figure 3. GPFS Resides on Each Node as a Kernel Extension and a GPFS Daemon

GPFS Daemon

The daemon is a multi-threaded process with some threads dedicated to specific functions. This insures that services requiring priority attention are not blocked because other threads are busy with routine work. The daemon also communicates with instances of the daemon on other nodes to coordinate configuration changes, recovery and parallel updates of the same data structures. Specific functions that execute on the daemon include:

1. Allocation of disk space to new files and newly extended files. This is done in coordination with the stripe group manager for the file system.
2. Management of directories including creation of new directories, insertion and removal of entries into existing directories, and searching of directories that require I/O.
3. Allocation of appropriate locks to protect the integrity of data and metadata. Locks affecting data that may be accessed from multiple nodes require interaction with the token manager function.

4. Disk I/O is initiated on threads of the daemon.
5. Security and quotas are managed on the daemon in conjunction with the stripe group manager.

Special Managers

In general, GPFS performs the same functions on all nodes. It services application requests on the node where the application exists. This provides maximum affinity of data to the application. There are five cases where one node provides a more global function affecting the operation of multiple nodes. These are nodes acting as:

1. The Configuration Manager
2. The Stripe Group Manager
3. The Token Manager
4. The Quota Manager
5. The Metanode

The Configuration Manager

The configuration manager is the GPFS node that selects stripe group managers. It also determines whether a quorum exists, which in turn, determines whether file system usage can continue. The oldest continuously operating node in the file system group, as monitored by the Group Services component of PSSP, is automatically assigned as configuration manager. If it should fail for any reason, the next oldest node takes its place.

The Stripe Group Manager

There is one stripe group manager per file system. A stripe group is the set of disks comprising the storage assigned to a file system. The stripe group manager is selected by the configuration manager. If a stripe group manager should fail for any reason, a new stripe group manager is selected by the configuration manager and all functions continue without disruption, except for the time required to accomplish the takeover.

The stripe group manager provides the following services for all nodes using the file system:

1. Processes changes to the state or description of the file system. These include:
 - Adding disks
 - Changing disk availability
 - Repairing the file system

Mount and unmount processing is performed on both the stripe group manager and the node requesting the service.

2. Controls which regions of disks are allocated to each node, allowing effective parallel allocation of space.

The Token Manager

Controls tokens relating to the operation of this file system. All nodes directly read and write data to the shared disks but these operations are controlled by tokens originating at the stripe group manager.

The Quota Manager

Manages quotas. Nodes request allocation of quotas from the quota manager in response to user needs on the node. The stripe group manager allocates blocks of quota larger than the immediate requirement of the node so that every new write will not need to send a message to the stripe group manager.

The Metanode

There is one metanode per open file. The metanode is responsible for maintaining file metadata integrity. In almost all cases, the node that has had the file open for the longest period of continuous time is the metanode.

All nodes accessing a file can read and write data directly using the shared disk capabilities, but updates to metadata are written only by the metanode. The metanode for each file is independent of that for any other file and can be moved to any node to meet application requirements.

GPFS and Memory

GPFS uses three areas of memory:

1. Memory allocated from the kernel heap
2. Memory allocated within the daemon segment
3. Shared segments accessed from both the daemon and the kernel

The kernel memory is used for control structures such as vnodes and related structures that establish the necessary relationship with AIX. The stripe group manager nodes requires more shared segment storage since token state for the entire file system is stored there.

The daemon memory is used for structures that persist for the execution of a command or I/O operation, and also for states related to other nodes. Stripe group manager functions use daemon storage.

Shared segments consist of pinned storage, which is allocated at daemon startup. Token manager functions use shared segments. There are two pools of memory, which are labeled **mallocpool** and **pagepool**. The size of these pools is controlled by configuration parameters.

pagepool is used for the storage of data and metadata in support of I/O operations. With some access patterns, increasing the amount of **pagepool** storage may increase I/O performance for file systems with the following operating characteristics:

- Heavy use of writes that can be overlapped with application execution
- Heavy reuse of files and sequential reads of a size such that prefetch will benefit the application

mallocpool is used for control structures such as control blocks related to open files. Increasing this storage may allow more of these control structures to be retained in memory and reused if the access patterns demand it.

Component Interfaces

There are several component interfaces that affect GPFS behavior. These are socket connections from user commands to the daemon, communications between the daemon and the kernel, and communications between instances of daemon code.

The most common interface is through normal file system calls to AIX, which are relayed to GPFS if data in a GPFS file system is involved. This uses GPFS code in a kernel extension, which attempts to satisfy the application request using data already in memory. If this can be accomplished, control is returned to the application through the AIX interface. If the request requires resources that are not available at the time, the request is transferred for execution by a daemon thread. The daemon threads wait for work in a system call in the kernel, and are scheduled as necessary. Services available at the daemon level are the acquisition of tokens and disk I/O.

In order for a node to relinquish a token, the daemon must give it up. First, the daemon must release any locks that are held using this token. This may involve waiting for I/O to complete.

The daemon uses sockets to communicate with other instances of the file system on other nodes. Specifically, the daemon on each node communicates with the stripe group manager for allocation of logs, allocation segments, and quotas, as well as for various recovery and configuration flows. The only use of daemon communications in main line application flows involves communication with the metanode.

Administrative Commands

Socket communications are also used to process administrative commands. Commands may be processed either on the node issuing the command or on the stripe group manager, depending on the nature of the command. The actual command processor merely assembles the input parameters and sends them along to the daemon on the local node using a socket.

If the command changes the state of a stripe group or its configuration, the command is processed at the stripe group manager. The results of the change are sent to all nodes and the status of the command processing is returned to the node, and eventually, to the process issuing the command. For example, a command to add a disk to a file system originates on a user process and,

1. Is sent to the daemon and validated.
2. If acceptable, it is forwarded to the stripe group manager, which updates the file system descriptors.
3. All nodes that have this file system are notified of the need to refresh their cached copies of the file system descriptor.
4. The return code is forwarded to the originating daemon and then to the originating user process.

Be aware that this chain of communication may allow faults related to the processing of a command to occur on nodes other than the node which issued the command.

Key Data Structures

Before addressing the internal GPFS flows, it is important to understand the building blocks and structures used by many application-initiated flows.

The GPFS Configuration

A configuration consists of a set of SP nodes that provide access to GPFS file systems. There can be multiple GPFS configurations per SP partition. However, each node in a configuration must be at the same level of GPFS code. Each configuration uses the two PSSP Group Services groups, **Gpfs.1** and **GpfsRec.1**, for purposes described in “Recovery” on page 14. The members of the configuration are defined in the **/etc/cluster.nodes** file on each node in the system. This file must be consistent on all nodes in general because the index to the node is used internally as an identifier of the node.

The **mmaddnode** and **mmdelnode** commands maintain this file in a synchronous manner, including insertion of blank slots for deleted nodes. The **/etc/cluster.nodes** file generates a lock structure within the kernel when the token manager kernel extension is loaded. It requires a GPFS call to create a new structure, which in turn, requires that all nodes stop and restart GPFS before it will operate on a newly defined node. Early definition of nonexistent nodes is acceptable because it merely uses additional lock space in memory.

The GPFS File System

A file system consists of a set of disks which are used to store user data. This set of disks is listed in a *file system descriptor* which is at a fixed position on each of the disks. Although this control structure is maintained on all disks, a set of disk descriptors which are listed in **/etc/filesystems** are considered the authoritative copy of the file system descriptors. In the case where there are different versions, the version on the majority of the disks listed in **/etc/filesystems** is used. In addition, the file system descriptor contains information about the state of the stripe group.

i-nodes

Having found the disks which make up the file system, and their size and order from the file system descriptor, it is possible to address any block in the file system. In particular, it is possible to find the first i-node, which describes the **i-node file**, and a small number of i-nodes which are the core of the rest of the file system. Specifically, there are fixed i-nodes within the i-node file for the allocation map, root directory of the stripe group, and the i-node allocation map. The data contents of each of these files are taken from the data space on the disks. These files are considered metadata and are allocated only on disks where metadata is allowed.

The i-node file is a collection of fixed length records that represent a single file, directory or link. The unit of locking is the single i-node because the i-node size must be a multiple of the sector size.

Block Allocation File

The block allocation file is a collection of bits that represent the availability of disk space within the stripe group. One unit in the allocation map represents a sub-block or 1/32 of the block size of the file system. The allocation map is broken into regions, which reside on disk sector boundaries. The number of regions is set at file system creation time by the parameter that specifies how many nodes will access this file system. The regions are separately locked and, as a result, different nodes can be allocating or deallocating space represented by different regions independently and concurrently.

I—Node Allocation File

The i-node allocation file represents the availability of i-nodes within the i-node file. This file represents all the files, directories, and links that can be created. The **mmchfs** command can be used to change the maximum number of files that can be created in the file system.

Log Files

Log files are created at file system creation. Additional log files may be created if needed. Log files are always replicated and are found through a pointer in the file system descriptor. The stripe group manager (see “The Stripe Group Manager” on page 4) assigns a log file to each node accessing the file system.

The remainder of the disk space consists of data, indirect blocks, and directories. It is allocated from the allocation map file as needed.

Logging

GPFS logging maintains the atomicity of the on-disk structures of a file through a combination of rigid sequencing of operations and logging. The data structures maintained are the i-node, the indirect block, the allocation map, and the data blocks. Data blocks are written to disk before any control structure that references the data is written to disk. This insures that the previous contents of a data block can never be seen in a new file. Allocation blocks, i-nodes, and indirect blocks are written and logged in such a way that there will never be a pointer to a block marked unallocated that is not recoverable from a log.

There are certain failure cases where blocks are marked allocated but not part of a file, and this can be recovered by running **mmfsck** online or offline. GPFS always replicates its log; there are two copies of the log for each executing node. Log recovery is run

1. As part of the recovery of a node failure affecting the objects that the failed node might have locked.
2. As part of mount after the file system has been unmounted everywhere.

Token Management

The token manager resides on all nodes in the GPFS daemon. For each mount point, there is a token server, which is located at the stripe group manager node. If the stripe group manager is switched to another node as a result of a failure, the token server moves with it. The status of each token is held in two places: on the token server and on the token client holding the token.

The normal flow for a token is a message to the token server, which returns either a granted token or a list of the nodes which have conflicting tokens. The token manager at the requesting node then has the responsibility to communicate with all nodes holding a conflicting token and get them to relinquish the token. This relieves the token server of having to deal with all nodes holding conflicting tokens.

Application Calls

There are three ways for an application to interact with the GPFS file system: it can issue commands (see “Administrative Commands” on page 6), it can issue a programming call (see the *GPFS: Installation and Administration Guide*), or it can issue system calls to AIX, which operate on GPFS-resident data in the following sequence.

1–Initialization

GPFS initialization can be done automatically as part of the node IPL sequence, or manually using SRC commands to start the daemon. The daemon startup process loads the necessary kernel extensions, if they have not been previously loaded by an earlier instance of the daemon subsequent to the current IPL of this node, and awaits membership in the **cssMembership** and **ha.vsd** group services groups, if not already established. The initialization sequence then waits for the configuration manager to declare that a quorum exists. Note that Group Services will queue the request to join these groups if a previous failure is still being recovered, which will delay initialization. This is crucial if the failure being recovered is a failure of this node. Completion of the group join means that all necessary failure recovery is complete.

If Group Services reports that this node is the first to join the GPFS group, this node becomes the configuration manager. Quorum is achieved when a majority of the nodes that have been defined to GPFS are members of the GPFS group as seen at the configuration manager. When this state has been achieved, the configuration manager changes the state of the group from *initializing* to *active* using Group Services interfaces. When this state change has occurred, each node uses Group Services notifications to determine if the IBM Virtual Shared Disk facilities are available. If so, the daemon is ready to accept mount requests. This transition is evident in a message to the GPFS console file (*/var/adm/ras*) and an entry in the AIX error log.

2–Mount

GPFS file systems are mounted using AIX **mount** commands, which build the AIX structures that serve as the path to this data. GPFS mount processing is performed on both the node requesting the mount and the stripe group manager node. If there is no stripe group manager, a call is made to the configuration manager, which appoints one. The stripe group manager will insure that the stripe group is ready to be mounted. This includes checking that there are no conflicting utilities being run by **mmfsck** or **mmcheckquota**, for example, and running any necessary log processing to insure that metadata on the file system is consistent.

On the local node the control structures required for a mounted file system are initialized and the token manager domains are created. In addition, paths to each of the disks which make up the file system are opened. Part of mount processing

involves unfencing the disks, which may be necessary if this node has previously failed. This is done automatically without user intervention.

Note: There is a maximum of 32 file systems that may exist within a GPFS configuration.

3–Open

The open of a GPFS file involves the application making a call to AIX specifying the name of the file. Processing of an open involves two stages: the directory processing required to identify the file specified by the application, and the building of the required data structures based on the i-node.

The kernel extension code will process the directory search for those directories, which reside in GPFS (part of the path to the file may be directories in other physical file systems). If the required information is not in memory, the daemon will be called to acquire the necessary tokens for the directory or part of the directory needed to resolve the lookup. It will also read the directory entry into memory.

The lookup process occurs one directory at a time in response to calls from AIX. In the final stage of open, the i-node for the file is read from disk and connected to the AIX vnode structure. This requires acquiring locks on the i-node, as well as a lock that indicates the presence to the metanode:

- If no other node has this file open, this node becomes the metanode
- If another node has a previous open, then that node is the metanode and this node will interface with the metanode for certain parallel write situations
- If the open involves creation of a new file, the appropriate locks are obtained on the parent directory and the i-node allocation file block

The directory entry is created, an i-node is selected and initialized and then open processing is completed.

4–Read

The GPFS read function is invoked in response to a read system call and a call through the AIX vnode interface to GPFS. Read processing falls into three levels of complexity based on system activity and status:

1. Buffer available in memory
2. Tokens available locally but data must be read
3. Data and tokens must be acquired

Buffer and Locks Available in Memory

The simplest read operation occurs when the data is already available in memory, either because it has been prefetched or because it has been read recently by another read call. In either case, the buffer is locally locked and the data is copied to the application data area. The lock is released when the copy is complete. Note that no token communication is required because possession of the buffer implies that we at least have a read token that includes the buffer. After the *copyin*, prefetch is initiated if appropriate.

Tokens Available Locally but Data Must Be Read

The second, more complex, type of read operation is necessary when the data is not in memory. This occurs under three conditions: when the token has been acquired on a previous read that found no contention, when the buffer has been stolen for other uses, and on some random reads. (In the first of a series of random reads the token will not be available locally, but in the second sequential read it might be available.) In such situations, the buffer is not found and must be read. No token activity has occurred because the node has a sufficiently strong token to lock the required region of the file locally. A message is sent to the daemon, which is handled on one of the waiting daemon threads. The daemon allocates a buffer, locks the file range that is required so the token cannot be stolen for the duration of the I/O, and initiates the I/O to the device holding the data. The originating thread waits for this to complete and is posted by the daemon upon completion.

Data and Tokens Must Be Acquired

The third, and most complex read operation requires that tokens as well as data be acquired on the application node. The kernel code determines that the data is not available locally and sends the message to the daemon waiting after posting the message. The daemon thread determines that it does not have the required tokens to perform the operation. In that case, a token acquire request is sent to the token server. The requested token specifies a required length of that range of the file, which is needed for this buffer. If the file is being accessed sequentially, a desired range of data, starting at this point of this read and extending to the end of the file, is specified. In the event that no conflicts exist, the desired range will be granted, eliminating the need for token calls on subsequent reads. After the minimum token needed is acquired, the flow proceeds as in “Tokens Available Locally but Data Must Be Read.”

At the completion of a read, a determination of the need for pre-fetch is made. GPFS computes a desired read-ahead for each open file based on the performance of the disks and the rate at which the application is reading data. If additional pre-fetch is needed, a message is sent to the daemon that will process it asynchronously with the completion of the current read.

5-Write

Write processing is initiated by a system call to AIX, which calls GPFS when the write involves data in a GPFS file system.

Like many open systems file systems, GPFS moves data from a user buffer into a file system buffer synchronously with the application write call, but defers the actual write to disk. This technique allows better scheduling of the disk and improved performance. The file system buffers come from the memory allocated by the **pagepool** parameter in the **mmconfig** or **mmchconfig** command. Increasing this value may allow more writes to be deferred, which improves performance in certain workloads.

A block of data is scheduled to be written to a disk in one of five situations:

1. Immediately and synchronously with the write call if the application has specified synchronous write
2. When the system needs the storage
3. When a token has been revoked

4. When the last byte of a block of a file being written sequentially is written
5. When a sync is done

Until one of these occurrences; the data remains in GPFS memory.

As in the case of a read operation, there are several paths of varying complexity through the write code:

1. Buffer already exists for this block of the file
2. Token exists but the buffer does not
3. No buffer or token exists at the local node

Metadata changes are flushed under a subset of the same conditions. They can be written either directly, if this node is the metanode, or through the metanode, which merges changes from multiple nodes. This last case occurs most frequently if processes on multiple nodes are creating new data blocks in the same region of the file.

Buffer Already Exists for This Block of The File

The simplest path involves a case where a buffer already exists for this block of the file but may not have a strong enough token. This occurs if a previous write call accessed the block and it is still resident in memory. The write token already exists from the prior call. In this case, the data is copied from the application buffer to the GPFS buffer. If this is a sequential write and the last byte has been written, an asynchronous message is sent to the daemon to schedule the buffer for writing to disk. This operation occurs on the daemon thread overlapped with the execution of the application.

Token Exists but The Buffer Does Not

There are two situations in which the token may exist but the buffer does not. Either the buffer has been stolen recently to satisfy other needs for buffer space, or a previous write obtained a desired range token for more than it needed. In either case, the kernel extension determines that the buffer is not available, suspends the application thread, and sends a message to a daemon service thread requesting the buffer. If the write call is for a full file system block, an empty buffer is allocated since the entire block will be replaced. If the write call is for less than a full block and the rest of the block exists, the existing version of the block must be read and overlaid. If the write call creates a new block in the file, the daemon searches the allocation map for a block that is free and assigns it to the file. With both a buffer assigned and a block on the disk associated with the buffer, the write proceeds as it would in “Buffer Already Exists for This Block of The File.”

No Buffer or Token at The Local Node

The third, and most complex path through write occurs when neither the buffer nor the token exists at the local node. Prior to the allocation of a buffer, a token is acquired for the area of the file which is needed. As was true for read, if sequential operations are occurring, a token covering a larger range than is needed will be obtained if no conflicts exist. If necessary, the token manager will revoke the needed token from another node holding the token. Having acquired and locked the necessary token, the write will continue as in “Token Exists but The Buffer Does Not.”

Command Processing

GPFS commands fall into two categories: those that are processed locally and those that are processed at the stripe group manager for the file system involved in the command. The stripe group manager is used to process any command that alters the state of the stripe group. When commands are issued but the stripe group is not mounted, a stripe group manager is appointed for the task. The following describes the processing of two commands typical of those processed at the stripe group manager.

The mmchdisk Command

The **mmchdisk** command is issued when a failure that caused the unavailability of one or more disks has been corrected. The need for the command can be determined by the output of the **mmlsdisk** command. **mmchdisk** performs three major functions:

- It changes the availability of the disk to **recovering**, and to **up** when all processing is complete. All GPFS utilities honor an availability of **down** and do not use the disk. **Recovering** means that recovery has not been completed but the user has authorized use of the disk.
- It restores any replicas of data and metadata to their correct value. This involves scanning all metadata in the system and copying the latest to the recovering disk. Note that this involves scanning large amounts of data and potentially rewriting all data on the disk. This can take a long time for a large file system with a great deal of metadata to be scanned.
- It stops or suspends usage of a disk. This merely involves updating a disk state and should execute quickly.

All of these functions operate in a mode that is mutually exclusive with other state change commands on the stripe group, although they co-exist with normal operations and changes that occur as a result of normal operations. Thus, the processing flows as follows:

1. Command processing occurs on the local node, which forwards the command to the stripe group manager. Status reports and error messages are shipped back through the command processor and reported via standard output and standard error.
2. The stripe group manager reads all metadata looking for replicated data or metadata that has a copy on the failed disks. When a block is found, the later copy is moved over the copy on the recovering disk. If required metadata cannot be read, the command exits leaving the disks with an availability of **recovering**. This can occur because the **mmchdisk** command did not specify all down disks or because additional disks have failed during the operation of the command.
3. When all required data has been copied to the recovering disk, the disk is marked **up** and is automatically used by the file system, if mounted.

Subsequent invocations of **mmchdisk** will attempt to restore the replicated data on any disk left in with an availability of **recovering**.

The mmfsck Command

The **mmfsck** command is a traditional UNIX command that repairs file system structures. **mmfsck** operates in two modes: online and offline. For performance reasons GPFS logging allows the condition where disk blocks are marked used but not actually part of a file after a node failure. The online version of **mmfsck** cleans up that condition. Running **mmfsck -o -n** scans to determine if correction might be useful. The online version of **mmfsck** runs on the stripe group manager and scans all i-nodes and indirect blocks looking for disk blocks which are allocated but not used. If authorized to repair the file system, it releases the blocks. If not authorized to repair the file system, it reports the condition to standard output on the invoking node.

The offline version of **mmfsck** is the last line of defense for a file system that cannot be used. It will most often be needed in the case where log files are not available because of disk media failures. **mmfsck** runs on the stripe group manager and reports status to the invoking node. It is mutually incompatible with any other use of the file system and checks for any running commands or any nodes with the file system mounted. It exits if any are found. It also exits if any disks are down and require the use of **mmchdisk** to change them to up or recovering. **mmfsck** performs a full file system scan looking for metadata inconsistencies. This process can be lengthy on large file systems. It seeks permission from the user to repair any problems that are found which may result in the removal of files or directories that are corrupt. The processing of this command is similar to those for other file systems

Recovery

In order to understand the GPFS recovery process, you must be familiar with the Group Services component of Parallel System Support Programs for AIX. See the *Group Services Programming Guide and Reference* for details. In particular, it should be noted that only one state change, such as the loss or initialization of a node, can be processed at a time and subsequent changes will be queued in the PSSP. This means that the entire failure processing must complete before the failed node can join the group again. PSSP also processes all failures first, which means that GPFS will handle all failures prior to completing any recovery.

GPFS uses two groups to process failures of nodes, or GPFS failure on other nodes. The primary group is used to process failures. The secondary group is used to restart a failure protocol if a second failure is reported. This may be an actual second failure, or one that occurred at the same time as the first but was detected by Group Services later. The only function of the second group is to abort the current protocol and restart a new one processing all known failures.

GPFS recovers from node failure using notifications provided by PSSP Group Services. When notified that a node has failed or that the GPFS daemon has failed on a node, GPFS invokes recovery for each of the file systems that were mounted on the failed node. If necessary, a new configuration manager is selected prior to the start of actual recovery, or new stripe group managers are selected for any file systems that no longer have one, or both. This processing occurs as the first phase of recovery and occurs on the configuration manager. This processing must complete before other processing can be attempted and is enforced using Group Services barriers.

The stripe group manager for each stripe group fences the failed node from the disks comprising the stripe group. If the stripe group manager is newly appointed as a result of this failure, it rebuilds token state by querying the other nodes of the group. This stripe group manager recovery phase is also protected by a Group Services barrier. After this is complete, the actual recovery of the log of the failed node proceeds. This recovery will rebuild the metadata that was being modified at the time of the failure to a consistent state with the possible exception that blocks may be allocated that are not part of any file and are effectively lost until **mmfsck** is run, online or offline. After log recovery is complete, the locks held by the failed node(s) are released for this stripe group. Completion of this activity for all stripe groups completes the failure processing. The completion of the protocol allows a failed node to rejoin the cluster. The failed node will unfence the disks after it has rejoined the group.

System Data Repository

GPFS administration commands save configuration and file system information in two files in the system data repository (SDR). These files are accessible from any node and contain the master copy of all configuration and file system information for **mm...** commands.

One file, **mmsdrfs**, contains the current administrative state of all file systems and configurations (clusters). The other, **mmsdrcfg1**, is an exact copy of the **mmfs.cfg** file for GPFS configuration *n*.

The **mmsdrfs** file is not intended to be modified manually. The GPFS administration commands should never corrupt the file. When migrating to a new release of PSSP, make a copy of the **mmsdrfs** file to maintain your GPFS file system configuration, and restore the contents to the SDR after migration has completed. To copy the **mmsdrfs** file:

```
SDRRetrieveFile mmsdrfs /tmp/filename
```

To write it back:

```
SDRCreateFile /tmp/filename mmsdrfs
```

or if it already exists:

```
SDRReplaceFile /tmp/filename mmsdrfs
```

The format of the **mmsdrfs** file looks like:

```

=====
| config 1
| cluster 1 r14sn01,r14sn02,129.40.99.76,r14sn04,129.40.99.139:L
| cluster 1 r14sn06,r14sn07
|
| /gpfs/gpfs1
|     dev          = /dev/gpfs1
|     vfs          = mmfs          nodename      = -
|     mount        = mmfs
|     type         = mmfs
|     options      = rw,disks=gpfs5n2;gpfs7n2;gpfs3n1;gpfs8n2;gpfs6n2;vsd1n 13;gpfs2n1
|     account      = false
|     vsddisk      = gpfs5n2,1074,21;gpfs6n2,1074,21;gpfs7n2,1074,22;vsd1n1 3,16,22;gpfs3n1,1074,17
|     vsddisk      = gpfs2n1,1074,22;gpfs8n2,1074,13
|     cluster      = 1 0 6460 1204 682752
|
=====

```

| *Figure 4. Example of mmsdrfs File*

The file begins with cluster information, followed by a list of filesystem stanzas. The header information contains cluster and configuration lines. Cluster lines show the cluster number and a list of nodes. This example shows five per line. Additional nodes appear on a new line, such as in cluster 1. Cluster 1 shows 6 nodes (one node specified as an IP address) and one deleted node (indicated by L). The deleted node will appear as localhost in the **/etc/cluster.nodes** file.

| Configuration lines contain token manager configuration values for the cluster. The
| filesystem stanzas are exactly as they would appear in **/etc/filesystems**, plus the
| addition of vsddisk lines and a cluster line.

vsddisk lines list all disks, even if the options line does not list them all. Each disk shows: *diskname,size (MB),failure_grp*. They are grouped five per line.

The cluster line shows: *cluster =cluster_no filesystem_size max_filesystem_size*, where the two sizes (in MB) are broken into two integers each: *high low*. The low value displays the lower 6 digits, the high value shows the rest. The numbers can be higher than 32 bits allow.

In summary, the one file system shown (gpfs1) contains 7 disks totalling 6460 MB, and has a maximum file system size of 1204682752 MB. Many disks may be added before reaching this limit.

Common GPFS Problems

This information describes errors that may occur during the use of GPFS. It provides information about conditions in the environment that can affect correct operation of the file system, and also suggests when to call IBM for help.

Problems are discussed as follows:

1. "Network Problems"
2. "Unable to Access Disks" on page 18
3. "Disk Media Failures" on page 20
4. "Incorrect VSD States After Failures in mmcrfs and mmdelfs" on page 21
5. "Multiple Stripe Group Manager Failures" on page 22
6. "I/O Operations Cannot Complete" on page 22
7. "Problems Reported by Group Services" on page 22
8. "Quorum Problems" on page 23
9. "Installation and Administration Problems" on page 24
10. "Possible Migration Problems" on page 24
11. "Mount Failures" on page 26
12. "Quota Problems" on page 27
13. "Loss of File System Access" on page 29
14. "GPFS Not Responding to Commands" on page 28
15. "Data Corruption Errors" on page 29
16. "sysctl Problems" on page 30
17. "The SDR is Locked" on page 30

Network Problems

GPFS is dependent on the correct operation of the switch network for full operation. This dependence is direct because various GPFS internal messages flow on this network, and indirect because IBM Virtual Shared Disk is also dependent on the network. The type of error reported is a function of when the network failed, and what GPFS was doing when it failed.

If a failure is detected first by the IBM Virtual Shared Disk, the error will usually be seen by GPFS as a series of disk errors. Follow the instructions in "Unable to Access Disks" on page 18. The problem can also be first detected by the PSSP High Availability services. In that case, instances of GPFS running on the nodes that have network isolation or css0 failures will exit leaving an error log entry of the MMFS_PHOENIX class. You should follow the PSSP problem determination procedures for debugging this type of problem.

GPFS may detect network problems first depending on the timing of its messages and the actual network problem. GPFS requires full connectivity between nodes to manage its file system and insure data integrity. When there is no communication

between two nodes, one of them must be considered failed. The node that detects the failure waits for a short period of time to allow the Group Services function to determine that there has been a failure on the far end of the connection, which appears as a network failure to GPFS. If Group Services reports a failure on another node, failure processing for that node occurs. If no group services report occurs, GPFS will exit on the detecting node on the assumption that the network failure is local to this node. If this occurs, an error log of the MMFS_GENERIC or MMFS_ENVIRON classes will be reported.

There are cases where the error may actually be on a partner node but the failure is partial and Group Services does not report the error. An example of an error that can cause this condition is a severe shortage of memory on a partner node.

If GPFS reports the error and exits, there are several possible explanations. One is that a partner node has failed in a way that Group Services could not detect and report. Determine whether all partner nodes are up and functioning correctly. Pinging all nodes is a start in this process, but **ping** will often work when other functions cannot. Use of **dsh** to execute commands on the other nodes will probably point you at one or more nodes that are not behaving correctly. Correction of the problem at those nodes will be necessary.

It is also possible that communications on the reporting node have failed. Examine the error log on this node for indications of switch or network errors and make necessary repairs.

GPFS is set up to respawn automatically after a failure. If GPFS exits on a node because of a perceived communication error and the error is local to the exiting node, it will fail similarly on the restart if the error persists. If the error is transient, the system may recover fully. In this case, examine error logs and console logs to determine the problem.

Unable to Access Disks

There are several situations that can occur in the SP I/O structure in which disks can appear to fail to GPFS. These involve the PSSP Group Services, IBM Virtual Shared Disk, and IBM Recoverable Virtual Shared Disk components, as well as the switch with its associated software, plus the disk software and hardware. The following information describes how GPFS reacts to these failures and how to find the cause.

GPFS will stop using a disk that is determined to have failed. This event is marked as MMFS_DISKFAIL in an error log entry (see "GPFS Error Logs" on page 34) or check the disk state using the **mmlsdisk** command. The consequences of stopping disk usage vary according to what is stored on the disk:

- Certain data blocks may be unavailable because the data residing on a stopped disk is not replicated.
- Certain data blocks may be unavailable because the controlling metadata resides on a stopped disk.
- In conjunction with other disks that have failed, all copies of critical data structures may be unavailable resulting in the unavailability of the entire file system.

The disk will remain unavailable until its status is explicitly changed through the **mmchdisk** command. After that command is issued, any replicas that exist on the failed disk are updated before the disk is used.

The simplest example of this situation is the single disk failure. Such a failure should be accompanied by error log entries on the IBM Virtual Shared Disk server for the failing disk. The GPFS error log on the node detecting the error will contain the identifier of the failed disk. The user action is to follow the problem determination and repair actions specified by the disk subsystem.

The failure of a storage adapter on the IBM Virtual Shared Disk server may result in loss of access to multiple disks.

If you attach your disks to multiple IBM Virtual Shared Disk servers, GPFS will not see the error because IBM Recoverable Virtual Shared Disk will switch to the backup server automatically. If there is no alternate path to these disks, multiple disks will be considered as having failed.

If you have configured your system so that there are other failure groups, all replicas will be configured in those other failure groups and this will be a simple multiple disk failure. If you have not specified multiple failure groups and replication of metadata, it is possible that GPFS will not be able to continue because it cannot write logs or other critical metadata. In this case, GPFS will forcibly unmount the file system. This will be indicated in the error log by records indicating exactly which access failed, with a SG_PANIC record indicating the forced unmount. The user response to this is to take the needed actions to restore the disk access and issue **mmchdisk** to disks that are shown as down in the return from **mmlsdisk**.

Large scale disk unavailability can result from failure of an IBM Virtual Shared Disk server. If you attach your disks to multiple IBM Virtual Shared Disk servers, such a failure will not affect your applications. Rather than leaving your applications unable to access data, they would only experience a delay while the servers were switched, with a possible impact on performance if the failover results in an increased load on the server.

If an IBM Virtual Shared Disk server fails and there is no alternate server, the failure appears to GPFS as the failure of all of the disks that are attached through that IBM Virtual Shared Disk. This is similar to an adapter failure, although multiple adapters may be involved. The use of multiple failure groups may limit the degree of unavailability of data due to the failure.

There are several cases where the state of a given virtual disk will prevent access by GPFS. This will be seen by the application as I/O errors of various types and will be reported in the error logs as SG_PANIC and/or MMFS_DISKFAIL records. This state can be found by the running **lsvsd -l** command as described in *Managing Shared Disks*. An IBM Virtual Shared Disk that is indicated as STP, SUS or fenced (indicated by an * in the line reporting state) is inaccessible to GPFS. Follow the directions in *Managing Shared Disks* for restoration of access.

Disk Media Failures

There are various ways to guard against the loss of data due to disk media failures. The use of a RAID controller, which masks disk failures with parity disks, or a twin-tailed disk, with or without mirroring data at the LVM level would prevent the need for using the following recovery steps.

GPFS offers a method of protection, replication, which overcomes disk failure at the expense of additional disk space. GPFS allows replication of data and metadata. This means that two instances of data, or metadata, or both can be automatically created and maintained for any file in a GPFS file system. If one instance becomes unavailable due to disk failure, another instance is used instead. You can set different replication specifications for each file, or apply default settings specified at file system creation. (See the GPFS: Installation and Administration Guide for details on setting your replications factor.)

Regardless of whether you have chosen additional hardware or replication to protect your data against media failures, you first need to determine if the disk has completely failed. If this is a complete failure of the disk you will need to take the following steps:

1. Delete the disk from the file system:

```
mmdeldisk fs0 gpfs1n12 -p
```

2. Restripe the file system to recreate replications:

```
mmrestripefs gpfs1n12 -r
```

Note: Insure there is sufficient space elsewhere in your filesystem for the data to be stored by using the **mmdf** command. See the *GPFS Installation and Administration Guide*.

3. Check the file system for metadata inconsistencies by executing the offline version of **mmfsck**:

```
mmfsck fs0
```

If **mmfsck** succeeds, you may still have errors that occurred. Check to verify no files were lost. If files containing essential data structures were lost, you will need to recreate your file system from the backup media. If files containing user data were lost, you will have to restore the files from the backup media.

If the consistency check fails, sufficient metadata was lost and you need to recreate your file system and restore the data from backup media.

If the disk has only partially failed and you have chosen not to implement hardware protection against media failures, the steps to restore your data depends on whether or not you have used replication. If you have replicated neither your data nor metadata, you will need to restore the lost information from the backup media and execute the offline version of **mmfsck** as stated above. If it is just the data which was not replicated, you will need to restore the data from the backup media. There is no need to run **mmfsck** if the metadata is intact.

If both you data and metadata have been replicated, implement the following recovery actions:

1. Suspend the disk to prevent further attempts at allocating additional space on the disk:

```
mmchdisk fs0 suspend -d gpfs1n12
```

2. Restripe the file system to rebalance the data onto other disks:

```
mmrestripefs gpfs1n12 -r
```

Note: Insure there is sufficient space elsewhere in your filesystem for the data to be stored by using the **mmdf** command. See the *GPFS Installation and Administration Guide*.

3. Delete the disk from the file system:

```
mmdeldisk fs0 gpfs1n12 -p
```

4. If you are replacing the disk, add the disk :

```
mmadddisk fs0 hdisk13
```

and then restripe the file system:

```
mmrestripefs fs0 -b
```

Incorrect VSD States After Failures in mmcrfs and mmdelfs

The **mmcrfs** command optionally allows you to create virtual shared disks as part of the file system creation. However, there are a number of factors that could cause the creation of **all** of the virtual shared disks to fail. These include:

1. The state of the disk
2. The path to the disk
3. The configuration data that was supplied

In this case, any virtual shared disks which were successfully created remain in existence.

The standard way of handling this error is to repair the specific LVM or virtual shared disk error and reissue the **mmcrfs -FDescFile** command specifying the disk descriptor file. For detailed information on the **mmcrfs** command, see the *GPFS: Installation and Administration Guide*.

However, if you would like to remove the successfully created virtual shared disks, this may be accomplished by issuing the **removevsd -f -vvsdname** command from the control workstation. For detailed information on the **removevsd** command, see the *Managing Shared Disks* manual.

The **mmdelfs** command will delete all the virtual shared disks which are part of the file system, unless you specify otherwise. However, there are possible states of the LVM and the virtual shared disk which will make it impossible to complete the deletion at a given time. If this is the situation, the file system will be deleted but the virtual shared disks in those states will remain. The command will output the names of the virtual shared disks which could not be deleted. These virtual shared disks can be deleted by issuing the **removevsd -f -vvsdname** command from the control workstation. If this command fails, the virtual shared disks could be fenced. Issue the **unfencevsdvsdname** command then reissue the **removevsd -f -vvsdname** command. For detailed information on the **unfencevsd** command or virtual shared disk problem determination, see the *Managing Shared Disks* manual.

Multiple Stripe Group Manager Failures

The correct operation of GPFS requires that one node function as the stripe group manager at all times. This instance of GPFS has additional responsibilities for coordinating usage of the file system. When the stripe group manager node fails, another stripe group manager is appointed in a manner that is not visible to applications except for the time required to switch over.

There are situations where it may be impossible to appoint a stripe group manager. Such situations involve the failure of paths to disk resources from many, if not all, nodes. In this event, the configuration manager nominates several host names to successively try to become the stripe group manager. If none succeed, the configuration manager unmounts the file system everywhere. The error logs on all nodes will reflect stripe group panics with 212 error codes, or the errors that occurred when they tried to assume management of the stripe group. Note that these errors apply to a specific file system although it is possible that disk/communication paths that are shared will cause the unmount of multiple file systems.

I/O Operations Cannot Complete

Certain conditions can occur in the IBM Virtual Shared Disk and SP switch communications software which cause I/O operations to take an unusually long time (tens of minutes) to complete. These conditions often involve excessive loads on the switch/server nodes, or failures in the switch. This happens when IBM Virtual Shared Disk retries operations and suspends specific disks that are waiting for failed components to recover or for memory resources on server nodes to become available.

In these situations, GPFS might present disk I/O operations to the IBM Virtual Shared Disk for execution and receive no response while retries and recovery proceed. If there is no error indication from GPFS, examine the state of the disks, the number of retries reported in the IBM Virtual Shared Disk statistics, and the number of retries reported at both the application node and the server in the communication subsystem. See the PSSP problem determination guides for additional information.

Problems Reported by Group Services

GPFS recovery is initiated by PSSP Group Services, which detects the failure of a node, or the file system daemon on a node, and reports this to the daemon on the partner nodes in the Group Services output. The presence of a node in GPFS groups can be detected using the following command:

```
lssrc -l -s hags
```

which should return output similar to the following


```

Subsystem      Group      PID      Status
hags           hags      17682    active
5 locally-connected clients. Their PIDs:
17196 13890 16986 18976 25838
HA Group Services domain information:
Domain established by node 0.
Number of groups known locally: 5

Group name      Number of   Number of local
                providers   providers/subscribers
cssMembership   16          1              1
Gpfs.1          16          1              0
GpfsRec.1       16          1              0
ha_em_peers     17          1              0
ha.vsd          16          1              0

```

Figure 5. Output from `Issrc Command`

This shows whether a node is a member of the GPFS groups, as well as the total membership of each group.

GPFS Join Problems

If for any reason, GPFS cannot join a group, it exits, restarts, and tries again to join the group. Error log entries will indicate GPFS failure and an error log entry is also generated by the SRC indicating the GPFS restart. This is not a GPFS failure, but a failure in Group Services or some other component upon which it is dependent. Follow the PSSP problem determination procedures.

If the PSSP Group Services daemon fails, then the GPFS daemon will also fail. Error log entries will indicate GPFS failure.

Quorum Problems

GPFS requires a quorum of nodes (one plus half of the number of nodes in the GPFS configuration) to be active members of a group before any operations can be honored. This requirement guarantees that a valid single token manager domain exists for each GPFS file system. Prior to the existence of a quorum, most requests are rejected with a message indicating that quorum does not exist. To check on a quorum, use the Group Services command:

```
Issrc -ls hags
```

The output, shown in Figure 5, displays how many members exist in each group. If more nodes have been started than are members of the group, follow the PSSP problem determination procedures.

A node indicates that it is part of a quorum by writing a line in the console and generating an error log entry. If you believe that sufficient nodes are active to achieve a quorum but no entry was written in the error log, follow the PSSP problem determination procedures. These procedures should indicate:

1. Group Services operations in process that would inhibit recognition of all nodes as members of the group
2. Communications failures that would inhibit the formation of the GPFS groups

If an existing quorum is lost, GPFS exits all nodes to protect the integrity of your data. This is marked by an error log entry on all nodes and an entry in the system

console file. In response to this exit, the AIX system resource controller restarts GPFS on all nodes without administrator intervention. This restart attempts to rebuild a quorum and allows operations when quorum is achieved. Automatic file system mounts are performed, if specified.

Any Group Services failure results in a loss of quorum, initiating these recovery procedures. If GPFS is started and Group Services is not available, it will repeatedly exit and be restarted by the system resource controller until either Group Services is available or the administrator issues a **stopsrc** command to terminate GPFS. In either case, when quorum is lost, applications with open files receive an ESTALE error return code until the files are closed and reopened. New file open operations will fail until quorum is restored and the file system is remounted. Applications accessing these files prior to GPFS return may receive a ENODEV return code from AIX.

Installation and Administration Problems

Most problems encountered during file system creation fall into three classes:

1. Invalid parameters for the specification of the disks to be used

If you predefined your IBM Virtual Shared Disks, this may mean that the disks are specified incorrectly to GPFS. If you are creating IBM Virtual Shared Disks with the **mmcrfs** command, it may mean that the disk descriptors are incorrect.

2. The creation operation cannot access the disk.

Follow the procedures for checking access to the disk. This can result from a number of factors including switch problems and those described in “Unable to Access Disks” on page 18.

3. Failure of the communications with the stripe group manager.

Keep in mind that file system creation runs on the stripe group manager. Failure of that node may result in failure of the **mmcrfs** command. It is also possible that execution of this command may continue, and even complete, when this node has failed.

Possible Migration Problems

The *GPFS: Installation and Administration Guide* provides the step by step procedure for migrating your GPFS configuration. Here is guidance for some errors which may occur when migrating from GPFS 1.1 to GPFS 1.2. The following problems will be addressed:

1. Unable to start GPFS after the installation of GPFS 1.2
2. Unable to access a file system
3. New GPFS 1.2 functions will not operate
4. Unable to access data outside the configuration

Unable to Start GPFS after the Installation of GPFS 1.2

If one or more nodes in the configuration will not start GPFS, there are two possibilities:

1. If message 6027–829 *Join rejected. Incompatible daemon version (number:number). Group vers (number:number)*, is displayed, incompatible versions of GPFS code exist on nodes within the same configuration.

You must have the same level of code installed on all nodes in a configuration to ensure correct locking exists to protect your data. You must bring all nodes up to the same level of GPFS and reboot them.

2. If any of several possible messages which report that function is not supported are displayed, it is probable you do not have the correct kernel extensions loaded.

Installing GPFS 1.2 and rebooting the nodes replaces the kernel extension. However, **any activity** involving the file system during this time period will cause the replacement of the kernel extensions to fail. You must reboot the affected nodes to load the correct kernel extensions.

Unable to Access a File System

If you have been operating with GPFS 1.2 code, but have chosen to access the file system through GPFS 1.1, there are two possible scenarios depending on whether you have issued the **mmchfs -V** command or not:

1. Until you issue the command, GPFS will not make any incompatible disk changes. You may access the file system through GPFS 1.1, if you reinstall it. See the *GPFS: Installation and Administration Guide* for instructions on reinstalling GPFS.
2. After you have issued the command, new data structures exist on the disk which cannot be understood by GPFS 1.1. You must recreate the file system from the backup medium and restore the content in order to access it through GPFS 1.1

New GPFS 1.2 Functions Will Not Operate

If you have successfully installed and are operating with GPFS 1.2, but cannot execute the new functions available, it is probable that you have not issued the **mmchfs -V** command to change the version of the file system. This command must be issued for **each** of your file systems.

Make sure you have operated with the new level of code for awhile and are sure you want to migrate to GPFS 1.2. Issue the **mmchfs -V** command only after you have permanently decided to accept GPFS 1.2, as this will cause disk changes incompatible with GPFS 1.1.

Unable to Access Data Outside the Configuration

If you have created more than one GPFS configuration, data may only be accessed from within an individual configuration. Nodes within a GPFS configuration **do not** interact with GPFS on nodes outside that configuration. However, these configurations do not restrict the operation of any other software, as a PSSP partition would.

The ability to create more than one configuration is intended for migration purposes only. Having multiple GPFS configurations allows testing of the new software

without inhibiting the main GPFS configuration running at production level. It is not the intention of the function to be the normal use of GPFS. Use of multiple GPFS configurations in that way may produce the following anomalies:

1. Users scheduled on nodes within one configuration can not access data in file systems in another configuration.
2. Parallel jobs which span configurations may have access to different file systems and different data. You may wish to limit parallel jobs to nodes within a single configuration.
3. Quorums are enforced within a configuration. This may cause different failure patterns than existed prior to the splitting of nodes into multiple configurations.
For example, if a 3 node configuration was created out of a 32 node SP, 2 of the 3 nodes must be available to satisfy quorum requirements. The main configuration would then require 15 of the 29 nodes, rather than 17 of the 32, to operate.
4. Commands to change state affect only those nodes within the configuration from which they are issued.

Mount Failures

Here are some conditions to check first if a mount failure occurs:

1. The GPFS file system code is not loaded
Check by issuing

```
lssrc -s mmfs
```

This should show that the GPFS subsystem is active. If not, follow the procedure in *GPFS Installation and Administration* for starting GPFS.
2. No quorum in the system includes this node
Check `/var/adm/ras/mmfs.log.date` to assure that a *daemon ready* message has been logged, and that no errors were reported on this or other nodes. Follow the problem determination procedures for the PSSP High Availability Group Services.
3. There is a conflicting command executing somewhere
A mount may not be issued while **mmfsck** or some GPFS debug commands are executing. The mount may not be issued until they complete. Note that interrupting **mmfsck** is not a solution because the file system will not be mountable until the command completes. Try again after the conflicting command has completed.
4. Disks in the file system cannot be accessed
GPFS expects a minimum of half the number of disks listed in `/etc/filesystems` to be available so it can find a current copy of the core metadata. If sufficient disks cannot be accessed, mount will fail. The corrective action is to fix the path to the disk. See “Unable to Access Disks” on page 18.
Missing disks can also cause GPFS to be unable to find critical metadata structures. The output of the **mmlsdisk** command will show unavailable disks. If you have not specified metadata replication, the failure of one disk may result in your file system being unmountable. If you have specified metadata replication,

it will require two disks in different failure groups to disable the entire file system.

5. Communications paths to the other nodes are unavailable

This may impede contact with the stripe group manager. Follow the problem determination procedures in the PSSP *Diagnosis and Messages Guide*.

6. The file system is already mounted
7. For mounts to succeed, one copy of critical metadata must be available. To determine availability, run **mmisdisk** and examine return for disks with availability of **down**. Bring these disks up. Refer to “Unable to Access Disks” on page 18.
8. The maximum number of 32 file systems for a GPFS configuration has been met.

Restricted Mode Mount

Some disk failures can result in the loss of enough metadata to render the entire file system unmountable. In that event it might be possible to preserve some user data through a *restricted mode mount*. This facility should only be used if a normal mount does not succeed, and should be considered a last resort to save some data after fatal disk failure.

Restricted mode mount is invoked by using the **-rs** flags on the **mount** command. After a restricted mode mount is done, some data MAY be sufficiently accessible to allow copying to another file system. The success of this technique depends on the actual disk structures damaged. Attempt this only after you've tried to repair the file system with **mmfsck**.

Quota Problems

Quota problems can cause a number of different failure symptoms.

Quota Problems that Interfere with File System Mount

The following symptoms indicate quota problems.

Error 237 (E_NO_QUOTA_INST)

Quota management was not correctly installed for the file system, resulting in inconsistent coding of the options stanza in the **/etc/filesystem** across the nodes.

User Response: Check the **/etc/filesystem** options stanza for the quota option (quota=userquota;groupquota) on all nodes.

Error 245 (E_NO_QMGR)

Quota management is installed but no node is acting as the quota manager.

User Response: This error indicates a problem with the quota file. Run **mmcheckquota** and then remount the file system.

If the file system must be repaired without quotas,

1. Uninstall quota management by deleting the quota option from **/etc/filesystems** stanza on all nodes.

2. Make repairs. Install quota file backup, taking care that the i-node number of the quota file is unchanged. The quota manager takes the quota file i-node number that is stored in the file system descriptor during initialization.
3. Restore quota management using **mmchfs** command.
4. Remount file system.

Quota is Denied

If allocation of data blocks or files, which quota limits should allow, fails, issue **mmisquota** for the user or group.

If usage is equal to or approaching the hard limit, or if the grace period has expired, make sure that no quotas are lost by checking *in doubt* values. Reduce quota usage by deleting or compressing files or moving them out of the file system. Consider increasing quota limit.

If too many quotas accumulate in the *in doubt* category, run **mmcheckquota**.

Note: There is no way to force the GPFS nodes to relinquish all their local shares in order to check for lost quotas. This can only be determined by running **mmcheckquota** immediately after mounting the file system, and before any allocations are made. In this case, the value *in doubt* is the amount lost.

To display the latest quota usage information, use the **-e** option with **mmisquota** or **mmrepquota**. Remember that **mmquotaon** and **mmquotaoff** do not install and uninstall quota management. These commands merely control enforcement of quota limits. Usage continues to be counted and recorded in quota file regardless of enforcement.

Cannot Write to Quota File

If an MMFS_QUOTA error log entry containing Error writing quota file...is generated, the quota manager continues operation if the next write for the user or group is successful. If not, further allocations to the file system will fail. Check the error code in the log and make sure that the disks containing the quota file are accessible. Run **mmcheckquota**.

GPFS Not Responding to Commands

This could be a temporary condition if the GPFS daemon is still initializing, or it might indicate a problem. To determine which is the case, run the following command on one or more nodes:

```
lssrc -ls hags
```

See Figure 5 on page 23. If the MmfsGroup state is initial, refer to “Quorum Problems” on page 23.

There are a number of conditions that can cause the GPFS daemon to exit. These are all conditions where the GPFS internal checking has determined that continued operation would be dangerous to the consistency of your data. Some of these conditions are errors within GPFS processing but most represent a failure of the surrounding environment. Examples are the lack of quorum or the failure of sufficient disks. In all cases the daemon will exit and restart after recovery has been performed on its behalf by another node. Applications executing at the time of the

failure will see either ENODEV or ESTALE errors. The ENODEV errors are generated by AIX until the daemon has restarted. The ESTALE error is generated by GPFS as soon as it restarts. Open requests are rejected with no such file or no such directory errors until the node has rejoined a valid quorum and mounted its file systems.

Begin problem determination for these errors with the system error log. If there are reports of errors from the PSSP Group Services, the communications network servicing locks, or other GPFS data, or the IBM Virtual Shared Disk, resolve these errors first following the problem determination procedures for these components. If messages indicate lack of quorum, follow the procedures in “Quorum Problems” on page 23. For all other errors follow the procedures for reporting the problem to IBM.

Data Corruption Errors

GPFS takes extraordinary care to ensure the integrity of customer data. However, certain hardware failures or in extremely unusual circumstances the occurrence of a programming error, can cause the loss of data in a file system. GPFS performs extensive checking to validate metadata and ceases using the file system should metadata be inconsistent. This can appear in two ways:

1. The file system will be unmounted and applications will begin seeing ESTALE return codes to file operations, and
2. Error log entries indicating a SG_PANIC and a corruption error are generated

If actual disk data corruption occurs, this will appear on each node in succession. If this happens, examine the error logs on the IBM Virtual Shared Disk servers for any indication of a disk error that has been reported and take appropriate disk problem determination and repair actions prior to continuing. After completing any required disk repair actions run the offline version of the **mmfsck** command on the file system. See “The mmfsck Command” on page 14.

If data corruption errors occur in only one node it is probable that memory structures within the node have been corrupted. In this case, the file system is probably good but a program error exists in GPFS or another authorized program with access to GPFS data structures. Follow the directions in “GPFS Problem Determination Tools” on page 33 and then reboot the node. This should clear the problem. If the problem repeats on one node without affecting other nodes check the GPFS, PSSP, and AIX code levels to determine that they are current and compatible and that no hardware errors were reported.

Loss of File System Access

Loss of file system access usually appears first as an error received by an application. Such errors are normally encountered when the application tries to access an unmounted file system. The most common reason for losing access to a single file system is a failure somewhere in the path to a large enough number of disks to jeopardize your data if operation continues. These errors may be reported in the error log on any node because they are logged in the first node to detect the error. Check all error logs for errors.

Internal errors in processing data on a single file system may also cause loss of file system access. These errors may clear with the execution of the **umount**

command, followed by a remount of the file system, but they should be reported as problems to IBM.

sysctl Problems

Sysctl is used by GPFS administration commands to perform operations on other nodes. The sysctl daemon on the remote node must recognize the command being run and must obtain authorization to execute it. Sysctl might fail due to a number of conditions. Problems involving insufficient authorization suggest numbers 1 and 2 below. An error such as `invalid command name` suggests numbers 3, 4, or 5.

1. A Kerberos ticket must be obtained on the control workstation and any node on which GPFS administration commands are run. Refer to the *PSSP Administration Guide*, or *IBM RS/6000 SP Management, Easy, Lean, and Mean* for information about sysctl and Kerberos.
2. The `/etc/sysctl.mmcmd.acl` file must be modified. The Kerberos realm must be changed and the corrected line(s) uncommented. The file must then be copied to all nodes, including the control workstation.
3. There should be an `mmfs` entry in `/etc/sysctl.conf`. The install step should have done this. The entry should appear as:

```
# Include mmfs sysctl commands
include /usr/lpp/mmfs/bin/mmcmdsystcl
```

4. Sysctl must be restarted after the `sysctl.conf` file has been modified. To do this, issue

```
sysctl svcrestart
```

on all nodes. If this results in a permission error, then perhaps the main sysctl `acl` file (`/etc/sysctl.acl`) was never customized exactly as `sysctl.mmcmd.acl` must be. Sysctl can also be restarted by issuing:

```
stopsrc -s sysctl
startsrc -s sysctl
```

5. `mmfs` must be installed everywhere. In particular, the `mmfs.gpfs` fileset must be installed on the control workstation.

The SDR is Locked

If repeated attempts to run a GPFS administration command fail, reporting that the SDR is locked, the lock can be manually removed by following this procedure.

1. On the control workstation, `cd` to `/var/mmfs/etc`. If the SDR is locked, there will be a directory `mmlock`. If the directory is not there, no lock exists and the problem is probably that sysctl cannot get through to the control workstation.
2. If the `mmlock` is there, examine `mmlock/haslock`. If the file does not exist, there is an internal error. Call service. The file should show a pid and a node name.
3. Go to the node shown and issue a `ps` command against the pid to determine who has the lock. If it is a GPFS administration command that is perhaps stuck, stopping the command (be sure you want to do this) will free the lock. If anything else has the pid, some serious error occurred to the original GPFS command allowing it to die without freeing the lock, AND this new process is using the same pid. The lock should be manually freed.

| 4. To manually remove the lock, on the control workstation issue:
| remove -rf mmlock
| from the **/var/mmfs/etc** directory.

GPFS Problem Determination Tools

This describes the problem determination tools provided with GPFS. These tools are intended for the use of AIX expert service personnel who know how to collect data and run debugging routines. The information is organized as follows:

- “GPFS Error Logs” on page 34
- “The Console Log” on page 38
- “Command and Application Returns” on page 39
- “GPFS Trace Facility” on page 40
- “The mmfsadm Command” on page 40
- “A Script for Generating GPFS Trace Reports” on page 41
- “A Script for Collecting GPFS Service Information” on page 42
- “mmsnap2 Utility” on page 42
- “Reporting a Problem to IBM” on page 44

Note: GPFS error log and messages contain the MMFS prefix. This is not an error. GPFS shares many components with the IBM Multi-Media LAN Server, a related product.

Non-Standard Errors

The following error numbers are used by GPFS in addition to the standard defined ones. These may be reported in error logs or returned to applications. The user action depends on the specific error

E_OFFLINE = 208, Operation failed because a disk is offline

This is most commonly returned when an open of a disk fails. Since GPFS will attempt to continue operation with failed disks, this will be returned when the disk is first needed to complete a command or application request. If this return code occurs, check your VSD for fenced or stopped states and check to determine if the network path exists. In rare situations, this will be reported if VSD definitions are incorrect.

E_NO_MGR = 212, The current stripe group manager failed and no new manager could be appointed.

This usually occurs when a large number of disks are unavailable or when there has been a major network failure. Run **mmlsdisk** to determine whether disks have failed and take corrective action if they have (**mmchdisk**). Check the error log on all nodes for indications of switch or RVSD errors.

E_VALIDATE = 214, Invalid checksum or other consistency check failure on disk data structure.

This indicates that internal checking has found an error in a metadata structure. The severity of the error depends on which data structure is involved. The cause of this is usually GPFS software, disk hardware or other software between GPFS and the disk. Running **mmfsck** should repair the error. The urgency of this depends on whether the error prevents access to some file or whether basic metadata structures are involved.

E_CONFIG = 215, Configuration invalid or inconsistent between different nodes.

This error is returned when the levels of software on different nodes cannot coexist. See the *GPFS Installation and Administration Guide* for the requirements of each release.

E_ALL_UNAVAIL = 218, A replicated read or write failed because none of the replicas were available.

Multiple disks in multiple failure groups are unavailable. Follow the procedures in this document for unavailable disks.

E_NO_QUOTA_INST = 237, Quota management was not correctly installed for the file system, resulting in inconsistent coding of the options stanza in the /etc/filesystem across the nodes.

Check the */etc/filesystem* options stanza for the quota option (**quota=userquota;groupquota**) on all nodes.

E_NO_QMGR = 245, Quota management is installed but no node is acting as the quota manager.

The error indicates a problem with the quota file. Run **mmcheckquota** and then remount the system.

See “Quota Problems” on page 27.

E_PANIC = 666, A file system has been forcibly unmounted because of an error. Most likely due to the failure of one or more disks containing the last copy of metadata.

See the error log for details.

E_NO_QUORUM = 668, There are no longer sufficient nodes available to form a quorum.

See “Quorum Problems” on page 23.

GPFS Error Logs

In normal operation GPFS writes error data to the system error log and to a console file. The console file is contained in the */var/adm/ras* directory on each node. The console file is named **mmfs.log.'date'** where *date* is the time stamp when the current instance of GPFS was started on this node. You may remove old copies of the console log if you are not actively pursuing a problem related to that GPFS startup.

The console log contains information that can be used should GPFS fail during initialization. When initialization completes and a quorum is achieved the message: **MMFS: 6027-300 mmfsd ready for sessions** is written in the console log. Should GPFS monitoring detect a serious error and terminate or panic a stripe group, diagnostic information is written to the console log. Have the console log available if you call IBM for assistance related to an error of this type. The error log contains information about several classes of events or errors. These classes are:

- “MMFS_LONGDISKIO” on page 35
- “MMFS_DISKFAIL” on page 35
- “MMFS_SGPANIC” on page 35
- “MMFS_ABNORMAL_SHUTDOWN” on page 35
- “MMFS_GENERIC” on page 36

- “MMFS_ENVIRON” on page 36
- “MMFS_PHOENIX” on page 36
- “MMFS_FSSTRUCT” on page 36
- “MMFS_RECOVERY” on page 37
- “MMFS_QUOTA” on page 37

MMFS_LONGDISKIO

The MMFS_LONGDISKIO indicates that GPFS is experiencing very long response time for disk requests. This is a warning message and may indicate that your disk system is overloaded or that a failing disk is requiring many I/O retries. Follow AIX instructions for monitoring the performance of your I/O subsystem on this node and on any disk server nodes that may be involved. The data portion of this error record specifies the disk involved. There may be related error log entries from the disk or IBM Virtual Shared Disk subsystems that will pinpoint the actual cause of the problem.

MMFS_DISKFAIL

The MMFS_DISKFAIL error log indicates that GPFS has detected the failure of a disk and forced the disk to the stopped state. “Unable to Access Disks” on page 18 describes the actions taken in response to this error. This is ordinarily not a GPFS error but a failure in the disk subsystem or the path to the disk subsystem.

MMFS_SGPANIC

The MMFS_SGPANIC error log means that GPFS has discovered a condition which may result in data corruption if operation with this file system continues from this node. This may be the result of an internal processing error within the file system or the loss of a path to all disks containing a critical data structure. GPFS has marked the file system as disconnected and applications accessing files within the file system will receive ESTALE errors.

Check the disk subsystem, the switch subsystem and the IBM Virtual Shared Disk using the procedures described in their problem determination information. If the problem is transient, normal operation can be restored by unmounting and remounting the file system on all nodes where this error occurred.

Many of the events that cause this class of error will occur on multiple nodes giving similar, but not identical, error reporting details. If there is no associated failure that would prevent access to disk structures or access to other GPFS nodes with the file system mounted, contact IBM service.

MMFS_ABNORMAL_SHUTDOWN

The MMFS_ABNORMAL_SHUTDOWN error log means that GPFS has determined that it must shutdown all operations on this node because of a problem. This is most likely caused by some interaction with the PSSP Group Services component. Group services failures will result in abnormal shutdown, as will loss of quorum. Unavailability of adequate memory on the node to handle critical recovery situations can also cause this error. In general there will be other error log entries from GPFS or some other component associated with this error log entry.

MMFS_GENERIC

The MMFS_GENERIC error log entry means that GPFS self diagnostics have detected an internal error, or that additional information is being provided with a SG_PANIC report. If the record is associated with an SG_PANIC report, the event code fields in the records will be the same. The error code and return code fields may describe the error. See “Messages” on page 47 for a listing of codes generated by GPFS.

If the error is generated by the self diagnostic routines, service personnel should interpret the return and error code fields since the use of these fields varies by the specific error. Errors caused by the self checking logic will result in the shutdown of GPFS on this node.

MMFS_GENERIC errors may result from an inability to reach a critical disk resource. These errors may look different depending on the specific disk resource (logs, allocation maps, etc.) that has become unavailable. This type of error will usually be associated with other error indications. Other errors generated by disk subsystems, high availability components and communications components at the same time or immediately preceding the GPFS error should be pursued first because they may be the cause of these errors. MMFS_GENERIC error indications without an associated error of those types represent a GPFS problem that requires IBM service.

MMFS_ENVIRON

These records are associated with other records of the MMFS_GENERIC or MMFS_SGPANIC types. They indicate that the root cause of the error is external to GPFS and usually in the network that supports GPFS. Check the network and its physical connections. The data portion of this record supplies the return code provided by the communications code.

MMFS_PHOENIX

The MMFS_PHOENIX error log entries reflect a failure in GPFS interaction with Group Services. See *IBM Parallel System Support Programs for AIX: Diagnosis and Messages Guide* for help in isolating the problem. These errors are usually not GPFS problems, although they may disrupt GPFS operation.

MMFS_FSSTRUCT

The MMFS_FSSTRUCT error log entry indicates that GPFS has detected a problem with the on-disk structure of the file system. The severity of these errors depends on the exact nature of the inconsistent data structure. If it is contained to a single file, EIO errors will be reported to the application and operation will continue. If the inconsistency affects vital metadata structures, operation will cease on this file system.

These errors are often associated with an SG_PANIC error log entry, and these will probably occur on all nodes. Collect trace data on the error as described in “GPFS Trace Facility” on page 40. If the error occurs on all nodes, some critical piece of the file system is inconsistent. This may occur as a result of a GPFS error or an error in the disk system. Execution of the **mmfsck** command may repair the error.

MMFS_QUOTA

Used when GPFS detects a problem in the handling of quota information. This error log is created when the quota manager has a problem reading or writing the quota file. If the quota manager cannot read all entries in the quota file when mounting a stripe group with quotas installed, the quota manager shuts down, but stripe group manager initialization continues. Client mounts will not succeed and will return an appropriate error message.

Check that the disks containing the quota file are available. It may be necessary to run offline quota-check (**mmcheckquota**) to repair or recreate the quota file.

MMFS_RECOVERY

These records are written to indicate that recovery of the specified failure is occurring on this node. It is informational only unless a failure occurs during recovery. These records are provided because recovery code may not generate an error log.

Error Log Entry Example

The following is an example of an error log entry:

LABEL: MMFS_ABNORMAL_SHUTD
IDENTIFIER: 1FB9260D

Date/Time: Mon Jun 29 13:25:58
Sequence Number: 105
Machine Id: 000038044C00
Node Id: c154n01
Class: S
Type: PERM
Resource Name: mmfs

Description
SOFTWARE PROGRAM ABNORMALLY TERMINATED

Probable Causes
SOFTWARE PROGRAM

Failure Causes
SOFTWARE PROGRAM

Recommended Actions
CONTACT APPROPRIATE SERVICE REPRESENTATIVE

Detail Data
COMPONENT ID
576527500
PROGRAM
mmfsd
DETECTING MODULE
/fs/mmfs/ts/phoenix/PhoenixInt.C
MAINTENANCE LEVEL
LABEL: MMFS_ABNORMAL_SHUTD
IDENTIFIER: 1FB9260D

```
Date/Time:      Mon Jun 29 13:25:58
Sequence Number: 105
Machine Id:     000038044C00
Node Id:        c154n01
Class:          S
Type:           PERM
Resource Name:  mmfs
Description
SOFTWARE PROGRAM ABNORMALLY TERMINATED
```

```
Probable Causes
SOFTWARE PROGRAM
```

```
Failure Causes
SOFTWARE PROGRAM
```

```
Recommended Actions
CONTACT APPROPRIATE SERVICE REPRESENTATIVE
```

```
Detail Data
      0
COMPONENT ID
      576527500
PROGRAM
mmfsd
DETECTING MODULE
/fs/mmfs/ts/phoenix/PhoenixInt.C
MAINTENANCE LEVEL
version1
LINE
      2323
RETURN CODE
      668
REASON CODE
0000 0000
EVENT CODE
```

The Console Log

Although it is not an error log, **var/adm/ras/mmfs.log**, the GPFS console log, can sometimes provide insight to help assess problems. This file has a timestamp appended to the filename when GPFS is active. If no timestamp appears as part of the filename, the GPFS daemons have not started. Look for the following if you cannot access the file system:

- mmfs is waiting for switch

If this message appears, the switch must be started. Procedures for switch startup vary according to environment and switch type (HPS or SPF), and involve using the **Eunfence** and **Estart** commands. Refer to the *Administration Guide* and the *IBM Parallel System Support Programs for AIX: Command and Technical Reference* for more information.

- mmfs is waiting for rvsd

This indicates that the IBM Virtual Shared Disks must be brought up. Procedures, which may vary according to environment, involve using the **ha.vsd reset**, **ha.vsd start**, and **startsrc -g rvsd** commands.

- VSD fence ioctl failed with err 46 and Disk list

This indicates that the IBM Virtual Shared Disks are in a stopped state and must be started. Procedures may vary according to environment but involve using the **ha.vsd reset** and **ha.vsd start** commands. Refer to *Managing Shared Disks* for details.

- My adapter css0 has failed. Shutting down the daemon.

A switch failure shut down the GPFS daemon on the node. This is recorded by closing the console file, after which GPFS recovery restarts the daemon and creates another new console file name **mmfs.log** (with no timestamp appended until the daemon is active).

Command and Application Returns

The messages returned by commands, whether those explicitly entered or those executed by an application, may contain some indication that file access has been lost.

Returns from User Applications

- write failed: I/O error OR read failed: I/O error

A disk may be broken or an IBM Virtual Shared Disk may be fenced, stopped, or suspended. Check the disk state using the **lsvsd -l** command. A state of ACT (active) indicates a communicating disk.

- write failed: Connection timed out OR read failed: Connection timed out

This is an IBM Virtual Shared Disk error. Consider tuning or evaluating the IBM Virtual Shared Disk configuration, including number and size of buddy buffers. Refer to *Managing Shared Disks* for more information.

Return from mfsadm Command

mmfs daemon not running

If this appears, start GPFS. Refer to *GPFS Installation and Administration Guide* for details.

Return from mmcrfs Command

MMFS: 6027-538 Error accessing disks

mmcrfs: tscrfs failed. Cannot create /dev/*stripe_group_name*

Check your IBM Virtual Shared Disks (**mmfsdisk can be used**) to determine whether they are in a stopped state. If so, follow procedures in *GPFS: Installation and Administration Guide* or *Managing Shared Disks* to bring them up. If not, refer to “Unable to Access Disks” on page 18.

Return from errpt Command

(rc=78) vsd_Timer: Timeout

Consider tuning or evaluating the IBM Virtual Shared Disk configuration, including number and size of buddy buffers. Refer to *Managing Shared Disks* for more information.

Return from df Command

df: /stripe_group_name: A file, file system, or message queue
is no longer available

If this message is displayed, try accessing the file system again after a reasonable period of time has passed. If the result is the same, check for an error log entry.

GPFS Trace Facility

The GPFS code is equipped with a large number of trace points to allow rapid problem determination of failures. GPFS tracing is based on the AIX trace facility and uses trace IDs 306,307,308, and 309. The granularity of the trace is controlled using “The mmfsadm Command.” If IBM service asks for a trace on a problem that you must recreate, issue a command similar to the following:

```
trace -a -l -L 4000000 -T 2000000 -j 306,307,308,309 &
```

The **-L** and **-T** parameters control the size of the trace file and trace buffers, and may need to be larger or smaller depending on the exact problem being debugged. When the event to be captured occurs, issue a **trcstop** command as soon as possible. Since GPFS is designed to recover from many errors and limit the impact of others, its continued operation will generate additional trace records that will overlay the ones of interest. This may mean that scripts, which invoke the application, must manipulate the trace facility.

Since the exact formatting of traces may vary by the service level of the system, you may be asked to format the trace on the generating system. This process consists of running the AIX **trcrpt** command using the format file associated with your exact level of the code:

```
trcrpt -0 exec=on,pid=on,tid=on -o /tmp/outputfile
```

IBM service personnel will give you specific instructions on validating the trace contents for the event to be captured. This should be checked because continued operation may overlay the needed trace entries.

Alternatively, a shell script has been provided to enable the AIX system trace facility for capturing GPFS trace events rather than issuing the individual trace commands. See “A Script for Generating GPFS Trace Reports” on page 41.

The mmfsadm Command

The **mmfsadm** command provides information display and change capabilities intended for use by trained service personnel. We suggest that this command not be run except under the direction of such personnel. Several options of this command, which may be required for use, are described as follows:

showtrace shows the current level of tracing for all of the subclasses of tracing available in GPFS. Trace level 9 provides the highest level of tracing for the class and trace level 0 provides no tracing. Intermediate values exist for most classes. More tracing requires more storage and results in a higher probability of overlaying the required event.

trace class n Sets the value of tracing to the value specified by *n*. Actual trace gathering only occurs when the AIX trace command has been issued.

dump *what* Dumps the state of a large number of internal state values that may be useful in determining the sequence of events. The *what* parameter may be set to **all** indicating that all available data should be collected, or another value indicating more restricted collection of data. The output is presented to stdout and should be collected by redirecting stdout.

quorum *n* | **reset** Changes the required value for quorum to *n* or to the calculated value (reset). This parameter should never be used except under service direction because it may affect the consistency of your data.

There are other options that provide interactive GPFS debugging, which are not described here. **mmfsadm** output will be required in almost all cases where a GPFS incident is being reported. **mmfsadm** collects state only on the node where it is executed. Depending on the nature of the problem, **mmfsadm** output may be required from several nodes, or all nodes. **mmfsadm** output from the stripe group manager is often required. To determine where the stripe group manager is, issue

```
mmfsadm dump cfgmgr
```

on any node. This will report the IP address of the configuration manager in output similar to this: Cluster configuration manager is 129.40.163.134. If you issue

```
mmfsadm dump cfgmgr
```

on the configuration manager node, the output will show the stripe group manager for each file system. This enables you to collect data for the stripe group managers.

A Script for Generating GPFS Trace Reports

The **mmtrace** shell script provides an alternative to issuing the individual AIX trace commands. The script enables the AIX trace facility for capturing GPFS trace events.

To use the trace:

1. Start the trace by issuing

```
/usr/lpp/mmfs/bin/mmtrace
```
2. Recreate the problem
3. When the event to be captured occurs, issue the following as soon as possible

```
/usr/lpp/mmfs/bin/mmtrace stop
```
4. If the problem results in a shutdown and restart of the file system, automatically stop the trace by uncommenting the following line in

```
/usr/lpp/mmfs/bin/runmmfs  
#MMFSDIR/bin/mmtrace stop
```

In order to restore the system to its normal operating state, once diagnosis of the problem is complete, comment out this line again .

If the problem requires more detailed tracing, IBM service personnel may ask you to modify the GPFS trace levels by modifying the **/usr/lpp/mmfs/bin/mmtrace** script.

A Script for Collecting GPFS Service Information

The various GPFS logs, traces, and dumps can easily be collected by using the **mmsnap** utility.

1. Ensure the SP Log Management (splm) facility is enabled. See “Enabling the SPLM Error Log for Use by GPFS.” If it is not possible to use splm, then GPFS service information can also be collected using **mmsnap2**. See “mmsnap2 Utility.”

2. Log into the failing node and issue the command

```
/usr/lpp/mmfs/bin/mmsnap
```

If service is needed on more than one node, use

```
/usr/lpp/mmfs/bin/mmsnap -a
```

to collect information on all GPFS nodes.

3. Send the resulting compressed tar files and the **splm.log** file to IBM Service.

By default, the service information is collected in the **/tmp/gpfs** directory.

Ensure that sufficient space is available in the directory for the information, or specify a different directory using

```
mmsnap -d my_dir
```

The suggested minimum amount of space is 20M.

Enabling the SPLM Error Log for Use by GPFS

The **mmsnap** utility uses the splm facility. This facility must be configured on a node before collecting service information on that node.

1. If the following line is not present in **/etc/sysctl.conf**, add it and restart **sysctld**

```
include /usr/lpp/ssp/sysctl/bin/logmgt.cmds
```

2. If a principal is not present in the splm ACL file, **/etc/logmgt.acl**, add one

```
_PRINCIPAL root.admin@PPD.POK.IBM.COM
```

3. If the **/etc/splm.allow** file exists, enable GPFS for use of the splm facility by adding the following line

```
/usr/lpp/mmfs/bin/mmfsadm
```

For details on managing error logs, see *Administration Guide*. For complete details on splm, see *IBM Parallel System Support Programs for AIX: Command and Technical Reference*

mmsnap2 Utility

If it is not possible to use the splm facility, GPFS service information can be collected using the **mmsnap2** utility.

1. Log into the failing node and issue the command

```
/usr/lpp/mmfs/bin/mmsnap2 -n node_list -T offon -R -M -I -s date_start -e date_end
```

where *node_list* is the node, or list of nodes separated by commas, requiring service

If service is needed on a node group, use

```
/usr/lpp/mmfs/bin/mmsnap2 -g SP_node_group -T offon -R -M -I -s date_start -e date_end
```

The **—n** and **—g** options cannot be used simultaneously. However, if they are both omitted, the default is all nodes in **/etc/clusters.nodes**.

You must specify at least one of the following options to collect data, or **mmsnap2** will collect no data

- T** traces **off | on | offon | log**
- R** error reports from each node with date/time range
- M** mmfs.log.XX.. files from each node
- I** internal dumps

You must specify the time period to collect the data, or **mmsnap2** will collect all of the data it can find

- s** starting date *mmddHHMMyy*
- e** ending date *mmddHHMMyy*. The default is the current date/time
- h** number of hours to collect data. Do not cross month boundaries with the **—h** option.

To collect data for the last 4 hours you would issue the command

```
mmsnap2 -n node_list -h 4 -d my_dir -T offon -M -R -I
```

Note: If GPFS is hung, the **—I** option will hang. Therefore, it is best to do one of the following:

- Issue the **—I** option in a separate command to ensure the **—M** and **—R** data is captured

```
mmsnap2 -n node_list -h 4 -d my_dir -T offon -M -R
```

```
mmsnap2 -n node_list -h 4 -d my_dir -I
```

- Issue the command with the **—e** option as **mmsnap2** will then collect old dumps without trying to create new ones

```
mmsnap2 -n node_list -h 4 -d my_dir -T offon -M -R -I -e date_end
```

2. Send the resulting files to IBM Service.

By default , the service information is collected in the **/u/DUMPS** directory. You may specify a different directory using

```
mmsnap2 -d my_dir
```

3. For further information, view the top of the **mmsnap2** script file supplied with GPFS.

Hints and Tips

The following topics may prove helpful if you encounter a problem using GPFS.

Cannot Unmount File System? (Device Busy)

If, while attempting to unmount a file system, you get the message:

```
MMFS: 6027-511 Cannot unmount /dev/filesystem_name: Device busy.
```

Try running the

```
ls -l | grep filesystem_name
```

This command returns the file system processes that are currently busy. After identifying the processes, you can kill them or contact the owner and request that they be killed. Once this is done, you should be able to unmount the file system.

Where to Get the `lsof` Command

The `lsof` (list open files) command can be acquired in prebuild binary format for aix from aixpdslib.seas.ucla.edu. You can get the latest version via anonymous ftp from vic.cc.purdue.edu (`cd` to `/pub/tools/unix/lsof`). The inventor of the `lsof` command is Victor A. Abell (abe@cc.purdue.edu), Purdue University Computing Center.

What Disk is Your IBM Virtual Shared Disk On?

If an error report contains a reference to a volume group or an IBM Virtual Shared Disk pertaining to GPFS, you can use the `lsvg` command to list volume groups. For a particular volume group, `g7t8v71g` for example, the command:

```
echo g7t8v71g | lsvg -i1
```

returns:

```
g7t8v71g:
LV NAME          TYPE      LPs  PPp  PVs  LV STATE      MOUNT POINT
lv7t8v71v1n7    jfs             537  537   1    open/syncd    N/A
```

the name of the logical volume, `lv7t8v71v1n7`, appears in row 3 column 1.

Use the `vsdata1st` command to cross reference the logical volume and IBM Virtual Shared Disk:

```
vsdata1st -v | grep lv7t8v71v1n7
```

which returns:

```
g7t8v71g          lv7t8v71v1n7          g7t8v71gn7b          1 nocache          2148
```

Reporting a Problem to IBM

The following data should be available when you report a problem to IBM :

1. A description of the problem
2. Any error logs relating to the event
3. **mmfsadm dump all** output on at least the failing node and the stripe group manager. Some errors require that this be available from all nodes.

To determine where the stripe group manager is, issue:

```
mmfsadm dump cfgmgr
```

on any node. This will report the IP address of the configuration manager in output similar to this :

```
Cluster configuration manager is 129.40.163.134
```

If you issue

```
mmfsadm dump cfgmgr
```

on the configuration manager node, the output will show you the stripe group manager for each file system. This enables you to collect data for the stripe group managers.

4. Console logs and actual application output on the failure
5. Traces on affected nodes, if available. In some cases this may mean all nodes

Messages

6027-300 mmfsd ready for sessions.

Explanation: The **mmfsd** server is up and running.

6027-301 Initialization file *name* could not be run: *error number*

Explanation: Trying to **execvp** the initialization shell script.

User Response: Check file existence and access permissions.

6027-302 Remount failed for mountid *number*.

Explanation: **mmfsd** restarts and tries to remount any filesystems that the VFS layer thinks are still mounted.

User Response: Check errors displayed and the **errno** description.

6027-303 Configuration manager node *name* not found in *filename*.

Explanation: The node list or primary node is not configured correctly.

User Response: Check the node list file and the **tscPrimary** parameter in the configuration file specified in the error message.

6027-304 Remount failed for *name* after daemon restart.

Explanation: The remount fails after daemon restart.

User Response: Perform **mmlsdisk** for any disk problems.

6027-305 Perform mmlsdisk for any disk failures and remount.

Explanation: The remount fails after daemon restart.

User Response: Perform **mmlsdisk** for any disk problems and remount.

6027-310 mmfsd initializing.

Explanation: The **mmfsd** server has started execution.

6027-311 mmfsd is shutting down.

Explanation: The **mmfsd** server or **mmcrfs** is about to terminate.

6027-312 Unknown trace class.

Explanation: The trace class is not recognized.

User Response: Specify a valid trace class.

6027-313 Cannot open configuration file *filename*.

Explanation: After last try the configuration file cannot be opened.

User Response: Verify that the configuration file is present and has the proper protection bits.

6027-314 Must run as root.

Explanation: `mmfsd` server is run without root authority.

User Response: Log in as the root user and retry the failing command.

6027-315 Bad config file entry in *filename*, line *number*.

Explanation: The `config.glb` file has a bad entry.

User Response: Fix the syntax error in the configuration file.

6027-316 Unknown config parameter in *filename*, line *number*.

Explanation: There is an unknown parameter in the configuration file.

User Response: Fix the syntax error in the configuration file.

6027-319 Creating shared segment

Explanation: The shared segment could not be created.

User Response: Check reason given by error message.

6027-320 Mapping shared segment.

Explanation: The shared segment could not be attached.

User Response: Check reason given by error message.

6027-321 Shared segment mapped at wrong address (is *value*, should be *value*).

Explanation: The shared segment did not get mapped to the expected address.

User Response: Contact IBM support.

6027-322 Registering shared segment

Explanation: The shared segment could not be registered.

User Response: Check reason given by error message.

6027-323 Detaching shared segment

Explanation: The shared segment could not be detached.

User Response: Check reason given by error message.

6027-324 Creating message queue for main process

Explanation: The message queue for the main process could not be created.

User Response: This is likely to be an operating system error. Contact IBM support.

6027-329 Cannot pin the main shared segment: *name*

Explanation: Trying to pin the shared segment during initialization.

User Response: Check the `mmfs.cfg` file. The pagepool size may be too large. It cannot be more than 80% of real memory. If a previous `mmfsd` crashed, check for processes that begin with the name "mmfs" which may be holding on to an old pinned shared segment. Issue `mmchconfig` command to change pagepool size.

6027-334 Error initializing internal communications.

Explanation: The mailbox system used by the daemon for communication with the kernel cannot be initialized.

User Response: Increase the size of available memory using **mmchconfig** command.

6027-335 Configuration error: check *name*.

Explanation: A configuration error is found.

User Response: Check **mmfs.cfg** file and other error messages.

6027-336 Value *value* for configuration parameter *name* is invalid. Check *name*.

Explanation: A configuration error is found.

User Response: Check **mmfs.cfg** file.

6027-337 Waiting for resources to be reclaimed before exiting.

Explanation: The **mmfsd** daemon is attempting to terminate, but cannot because data structures in the daemon shared segment may still be referenced by kernel code. This message may be accompanied by other messages that show which disks still have I/O in progress.

6027-338 Waiting for *number* user(s) of shared segment to release it.

Explanation: The **mmfsd** daemon is attempting to terminate, but cannot because some process is holding the shared segment while in a system call. The message will repeat every 30 seconds until the count drops to zero.

User Response: Find stuck process, and find a way to get it out of its system call.

6027-340 Could not retrieve network interface configuration.

Explanation: The socket or ioctl call to retrieve the interface configuration failed.

User Response: Check reason given by error message.

6027-341 Invalid *hostname* in *filename*.

Explanation: The IP address of a host listed in the node list file could not be retrieved.

User Response: Check node list file specified in error message.

6027-342 Duplicate address in *filename*.

Explanation: The IP address appears more than once in the node list file.

User Response: Check node list file specified in error message.

6027-343 Local host not found in *filename*.

Explanation: The local host specified in the node list file could not be found.

User Response: Check node list file specified in error message.

6027-344 Empty node list file *filename*.

Explanation: The node list file is empty.

User Response: Check node list file specified in error message.

6027-345 Fatal Error from the network rc=*number* Check your network connectivity.

Explanation: A TCP error has caused GPFS to exit due to a bad return code from an error. Exiting allows recovery to proceed on another node and resources are not tied up on this node.

User Response: Follow network problem determination procedures.

6027-346 Incompatible daemon version. My version =*number*, repl.my_version = *number*.

Explanation: The GPFS daemon tried to make a connection with another GPFS daemon. However, the other GPFS daemon is not the same version and sends a reply indicating its version number is incompatible.

User Response: Verify your GPFS daemon version.

6027-347 Remote host *name* refused connection because IP address *number* was not in the node list file.

Explanation: The GPFS daemon tried to make a connection with another GPFS daemon. However, the other GPFS daemon sent a reply indicating it did not recognize the IP address of the connector.

User Response: Add the IP address of the local host to the node list file on the remote host.

6027-362 Warning: No disks were deleted, but some data was migrated. The file system may no longer be properly balanced.

Explanation: **mmdeledisk** command failed after it started migrating data off disks being deleted. The disks were restored to normal ready, status, but the migration has left the file system unbalanced.

User Response: Run the **mmrestripefs** command with the **-b** option to restore proper balance if rebalancing required.

6027-363 Error writing disk descriptor for disk *name*.

Explanation: I/O error occurred when **mmadddisk** command was writing a disk descriptor on a disk.

User Response: See error code information that follows. Then reissue command.

6027-364 Error accessing disks.

Explanation: I/O error occurred when **mmadddisk** command was reading a disk descriptor on a disk.

User Response: See error code information that follows. Then reissue command.

6027-365 Too many disks being added.

Explanation: More disks are being added with **mmadddisk** command than can fit in a file system stripe group.

User Response: Reduce the number of disks.

6027-369 Error writing stripe group descriptor for disk *name*.

Explanation: **mmadddisk** detected an I/O error while writing a stripe group descriptor on a disk.

User Response: See error code information that follows. Then reissue command.

6027-370 mmdeldisk completed.

Explanation: The **mmdeldisk** command has completed.

6027-371 Cannot delete all disks in the file system

Explanation: Trying to delete all the disks in a file system.

User Response: Either reduce the number of disks to be deleted or use the **mmdelfs** command to delete the file system.

6027-372 Replacement disk must be in the same failure group as the disk being replaced.

Explanation: An improper failure group was specified for **mmrpldisk**.

User Response: Specify a failure group in the disk descriptor for the replacement disk that is the same as the failure group of the disk being replaced.

6027-373 Disk *name* is being replaced, so status of disk must be replacement.

Explanation: Failed replace operation because the new disk does not have the correct status specified in **mmrpldisk** command.

User Response: Use the **mmlsdisk** command to display disk status, or **mmchdisk** to change it. Specify a new disk that has a status of **replacement**.

6027-374 Disk *name* may not be replaced.

Explanation: A disk being replaced with **mmrpldisk** does not have a status of **ready** or **suspended**.

User Response: Use the **mmlsdisk** command to display disk status. Specify a disk that has a status of **ready** or **suspended**.

6027-375 Disk *name* already in stripe group.

Explanation: Replacement disk name specified in **mmrpldisk** command already exists in the stripe group.

User Response: Rename or remove duplicate disk.

6027-376 Previous replace command must be completed before starting a new one.

Explanation: Status of other disks specified in **mmrpldisk** command shows that a replace command did not complete.

User Response: Use the **mmlsdisk** command to display disk status. Retry the failed **mmrpldisk** command or use **mmchdisk** change the status of the disks that have a status of **replacing** or **replacement**.

6027-377 Cannot replace a disk that is in use.

Explanation: Attempting to replace a disk in place, but the disk specified in **mmrpldisk** command is still available for use.

User Response: Use the **mmchdisk** command to stop use of the disk.

6027-378 I/O still in progress near sector *number* on disk *name*.

Explanation: The `mmfsd` daemon is attempting to terminate, but cannot because data structures in the daemon shared segment may still be referenced by kernel code. In particular, the daemon has started an I/O that has not yet completed. It is unsafe for the daemon to terminate until the I/O completes, because of asynchronous activity in the device driver that will access data structures belonging to the daemon.

User Response: Either wait for the I/O operation to time out, or issue a device-dependent command to terminate the I/O.

6027-379 Could not invalidate disk(s).

Explanation: Trying to delete a disk and it could not be written to in order to invalidate its contents.

User Response: No action needed if removing that disk permanently, however if the disk is ever to be used again the `-v` flag must be specified with a value of `no` when using either the `mmcrfs` or `mmaddisk` command.

6027-380 Disk name missing from disk descriptor list entry *number*.

Explanation: No disks were named.

User Response: Check argument list.

6027-381 Too many items in disk descriptor list entry *number*.

Explanation: Too many items with colons given for one disk list.

User Response: Each disk specification has a maximum of 6 components. Check argument list syntax.

6027-383 Disk *name* appears more than once.

Explanation: A duplicate name was found when parsing disk lists.

User Response: Remove duplicate

6027-384 Disk *name* already in stripe group.

Explanation: A disk name was found to already exist in the stripe group when parsing disk lists.

User Response: Rename or remove duplicate disk.

6027-387 *value* for the 'failure group' option for disk *name* is out of range. Valid values are *number* through *number*.

Explanation: The failure group value given is invalid.

User Response: Specify correct failure group.

6027-388 *value* for the 'failure group' option for disk *name* is invalid.

Explanation: The failure group value given is invalid.

User Response: Specify correct failure group.

6027-389 *value* for the 'has metadata' option for disk *name* is out of range. Valid values are *number* through *number*.

Explanation: The `disk usage` parameter specified is invalid.

User Response: Specify correct `disk usage` parameter.

6027-390 *value for the 'has metadata' option for disk name is invalid.*

Explanation: The *disk usage* parameter specified is invalid.

User Response: Specify correct *disk usage* parameter.

6027-391 *value for the 'has data' option for disk name is out of range. Valid values are number through number.*

Explanation: The *disk usage* parameter specified is invalid.

User Response: Specify correct *disk usage* parameter.

6027-392 *value for the 'has data' option for disk name is invalid.*

Explanation: The *disk usage* parameter specified is invalid.

User Response: Specify correct *disk usage* parameter.

6027-393 *Either the 'has data' option or the 'has metadata' option must be '1' for disk name.*

Explanation: The *disk usage* parameter specified is invalid.

User Response: Specify correct *disk usage* parameter.

6027-414 *command is not defined.*

Explanation: Trying to execute a command on an undefined file system.

User Response: Verify that the spelling is correct and/or add the file system to the system.

6027-416 **Incompatible file system stripe group descriptor version or not formatted.**

Explanation: An invalid file system stripe group descriptor was encountered

User Response: Check disk hardware. If the problem persists, run offline **mmfsck**. If the problem still persists, reformat the file system.

6027-417 **Bad file system stripe group descriptor.**

Explanation: An invalid file system stripe group descriptor was encountered

User Response: Check disk hardware. If the problem persists, run **mmfsck** with the file system unmounted. If the problem still persists, reformat the file system.

6027-418 **Inconsistent file system stripe group quorum. readQuorum=number writeQuorum=number quorumSize=number.**

Explanation: An invalid file system stripe group descriptor was encountered

User Response: Start any disks that have been stopped by the **mmchdisk** command or by hardware failures. If the problem persists, run **mmfsck** with the file system unmounted. If the problem persists, reformat the file system.

6027-419 **Failed to read a stripe group descriptor.**

Explanation: Not enough valid replicas of the file system stripe group descriptor could be read from the file system stripe group.

User Response: Start any disks that have been stopped by the **mmchdisk** command or by hardware failures. If the problem persists, run **mmfsck** with the file system unmounted. If the problem persists, reformat the file system.

6027-427 Indirect block size must be at least as large as the minimum fragment size.

Explanation: The indirect block size was specified as a value smaller than 512 bytes.

User Response: Specify a new indirect block size that is a multiple of 512 bytes.

6027-428 Indirect block size must be a multiple of the minimum fragment size.

Explanation: The indirect block size specified is not a multiple of 512 bytes.

User Response: Specify a new indirect block size that is a multiple of 512 bytes.

6027-429 Indirect block size must be less than full data block size.

Explanation: The indirect block size specified is not smaller than data block size.

User Response: Specify a new indirect block size that is smaller than the data block size and is a multiple of 512 bytes.

6027-430 Default metadata replicas must be less than or equal to default maximum number of metadata replicas.

Explanation: The default metadata replicas specified is larger than the value specified for maximum metadata replicas.

User Response: Specify a value for default metadata replicas that is equal to or smaller than the value specified for maximum metadata replicas.

6027-431 Default data replicas must be less than or equal to default maximum number of data replicas.

Explanation: The default data replicas specified is larger than the value specified for maximum data replicas.

User Response: Specify a value for default data replicas that is equal to or smaller than the value specified for maximum data replicas.

6027-432 Default maximum metadata replicas must be less than or equal to *number*.

Explanation: The maximum metadata replicas specified is larger than 2.

User Response: Specify a value of 1 or 2 for maximum metadata replicas.

6027-433 Default data replicas must be less than or equal to *number*.

Explanation: The default data replicas specified is larger than 8.

User Response: Specify a value of 8 or less for default data replicas.

6027-434 Indirect blocks must be at least as big as i-nodes.

Explanation: The indirect block size was specified as a value smaller than 512 bytes.

User Response: Specify a new indirect block size that is a multiple of 512 bytes.

6027-438 Duplicate disk name *name*.

Explanation: A duplicate name was found when parsing disk lists.

User Response: Remove duplicate.

6027-441 Unable to open disk *name*.

Explanation: Invalid disk name specified in `mmcrfs` command.

User Response: Correct the indicated command parameters.

6027-445 Value for option '-m' cannot exceed the number of metadata failure groups.

Explanation: The current number of replicas of metadata cannot be larger than the number of failure groups that are enabled to hold metadata.

User Response: Use a smaller value for **-m** in the **mmchfs** command, or increase the number of failure groups by adding disks to the file system stripe group.

6027-446 Value for option '-r' cannot exceed the number of data failure groups.

Explanation: The current number of replicas of data cannot be larger than the number of failure groups that are enabled to hold data.

User Response: Use a smaller value for **-r** in the **mmchfs** command, or increase the number of failure groups by adding disks to the file system stripe group.

6027-451 No disks= list found in mount options.

Explanation: No "disks=" clause found in the mount options list when opening a file system.

User Response: Check **/etc/filesystems** stanza for this file system.

6027-452 No disks found in disks= list.

Explanation: Opening a file system, no disks listed.

User Response: Check **/etc/filesystems** stanza for this file system.

6027-453 No disk name found in a clause of the disks= list.

Explanation: When opening a file system, no disk name given.

User Response: Check **/etc/filesystems** stanza for this file system.

6027-461 Unable to find *name* device.

Explanation: Opening a file system.

User Response: There must be a **/dev/*name*** special device defined. Check error code.

6027-462 *name* must be a char or block special device.

Explanation: Opening a file system.

User Response: There must be a **/dev/*name*** special device defined.

6027-468 Disk *variable* listed in **/etc/filesystems not found in device *variable*.**

Explanation: Trying to access a file system and the disks listed in the stanza in **/etc/filesystems** for the device do not exist in the file system.

User Response: Check the validity of the stanza in **/etc/filesystems**.

6027-469 File system *name* does not match *descriptor*.

Explanation: The file system name found in the descriptor on disk does not match the corresponding device name in **/etc/filesystems**.

User Response: Check **/etc/filesystems**

6027-470 Disk *name* may still belong to an active stripe group.

Explanation: The disk being added by the **mmcrfs**, **mmaddisk**, or **mmrpldisk** command appears to still belong to some stripe group.

User Response: Verify that the disk(s) you are adding do not belong to an active stripe group, and use the **-v no** option to bypass this check. Use this option only if you are certain that no other stripe group has this disk configured because you can cause data corruption in both file systems if this is not the case.

6027-471 You have requested the file system to be upgraded to version *number*. This will enable new functionality but will prevent you from using the file system with earlier releases of GPFS. Do you want to continue?

Explanation: Verification request in response to the **mmchfs -V** command. This is a request to upgrade the file system and activate functions which are incompatible with the previous release of GPFS.

User Response: Enter **yes** if you want the conversion to take place.

6027-472 Failed to enable extended attribute support.

Explanation: An error occurred while attempting to upgrade the file system to a version that supports extended attributes. The file to store extended attribute data could not be created. An additional GPFS message will be issued providing more specific information.

User Response: Follow the recommended action for the other message that occurs with this one.

6027-473 *file system*: no *quota type* quota management installed.

Explanation: *quota type* quotas are not established for *file system*.

6027-474 Recovery Log I/O Failed, unmounting *file system* RC = *number*.

Explanation: Log I/O could not be completed in writelog.C or recoverlog.C.

User Response: Check the paths to all disks making up the file system. Run **mmfsdisk** to determine if GPFS has declared the disks unavailable. Repair any paths to the disks which have failed. Remount the file system.

6027-475 The option -F is not enabled. Use option -V to enable the most recent features.

Explanation: **mmchfs -F** is not enable under the current file system format version.

User Response: Issue **mmchfs -V** to change the file system format to the latest format supported.

Note: This is a request to upgrade the file system and activate functions which are incompatible with the previous release of GPFS.

6027-476 Volume label of disk *name* is invalid.

Explanation: A **mount** command found a disk with an invalid disk descriptor.

User Response: Examine the console log of this node for more information on the bad volume label. See message **6027-1501** or **6027-1502**.

6027-500 *name* is loaded and configured.

Explanation: The kernel extension is loaded and configured.

6027-502 Invalid parameter: *name*.

Explanation: `mmfsmnthelp` is called with an invalid parameter.

User Response: Contact IBM support.

6027-504 Not enough memory to allocate internal data structure.

User Response: increase ulimit or paging space

6027-505 Internal error, aborting.

User Response: Report error to IBM support.

6027-506 *name* is already loaded at *address*.

Explanation: If a program is already loaded, the address is displayed.

6027-508 `installmmfs: [-slq] [-L library] {module | module_id}`

Explanation: `installmmfs` was improperly used.

User Response: Contact IBM support.

6027-509 `installmmfs: [-slq] [-L library] -p NFSWatchPid {module | module_id}`

Explanation: `installmmfs -r` was improperly used.

User Response: Contact IBM support.

6027-510 Cannot mount *file system* on *mountpoint*.

Explanation: There is an error mounting the GPFS file system. There is a maximum limit of 32 file systems that may exist within a GPFS configuration.

User Response: Determine action indicated by error description. If the error description is *A system call does not exist*, GPFS was not started before it was configured on this node. You will need to

1. Stop GPFS
2. Configure GPFS by issuing `mmaddnode`
3. Reboot the affected nodes

For more information see the *GPFS: Installation and Administration Guide* section on *Starting and Stopping GPFS*.

6027-511 Cannot unmount *file system*.

Explanation: There is an error unmounting the GPFS file system.

User Response: Determine action indicated by errno description.

6027-512 *name* not listed in `/etc/vfs`

Explanation: Error occurred while installing the GPFS kernel extension, or when trying to mount a filesystem.

User Response: Check for "mmfs" entry in `/etc/vfs`

6027-513 Mount of *file system* on *mount point* partially completed. Use unmount command to clean up.

Explanation: Mounting a file system and a signal causes early termination.

User Response: Issue appropriate unmount command.

6027-514 Cannot mount *file system* on *mount point*. Already mounted.

Explanation: An attempt to mount a file system that is already mounted has been made.

6027-531 The following disks of *name* will be formatted: *list*.

Explanation: Output in response to the **mmcrfs** command.

6027-538 Error accessing disks.

Explanation: **mmcrfs** command encountered an error accessing one or more of the disks named in the **-d** option

User Response: Verify that the **-d** option was coded correctly, and that all named disks exist and are online.

6027-539 Unable to clear descriptor areas for *name*.

Explanation: **mmcrfs** command encountered an error overwriting the descriptor of the disks named in the **-d** option.

User Response: Verify that the **-d** option was coded correctly, and that all named disks exist and are online.

6027-541 Error formatting file system.

Explanation: **mmcrfs** command encountered an error while formatting a new file system.

User Response: Follow instructions for other messages that appear with this one.

6027-543 Error writing the descriptor for *name*.

Explanation: **mmcrfs** could not successfully write the stripe group descriptor in a particular file system.

User Response: Check system error log, rerun **mmcrfs**.

6027-544 Could not invalidate *disk* of *file system*.

Explanation: A disk could not be written to invalidate its contents.

User Response: Make sure the indicated logical volume is writable.

6027-545 Error writing disk descriptor for disk *name*.

Explanation: An error occurred while formatting a particular disk.

User Response: Check system error log, rerun **mmcrfs**

6027-549 Failed to open *name*.

Explanation: The mount command is unable to access a file system.

User Response: Follow the recommended actions for the other messages that occur with this one.

6027-551 *name is still in use.*

Explanation: The `mmdelfs` OR `mmcrfs` command found that the named file system is still mounted or that another GPFS command is running against the file system.

User Response: Unmount the file system if it is mounted, or wait for GPFS commands in progress to terminate before retrying the command.

6027-553 **Scan failed on user or system file:** *filename.*

Explanation: Data may be lost as a result of invalid pointers.

User Response: Some files may have to be restored from backup copies.

6027-554 **Scan failed on** *number* **out of** *number* **user or system files.**

Explanation: Data may be lost as a result of invalid pointers.

User Response: Some files may have to be restored from backup copies.

6027-555 **The desired replication factor exceeds the number of available failure groups.**

Explanation: You have specified a number of replicas which exceeds the number of failure groups available.

User Response: Reissue command with fewer replicas.

6027-556 **Not enough space for the desired number of replicas.**

Explanation: In attempting to restore the correct replication, MMFS ran out of space in the file system. Operation can continue but some data is not fully replicated.

6027-557 **Not enough space or available disks to properly balance the file.**

Explanation: In attempting to stripe data within the file system; data was placed on a disk other than the desired one. This is normally not a problem. `mmrestripe` can be run to rebalance all files.

6027-558 **Some data are unavailable.**

Explanation: An I/O error has occurred or some disks are in the stopped state.

User Response: Check the availability of all disks and the path to all disks and re-execute the command.

6027-559 **Some data could not be read or written.**

Explanation: An I/O error or some disks are in the stopped state.

User Response: Check the availability of all disks and the path to all disks and re-execute the command.

6027-561 **Error migrating log.**

Explanation: There are not sufficient available disks to continue operation.

User Response: Restore the unavailable disk and re-execute the command.

6027-562 **Error processing i-nodes.**

Explanation: There is not an I/O path to critical metadata or metadata has been corrupted.

User Response: Validate that the I/O paths to all disks are valid and that all disks are either in the recovering or up availability.

6027-563 Error processing allocation map.

Explanation: There is not an I/O path to critical metadata or metadata has been corrupted.

User Response: Validate that the I/O paths to all disks are valid and that all disks are either in the recovering or up availability.

6027-564 Error processing i-node allocation map.

Explanation: There is not an I/O path to critical metadata or metadata has been corrupted.

User Response: Validate that the I/O paths to all disks are valid and that all disks are either in the recovering or up availability.

6027-569 Incompatible parameters: Unable to allocate space for file system metadata. Change one or more of the following as suggested and try again.

Explanation: Parameters entered require more space than is available.

User Response: Reissue command with changes suggested in message.

6027-570 Incompatible parameters: Unable to create file system. Change one or more of the following as suggested and try again.

Explanation: `mmcrfs` was unable to create the low level file structures for the file system.

User Response: Check other error messages and the error log. If none exist call your service representative.

6027-571 Logical sector size *number* must be the same as disk sector size.

Explanation: The `mmcrfs` command found that the sector size given by the `-l` option is not the same as the sector size given for disks in the `-d` option

User Response: Correct the command line entry.

6027-575 Unable to complete low level format for *name*

Explanation: The `mmcrfs` command was unable to create the low level file structures for the file system.

User Response: Check other error messages and the error log if none exist call your service representative.

6027-576 Unable to allocate space for file system metadata. Decrease the value for *name* and try again.

Explanation: The `mmcrfs` command did not find sufficient space to create the file system reserved files.

User Response: Refer the *GPFS Installation and Administration Guide* for valid parameter values.

6027-577 Warning: user or system file is not properly replicated.

Explanation: MMFS detected a file that is not replicated correctly due to a previous failure.

6027-578 Warning: *number* out of *number* user or system files are not properly replicated.

Explanation: MMFS detected a file that is not replicated correctly due to a previous failure.

6027-579 Some unreplicated file system metadata has been lost. File system usable only in restricted mode.

Explanation: A disk was deleted that contained vital file system metadata that was not replicated.

User Response: Mount the file system in restricted mode (`-o rs`) and copy any user data that may be left on the file system. Then delete the file system.

6027-580 Unable to access vital system metadata, too many disks are unavailable.

Explanation: Metadata is unavailable because the disk(s) on which the data reside are stopped, or an attempt was made to delete them.

User Response: Either start the stopped disks, try to delete the disk(s) again or recreate the file system.

6027-581 Unable to access vital system metadata, file system corrupted.

Explanation: Trying to access the file system but the metadata is unavailable because the disk was deleted.

User Response: Recreate the file system.

6027-582 Some data has been lost.

Explanation: An I/O error or some disks are in the stopped state.

User Response: Check the availability of all disks and the path to all disks and re-execute the command.

6027-583 Disk *name* cannot be added. Allocation map cannot accommodate it.

Explanation: Cannot add a disk that is so much larger than the disks initially used to create the file system that the allocation map cannot accommodate the new disk.

User Response: Break the disk into smaller pieces by creating logical volumes or recreate the file system.

6027-584 Incompatible parameters: Unable to allocate space for root directory. Change one or more of the following as suggested and try again

Explanation: An I/O error occurred while attempting to write to disk.

User Response: Check disk for errors.

6027-585 Incompatible parameters: Unable to allocate space for ACL data. Change one or more of the following as suggested and try again: *disk name/size* as shown by `mmcrfs`.

Explanation: Parameters entered require more space than is available.

User Response: Reissue command with changes suggested in message.

6027-586 Incompatible parameters: Unable to allocate space for ACL data. Change one or more of the following as suggested and try again.

Explanation: An I/O error occurred while attempting to write to disk.

User Response: Check disk for errors.

6027-587 Unable to initialize quota client because there is no quota server. Please check error log on SG manager. Eventually offline quota-check must be started before retrying the command.

Explanation: startQuotaClient failed.

User Response: If the quota file could not be read (check error log on SG mgr), **mmcheckquota** must be run.

6027-588 No more than *number* nodes can mount a stripe group.

Explanation: The limit of the number of nodes that can mount a stripe group is exceeded.

User Response: Reissue command with smaller number of nodes.

6027-590 Quota server initialization failed.

Explanation: The quota server could not start execution.

User Response: Run **mmcheckquota** then check the quota limits using **mmrepquota**.

6027-591 Error writing user.quota file.

Explanation: The user.quota file could not be written to. The quota manager continues operation if the next write is successful. If not, further allocations to the file system will fail.

User Response: If further allocations to the file system fail, check the error code in the MMFS_QUOTA error log and make sure that the disks containing the quota file are accessible. Run **mmcheckquota**.

6027-592 Error writing group.quota file.

Explanation: The group.quota file could not be written to. The quota manager continues operation if the next write is successful. If not, further allocations to the file system will fail.

User Response: If further allocations to the file system fail, check the error code in the MMFS_QUOTA error log and make sure that the disks containing the quota file are accessible. Run **mmcheckquota**.

6027-600 Pool size changed to maximum available/allowed: *number K* = *number M*.

Explanation: **mmchconfig** could not allocate the entire amount of additional memory requested for the pagepool.

User Response: Leave the state as is, or free up some memory and retry the command.

6027-601 Error changing pool size.

Explanation: **mmchconfig** failed to change the pool size to the requested value.

User Response: Follow the recommended actions in the other messages that occur with this one.

6027-604 Parameter incompatibility: either the values of maxblocksize and minpagepool are incompatible or there is a mounted file system with a bigger block size.

Explanation: **mmchconfig** received incompatible parameters.

User Response: Correct the command line and try again.

6027-605 Minimum page pool size changed to *number*.

Explanation: Minimum page pool size successfully changed.

6027-606 Maximum block size changed to *number* K.

Explanation: Maximum block size successfully changed.

6027-608 Stripe Group Takeover Failed. Error = *number*

Explanation: An attempt to takeover as stripe group manager failed. The stripe group is unmounted to allow another node to try.

User Response: Check the return code. This is usually due to network or disk connectivity problems. Run **mmlsdisk** to determine if the paths to the disk are unavailable, and run **mmchdisk** if necessary.

6027-609 File system *name* unmounted because it does not have a stripe group manager.

Explanation: The stripe group has to be unmounted because a stripe group manager cannot be assigned. An accompanying message tells which node was the last stripe group manager.

User Response: Examine error log on the last SG manager. Run **mmlsdisk** to determine if a number of disks are down. Examine other error logs for an indication of network, disk, or virtual shared disk problems. Repair the base problem, run **mmchdisk** if required.

6027-610 Cannot mount File system *name* because it does not have a stripe group manager.

Explanation: The stripe group has to be unmounted because a stripe group manager cannot be assigned. An accompanying message tells which node was the last stripe group manager.

User Response: Examine error log on the last SG manager. Run **mmlsdisk** to determine if a number of disks are down. Examine other error logs for an indication of network, disk, or virtual shared disk problems. Repair the base problem, run **mmchdisk** if required.

6027-614 Value *number* for option *name* is out of range. Valid ranges are *number* through *number*.

Explanation: Values out of range when parsing command line arguments.

User Response: Correct the command line and try again.

6027-619 Negative grace times are not allowed

Explanation: The **mmedquota** command received an invalid parameter for the **-t** option.

User Response: Reissue the **mmedquota** command with a non-negative value for grace time.

6027-620 Hard quota limit must not be less than soft limit.

Explanation: The **mmedquota** command received invalid parameters in the edited quota information. The hard quota limit must be greater than or equal to the soft quota limit.

User Response: Reissue the **mmedquota** command and enter valid values when editing the information.

6027-621 Negative quota limits are not allowed.

Explanation: The **mmedquota** command received invalid parameters in the edited quota information. The quota value must be a non-negative value.

User Response: Reissue the **mmedquota** command and enter valid values when editing the information.

6027-650 The mmfs daemon is shutting down abnormally.

Explanation: The GPFS daemon is shutting down as a result of an unrecoverable condition, typically a resource shortage.

User Response: Review error log entries, correct resource shortage condition and restart the GPFS daemon.

6027-660 Error displaying message from mmfsd

Explanation: GPFS command cannot properly display an output string sent from the **mmfsd** daemon. This string is the label on perror output.

User Response: Check proper installation of **mmfsd**.

6027-663 Lost connection to file system daemon.

Explanation: The connection between a GPFS command and the **mmfsd** daemon has broken. The daemon has probably crashed.

User Response: Insure that the **mmfsd** daemon is running. Check the error log.

6027-665 Failed to connect to file system daemon (error code *number*).

Explanation: Error occurred while trying to create a session with **mmfsd**.

User Response: Insure that the **mmfsd** daemon is running. Also, only root can execute most GPFS commands.

6027-666 Failed to determine file system manager.

Explanation: While running a GPFS command in a multi-node configuration, the local file system daemon is unable to determine which node is managing the file system affected by the command.

User Response: Check inter-node communication configuration; make sure enough GPFS nodes are up to make a quorum.

6027-667 Could not set up socket

Explanation: One of the calls to create or bind the socket used for sending parameters and messages between the command and the daemon failed.

User Response: Check additional error message printed.

6027-668 Could not send message to file system daemon

Explanation: Attempt to send a message to the file system failed.

User Response: Check for additional error messages.

6027-669 Could not connect to file system daemon.

Explanation: The tcp connection between the command and the daemon could not be established.

User Response: Check additional error message printed.

6027-670 Value for *option* is invalid. Valid values are *list*.

Explanation: Parsing a command option that can be followed by one of several keywords. The remainder of the line will list the valid keywords.

User Response: Correct the command line.

6027-671 Keyword missing or invalid.

Explanation: Parsing command line arguments

User Response: Correct the command line.

6027-672 Too few arguments specified.

Explanation: Parsing command line arguments.

User Response: Correct the command line.

6027-673 Too many arguments specified.

Explanation: Parsing command line arguments.

User Response: Correct the command line.

6027-674 Too many values specified for option *name*.

Explanation: Parsing command line arguments.

User Response: Correct the command line.

6027-675 Required value for *option* is missing.

Explanation: Parsing command line arguments.

User Response: Correct the command line.

6027-676 *option* specified more than once.

Explanation: Parsing command line arguments.

User Response: Correct the command line.

6027-677 *option* is invalid.

Explanation: Parsing command line arguments.

User Response: Correct the command line.

6027-678 Misplaced or invalid parameter *name*.

Explanation: Parsing command line arguments.

User Response: Correct the command line.

6027-679 Device *name* is invalid.

Explanation: Parsing command line arguments.

User Response: Correct the command line.

6027-681 Required option *name* was not specified.

Explanation: Parsing command line arguments.

User Response: Correct the command line.

6027-682 Device argument is missing.

Explanation: Parsing command line arguments.

User Response: Correct the command line.

6027-683 Disk *name* is invalid.

Explanation: Parsing command line arguments.

User Response: Correct the command line.

6027-684 *variable for option* is invalid.

Explanation: Parsing command line arguments.

User Response: Correct the command line.

6027-685 *variable for option number* is out of range. Valid values are *number* through *number*.

Explanation: Parsing command line arguments.

User Response: Correct the command line.

6027-686 *variable (number)* exceeds (*number*).

Explanation: Parsing command line arguments.

User Response: Correct the command line.

6027-687 Disk *name* is specified more than once.

Explanation: Parsing command line arguments.

User Response: Correct the command line.

6027-688 Failed to read file system stripe group descriptor.

Explanation: The disk block containing critical information about the file system stripe group could not be read from disk.

User Response: This is a serious error, which may leave the file system in an unusable state. Correct any I/O errors, then run **mmfsck** with the file system unmounted to make repairs.

6027-689 Failed to update file system stripe group descriptor.

Explanation: The disk block containing critical information about the file system stripe group could not be written to disk.

User Response: This is a serious error, which may leave the file system in an unusable state. Correct any I/O errors, then run **mmfsck** with the file system unmounted to make repairs.

6027-690 Failed to allocate I/O buffer.

Explanation: Sufficient memory (RAM) could not be obtained to perform an operation.

User Response: Either retry the operation when the **mmfsd** daemon is less heavily loaded, or increase the size of one or more of the memory pool parameters by running **mmchconfig**.

6027-691 Failed to send message to node *name*.

Explanation: A message to another file system node could not be sent.

User Response: Check additional error message; check inter-node communication configuration.

6027-692 Value for *option* is invalid. Valid values are yes, no.

Explanation: Parsing command line arguments, and an option that is required to be **yes** or **no** (suitably translated) is neither.

User Response: Correct the command line.

6027-693 Cannot open disk *name*.

Explanation: Error detected by the **start** option of the **mmchdisk** command. The disk given could not be accessed.

User Response: Check the disk hardware and path to the disk. Refer to "Unable to Access Disks" on page 18.

6027-695 File system is read-only.

Explanation: An operation is attempted that would require modifying the contents of a file system, and the file system is read-only.

User Response: Make the file system R/W before retrying the operation.

6027-696 Too many disks are unavailable.

Explanation: A file system operation failed because all replicas of a data or metadata block are currently unavailable.

User Response: Run the **mmlsdisk** command to check the availability of the disks in the file system stripe group; correct disk hardware problems, and then run the **mmchdisk** command with the **start** option to inform the file system that the disk or disks are available again.

6027-697 No log available.

Explanation: A file system operation failed because no space for logging metadata changes could be found.

User Response: Check additional error message. A likely reason for this error is that all disks with available log space are currently unavailable.

6027-698 Not enough memory to allocate internal data structure.

Explanation: A file system operation failed because no memory is available for allocating internal data structures.

User Response: Stop other processes that may have main memory pinned for their use.

6027-699 Inconsistency in file system metadata.

Explanation: File system metadata on disk has been corrupted.

User Response: This is an extremely serious error that may cause loss of data. Run **mmfsck** with the file system unmounted to make repairs. There will be a "POSSIBLE FILE CORRUPTION" entry in the system error log that should be forwarded to IBM support.

6027-700 Log recovery failed.

Explanation: An error is encountered while restoring file system metadata from the log.

User Response: Check additional error message. A likely reason for this error is that none of the replicas of the log could be accessed because too many disks are currently unavailable. If the problem persists, run **mmfsck** with the file system unmounted.

6027-701 Some file system data are inaccessible at this time.

Explanation: This message appears when the file system has encountered an error that is serious enough to make some or all data inaccessible.

User Response: This message indicates that an error had occurred that left the file system in an unusable state. Possible reasons include too many unavailable disks or insufficient memory for file system control structures. Check other error messages as well as the error log for additional information. Unmount the file system and correct any I/O errors. Then remount the file system and try the operation again. If the problem persists, run **mmfsck** with the file system unmounted to make repairs.

6027-702 Some file system data are inaccessible at this time. Check error log for additional information. After correcting the problem, the file system must be unmounted and then mounted to restore normal data access.

Explanation: This message appears when the file system has encountered an error that is serious enough to make some or all data inaccessible.

User Response: This message indicates that an error had occurred that left the file system in an unusable state. Possible reasons include too many unavailable disks or insufficient memory for file system control structures. Check other error messages as well as the error log for additional information. Unmount the file system and correct any I/O errors. Then remount the file system and try the operation again. If the problem persists, run **mmfsck** with the file system unmounted to make repairs.

6027-703 Some file system data are inaccessible at this time. Check error log for additional information.

Explanation: This message appears when the file system has encountered an error that is serious enough to make some or all data inaccessible.

User Response: This message indicates that an error had occurred that left the file system in an unusable state. Possible reasons include too many unavailable disks or insufficient memory for file system control structures. Check other error messages as well as the error log for additional information. Unmount the file system and correct any I/O errors. Then remount the file system and try the operation again. If the problem persists, run **mmfsck** with the file system unmounted to make repairs.

6027-704 Warning: Due to an earlier error normal access to this file system has been disabled. Check error log for additional information. After correcting the problem, the file system unmounted and then mounted again to restore normal data access.

Explanation: This message appears when the file system has encountered an error that is serious enough to make some or all data inaccessible.

User Response: This message indicates that an error had occurred that left the file system in an unusable state. Possible reasons include too many unavailable disks or insufficient memory for file system control structures. Check other error messages as well as the error log for additional information. Unmount the file system and correct any I/O errors. Then remount the file system and try the operation again. If the problem persists, run **mmfsck** with the file system unmounted to make repairs.

6027-705 Error code *number*.

Explanation: Provides additional information about an error.

6027-706 The device *name* has no corresponding entry in `/etc/filesystems`.

Explanation: Running a command requiring a device that has a file system associated with it.

User Response: Check the `/etc/filesystems` file for a valid stanza entry for that device.

6027-707 Unable to open file *name*.

Explanation: A given file cannot be opened.

User Response: Check to make sure the file exists and has the correct permissions.

6027-708 Keyword *name* is invalid. Valid values are *list*.

Explanation: Parsing a command that includes a selection of one of several keywords. The remainder of the line will list the valid keywords.

User Response: Correct the command line.

6027-709 Invalid response. Valid responses are yes or no.

Explanation: A question is asked which requires a yes or no answer and the answer entered is neither yes or no.

6027-711 Specified entity, such as a disk or stripe group, does not exist.

Explanation: A file system operation failed because the specified entity, such as a disk or stripe group, could not be found.

User Response: Specify existing disk, stripe group, etc.

6027-712 Error in communications between mmfsd daemon and client program.

Explanation: A message sent between the `mmfsd` daemon and the client program has an incorrect format or content.

User Response: Assure `mmfsd` daemon is running.

6027-714 Terminating because conflicting program name is running.

Explanation: A program detects that it must terminate because a conflicting program is running.

User Response: Retry the command once the conflicting program has ended.

6027-716 Some file system data or metadata has been lost.

Explanation: Unable to access some piece of file system data that has been lost due to the deletion of disks beyond the replication factor.

User Response: If the function did not complete try to mount the file system in restricted mode.

6027-717 Must execute mmfsck before mount.

Explanation: `mmfsck` has started on a file system, but crashed prior to completion. Message is displayed when attempting to mount this file system.

User Response: Run `mmfsck` to repair file system, then reissue `mount` command.

6027-718 The mmfsd daemon is not ready to handle commands yet.

Explanation: The `mmfsd` is not accepting messages because it is restarting or stopping.

6027-719 Device type not supported.

Explanation: A disk being added to a stripe group with the **mmadddisk** or **mmcrfs** command is not a character mode special file, or has characteristics not recognized by GPFS.

6027-720 Actual sector size does not match given sector size.

Explanation: A disk being added to a stripe group with the **mmadddisk** or **mmcrfs** command has a physical sector size that differs from that given in the disk description list.

6027-721 Invalid host *name* in *name*.

Explanation: An invalid host name or IP address was found in a configuration file.

User Response: Check the configuration file specified in the error message.

6027-722 Warning: Due to an earlier error normal access to this file system has been disabled. Check error log for additional information. The file system must be mounted again to resume data access.

Explanation: This message is generated when the file system has encountered an error that is serious enough to make some or all data inaccessible.

User Response: This message indicates that an error had occurred that left the file system in an unusable state. Possible reasons include too many unavailable disks or insufficient memory for file system control structures. Check other error messages as well as the error log for additional information. Correct any I/O errors. Then remount the file system and try the operation again. If the problem persists, run **mmfsck** with the file system unmounted to make repairs.

6027-723 Warning: Due to an earlier error normal access to this file system has been disabled. Check error log for additional information. After correcting the problem, the file system be mounted again to restore normal data access.

Explanation: This message is generated when the file system has encountered an error that is serious enough to make some or all data inaccessible.

User Response: This message indicates that an error occurred that left the file system in an unusable state. Possible reasons include too many unavailable disks or insufficient memory for file system control structures. Check other error messages as well as the error log for additional information. Correct any I/O errors. Then remount the file system and try the operation again. If the problem persists, run **mmfsck** with the file system unmounted to make repairs.

6027-724 Incompatible file system format.

Explanation: Attempting to access a file system that was formatted with an older version of the product that is no longer compatible with the version currently running.

User Response: Check release notes for instructions on how to migrate old file systems to the current version.

6027-725 The mmfsd daemon is not ready to handle commands yet. Waiting for quorum.

Explanation: The **mmfsd** is not accepting messages, because it is waiting for quorum.

6027-726 Quota initialization/start-up failed.

Explanation: Quota manager initialization was unsuccessful. The stripe group manager finished without quotas. Subsequent client mount requests will fail.

User Response: Check the error log and correct I/O errors. It may be necessary to run **mmcheckquota** with the file system unmounted.

6027-729 Warning: No changes for disk *name* were specified.

Explanation: The disk descriptor in the **mmchdisk** command does not specify that any changes are to be made to the disk.

User Response: Check disk descriptor to determine whether changes are needed.

6027-731 Error while performing *command* for *name* quota on file system.

Explanation: An error occurred when switching quotas of a certain type on or off.

User Response: Check the associated error code to determine the reason. If error returned for multiple file systems, only the error code for the last failure is returned.

6027-732 Error while performing *command* on file system.

Explanation: An error occurred when listing or reporting quotas.

6027-733 Edit quota: Bad format!

Explanation: When editing quota limits the format of one or more entries was changed.

User Response: Run **mmedquota** again. Change only the values for the limits and follow the instructions given.

6027-734 *name* : cannot change current allocation

Explanation: The user attempted to change the quota usage when editing quota entries with the **mmedquota** command.

6027-736 Warning: Due to an earlier error normal access to this file system has been disabled. Check error log for additional information. The file system must be unmounted and then restore normal data access.

Explanation: This message appears when the file system has encountered an error that is serious enough to make some or all data inaccessible.

User Response: This message indicates that an error had occurred that left the file system in an unusable state. Possible reasons include too many unavailable disks or insufficient memory for file system control structures. Check other error messages as well as the error log for additional information. Unmount the file system and correct any I/O errors. Then remount the file system and try the operation again. If the problem persists, run **mmfsck** with the file system unmounted to make repairs.

6027-739 Warning: Due to an earlier configuration change the file system is no longer properly balanced.

Explanation: The **mmfsdisk** command when the file system is not properly balanced.

User Response: run the **mmrestripefs -b** command at your convenience

6027-740 Warning: Due to an earlier configuration change the file system is no longer properly replicated.

Explanation: The `mmfsdisk` command found that the file system is not properly replicated.

User Response: Run the `mmrestripefs -r` command at your convenience

6027-741 Warning: Due to an earlier configuration change the file system may contain data that is at risk of being lost.

Explanation: The `mmfsdisk` command found that critical data resides on disks that are suspended and/or being deleted.

User Response: Run the `mmrestripefs -m` command as soon as possible.

6027-743 Disk state was updated successfully, but a subsequent metadata update could not be completed at this time.

Explanation: The `mmchdisk` command encountered an error after the disk status or availability change was already recorded in the file system configuration.

User Response: Most likely reason for this failure is that too many disks have become unavailable or still are unavailable after the disk state change. Issue a `mmchdisk start` command when more disks are available.

6027-745 Warning: Failed to create spare log file(s). A subsequent disk failure may make the file system unavailable on some nodes.

Explanation: A disk state change command was not able to create as many spare log files as desired.

User Response: Free some space in the file system and rerun the command, or restripe the file system.

6027-746 Editing quota limits for this user or group not permitted.

Explanation: The `root` user or `system` group was specified for quota limit editing in the `mmedquota` command. These names are not allowed.

User Response: Specify a valid name in the `mmedquota` command.

6027-749 Pool size changed to *number K* = *number M*.

Explanation: Pool size successfully changed.

6027-755 Limit of *number* not identical data block replicas exceeded. Re-execute command to find additional not identical replicas.

Explanation: `mmfsck` has reached its limit of the number of data blocks whose replicas do not match.

User Response: After letting `mmfsck` repair/delete files with detected non-identical data block replicas, re-execute `mmfsck` to find additional non-identical data block replicas.

6027-756 Error reading i-node file

Explanation: `mmfsck` command cannot read an i-node from disk.

User Response: Make sure all disks are available. Otherwise, reformat file system.

6027-757 Error writing i-node file.

Explanation: `mmfsck` command cannot write an i-node to disk.

User Response: Make sure disk is writable. Otherwise, reformat file system.

6027-758 Error reading indirect block.

Explanation: An indirect block cannot be read from a file system.

User Response: Make sure all disks are available. Otherwise, reformat file system.

6027-759 Error writing indirect block.

Explanation: An indirect block cannot be written to a file system.

User Response: Make sure disk is writable. Otherwise, reformat file system.

6027-760 Error reading block allocation map.

Explanation: **mmfsck** command cannot read a block allocation map segment from disk.

User Response: Make sure all disks are available. Otherwise, reformat file system.

6027-761 Error writing i-node allocation map.

Explanation: **mmfsck** command cannot write an i-node allocation region to disk.

User Response: Make sure disk is writable. Otherwise, reformat file system.

6027-762 Error reading i-node allocation map

Explanation: **mmfsck** command cannot read an i-node allocation region from disk.

User Response: Make sure all disks are available. Otherwise, reformat file system.

6027-763 Limit of *number* multiply referenced fragments exceeded. Re-execute command to find additional multiply referenced fragments.

Explanation: **mmfsck** command internal limit on number of multiply referenced fragments is reached.

User Response: Let **mmfsck** command execute to completion and reexecute command to detect additional multiply referenced fragments.

6027-764 Limit of *number* damaged i-nodes exceeded. Re-execute command to find additional damaged i-nodes.

Explanation: **mmfsck** command internal limit on number of damaged i-nodes is reached.

User Response: Let **mmfsck** command execute to completion and reexecute command to detect additional damaged i-nodes.

6027-765 Cannot check. *variable* is mounted on *number* node(s).

Explanation: Error occurred when file system was mounted.

User Response: Unmount file system on all nodes, then re-execute command.

6027-766 mmfsck already in progress for *name*.

Explanation: **mmfsck** command is already running

User Response: Wait for currently executing command to complete and re-execute command.

6027-767 Cannot allocate storage

Explanation: Sufficient memory cannot be allocated to execute **mmfsck** command.

User Response: Make more memory available.

6027-768 Cannot allocate storage for data comparison

Explanation: Sufficient memory cannot be allocated to perform data comparison. **mmfsck** will still continue to execute, but without reading, comparing file data.

User Response: Make more memory available.

6027-769 Error accessing i-node file

Explanation: **mmfsck** command cannot open the i-node file.

User Response: Check availability of all disks in file system.

6027-770 Error accessing block allocation map file.

Explanation: **mmfsck** command cannot open the block allocation map file.

User Response: Check availability of all disks in file system.

6027-771 Error accessing i-node allocation map file.

Explanation: **mmfsck** command cannot open the i-node allocation map file.

User Response: Check availability of all disks in file system.

6027-772 Invalid record size.

Explanation: **mmfsck** checked the record size for one of the files it is about to open and record size is not correct.

User Response: Check availability of all disks in file system.

6027-773 Error reading root directory file.

Explanation: **mmfsck** command cannot read the root directory file from disk.

User Response: Check availability of all disks in file system.

6027-774 Unable to create a temporary storage file in *name* on node *name*.

Explanation: **mmfsck** is unable to create a file in the specified directory. The error could be permission, no space, or no memory.

6027-775 Unable to reference the root directory.

Explanation: **mmfsck** is unable to access the file information for the root directory.

6027-776 Unable to reference i-node *number*.

Explanation: **mmfsck** is unable to get the i-node from disk, usually an error from `getinode()`.

User Response: Check availability of all disks in file system.

6027-777 Unable to lock or unlock i-nodes.

Explanation: **mmfsck** is unable to lock a range of i-nodes. Error from `LockRange()`.

User Response: Report problem to IBM support.

6027-778 Unable to broadcast allocation messages.

Explanation: **mmfsck** is unable to notify other nodes of changes to the allocation maps because the communication path between nodes has failed.

User Response: Report problem to IBM support.

6027-779 Unable to deallocate data blocks.

Explanation: `mmfsck` is unable deallocate data blocks it found not to belong to any file but were marked allocated in the block allocation map.

6027-780 Error reading the ACL data file.

Explanation: `mmfsck` command cannot read the ACL data file from disk.

User Response: Verify all disks are available. Otherwise, reformat the file system.

6027-781 Unable to reset the ACL data file.

Explanation: `mmfsck` command is unable to repair the ACL data file.

User Response: If the file system is not mountable, reformat the file system.

6027-799 Error writing block allocation map.

Explanation: `mmfsck` command cannot write a block allocation map segment to disk.

User Response: Make sure disk is writable. Otherwise, reformat file system.

6027-800 Could not run command *name*.

Explanation: The GPFS daemon failed to open the specified command.

User Response: Verify correct installation.

6027-801 *filename* fgets failed

Explanation: The GPFS daemon failed to read the specified file.

User Response: Verify correct installation and configuration.

6027-802 Invalid request type for MmfsGrApprCb callback function: Protocol type = *type* notification type = *type*.

Explanation: Group Services returned invalid notification.

User Response: Follow problem determination procedures in *PSSP Diagnosis Guide*.

6027-804 Lost quorum. Shutting down the daemon.

Explanation: The mmfs group lost the quorum.

6027-805 The node number *number* is not defined in *filename*

Explanation: When queried, a given node number did not exist in the `cluster.nodes` file

User Response: Verify `cluster.nodes` file.

6027-806 Failed to open `/dev/VSD0`. Err = *number*

Explanation: Error detected when trying to open `/dev/VSD0`

User Response: Follow IBM Recoverable Virtual Shared Disk problem determination procedures.

6027-807 `vsd fence ioctl` failed with err *number* and return code

Explanation: Error detected when trying to fence or unfence virtual shared disks.

User Response: Follow IBM Recoverable Virtual Shared Disk problem determination procedures.

6027-808 The node addr *number* is not defined in *filename*

Explanation: Error detected when querying a given address. Does not exist in the **cluster.nodes** file.

User Response: Verify **cluster.nodes** file.

6027-809 Failed to subscribe to group; rc = *number*.

Explanation: Error detected when trying to subscribe to ha.vsd or cssMembership group

User Response: None

6027-810 Could not connect to Phoenix, rc = *number*.

Explanation: Error detected when trying to connect to Group Services.

User Response: Verify **hagsd** daemon is active.

6027-811 Group Services has failed. Shutting down the daemon. Rc = *number*.

Explanation: The **mmfsd** daemon could not connect to HA.

User Response: Follow Group Services problem determination procedures.

6027-812 Node *number* is not part of the group.

Explanation: Error detected when the local failed to join the cluster.

User Response: Verify **cluster.nodes** file.

6027-813 Recovering nodes *list*.

Explanation: Written to the error log when recovery for a node has begun.

User Response: None if followed by recovered nodes entry specifying this node. If not followed by such a message, determine why recovery did not complete.

6027-814 Recovered nodes *list* .

Explanation: Written to the error log when recovery for a node has completed..

6027-815 Failed to join Phoenix groups

Explanation: Error detected when the **mmfsd** daemon failed to join Group Services groups.

User Response: Verify Group Services is installed and make sure **HA_SYSPAR_NAME** is set properly.

6027-816 MMFSGR group join rejected. Verify group

Explanation: Error detected when Group Services rejected a join from mmfs on this node

User Response: Follow problem determination procedures in *PSSP Diagnosis Guide*.

6027-817 Join failed. Name *node code number*.

Explanation: Error detected when the **mmfsd** daemon failed to join Group Services groups.

User Response: Follow Group Services problem determination procedures.

6027-818 Subscription failed: Group *number* rc *number*

User Response: Determine why prerequisite subsystem is not available. If unavailable group is ha.vsd, follow problem determination procedures in *PSSP Diagnosis Guide*.

6027-819 State change failed. new state = *state*

User Response: React to other associated errors that caused the state change to fail.

6027-820 My adapter *adapter* has failed. Shutting down the daemon

Explanation: Group Services reported that required adapters failed.

User Response: Follow Group Services and adapter problem determination procedures.

6027-821 mmfsd ready for sessions.

Explanation: The **mmfsd** daemon has joined an active Group Services group with a quorum. This means that commands can be issued and mounts can be processed.

6027-822 I am not in subscriber list of group; Verify adapter

Explanation: Error detected when the **mmfsd** daemon subscribed to the adapter group, it did not see its node number in the subscriber list. This means that the node is not provider list.

User Response: If this is a switch adapter, determine whether the adapter is up or down

6027-823 The adapter group dissolved; Shutting down the daemon.

Explanation: The **mmfsd** daemon receives adapter group dissolved notification from group services.

User Response: Follow Group Services problem determination procedures.

6027-824 rvsd daemon is not active. SRC rc=*number*.

Explanation: The **mmfsd** daemon tries to fence or unfence virtual shared disks.

User Response: Follow IBM Recoverable Virtual Shared Disk problem determination procedures.

6027-829 Join rejected. Incompatible daemon version (*number:number*). Group vers(*number:number*)

Explanation: The **mmfsd** daemon is not compatible the the group version.

User Response: Verify daemon version.

6027-830 Could not get minor number for */dev/r name* in *name*, err = *number*.

Explanation: The **mmfsd** daemon failed to obtain a virtual shared disk minor number.

User Response: Follow IBM Virtual Shared Disk problem determination procedures.

6027-943 -M and -R are only valid for zero length files.

Explanation: The **mmchatr** command received invalid command line arguments.

User Response: Correct command line and try again.

6027-944 -m value exceeds number of failure groups for metadata.

Explanation: The **mmchattr** command received invalid command line arguments.

User Response: Correct command line and try again.

6027-945 -r value exceeds number of failure groups for data.

Explanation: The **mmchattr** command received invalid command line arguments.

User Response: Correct command line and try again.

6027-961 Cannot execute command.

Explanation: The **mmeditACL** command cannot invoke the **mmgetACL** or **mmputACL** command.

User Response: Contact your system administrator

6027-972 Could not run command command.

Explanation: The daemon failed to **popen** the specified command.

User Response: Verify the file.

6027-973 Cannot allocate number byte buffer for ACL.

Explanation: There is not enough available memory to process the request.

User Response: Contact your system administrator

6027-974 Failure reading ACL (rc=number).

Explanation: An unexpected error was encountered by **mmgetACL** or **mmeditACL**.

User Response: Examine the return code, contact IBM support if necessary.

6027-976 Failure writing ACL (rc=number).

Explanation: An unexpected error encountered by **mmputACL** or **mmeditACL**.

User Response: Examine the return code, contact IBM support if necessary.

6027-977 Authorization failure

Explanation: Attempting to create or modify the ACL for a file that you do not own.

User Response: Only the owner of a file or the root user can create or change the access control list for a file.

6027-978 Invalid, duplicate, or missing access control entry detected.

Explanation: An access control entry in the ACL that you are creating has incorrect syntax, or one of the required access control entries is missing, or the ACL contains duplicate access control entries.

User Response: Correct the problem and reissue the command

6027-985 Failure deleting the ACL (rc=number).

Explanation: An unexpected error was encountered by **mmdelACL**.

User Response: Examine the return code, contact IBM support if necessary.

6027-1000 Usage: `mmawk [Fdelim] field_number <infile >outfile or: mmawk [Fdelim] NF <infile >outfile`

Explanation: An invalid option was specified

User Response: None, informational message only. If this message appears in another command, it is an internal processing error. Contact IBM support.

6027-1001 Invalid digit in *field_number*.

Explanation: Something other than digits was specified for *field number*.

User Response: Use valid integer in last **mmawk** argument. If this message appears in another command, it is an internal processing error. Contact IBM support.

6027-1002 Invalid positive integer in *field_number*.

Explanation: *field number* is 0 or overflows.

User Response: Specify an integer from 1 to 2G. If this message appears in another command, it is an internal processing error. Contact IBM support.

6027-1003 Usage: `mmfixcfg [attribute newvval] [nodelist] <infile >outfile`

Explanation: Incorrect arguments were specified. Note that **mmfixcfg** does not return in this case. It passes stdin unchanged to stdout.

User Response: Use correct syntax. If this message appears in another command, it is an internal processing error. Contact IBM support.

6027-1004 Invalid [nodelist] format in file: *list*

Explanation: A [nodelist] line in the input stream is not a comma-separated list of nodes.

User Response: Configuration data contained in the SDR is incorrect. Correct any user modifications that have been made. If no user modifications have been made to the SDR, contact IBM support.

6027-1005 Common is not sole item on [] line *number*.

Explanation: A [nodelist] line in the input stream contains common plus any other names.

User Response: Configuration data contained in the SDR is incorrect. Correct any user modifications that have been made. If no user modifications have been made to the SDR, contact IBM support.

6027-1006 Invalid custom [] line *number*.

Explanation: A [nodelist] line in the input stream is not of the format: [nodelist]. This covers syntax errors not covered by messages 6027-1004 and 6027-1005.

User Response: Configuration data contained in the SDR is incorrect. Correct any user modifications that have been made. If no user modifications have been made to the SDR, contact IBM support.

6027-1007 attribute found in common multiple times: *attribute*.

Explanation: The attribute specified on the command line is in the main input stream multiple times. This is occasionally legal, such as with the trace attribute. These attributes, however, are not meant to be repaired by **mmfixcfg**.

User Response: Fix the configuration file (**mmfs.cfg** or **mmfscfg1** in the SDR). All attributes modified by GPFS configuration commands may appear only once in common sections of the configuration file.

6027-1008 Attribute found in custom multiple times: *attribute.*

Explanation: The attribute specified on the command line is in a custom section multiple times. This is questionably legal. These attributes are not meant to be repaired by **mmfixcfg**.

User Response: Fix the configuration file (**mmfs.cfg** or **mmfscfg1** in the SDR). All attributes modified by GPFS configuration commands may appear only once in common sections of the configuration file.

6027-1009 Usage: **mmsdref** [**a** *sdrfile filename*] [**c** *sdrfile clustno*] [**d** *sdrfile clustno device*] [**l** *sdrfile clustno*] [**g** *sdrfile clustno*] [**o** *stanza clustno*] [**s** *sdrfile clustno*] [**u** *sdrfile clustno*] [**x** *sdrfile clustno device*]

Explanation: Incorrect arguments were given.

User Response: Use correct syntax. If the error appears in another command, it is an internal processing error. Contact IBM support.

6027-1010 Cluster *number* not found in mmsdrfs.

Explanation: The configuration id specified is not found in the SDR file **mmsdrfs**.

User Response: Specify a configuration id that exists in the file. If the error appears in another command, it is an internal processing error. Contact IBM support.

6027-1011 Invalid configuration id in header of mmsdrfs: *number.*

Explanation: The number in a cluster line of SDR file is not an integer.

User Response: Configuration data contained in the SDR is incorrect. Correct any user modifications that have been made. If no user modifications have been made to the SDR, contact IBM support.

6027-1012 Invalid stanza header: *value.*

Explanation: The first line after the header of sdrfile is not the beginning of a stanza, which should look like 'name:'.

User Response: Configuration data contained in the SDR is incorrect. Correct any user modifications that have been made. If no user modifications have been made to the SDR, contact IBM support.

6027-1013 Unexpected EOF in input.

Explanation: The last line of the SDR file is not terminated.

User Response: Configuration data contained in the SDR is incorrect. Correct any user modifications that have been made. If no user modifications have been made to the SDR, contact IBM support.

6027-1014 Invalid configuration id in stanza: *number.*

Explanation: The number field in a cluster line of a stanza in SDR file is not a valid integer.

User Response: Configuration data contained in the SDR is incorrect. Correct any user modifications that have been made. If no user modifications have been made to the SDR, contact IBM support.

6027-1015 Warning: dev entry is not in /dev: number.

Explanation: The device line of a stanza specifies a path not in */dev/...*

User Response: Configuration data contained in the SDR is incorrect. Correct any user modifications that have been made. If no user modifications have been made to the SDR, contact IBM support.

6027-1016 Could not open input file value.

Explanation: One of the files specified could not be read.

User Response: Configuration data contained in the SDR is incorrect. Correct any user modifications that have been made. If no user modifications have been made to the SDR, contact IBM support.

6027-1017 Could not open /etc/filesystems.

Explanation: */etc/filesystems* cannot be read.

User Response: Configuration data contained in the SDR is incorrect. Correct any user modifications that have been made. If no user modifications have been made to the SDR, contact IBM support.

6027-1018 No stanzas in input file.

Explanation: Use the add option with a file containing stanzas.

6027-1019 Invalid configuration id argument: value.

Explanation: The configuration id command argument is longer than eight characters.

User Response: Specify a valid configuration id. If the error appears in another command, it is an internal processing error. Contact IBM support.

**6027-1020 New stanza matches an old stanza in device or mountpoint but not both:
New stanza: value Old stanza: value**

Explanation: The two specified stanza names match incompletely.

User Response: The name of a new stanza should not match the name of an old one unless it is replacing the old stanza. For replacement, the new stanza must match the old in both device and mountpoint path. **mmcrfs** should detect this error before it reaches **mmsdref**; contact IBM support.

6027-1021 Usage: **mmcrfs** *MountPoint Device {DiskDescList | -F DescFile } [-A yes|no] [-B BlockSize] [-i InodeSize] [-I IndirectSize] [-k yes|no] [-m DefaultMetadataReplicas] [-M MaxMetadataReplicas] [-N NumberOfNodes] [-n NumNodes] [-Q yes|no] [-r DefaultDataReplicas] [-R MaxDataReplicas] [-s StripeMethod] [-v yes|no]*

Explanation: Incorrect arguments are given to **mmcrfs**. Note that this is not the usage of **mmcrfsc**, which add some arguments to the front. **mmcrfsc** is never meant to be executed directly. If it is, the usage is at the top of the script.

User Response: Use correct syntax to **mmcrfs**.

6027-1022 Missing mandatory arguments on command line.

Explanation: Some but not enough arguments are specified to **mmcrfsc**.

User Response: Specify all arguments as per usage statement following.

6027-1023 File system size args must be an integer: *value*.

Explanation: The first two arguments specified to **mmcrfsc** are not integers.

User Response: File system size is an internal argument. **mmcrfs** should never call **mmcrfsc** without a valid fssize argument. Call IBM support.

6027-1024 Cannot open *file*

Explanation: An input file specified to **mmcrfsc** cannot be read or write.

User Response: This may be an internal file, or possibly the descriptor file. **mmcrfs** should not pass this error through to **mmcrfsc**.

6027-1025 No disks specified.

Explanation: There are no disks in the descriptor list or file.

User Response: Specify at least one disk to **mmcrfs**.

6027-1026 Missing file argument for -F option.

Explanation: -F is specified and is the last item on the command line.

User Response: Follow -F with the name of the descriptor file.

6027-1027 Mountpoint must not be a relative path name.

Explanation: The mountpoint does not begin with '/'.

User Response: Specify the full path for mountpoint.

6027-1028 Invalid value for flag.

Explanation: An incorrect argument is specified with a flag requiring one of a limited number of legal options. (-s or any of the **yes|no** flags)

User Response: Use one of the legal values for the indicated flag.

6027-1029 Invalid characters in integer field for *flag*.

Explanation: An invalid integer is specified with the indicated flag.

User Response: Use a valid integer for the indicated flag.

6027-1030 Value below minimum for *flag*.

Explanation: The valued specified with a flag is below the minimum.

User Response: Use an integer in the legal range for the indicated flag.

6027-1031 Value above maximum for *flag*.

Explanation: The valued specified with a flag is above the maximum.

User Response: Use an integer in the legal range for the indicated flag.

6027-1032 Invalid *flag*.

Explanation: An unknown flag is specified.

User Response: Use only flags shown in the syntax.

6027-1033 *flag* specified twice.

Explanation: Any flag is specified more than once on the command line.

User Response: Use flags only once.

6027-1034 Missing argument after *flag*.

Explanation: A flag is not followed by a argument.

User Response: All flags need an argument. Specify one.

6027-1035 *flag* is mandatory.

Explanation: A mandatory input flag is not specified.

User Response: Specify all mandatory flags.

6027-1036 Flag expected at *string*.

Explanation: Something other than an expected flag was encountered on the latter portion of the command line.

User Response: Follow the syntax shown. No flag may have multiple values following.

6027-1037 No room for all i-nodes. Increase maximum file size.

Explanation: Too many i-nodes are specified for the configuration determined by the command line flags.

User Response: Increase the maximum file size by specifying a larger value for **-B**, **-I**, or **-i**, OR by specifying a smaller value for **-M** or **-R**.

6027-1038 IndirectSize must be equal to or less than BlockSize and must be multiple of BlockSize / 32.

Explanation: IndirectSize specified is not a multiple of BlockSize / 32 or if IndirectSize specified is larger than BlockSize.

User Response: Use valid values for IndirectSize and BlockSize.

6027-1039 InodeSize must be multiple of LogicalSectorSize (512).

Explanation: InodeSize is not a multiple of 512

User Response: Use valid value for InodeSize.

6027-1040 InodeSize must be equal to or less than BlockSize.

Explanation: InodeSize is not equal to or less than BlockSize.

User Response: Use valid value for InodeSize.

6027-1041 TrailerSize must be 0 or greater than 63.

Explanation: TrailerSize is not a valid value.

User Response: Specify a valid value for TrailerSize. This is a single node error only. Multinode file systems do not use this flag.

6027-1042 DefaultMetadataReplicas must be equal to or less than MaxMetadataReplicas.

Explanation: DefaultMetadataReplicas is greater than MaxMetadataReplicas.

User Response: Specify a valid value for DefaultMetadataReplicas.

6027-1043 DefaultDataReplicas must be equal to or less than MaxDataReplicas.

Explanation: DefaultDataReplicas is greater than MaxDataReplicas.

User Response: Specify a valid value for DefaultDataReplicas.

6027-1044 Cannot specify backup if server is not specified.

Explanation: A descriptor specifies a backup node and not a server node.

User Response: To create an IBM Virtual Shared Disk, specify the server node. If the IBM Virtual Shared Disk exists, leave off the backup node.

6027-1045 Missing mmfs disk name in descriptor.

Explanation: A descriptor does not specify disk_name.

User Response: Fix the descriptor. disk_name is always required.

6027-1046 Invalid diskusage in descriptor: value.

Explanation: The diskusage field in a descriptor has a invalid value.

User Response: Fix the descriptor. disk_usage must be one of: **dataOnly**, **metadataOnly**, **dataAndMetadata**, or left unspecified.

6027-1047 Invalid integer specified for failure group.

Explanation: An invalid integer is specified for failure group in a descriptor.

User Response: Fix the descriptor. Use a valid integer for failure group.

6027-1048 Failure group must be equal to or greater than -1: value.

Explanation: The failure group in a descriptor is less than -1.

User Response: Fix the descriptor. A user specified failure group may range from -1 to 4000.

6027-1049 Disk descriptor too long.

Explanation: A disk descriptor is over 4K bytes.

User Response: Disk descriptors usually should be about 40 chars max. Use shorter names in the descriptor.

6027-1050 Invalid character found in disk descriptor. No special character, comments, or whitespace is allowed.

Explanation: A disk descriptor file contains special characters.

User Response: Remove special characters from descriptor file.

6027-1051 malloc failed.

Explanation: Space cannot be allocated for a descriptor.

User Response: Try again, or run with fewer/shorter descriptors.

6027-1053 Unable to create the *disk* on *server*.

Explanation: `mmkvds` failed to create an IBM Virtual Shared Disk.

User Response: The command will continue. Restart the command when it finishes, using the `restart.desc` file created as the descriptor file. This message should be preceded by a fairly verbose error message from `createvds`. Fix errors in the disk descriptor, if any, and run the command again using the `restart.desc` file.

6027-1054 Unable to allocate space for file system metadata. Increase `-i`, `-l`, `-B`, or decrease `-M` or `-R`, and try again.

Explanation: Data structures are too small to contain the space represented by all the disks specified.

User Response: Increase or decrease respectively the flags shown in the error message.

6027-1060 BlockSize must be 16K, 64K, or 256K.

Explanation: The specified value for `BlockSize` is not valid.

User Response: Specify a valid `BlockSize` value

6027-1101 Usage: `mmadddisk device {disk_desc_list | -F desc_file} [-r] [-a] [-v yes|no]`

Explanation: The `mmadddisk` command received invalid command line arguments.

User Response: Correct command line and try again.

6027-1102 Usage: `mmaddnode nodelist`

Explanation: The `mmaddnode` command received invalid command line arguments.

User Response: Correct command line and try again.

6027-1103 Usage: `mmchconfig attr=value [nodelist] [-i]`

Explanation: The `mmchconfig` command received invalid command line arguments.

User Response: Correct command line and try again.

6027-1104 Usage: `mmchfs device [-A yes|no] [-Q yes|no] [-m DefaultMetadataReplicas] [-r DefaultDataReplicas] [-s StripeMethod] [-T mountpoint] [-F MaxNumberOfNodes] [-V]`

Explanation: The `mmchfs` command received invalid command line arguments.

User Response: Correct command line and try again.

6027-1105 Usage: `mmconfig {-a | -n nodefile} [-A] [-c cfgfile] [-C configurationId] [-D dataStructureDump] [-i] [-m mallocsize] [-M maxFilesToCache] [-p pagepool] [-r priority]`

Explanation: The `mmconfig` command received invalid command line arguments.

User Response: Correct command line and try again.

6027-1106 Usage: `mmdeldisk device disk_list [-a] [-n] [-p] [-r]`

Explanation: The `mmdeldisk` command received invalid command line arguments.

User Response: Correct command line and try again.

6027-1107 Usage mmdelfs device [-n]

Explanation: The **mmdelfs** command received invalid command line arguments.

User Response: Correct command line and try again.

6027-1108 Usage: mmdelnode [-c] {*nodelist* | -a}

Explanation: The **mmdelnode** command received invalid command line arguments.

User Response: Correct command line and try again.

6027-1109 Usage: mmrpldisk device diskname disk_desc [-n] [-v yes|no]

Explanation: The **mmrpldisk** command received invalid command line arguments.

User Response: Correct command line and try again.

6027-1110 Usage *command*

Explanation: The **mmdevlist** or **mmrecover** command received command line arguments.

User Response: These commands must not be called with arguments. **mmdevlist** is a routine called by **smit**. **mmrecover** is used on the control workstation by **mmremote**. The user need never use either command. If this message appears in another command, it is an internal processing error. Contact IBM support.

6027-1111 Usage: mmmkvsd diskname rtnfile server numlist [backup]

Explanation: The **mmkvsds** command received invalid command line arguments.

User Response: **mmkvsds** is a routine called internally by other GPFS administration commands (mm...). The user need never use **mmkvsds**. If another command returns this message, check that GPFS has the same level of code installed on the calling node, and on the control workstation. Otherwise, it is an internal processing error and you should contact IBM support.

6027-1112 Usage: mmremote {[boot | single | cfg | active | fs | gone | nofs | unlock | vfs | freepv | clmgr] | [lock | pid] *pid*] [sgmgr | mount] *device* | [undo] *type,device,more* | vsd *vsdname* | hdsiz *vg_name*}

Explanation: The **mmremote** command received invalid command line arguments.

User Response: **mmremote** is a routine called internally by other GPFS administration commands (mm...). The user need never use **mmremote**, but may if he wishes. (Check command line arguments.) If another command returns this message, check that GPFS has the same level of code installed on all nodes. Otherwise, it is an internal processing error. Contact IBM support.

6027-1113 Invalid *option*.

Explanation: A command received an invalid command option.

User Response: Correct command line and try again.

6027-1114 Get adapter failed for IPA: *address, node*

Explanation: A node being processed (from the command line or from another source) is recognized by the host command and yet whose IP address does not have a valid adapter in the SDR adapter table.

User Response: If the IP address is invalid, something is wrong with the indicated node. If the IP address is valid, and should have an adapter, check the integrity of the SDR. If the node name is 'localhost', the IP address will be valid, but will have no adapter. Localhost should not be entered from the command line.

6027-1115 Cannot find any GPFS configuration in this partition.

Explanation: The **mmsdrfs** file in the SDR is not found.

User Response: If a GPFS configuration should exist in this partition, the SDR may have been corrupted. Determine and correct the problem that caused the loss of the **mmsdrfs** file. New GPFS configurations can be created with the **mmconfig** command.

6027-1116 Node represented in multiple GPFS configurations.

Explanation: The **mmsdrfs** file in the SDR lists a node in more than one GPFS configuration.

User Response: Configuration data contained in the SDR is incorrect. Correct any user modifications that have been made. If no user modifications have been made to the SDR, contact IBM support.

6027-1117 node represented multiple times in the same GPFS configuration.

Explanation: The **mmsdrfs** file in the SDR lists a node more than once in the same GPFS configuration.

User Response: Configuration data contained in the SDR is incorrect. Correct any user modifications that have been made. If no user modifications have been made to the SDR, contact IBM support.

6027-1118 This node does not belong to any GPFS configuration.

Explanation: A GPFS administration command (mm...) is run from a node which is not part of a GPFS configuration.

User Response: Run the command from a node in the GPFS configuration on which the command should act. This is the only way the command can tell which partition is being modified.

6027-1119 sysctl failed. Fix /etc/sysctl.mmcmd.acl, restart sysctl on node, or run kinit.

Explanation: sysctl is not set up on a node being used or you do not have a kerberos ticket for your local node.

User Response: Fix sysctl on the target node. sysctl must have permission. This is controlled by the **/etc/sysctl.mmcmd.acl** file. sysctl must know about GPFS administration commands (mm...). The install step should add a line about mmfs in **/etc/sysctl.conf**. If it is there, the sysctl daemon must be restarted. A kerberos ticket is also needed locally, obtained by **kinit**

6027-1120 Interrupt received: No changes made.

Explanation: A GPFS administration command (mm...) receives an interrupt before committing any changes.

User Response: No action needed.

6027-1121 MMFS is not active.

Explanation: A GPFS administration command (mm...) is run from a node on which the **mmfsd** daemon is not active.

User Response: Start **mmfsd** on this node or use another node.

6027-1122 There are not enough active nodes for quorum.

Explanation: A GPFS administration command (mm...) is run from a GPFS configuration that does not have enough active nodes to achieve quorum.

User Response: Start mmfs on enough nodes to achieve quorum.

6027-1123 Disk name must be specified in a descriptor.

Explanation: The disk name positional parameter (field 1) in a disk descriptor is empty. The bad disk descriptor is displayed following this message.

User Response: Correct the input and rerun the command.

6027-1124 Disk usage must be dataOnly, metadataOnly, or dataAndMetadata.

Explanation: The disk usage positional parameter (field 4) in a disk descriptor has an invalid value. The bad disk descriptor is displayed following this message.

User Response: Correct the input and rerun the command.

6027-1125 Backup server name must be omitted if server name is omitted.

Explanation: The server name positional parameter (field 2) is omitted, but the backup server name parameter (field 3) has a value. The bad disk descriptor is displayed following this message.

User Response: Specify the server node, or remove the backup server name.

6027-1126 The specified vsd *name* does not exist or is not available.

Explanation: The internally issued **lsvsd** command fails for a virtual shared disk that should exist.

User Response: Specify an existing virtual shared disk. If the virtual shared disk does already exist, make sure the disk is available. Also check that the virtual shared disk is configured and started.

6027-1127 Unable to allocate space for the file system metadata. There would be too many disks in the file system.

Explanation: The file system data structures are not large enough to handle all of the space represented by the new disks.

User Response: Use fewer disks. If you really need that much space, you must create a new file system with larger data structures. This can be achieved by specifying, at file system create time, larger values for **-B**, **-i**, or **-I**, or smaller values for **-M** or **-R**.

6027-1128 Invalid *value* server node *value* in disk descriptor.

Explanation: A disk descriptor specifies a server node that is invalid.

User Response: Correct the input and rerun the command.

6027-1129 Cannot validate physical volume *name*.

Explanation: The size of the passed physical volume cannot be determined or if there is already a volume group on the specified physical volume.

User Response: Use a valid and unused physical volume.

6027-1130 User specified failure group must be in the range from -1 to 4000.

Explanation: A disk descriptor specifies a failure group out of range. The bad disk descriptor is displayed following this message.

User Response: Correct the input and rerun the command.

6027-1131 Cannot determine the primary server for vsd *name*.

Explanation: The internally issued **vsdata1st** command does not show the primary server of the passed virtual shared disk.

User Response: The virtual shared disk may not be valid. Try **vsdata1st -v** and **vsdata1st -g**, looking for the virtual shared disk in question and its volume group. Correct the input and rerun the command.

6027-1132 Interrupt received: changes not propagated.

Explanation: An interrupt was received after changes were committed but before the changes could be propagated to all the nodes.

User Response: All changes will eventually propagate as nodes recycle or other GPFS administration commands (mm...) are issued. Changes can be activated now by manually restarting the **mmfs** daemons.

6027-1133 Interrupt received. Only -A, -Q, and -T were changed.

Explanation: An interrupt was received in **mmchfs**. Any changes to -A, -Q, and -T attributes were successfully completed. No other file system attributes were changed.

User Response: Repeat the command if you want to change additional attributes of the file system. The changes can be undone by running **mmchfs** with the original values for the affected attributes.

6027-1134 Interrupt: Disks removed from *file system*, but not all deleted.

Explanation: An interrupt was received in **mmdeldisk** after the specified disks were successfully removed from the file system. The virtual shared disks may not all have been deleted. If restriping was requested, it may have been interrupted.

User Response: Delete any remaining virtual shared disks manually. Issue the command **mmrestripe device** to restart the restriping if necessary.

6027-1135 Restriping may not have finished.

Explanation: An interrupt occurred during restriping.

User Response: Restart the restripe. Verify that the file system was not damaged by running the **mmfsck** command.

6027-1136 *flag* flag specified twice.

Explanation: A flag is specified multiple times on a command line

User Response: Correct the error and reissue the command.

6027-1137 *option* value must be yes or no.

Explanation: A yes or no flag is used with something other than yes or no.

User Response: Correct the error and reissue the command.

6027-1138 Invalid extra argument: *argument*

Explanation: Non-flag arguments follow the mandatory arguments.

User Response: Unlike most posix commands, the main arguments come first, followed by the optional flags. Correct the error and reissue the command.

6027-1139 Insufficient number of arguments.

Explanation: Not all of the mandatory arguments are specified.

User Response: Correct the error and reissue the command.

6027-1140 Invalid integer for flag: *number*

Explanation: A flag requiring an integer argument is followed by something that cannot be parsed as an integer.

User Response: Specify an integer with the indicated flag.

6027-1141 No disk descriptor file specified.

Explanation: An **-F** flag is not followed by the path name of a disk descriptor file.

User Response: Specify a valid disk descriptor file.

6027-1142 The configuration file for this GPFS configuration cannot be found in the SDR.

Explanation: The command cannot locate in the SDR the configuration file (*mmsdrcfgconfigld*) for the GPFS configuration affected by the command.

User Response: The SDR is corrupt. Determine what caused the loss of the file and correct the problem. The file can be restored by copying the **mmfs.cfg** file into the SDR as file *mmsdrcfgconfigld*.

6027-1143 Cannot open file.

Explanation: A file cannot be opened.

User Response: Make sure that the specified file exists and that you have the proper authorizations.

6027-1144 Cannot write to file.

Explanation: A file cannot be written to.

User Response: Make sure that the specified file exists, that you have the proper authorizations, and that the file system is not full.

6027-1145 parameter must be greater than 0: *value*

Explanation: A negative value has been specified for the named parameter which requires a positive value.

User Response: Correct the input and rerun the command.

6027-1146 node_number failed. command must be run on an SP.

Explanation: The *node_number* command failed.

User Response: Make sure PSSP is installed. If the error persists, follow problem determination procedures in *PSSP Diagnosis Guide*.

6027-1147 *command must be run on a node that belongs to a GPFS configuration.*

Explanation: A GPFS administration command (mm...) is run on the control workstation.

User Response: Run the command on a node that belongs to the GPFS configuration affected by the command.

6027-1148 *command must be run on a node in the GPFS configuration containing device.*

Explanation: A GPFS administration command (mm...) taking a device argument is run on the control workstation.

User Response: Run on a node in the GPFS configuration containing the device.

6027-1149 *Invalid node name name found in mmsdrfs file.*

Explanation: The **mmsdrfs** file in the SDR contains an invalid node name.

User Response: Configuration data contained in the SDR is incorrect. Correct any user modifications that have been made. If no user modifications have been made to the SDR, contact IBM support.

6027-1150 *Invalid cluster line in the mmsdrfs file.*

Explanation: The **mmsdrfs** contains a bad cluster line. (not 3 fields)

User Response: Configuration data contained in the SDR is incorrect. Correct any user modifications that have been made. If no user modifications have been made to the SDR, contact IBM support.

6027-1151 *Cannot create file name in the SDR.*

Explanation: A GPFS administration command (mm...) cannot create the specified file in the SDR.

User Response: Verify that you are authorized to use the **SDRCreatFile** command. If the problem persists, follow problem determination procedures in *PSSP Diagnosis Guide* and contact IBM support.

6027-1152 *Cannot replace file name in the SDR.*

Explanation: A GPFS administration command (mm...) cannot replace the specified file in the SDR.

User Response: Verify that you are authorized to use the **SDRReplaceFile** command. If the problem persists, follow problem determination procedures in *PSSP Diagnosis Guide* and contact IBM support.

6027-1153 *Warning: mmsdrfs file has nonstandard cluster lines.*

Explanation: Not all but the last cluster line for a GPFS configuration have exactly five nodes listed.

User Response: Although it is legal to have fewer than five nodes in a cluster line in the **mmsdrfs** file, when nodes are added using the **mmaddnode** command, it is possible that they will be inserted out of order. This forces a restart of all daemons that would otherwise have not been needed. This condition occurs only upon manual intervention in the **mmsdrfs** file. The **mmdelnode -c** command can be used to fix this condition.

6027-1154 Invalid node name: *name* specified for command.

Explanation: The user specified an invalid node name.

User Response: Use a valid name.

6027-1155 Node *name* is already represented in some GPFS configuration.

Explanation: The user specified a node that is already in use in an existing GPFS configuration.

User Response: Nodes can belong to only one GPFS configuration at a time. Specify a free node or delete the named node from its current GPFS configuration.

6027-1156 *node name* contains a single node GPFS configuration.

Explanation: The user specified a node containing a single node GPFS configuration.

User Response: Single node GPFS configurations are not listed in the SDR since they operate outside of the SP environment. If you want to use this node, run the command **mmdelnode *node name*** from the named node. Single node GPFS configurations cannot be deleted remotely.

6027-1157 This node contains a single node GPFS configuration. Run the command from another node or from the control workstation.

Explanation: The **mmdelnode** command was executed on a node containing a single node GPFS configuration.

User Response: The **mmdelnode** command cannot be run from a node which contains a single node GPFS configuration. Run it from another node in the affected GPFS configuration or from the control workstation.

6027-1158 Node *name* represented twice on the command line.

Explanation: The same node is referenced more than once on the command line.

User Response: All nodes specified must be unique. Note that even though two node names might appear to be different on the command line, both names may still refer to the same node.

6027-1159 GPFS configuration *name* not found in mmsdrfs second time.

Explanation: The temporary copy of **mmsdrfs** seems to have changed during execution.

User Response: Try the command again. If the problem persists contact IBM support.

6027-1160 *Command* is not valid for single node GPFS configurations.

Explanation: An attempt was made to add nodes to a single node GPFS configuration.

User Response: To convert a single node GPFS configuration into a multinode configuration, the single node configuration must be deleted first.

6027-1161 This GPFS configuration contains file systems. You cannot delete the last node.

Explanation: An attempt was made to delete a GPFS configuration that still has one or more file systems associated with it.

User Response: Before deleting the last node of a GPFS configuration, delete all file systems that are associated with it or move the file systems to other GPFS configurations.

6027-1162 *node* is not in any GPFS configuration.

Explanation: An attempt was made to work with a node that is not part of any GPFS configuration.

User Response: Issue the command from a node that belongs to the affected GPFS configuration.

6027-1163 MMFS is still active on *node*.

Explanation: An attempt was made to delete a node that has the mmfs daemon still active.

User Response: Stop the daemon on the nodes to be deleted. If **-c** is specified, then the daemon must be stopped on all nodes in the GPFS configuration.

6027-1164 *Node* is represented in no cluster, a different cluster, or matches another node on the command line.

Explanation: An invalid list of nodes was specified.

User Response: All nodes specified must be in the cluster and must be unique.

6027-1165 Warning: mmfixcfg failed while removing custom changes.

Explanation: **mmfixcfg** fails for some reason.

User Response: Custom changes will not be deleted from the configuration file. This is of consequence only if there WERE custom changes for some of the nodes being deleted, (put there by **mmchconfig**), AND if the nodes are eventually added back. The custom fixes will remain. Normally custom changes are lost. This error occurs if there is something wrong with the format of the **mmfs.cfg** file.

6027-1166 Not all nodes found in second pass.

Explanation: The **mmsdrfs** file is not successfully parsed on pass2.

User Response: Configuration data contained in the **mmsdrfs** is incorrect. Correct any user modifications that have been made. If no user modifications have been made to **mmsdrfs**, contact IBM support.

6027-1167 *name* flag is not valid in single node mode of operation.

Explanation: The **-n** flag was specified with the **mmconfig** command in single node operation mode. Single node mode of operation is in effect when the command runs on a non-SP or with the **-M** flag specified.

User Response: Omit the **-n** (nodelist) flag when running in single node mode.

6027-1169 No nodes specified for this GPFS configuration.

Explanation: File passed with the **-n** option contains no node names.

User Response: Place at least one node name in the node file.

6027-1170 Could not generate list of nodes for the GPFS configuration.

Explanation: A **/usr/lpp/ssp/bin/splstdata** command failed or it showed no lines for adapter type **css0**.

User Response: SDR tables might not contain switch names. This should be fixed. If the **/usr/lpp/ssp/bin/splstdata** command does not exist, you need to install PSSP. If there are no switch adapters, generate your own list with some other type of name and use it with the **-n** flag. Follow problem determination procedures in *PSSP Diagnosis Guide*.

6027-1171 *node* is already in a single node GPFS configuration.

Explanation: The node is configured as a single node GPFS configuration.

User Response: Nodes can belong to only one GPFS configuration at a time. Run the command from another node or delete the current GPFS configuration.

6027-1172 This node is not configured as a single node configuration.

Explanation: The node is not configured as a single node configuration and the **SDRGetObjects** command fails.

User Response: Run the command on a properly configured node. If this is supposed to be a multinode GPFS configuration, then the SDR commands are inaccessible. Follow problem determination procedures.

6027-1173 Specify only one node per line in *filename*.

Explanation: The node file contains lines with multiple host names.

User Response: The node file should contain one node name per line. Comments or whitespace are not allowed.

6027-1174 The GPFS configuration must be created on the node being configured.

Explanation: An attempt was made to create a single node GPFS configuration from a remote node.

User Response: Run the command from the node being configured.

6027-1175 Node *name* contains GPFS file systems that do not belong to a GPFS configuration. Such file systems must be deleted before running **mmconfig.**

Explanation: The **mmconfig** command detects GPFS file systems on a node that is not part of a GPFS configuration.

User Response: If these file systems were created outside the GPFS environment (not with a GPFS administration command), they need to be removed before running the **mmconfig** command. Otherwise, the data in the `/etc/filesystems` file may be incorrect and you might have to contact IBM support. The **mmcongif** command can be forced to ignore this error condition by specifying the `-f` option.

6027-1176 Warning: Node *name* contains GPFS file systems that do not belong to a GPFS configuration. References to these file systems in `/etc/filesystems` file will be lost.

Explanation: The **mmconfig** command detects GPFS file systems on a node that is not part of a GPFS configuration. Processing continues.

User Response: The subsequent use of a GPFS administration command (`mm...`) will cause references to these file systems to be deleted from the `/etc/filesystems`.

6027-1178 *parameter* must be from *number* to *number*: *value specified*

Explanation: A parameter value specified is out of range.

User Response: Keep the specified value within the range shown.

6027-1179 mmfixcfg failed while changing file.

Explanation: The **mmfixcfg** command fails to modify the **mmfs.cfg** file.

User Response: The **mmfs.cfg** file, or its copy in the SDR, may have illegal format. Check the preceding messages for more information. Correct any user modifications that have been made. If no user modifications have been made to the SDR, contact IBM support.

6027-1181 Cannot specify nodelist parameter on a single node GPFS configuration.

Explanation: The **mmchconfig** command is used with the **nodelist** parameter on a single node GPFS configuration.

User Response: Omit the **nodelist** parameter and reissue the command.

6027-1182 The nodelist parameter must contain fewer nodes than are in the entire GPFS configuration.

Explanation: Too many nodes are specified in the **nodelist**.

User Response: The idea of **nodelist** is to customize a subset of nodes. Specify only that subset. If the change is for all nodes, do not use **nodelist**.

6027-1184 Node *number* does not belong to any GPFS configuration, is not in the same GPFS configuration as previous nodes, or is represented multiple times on the command line.

Explanation: An invalid list of nodes is specified.

User Response: All specified nodes must belong to the same GPFS configuration. A node can be specified only once.

6027-1186 Invalid attribute: *value*.

Explanation: An unrecognized attribute is specified on the command line.

User Response: Correct the command line and try again.

6027-1187 disk not found in stanza for *command*.

Explanation: A disk specified for deletion or replacement does not exist.

User Response: Configuration data contained in the SDR is incorrect. Correct any user modifications that have been made. If no user modifications have been made to the SDR, contact IBM support.

6027-1188 Duplicate disk specified: *disk*.

Explanation: A disk specified for deletion is listed more than once on the command line.

User Response: Specify each disk to be deleted only once.

6027-1189 You cannot delete all the disks.

Explanation: The number of disks to delete is greater than or equal to the number of disks in the file system.

User Response: Delete only some of the disks. If you want to delete them all, use **mmdelfs**.

6027-1190 Cannot find stanza for *device* in the *mmsdrfs*.

Explanation: The file system device name passed to the command does not exist in the *mmsdrfs* file in the SDR.

User Response: Configuration data contained in the SDR is incorrect. Correct any user modifications that have been made. If no user modifications have been made to the SDR, contact IBM support.

6027-1191 Unexpected options line in stanza: *disks=* should be last option.

Explanation: The options line in the stanza for the device does not list the *disks=* option last.

User Response: Configuration data contained in the SDR is incorrect. Correct any user modifications that have been made. If no user modifications have been made to the SDR, contact IBM support.

6027-1192 Options line in stanza does not begin with '*rw*': *value present*

Explanation: The options line in the stanza for the device does not list *rw* as the first option.

User Response: Configuration data contained in the SDR is incorrect. Correct any user modifications that have been made. If no user modifications have been made to the SDR, contact IBM support.

6027-1200 *tscrfs* failed. Cannot create *device*.

Explanation: *tscrfs* failed.

User Response: See the MMFS error message from *tscrfs*.

6027-1201 *command of device* failed. Nothing is being changed.

Explanation: *tsdelfs* or *tsadddisk* failed.

User Response: See the MMFS error message.

6027-1202 *mmsdref* failed for *device*.

Explanation: *mmsdref* failed.

User Response: Configuration data contained in the SDR is incorrect. Correct any user modifications that have been made. If no user modifications have been made to the SDR, contact IBM support.

6027-1203 *mmkvsds* failed.

Explanation: One of the virtual shared disks to be created failed. An error message from *mmkvsds* should follow.

User Response: See the error message from *mmkvsds*.

6027-1204 *command* failed.

Explanation: An internal command fails. This is usually a call to the *mmfsd* daemon.

User Response: See the error message from the command that failed.

6027-1205 *-s* argument must be *balancedRandom*, *roundRobin*, or *random*.

Explanation: The stripe method argument specifies in invalid method.

User Response: Choose one of the types shown for *-s*.

6027-1206 *command currently runs only on an SP system.*

Explanation: A GPFS administration command (mm...) is run from a non-SP system.

User Response: Use these commands only on SP systems.

6027-1207 *There is already an existing file system using value.*

Explanation: The mountpoint or device specified matches that of an existing file system. The mountpoint must be unique within a given configuration. The device name must be unique across all GPFS configurations.

User Response: Choose an unused name or path.

6027-1208 *Each vsd specifies attribute for diskusage. This file system could not store value.*

Explanation: All descriptors specify the same diskusage.

User Response: Specify storage for data and metadata on at least one disk in the file system.

6027-1209 *command can be restarted using descriptors in restart.desc.*

Explanation: One of the disk descriptors has an error.

User Response: Fix the error, if any, on the descriptor(s) with errors. Some errors are transient. If a descriptor must be changed, change it in the original and also in the file **restart.desc**, which is created by **mmcrfs**. Try the command again, but this time use the new descriptor file: "-F **restart.desc**". This file will be the same as the old list, except all successful IBM Virtual Shared Disk creations will have the descriptor changed to "existing IBM Virtual Shared Disk style."

6027-1210 *Cannot understand return info from mmcrfsc.*

Explanation: **mmcrfsc**, the internal portion of **mmcrfs**, failed to return all data needed by the **mmcrfs** wrapper script.

User Response: This should only happen if the **mmcrfsc** is killed before writing all information. Restart the command. If this problem persists it is an internal processing error. Call IBM support.

6027-1211 *Cannot find mountpoint in filesystems stanza. Cannot remove device.*

Explanation: The device is found in **mmsdrfs**, but the stanza does not show a mountpoint as the first line.

User Response: Configuration data contained in the SDR is incorrect. Correct any user modifications that have been made. If no user modifications have been made to the SDR, contact IBM support.

6027-1212 *file system is still mounted on mountpoint.*

Explanation: Trying to delete a file system that is still mounted somewhere.

User Response: Unmount the file system on all nodes before deleting.

6027-1214 *Cannot make vsds on a non-SP system.*

Explanation: Running **mmkvsds** on an environment with no control workstation.

User Response: **mmkvsds** should not be called in this situation.

6027-1215 *disk on node (number megabytes) needs a partition size larger than the maximum of 256.*

Explanation: A disk appears to be larger than the max of 1016 partitions X 256M partition size = about 260 GB.

User Response: **mmkvsds** cannot handle this disk. Create the virtual shared disk manually if you can, and pass it in as existing.

6027-1216 *createvsd did not create the vsd.*

Explanation: **createvsd** should show the error message.

User Response: See accompanying error messages reported by **createvsd**.

6027-1217 *createvsd failed.*

Explanation: The error is followed by the output of **createvsd**, and then the contents of the IBM Virtual Shared Disk rollback file. The problem should be shown in this output.

User Response: See accompanying error messages reported by **createvsd**.

6027-1218 *No live nodes in GPFS configuration value.*

Explanation: Recovery needs a live node to repair the previous GPFS administration command (mm...).

User Response: Fix the state manually, or use the backout file once mmfs is up in the indicated GPFS configuration.

6027-1219 *Backout file moved to hostname.*

Explanation: Recovery cannot complete at this time.

User Response: The backout file is on the control workstation. If it is moved back to the original name (*/var/mmfs/etc/mmlog*) the recovery function will attempt to recover again at the next GPFS administration command (mm...).

6027-1220 *Inconsistent log types in logfile.*

Explanation: The **mmlog** file has more than one type of log line.

User Response: The backout file was corrupt and was deleted. This is an internal processing error. Contact IBM support.

6027-1221 *Cannot determine stripe group mgr for device.*

Explanation: The stripe group manager for a device is unavailable.

User Response: Fix the state manually, or use the backout file once mmfs has assigned a Stripe Group Manager to the indicated device.

6027-1222 *Cannot work on stripe group manager for device.*

Explanation: The stripe group manager for a device is unavailable.

User Response: Fix the state manually, or use the backout file once GPFS has assigned a live node as stripe group manager for the indicated device.

6027-1223 *Cannot do mmlsdisk for stripe group device.*

Explanation: **mmlsdisk** fails on the stripe group manager for the device.

User Response: See the error message from **mmlsdisk** or try the backout file again later.

6027-1224 *command must be run on the control workstation.*

Explanation: `mmrecover` is run on a node other than the control workstation.

User Response: Run `mmrecover` on the control workstation. If this message appears in any other command, call Service.

6027-1225 *Cannot access SDR.*

Explanation: The SDR is unavailable to a remote node.

User Response: Follow problem determination procedures in *PSSP Diagnosis Guide*.

6027-1226 *Could not command vsd name on node.*

Explanation: `cfgvsd` or `startvsd` failed on a remote node.

User Response: Refer to *Managing Shared Disks* for diagnosis information.

6027-1229 *Timed out waiting for lock: Try again later.*

Explanation: Another GPFS administration command (mm...) keeps the SDR lock for over a minute.

User Response: Try again later. If you think that no other GPFS administration command is using the SDR, it can be unlocked on the control workstation. See the section on SDR locks.

6027-1230 *Invalid undo type in argument.*

Explanation: Recovery code is calling `mmremote` with invalid arguments.

User Response: This is an internal processing error. Contact IBM support. Save the backout file if the command kept it.

6027-1231 *Unexpected second argument to mmremote: value.*

Explanation: `mmremote` is called with too many arguments.

User Response: See syntax shown. If this comes from inside another GPFS administration command (mm...), then check that all nodes including the control workstation have the same level of code in `/usr/lpp/mmfs/bin`. Otherwise, this is an internal processing error. Contact IBM support.

6027-1232 *Missing argument for keyword.*

Explanation: `mmremote` is called with a keyword that needs an argument.

User Response: Specify the keyword with an argument. If this comes from inside another GPFS administration command (mm...), then check that all nodes including the control workstation have the same level of code in `/usr/lpp/mmfs/bin`. Otherwise, this is an internal processing error. Contact IBM support.

6027-1233 *Invalid keyword: value.*

Explanation: `mmcommon` or `mmremote` is called with an invalid keyword.

User Response: See syntax shown. If this comes from inside another GPFS administration command (mm...), then check that all nodes including the control workstation have the same level of code in `/usr/lpp/mmfs/bin`. Otherwise, this is an internal processing error. Contact IBM support.

6027-1234 vfs number for mmfs in /etc/vfs is out of range.

Explanation: `mmremote` finds an invalid vfs number for mmfs.

User Response: The valid range is 8-15. Check `/etc/vfs` on the node in question. Perhaps there is something that can be released. If so, change the mmfs number to the available value.

6027-1235 Warning: there is no valid value available for mmfs in /etc/vfs.

Explanation: `mmremote` cannot find an unused vfs number for mmfs.

User Response: The valid range is 8-15. Check `/etc/vfs` on the node in question. Perhaps there is something that can be released.

6027-1236 Multiple config lines found for GPFS configuration *device*.

Explanation: The `mmsdrfs` has more than one config line for a GPFS configuration.

User Response: Configuration data contained in the SDR is incorrect. Correct any user modifications that have been made. If no user modifications have been made to the SDR, contact IBM support.

6027-1237 No config lines found for GPFS configuration *device*.

Explanation: The `mmsdrfs` has no config line for a GPFS configuration.

User Response: Configuration data contained in the SDR is incorrect. Correct any user modifications that have been made. If no user modifications have been made to the SDR, contact IBM support.

6027-1238 Unexpected number of config fields found for GPFS configuration *device*.

Explanation: The `mmsdrfs` has an invalid configuration line for a GPFS configuration.

User Response: Configuration data contained in the SDR is incorrect. Correct any user modifications that have been made. If no user modifications have been made to the SDR, contact IBM support.

6027-1239 *node* is still part of a GPFS configuration.

Explanation: `mmremote` gone is run on a node in a GPFS configuration.

User Response: If this message comes from some other GPFS administration command (`mm...`) it is an internal processing error. Contact IBM support.

6027-1240 Cannot find *disk* with *lspv*.

Explanation: The `hdisk` specified does not exist on the server specified.

User Response: Choose a disk that exists.

6027-1241 *disk* already has a volume group.

Explanation: The `hdisk` specified already has a volume group assigned.

User Response: When creating virtual shared disks you must choose physical volumes that are not a member of any volume group.

6027-1243 Diskusage *value* does not match that of disk being replaced *value*.

Explanation: The diskusage field specified to **mmrpldisk** does not match the disk being replaced.

User Response: Specify the same diskusage as the disk being replaced or leave the field blank. It will default to the correct value.

**6027-1244 Failure group *number* does not match that of disk. being replaced *number*.
To match it, let the failure group field default and use a disk on the same primary server as the disk being replaced.**

Explanation: The user specifies a failure group to **mmrpldisk** and the old disk has a failure group greater than 4000.

User Response: Use a disk on the same node and let the failure group default.

6027-1245 Failure group *number* does not match that of disk being replaced (*number*).

Explanation: The failure group field specified to **mmrpldisk** does not match the disk being replaced.

User Response: Specify the same failure group or let it default.

6027-1246 Primary vsd server node of new disk does not match that of disk being replaced. Use a disk on the same server to ensure that the disks are members of the same failure group.

Explanation: Default failure groups are used for both disks specified to **mmrpldisk**, and the disks do not share the same primary virtual shared disk server node.

User Response: Specify a disk on the same node as the disk being replaced.

6027-1248 Mountpoint must not be a relative path name: *path*.

Explanation: The mountpoint does not begin with `/`.

User Response: Specify the full path for mountpoint.

6027-1249 dataStructureDump must not specify a relative path name: *path*.

Explanation: The pathname specified for **dataStructureDump** (**-D** flag in **mmconfig**) does not begin with `/`.

User Response: Specify the full path for **dataStructureDump**.

6027-1250 -i cannot be used with attribute *name*.

Explanation: **-i** is specified with a configuration attribute which does not allow the option.

User Response: If the change must be made immediately, reissue the command leaving off the **-i** option, then recycle the nodes to activate the new value.

6027-1251 pagepool (*number*) must be greater than minpagepool (*number*).

Explanation: An incompatible change in a configuration attribute for **pagepool** has been specified.

User Response: Use valid relative values for **pagepool** attributes.

6027-1252 minpagepool (*number*) must be \geq maxblocksize (*number*) \times 5 + 640K.

Explanation: An incompatible change in the configuration attribute **minpagepool** or **maxblocksize**, relative to one another, has been specified.

User Response: Use valid relative values for the **minpagepool** and **maxblocksize** attributes.

6027-1253 Invalid value for *flag*.

Explanation: An incorrect argument is specified with a flag requiring one of a limited number of legal options.

User Response: Use one of the legal values for the indicated flag.

6027-1254 Usage: mmcommon { convhr | convhn | convnr | convin } *argument*

mmcommon { create | delete | read | write} *filename* [*pathname*]

mmcommon init *filename* [*pid* [*configid*]]

mmcommon on1 *hostname arguments*

mmcommon oncw *arguments*

mmcommon onall *pathname arguments*

Explanation: The **mmcommon** command received invalid command line arguments.

User Response: **mmcommon** is a routine called internally by other GPFS administration commands (mm...). The user need never use **mmcommon**, but may choose to. If another command returns this message, check that GPFS has the same level of code installed on all nodes. If the user has issued this command, verify the arguments. Otherwise, it is an internal processing error. Contact IBM support.

6027-1255 Usage: mmcommon init *filename* [*pid* [*cluster num*]]

Explanation: The **mmcommon** command with the **init** flag received invalid command line arguments

User Response: Correct the command line and try again.

6027-1256 mmcommon init must have MMMODE set to sp, avs, or single when configuration id is provided.

Explanation: The **mmcommon init** command received an invalid environment variable for **MMMODE** when a configuration id was specified.

User Response: Correct the command line and try again.

6027-1257 MMMODE must be set to sp, avs, or single.

Explanation: The **mmcommon** command received an invalid environment variable for **MMMODE**

User Response: Correct the command line and try again.

6027-1258 Usage: mmcommon *option argument*.

Explanation: The **mmcommon** command specified with *option* received an invalid command line argument.

User Response: Correct the command line and try again

6027-1259 Get adapter failed for IPA *number*.

Explanation: The IP address being converted does not have a valid adapter in the SDR adapter table.

User Response: If the IP address is invalid, the calling command is in error. If the IP address is valid, and should have an adapter, check the integrity of the SDR. If the node name is **localhost**, the IP address is valid but has no adapter. **Localhost** should not be entered from the command line.

6027-1260 Invalid IP address *number* specified for command.

Explanation: The IP address being converted is not recognized by the host command.

User Response: If the IP address is invalid, the calling command is in error. If the IP address is valid, the host should recognize it.

6027-1261 Usage: mmcommon delete *name*.

Explanation: The **mmcommon delete** command received invalid command line arguments.

User Response: Correct command line and try again.

6027-1262 Usage: mmcommon *argument name pathname*.

Explanation: The **mmcommon** command is issued with an invalid argument. Valid arguments are **create**, **read**, and **write**.

User Response: Correct the command and try again.

6027-1263 *filename* already exists.

Explanation: Attempting to create a file that already exists.

User Response: Replace the file instead of creating it with the **mmcommon** command.

6027-1264 *filename* does not exist.

Explanation: An attempt has been made to read, delete, or replace a file that does not exist

User Response: The configuration file has been deleted. If the file was unexpectedly deleted while the command was running, it is an internal processing error. Contact IBM support. Files on a single node system cannot be recovered.

6027-1265 Cannot perform remote operations in a single node environment.

Explanation: **mmcommon** was called to perform a remote operation in a single node environment.

User Response: If this error appears in another GPFS administration command, it is an internal processing error. Contact IBM support.

6027-1266 *name* flag and *name* flag cannot both be specified.

Explanation: A GPFS administration command has been issued with conflicting options.

User Response: Correct command line and try again.

6027-1267 rsh on *node* failed. Fix access permissions.

Explanation: A remote operation in AVS mode failed.

User Response: Give rsh authority to run on the remote node

6027-1268 Missing arguments.

Explanation: A GPFS administration command received an insufficient number of arguments.

User Response: Correct command line and try again.

6027-1269 The device name *device* starts with a slash, but not /dev/.

Explanation: The device name was specified incorrectly.

User Response: Correct the device name and try again.

6027-1270 The device name *device* contains a slash, but not in the first position of the name.

Explanation: The device name was specified incorrectly.

User Response: Correct the device name and try again.

6027-1271 Unexpected error from *command*. Return code: *number*.

Explanation: An internally called GPFS administration command (mm...) received an unexpected error code.

User Response: Perform problem determination.

6027-1272 Ping of *node* failed. The node must be up to be included in a partition.

Explanation: Command fails to communicate with a node being added to a cluster. The node must be up to perform certain checks.

User Response: Remove this node from the node list and add it to the cluster later when it is up.

6027-1273 Warning: An old MMFS/GPFS kernel extension is still loaded and cannot be unloaded. Please reboot before using newly installed version of the product.

Explanation: An old kernel extension was detected during the installation of the latest level of GPFS.

User Response: Reboot the nodes to allow loading of the new kernel extension.

6027-1274 Cannot retrieve the mmsdrfs file from the SDR.

Explanation: The mmsdrfs file in the SDR does not exist or cannot be retrieved.

User Response: This message is normal if no GPFS configurations have been created yet. However, if the **mmconfig** command has been successfully executed in the past, perform problem determination and contact IBM support.

6027-1275 *node* does not belong to any GPFS configuration.

Explanation: **mmlsnode** cannot find a GPFS configuration containing the specified node.

6027-1276 There are no known GPFS configurations.

Explanation: **mmlsnode** cannot find any GPFS configurations in the SDR file mmsdrfs.

6027-1277 GPFS configuration *number* does not exist.

Explanation: The specified configuration number is not found in the SDR file mmsdrfs.

6027-1281 *file system* does not belong to any GPFS configuration.

Explanation: The command which was issued cannot find a GPFS configuration containing the specified file system.

6027-1282 Usage: *Device* [-a] [-B] [-C] [-d] [-f] [-F] [-i] [-I] [-m] [-M] [-n] [-Q] [-r] [-R] [-s] [-V].

Explanation: Usage statement for **mmlsfs**

6027-1283 Invalid node name found in the mmsdrfs file.

Explanation: An IP address being converted does not have a valid adapter in the SDR adapter table.

User Response: Configuration data contained in the SDR is incorrect. Correct any user modifications that have been made. If no user modifications have been made to the SDR, contact IBM support.

6027-1284 Usage: *node* { *configId* | -a | -n *nodename* }.

Explanation: **mmlsnode** received invalid command line arguments.

User Response: Correct command line and try again.

6027-1285 Usage: *file system device* -C *ConfigurationId*.

Explanation: **mmmvfs** received invalid command line arguments.

User Response: Correct command line and try again.

6027-1286 Cannot generate a list of node names.

Explanation: The mmsdrfs file in the SDR does not exist or does not contain valid node names.

User Response: This message is normal if no GPFS configurations have been made yet. Otherwise, if the **mmconfig** command has been successfully executed in the past, perform problem determination and contact IBM support.

6027-1287 The format of the file systems stanza in the SDR is invalid.

Explanation: A file system stanza in the mmsdrfs does not have a valid format.

User Response: Configuration data contained in the SDR is incorrect. Correct any user modifications that have been made. If no user modifications have been made to the SDR, contact IBM support.

6027-1288 *file system* is still mounted on: *node list*.

Explanation: Attempting to delete or move a file system that is still mounted on the nodes listed.

User Response: Unmount the file system and retry.

6027-1289 GPFS configuration *name* already exists.

Explanation: The **mmconfig -C** command received a configuration id that already exists

User Response: Specify a different configuration id, or allow the system to generate one. The **mmlsnode -a** command can be used to display all existing GPFS configurations.

6027-1290 Invalid value for GPFS configuration id: *name*

Explanation: The specified GPFS configuration id is invalid.

User Response: Specify a valid GPFS configuration id.

6027-1291 Options *name* and *name* cannot be specified at the same time.

Explanation: Incompatible options were specified on the command line.

User Response: Select one of the options and reissue the command.

6027-1401 Recovery log I/O failed: unmounting file system *name*, RC=*number*.

Explanation: Recovery log I/O could not complete.

User Response: Check the paths to all disks that make up the file system. Run `mmlsdisk` to determine whether GPFS has declared any disks unavailable. Repair any failing paths to disks. Remount the file system.

6027-1500 Open of *stripe group* failed with error *number*.

Explanation: A device failed. Operation of the stripe group will continue unless this disk is needed for an operation. Often, in replicated disk situations, it will not be needed.

User Response: Take needed repair action on the specified disk or the path to the specified disk.

6027-1501 Volume label of disk *name* is *name*, should be *name*.

Explanation: The UID in the disk descriptor does not match the expected value from the stripe group descriptor. This could occur if a disk was overwritten by another application or if IBM Virtual Shared Disk incorrectly identified the disk.

User Response: Check the disk configuration.

6027-1502 Volume label of disk *name* is corrupt.

Explanation: The disk descriptor has a bad magic number, version, or checksum. This could occur if a disk was overwritten by another application or if IBM Virtual Shared Disk incorrectly identified the disk.

User Response: Check the disk configuration.