

# GPFS Questions and Answers



## Overview

General Parallel File System™ (GPFS™) is a high performance shared-disk file management solution that provides fast, reliable access from nodes in a cluster environment. Parallel and serial applications can readily access shared files using standard UNIX® file system interfaces, and the same file can be accessed concurrently from multiple nodes. GPFS is designed to provide high availability through logging and replication, and can be configured for failover from both disk and server malfunctions. GPFS scalability and performance are designed to meet the needs of data intensive applications such as engineering design, digital media, data mining, relational databases, financial analytical, seismic data processing, scientific research and scalable file serving.

GPFS for POWER™ is supported on both AIX® and Linux®. GPFS for AIX runs on the IBM® eServer™ Cluster 1600 as well as clusters of IBM Power, IBM System p™, IBM eServer p5, IBM BladeCenter® servers. GPFS for Linux runs on select IBM Power, System p, eServer p5, BladeCenter and IBM eServer OpenPower® servers. The GPFS Multiplatform product runs on the IBM System Cluster 1350™ as well as Linux clusters based on selected IBM x86 xSeries® rack-optimized servers, select IBM BladeCenter servers, or select IBM for AMD processor-based servers.

Additionally, GPFS Multiplatform V3.2.1 is supported on nodes running Windows® Server 2003 R2 on 64-bit architectures (AMD x64 / EM64T) in an existing GPFS V3.2.1 cluster of AIX and/or Linux (32-bit or 64-bit) where all nodes are at service level 3.2.1-5 or later.

For further information regarding the use of GPFS in your clusters, see the *GPFS: Concepts, Planning, and Installation Guide*.

Updates to this FAQ include:

*Table 1.*

December 2008	
	2.3 What are the latest distributions and kernel levels that GPFS has been tested with?
	7.2 What is the current service information for GPFS?

# Questions & Answers

## 1. General questions:

- 1.1 How do I order GPFS?
- 1.2 How is GPFS priced?
- 1.3 Where can I find the documentation for GPFS?
- 1.4 What resources beyond the standard documentation can help me learn and use GPFS?
- 1.5 How can I ask a more specific question about GPFS?
- 1.6 Does GPFS participate in the IBM Academic Initiative Program?

## 2. Software questions:

- 2.1 What levels of the AIX O/S are supported by GPFS?
- 2.2 What Linux distributions are supported by GPFS?
- 2.3 What are the latest distributions and kernel levels that GPFS has been tested with?
- 2.4 What levels of the Windows O/S are supported by GPFS?
- 2.5 Can different GPFS maintenance levels coexist?
- 2.6 Are there any requirements for Clustered NFS (CNFS) support in GPFS V3.2?
- 2.7 Are there any requirements for Persistent Reserve support in GPFS V3.2?
- 2.8 Are there any considerations when utilizing the Simple Network Management Protocol (SNMP)-based monitoring capability in GPFS V3.2?

## 3. Machine questions:

- 3.1 What are the minimum hardware requirements for a GPFS cluster?
- 3.2 Is GPFS for POWER supported on IBM System i<sup>™</sup> servers?
- 3.3 What machine models has GPFS for Linux been tested with?
- 3.4 Is GPFS for Linux supported on all IBM ServerProven<sup>®</sup> servers?
- 3.5 What interconnects are supported for GPFS daemon-to-daemon communication in my GPFS cluster?
- 3.6 Does GPFS support exploitation of the Virtual I/O Server (VIOS) features of POWER5<sup>™</sup> processors?

## 4. Disk questions:

- 4.1 What disk hardware has GPFS been tested with?
- 4.2 What Fibre Channel (FC) Switches are qualified for GPFS usage and is there a FC Switch support chart available?
- 4.3 Can I concurrently access disks from both AIX and Linux nodes in my GPFS cluster?
- 4.4 What disk support failover models does GPFS support for the IBM TotalStorage<sup>®</sup> DS4000<sup>™</sup> family of storage controllers with the Linux operating system?
- 4.5 What devices have been tested with SCSI-3 Persistent Reservations?
- 4.6 Are there any special considerations when my cluster consists of two nodes?

## 5. Scaling questions:

- 5.1 What are the GPFS cluster size limits?
- 5.2 What are the current file system size limits?
- 5.3 What is the current limit on the number of mounted file systems in a GPFS cluster?
- 5.4 What is the architectural limit of the number of files in a file system?

- 5.5 What is the current limit on the number of nodes that may concurrently join a cluster?
- 5.6 What are the limitations on GPFS disk size?
- 6. Configuration and tuning questions:
  - 6.1 What specific configuration and performance tuning suggestions are there?
  - 6.2 What configuration and performance tuning suggestions are there for GPFS when used primarily for Oracle databases?
  - 6.3 What configuration requirements exist for utilizing Remote Direct Memory Access (RDMA) offered by InfiniBand?
  - 6.4 What Linux configuration settings are required when NFS exporting a GPFS filesystem?
  - 6.5 Sometimes GPFS appears to be handling a heavy I/O load, for no apparent reason. What could be causing this?
  - 6.6 What considerations are there when using IBM Tivoli® Storage Manager with GPFS?
  - 6.7 How do I get OpenSSL to work on AIX and SLES8/ppc64?
  - 6.8 What ciphers are supported for use by GPFS?
  - 6.9 When I allow other clusters to mount my file systems, is there a way to restrict access permissions for the root user?
- 7. Service questions:
  - 7.1 What support services are available for GPFS?
  - 7.2 What is the current service information for GPFS?
  - 7.3 How do I download fixes for GPFS?
  - 7.4 What are the current GPFS advisories?
  - 7.5 What Linux kernel patches are provided for clustered file systems such as GPFS?
  - 7.6 What Windows hotfix updates are required for GPFS?
  - 7.7 Where can I find licensing and ordering information for GPFS?

---

## General questions

### Q1.1: How do I order GPFS?

#### A1.1: To order GPFS:

- To order GPFS on POWER for AIX or Linux, find contact information for your country at <http://www.ibm.com/planetwide/>
- To order GPFS Multiplatform for Linux or Windows, go to the Passport Advantage® site at <http://www.ibm.com/software/lotus/passportadvantage/>

### Q1.2: How is GPFS priced?

#### A1.2: The price for GPFS for POWER is based on the number of processors active on the server where GPFS is installed.

The price for GPFS Multiplatform is based on a Processor Value Unit metric. A Value Unit is a pricing charge metric for program license entitlements which is based upon the quantity of a specifically designated measurement used for a given program, in this case processors or processor cores. Under the processor Value Unit licensing metric, each processor core is assigned a specific number of Value Units. You must acquire the total number of processor Value Units for each processor core on which the software program is deployed. IBM continues to define a processor to be each processor core on a chip. For example, a dual-core chip contains two processor cores.

A processor core is a functional unit within a computing device that interprets and executes instructions. A processor core consists of at least an instruction control unit and one or more arithmetic or logic unit. Not all processor cores require the same number of Value Unit entitlements. With multi-core technology, each core is considered a processor.

See [http://www.ibm.com/software/lotus/passportadvantage/pvu\\_licensing\\_for\\_customers.html](http://www.ibm.com/software/lotus/passportadvantage/pvu_licensing_for_customers.html)

Each software program has a unique price per Value Unit. The number of Value Unit entitlements required for a program depends on how the program is deployed in your environment and must be obtained from a Value Unit table. GPFS Multiplatform is grouped into packs of 10 processor Value Units as the minimum order quantity. For example, when you need 50 processor Value Units, you will order 5 of these 10 processor Value Unit part numbers to get the required 50 processor Value Units. To determine the total cost of deploying GPFS, multiply the program price per Value Unit by the total number of processor Value Units required. To calculate the number of Value Unit entitlements required, refer to the Value Unit Table at

[http://www.ibm.com/software/lotus/passportadvantage/pvu\\_table\\_for\\_customers.html](http://www.ibm.com/software/lotus/passportadvantage/pvu_table_for_customers.html)

and the Value Unit Calculator at

<https://www-112.ibm.com/software/howtobuy/passportadvantage/valueunitcalculator/vucalc.wss>

For further information:

- In the United States, please call 1-888-SHOP-IBM
- In all other locations, please contact your IBM Marketing Representative. For a directory of worldwide contact, see [www.ibm.com/planetwide/index.html](http://www.ibm.com/planetwide/index.html)

### Q1.3: Where can I find the documentation for GPFS?

#### A1.3: The GPFS documentation is available in both PDF and HTML format on the Cluster Information Center at [publib.boulder.ibm.com/infocenter/clresctr/vxrx/index.jsp?topic=/com.ibm.cluster.gpfs.doc/gpfsbooks.html](http://publib.boulder.ibm.com/infocenter/clresctr/vxrx/index.jsp?topic=/com.ibm.cluster.gpfs.doc/gpfsbooks.html).

**Q1.4: What resources beyond the standard documentation can help me learn about and use GPFS?**

**A1.4:** For additional information regarding GPFS see:

- The GPFS section of the HPC wiki at <http://www.ibm.com/developerworks/wikis/display/hpccentral/General+Parallel+File+System+%28GPFS%29>
- IBM Training course GPFS 3.1 System Administration [http://www-304.ibm.com/jct03001c/services/learning/ites.wss/us/en?pageType=course\\_description&courseCode=AU310](http://www-304.ibm.com/jct03001c/services/learning/ites.wss/us/en?pageType=course_description&courseCode=AU310)
- The Clusters Literature site for AIX 5L™ at [www.ibm.com/servers/eserver/clusters/library/wp\\_aix\\_lit.html](http://www.ibm.com/servers/eserver/clusters/library/wp_aix_lit.html)
- The Clusters Literature site for Linux at [www.ibm.com/servers/eserver/clusters/library/wp\\_linux\\_lit.html](http://www.ibm.com/servers/eserver/clusters/library/wp_linux_lit.html)
- The IBM Redbooks® and Redpapers site at [www.redbooks.ibm.com](http://www.redbooks.ibm.com)
  - recently published *A Guide to the IBM Clustered Network File System*
- The IBM Almaden Research GPFS web page at [www.almaden.ibm.com/StorageSystems/file\\_systems/GPFS/index.shtml](http://www.almaden.ibm.com/StorageSystems/file_systems/GPFS/index.shtml)
- Go to the IBM Systems Magazine at <http://www.ibmsystemsmag.com/> and search on GPFS.

**Q1.5: How can I ask a more specific question about GPFS?**

**A1.5:** Depending upon the nature of your question, you may ask it in one of several ways.

- If you want to correspond with IBM regarding GPFS:
  - If your question concerns a potential software error in GPFS and you have an IBM software maintenance contract, please contact 1-800-IBM-SERV in the United States or your local IBM Service Center in other countries. IBM Scholars Program users should notify the GPFS development team of potential software bugs through [gpfs@us.ibm.com](mailto:gpfs@us.ibm.com).
  - If you have a question that can benefit other GPFS users, you may post it to the GPFS technical discussion forum at [www.ibm.com/developerworks/forums/dw\\_forum.jsp?forum=479&cat=13](http://www.ibm.com/developerworks/forums/dw_forum.jsp?forum=479&cat=13)
  - This FAQ is continually being enhanced. To contribute possible questions or answers, please send them to [gpfs@us.ibm.com](mailto:gpfs@us.ibm.com)
- If you want to interact with other GPFS users, the San Diego Supercomputer Center maintains a GPFS user mailing list. The list is [gpfs-general@sdsc.edu](mailto:gpfs-general@sdsc.edu) and those interested can subscribe to the list at [lists.sdsc.edu/mailman/listinfo/gpfs-general](http://lists.sdsc.edu/mailman/listinfo/gpfs-general)

If your question does not fall into the above categories, you can send a note directly to the GPFS development team at [gpfs@us.ibm.com](mailto:gpfs@us.ibm.com). However, this mailing list is informally monitored as time permits and should not be used for priority messages to the GPFS team.

**Q1.6: Does GPFS participate in the IBM Academic Initiative Program?**

**A1.6:** GPFS no longer participates in the IBM Academic Initiative Program.

If you are currently using GPFS with an education license from the Academic Initiative, we will continue to support GPFS 3.2 on a best-can-do basis via email for the licenses you have.

However, no additional or new licenses of GPFS will be available from the IBM Academic Initiative program. You should work with your IBM client representative on what educational discount may be available for GPFS. See [www.ibm.com/planetwide/index.html](http://www.ibm.com/planetwide/index.html)

## Software questions

**Q2.1: What levels of the AIX O/S are supported by GPFS?**

**A2.1:** GPFS supports AIX V6.1, AIX V5.3 and V5.2 nodes in a homogenous or heterogeneous cluster running either the AIX or the Linux operating system.

Table 2. GPFS for AIX

GPFS	AIX V6.1	AIX V5.3	AIX V5.2
GPFS V3.2	X	X	X
GPFS V3.1		X	X
GPFS v2.3		X	X

### Notes:

1. The following additional filesets are required by GPFS V3.2:

- xlc.aix50.rte (C Set ++<sup>®</sup> Runtime for AIX 5.0), version 8.0.0.0 or later
- xlc.rte (C Set ++ Runtime), version 8.0.0.0 or later

These can be downloaded from Fix Central at <http://www.ibm.com/eserver/support/fixes/fixcentral>

2. Enhancements to the support of Network File System (NFS) V4 in GPFS V3 are only available on AIX V5.3 systems with the minimum technology level of 5300-04 applied or on AIX V6.1 with GPFS V3.2 .
3. The version of OpenSSL shipped with AIX V6.1 will not work with GPFS due to a change in how the library is built. To obtain the level of OpenSSL which will work with GPFS, see the question *How do I get OpenSSL to work on AIX and SLES8/ppc64?*
4. For additional support information, please also see the question, *What is the current service information for GPFS?*
5. Customers should consider the support plans for AIX V5.2 in their operating system decision.
6. For the latest GPFS fix level, go to <https://www14.software.ibm.com/webapp/set2/sas/f/gpfs/home.html>

**Q2.2: What Linux distributions are supported by GPFS?**

**A2.2:** GPFS supports the following distributions:

**Note:** For kernel level support, please see question *What are the latest kernel levels that GPFS has been tested with?*

Table 3. Linux distributions supported by GPFS

	RHEL 5 <sup>2</sup>	RHEL 4	RHEL 3	SLES 10 <sup>1,4</sup>	SLES 9	SLES 8
GPFS Multiplatform V3.2	X	X		X	X	
GPFS for POWER V3.2	X <sup>3</sup>	X		X	X	
GPFS Multiplatform V3.1		X	X	X	X	X
GPFS for POWER V3.1		X		X	X	
GPFS for POWER V2.3		X			X	X
GPFS Multiplatform V2.3		X	X		X	X

Table 3. Linux distributions supported by GPFS (continued)

	RHEL 5 <sup>2</sup>	RHEL 4	RHEL 3	SLES 10 <sup>1,4</sup>	SLES 9	SLES 8
1. There is required service for GPFS V3.1 support of SLES 10. Please see question <i>What is the current service information for GPFS?</i> 2. RHEL 5.0 and later on POWER requires GPFS V3.2.0.2 or later 3. GPFS V3.2 for Linux on POWER does not support mounting of a file system with a 16KB block size when running on RHEL 5. 4. The GPFS GPL build requires imake. If imake was not installed on the SLES 10 system, install <code>xorg-x11-devel-*.rpm</code> .						

**Q2.3: What are the latest kernel levels that GPFS has been tested with?**

**A2.3:** While GPFS runs with many different AIX fixes and Linux kernel levels, it is highly suggested that customers apply the latest fix levels and kernel service updates for their operating system. To download the latest GPFS service updates, go to <https://www14.software.ibm.com/webapp/set2/sas/f/gpfs/home.html>

GPFS does not currently support the following kernels:

- RHEL hugemem kernel
- RHEL largesmp
- RHEL uniprocessor (UP) kernel
- SLES xen kernel

Table 4. GPFS for Linux V3.2

	Linux Distribution	Kernel Level
	POWER	
	Red Hat EL 5.2 <sup>1,2,3</sup>	2.6.18-92.1.10
	Red Hat EL 4.6	2.6.9-67.0.7
	SUSE Linux ES 10 SP2	2.6.16.60-0.27
	SUSE Linux ES 9 SP4	2.6.5-7.312
	x86_64	
	Red Hat EL 5.2 <sup>2,3</sup>	2.6.18-92.1.10
	Red Hat EL 4.6	2.6.9-67.0.7
	SUSE Linux ES 10 SP2	2.6.16.60-0.27
	SUSE Linux ES 9 SP4	2.6.5-7.312
	i386	
	Red Hat EL 5.2 <sup>2,3</sup>	2.6.18-92.1.10
	Red Hat EL 4.6	2.6.9-67.0.7
	SUSE Linux ES 10 SP2	2.6.16.60-0.27
	SUSE Linux ES 9 SP4	2.6.5-7.312
	Itanium® 2 <sup>4</sup>	
	Red Hat EL 4.5	2.6.9-55.0.6
	SUSE Linux ES 10 SP1	2.6.16.53-0.8
	SUSE Linux ES 9 SP3	2.6.5-7.286

Table 4. GPFS for Linux V3.2 (continued)

Linux Distribution	Kernel Level
1. RHEL 5.0 and later on POWER requires GPFS V3.2.0.2 or later 2. With RHEL5.1, the automount option is slow. This issue should be addressed in the 2.6.18-53.1.4 kernel when it is available. 3. GPFS V3.2.1-3 or later supports the RHEL xen kernel. 4. GPFS for Linux on Itanium Servers is available only through a special Programming Request for Price Quotation (PRPQ). The install image is not generally available code. It must be requested by an IBM client representative through the RPQ system and approved before order fulfillment. If interested in obtaining this PRPQ, reference PRPQ # P91232 or Product ID 5799-GPS.	

Table 5. GPFS for Linux V3.1

Linux Distribution	Kernel Level
POWER	
Red Hat EL 4.6	2.6.9-67.0.7
SUSE Linux ES 10 SP2	2.6.16.60-0.27
SUSE Linux ES 9 SP4	2.6.5-7.312
x86_64	
Red Hat EL 4.6	2.6.9-67.0.7
Red Hat EL 3.8	2.4.21-47.0.1
SUSE Linux ES 10 SP2	2.6.16.60-0.27
SUSE Linux ES 9 SP4	2.6.5-7.312
SUSE Linux ES 8 SP4	2.4.21-309
i386	
Red Hat EL 4.6	2.6.9-67.0.7
Red Hat EL 3.8	2.4.21-47.0.1
SUSE Linux ES 10 SP2	2.6.16.60-0.27
SUSE Linux ES 9 SP4	2.6.5-7.312
SUSE Linux ES 8 SP4	2.4.21-309

Table 6. GPFS for Linux V2.3

Linux Distribution	Kernel Level
POWER	
Red Hat EL 4.4	2.6.9-42.EL
SUSE Linux ES 9 SP3	2.6.5-7.282
SUSE Linux ES 8 SP4	2.4.21-309
x86_64	
Red Hat EL 4.3	2.6.9-34.0.2
Red Hat EL 3.8	2.4.21-47.EL
SUSE Linux ES 9 SP3	2.6.5-7.276-smp
SUSE Linux ES 8 SP4	2.4.21-309
i386	
Red Hat EL 4.3	2.6.9-34.0.2.EL
Red Hat EL 3.7	2.4.21-40.EL



Table 6. GPFS for Linux V2.3 (continued)

Linux Distribution	Kernel Level
SUSE Linux ES 9 SP3	2.6.5-7.282-smp
SUSE Linux ES 8 SP4	2.4.21-309

**Q2.4: What levels of the Windows O/S are supported by GPFS?**

**A2.4:** GPFS Multiplatform V3.2.1-5 and later, is supported on nodes running Windows Server 2003 R2 on 64-bit architectures (AMD x64 / EM64T) in an existing GPFS V3.2.1 cluster of AIX and/or Linux at V3.2.1-5 or later.

**Q2.5: Can different GPFS maintenance levels coexist?**

**A2.5:** Certain levels of GPFS can coexist, that is, be active in the same cluster and simultaneously access the same file system. This allows upgrading GPFS within a cluster without shutting down GPFS on all nodes first, and also mounting GPFS file systems from other GPFS clusters that may be running a different maintenance level of GPFS. The current maintenance level coexistence rules are:

- All GPFS V3.2 maintenance levels can coexist with each other and with GPFS V3.1 Maintenance Level 13 or later, unless otherwise stated in this FAQ.

See the *Migration, coexistence and compatibility* information in the GPFS V3.2 *Concepts, Planning, and Installation Guide*

- The default file system version was incremented in GPFS 3.2.1-5. File systems created using GPFS v3.2.1.5 code without using the `--version` option of the `mmcrfs` command will not be mountable by earlier code.
- GPFS V3.2 maintenance levels 3.2.1.2 and 3.2.1.3 have coexistence issues with other maintenance levels.

Customers using a mixed maintenance level cluster that have some nodes running 3.2.1.2 or 3.2.1.3 and other nodes running other maintenance levels should uninstall the `gpfs.msg.en_US rpm/fileset` from the 3.2.1.2 and 3.2.1.3 nodes. This should prevent the wrong message format strings going across the mixed maintenance level nodes.

- **Attention:** Do not use the `mmrepquota` command if there are nodes in the cluster running a mixture of 3.2.0.3 and other maintenance levels. A fix will be provided in APAR #IZ16367. A fix can be provided for 3.2.0.3 upon request prior to APAR availability in the March service level available at <https://www14.software.ibm.com/webapp/set2/sas/f/gpfs/home.html>
  - All GPFS V3.1 maintenance levels can coexist with each other, unless otherwise stated in this FAQ.
- Attention:** GPFS V3.1 maintenance levels 10 (GPFS-3.1.0.10) thru 12 (GPFS-3.1.0.12) do not coexist with other maintenance levels

All nodes in the cluster must conform to one of these maintenance level compatibility restrictions:

- all nodes must be at maintenance levels 1-9 or 13 and later (GPFS-3.1.0.1 thru GPFS-3.1.0.9 or GPFS-3.1.0.13 and later)
- all nodes must be at maintenance levels 10-12 (GPFS-3.1.0.10 - GPFS-3.1.0.12)
- Starting with GPFS 2.3.0.6, different maintenance levels of GPFS 2.3 can coexist. Prior to GPFS 2.3.0.6 all nodes must be at the same maintenance level.
- GPFS V2.3 is completely incompatible and can not coexist with either GPFS V3.1 or GPFS V3.2..

**Q2.6: Are there any requirements for Clustered NFS (CNFS) support in GPFS V3.2?**

**A2.6:** GPFS V3.2 Clustered NFS (CNFS) support requirements:

- If NLM locking is required, until the code is included in the kernel, a kernel patch for lockd must be applied. This patch is currently available at [http://sourceforge.net/tracker/?atid=719124&group\\_id=130828&func=browse](http://sourceforge.net/tracker/?atid=719124&group_id=130828&func=browse)

The required lockd patch is not supported on RHEL 4 ppc64.

- For SUSE distributions, use of the common NFS utilities (sm-notify in user space) is required. The specific patches required within util-linux are:
  - support statd notification by name (patch-10113) <http://support.novell.com/techcenter/psdb/2c7941abcdf7a155ecb86b309245e468.html>
  - specify a host name for the -v option (patch-10852) <http://support.novell.com/techcenter/psdb/e6a5a6d9614d9475759cc0cd033571e8.html>
  - allow selection of IP source address on command line (patch-9617) <http://support.novell.com/techcenter/psdb/c11e14914101b2debe30f242448e1f5d.html>
- For Red Hat distributions, use of nfs-utils 1.0.7 is required for rpc.statd fixes. Go to <https://www.redhat.com/>

Table 7. CNFS requirements

	lockd patch required	sm-notify required	rpc.statd required
SLES 10	X	X	not required
SLES 9	X	X	not required
RHEL 5	X (not available for ppc64)	included in base distribution	X
RHEL 4	X (not available for ppc64)	included in base distribution	X

See also *What Linux kernel patches are provided for clustered file systems such as GPFS?*

**Q2.7: Are there any requirements for the use of the Persistent Reserve support in GPFS V3.2?**

**A2.7:** GPFS V3.2 supports Persistent Reserve on AIX and requires:

- For AIX 5L<sup>™</sup> V5.2 APAR IZ00673
- For AIX 5L V5.3 IZ01534, IZ04114

**Q2.8: Are there any considerations when utilizing the Simple Network Management Protocol (SNMP)-based monitoring capability in GPFS V3.2?**

**A2.8:** Considerations for the use of the SNMP-based monitoring capability in GPFS include:

- Currently, the SNMP collector node must be a Linux node in your GPFS cluster. GPFS utilizes Net-SNMP which is not supported by AIX.
- Support for ppc64 requires the use of Net-SNMP 5.4.1. Binaries for Net-SNMP 5.4.1 on ppc64 are not available. You will need to download the source and build the binary. Go to <http://net-snmp.sourceforge.net/download.html>
- If the monitored cluster is relatively large, you need to increase the communication time-out between the SNMP master agent and the GPFS SNMP subagent. In this context, a cluster is considered to be large if the number of nodes is greater than 25, or the number of file systems is greater than 15, or the total number of disks in all file systems is greater than 50. For more information see *Configuring Net-SNMP* in the *GPFS: Advanced Administration Guide*.

---

## Machine questions

**Q3.1: What are the minimum hardware requirements for a GPFS cluster?**

**A3.1:** The minimum hardware requirements are:

- GPFS for POWER: IBM POWER3™ or newer processor, 1 GB of memory
- GPFS Multiplatform for Linux:
  - Intel® Pentium® 3 or newer processor, with 512 MB of memory
  - AMD Opteron™ processors, with 1 GB of memory
  - Intel Itanium 2 processor with 1 GB of RAM<sup>1</sup>
- GPFS Multiplatform for Windows:
  - Intel EM64T processors, with 1GB of memory
  - AMD Opteron processors, with 1 GB of memory

**Note:** Due to issues found during testing, GPFS for Windows is not supported on e325 servers

Additionally, it is highly suggested that a sufficiently large amount of swap space is configured. While the actual configuration decisions should be made taking into account the memory requirements of other applications, it is suggested to configure at least as much swap space as there is physical memory on a given node.

GPFS is supported on systems which are listed in, or compatible with, the IBM hardware specified in the *Hardware requirements* section of the Sales Manual for GPFS. If you are running GPFS on hardware that is not listed in the *Hardware Requirements*, should problems arise and after investigation it is found that the problem may be related to incompatibilities of the hardware, we may require reproduction of the problem on a configuration conforming to IBM hardware listed in the sales manual.

To access the Sales Manual for GPFS:

1. Go to <http://www-306.ibm.com/common/ssi/OIX.wss>
2. From **A specific type** menu, choose **HW&SW Desc (Sales Manual,RPQ)**.
3. To view a GPFS sales manual, choose the corresponding product number to enter in the **keyword** field then click on **Go**
  - For General Parallel File System for POWER V3.2.1, enter **5765-G66**
  - For General Parallel File System Multiplatform V3.2.1, enter **5724-N94**
  - For General Parallel File System for AIX 5L V3.1, enter **5765-G66**
  - For General Parallel File System for Linux on POWER V3.1, enter **5765-G67**
  - For General Parallel File System Multiplatform V3.1 for Linux, enter **5724-N94**
  - For General Parallel File System for AIX 5L V2.3, enter **5765-F64**
  - For General Parallel File System for Linux on POWER V2.3, enter **5765-G20**
  - For General Parallel File System Multiplatform V2.3 for Linux, enter **5765-G23**

**Q3.2: Is GPFS for POWER supported on IBM System i servers?**

**A3.2:** GPFS for POWER extends all features, function, and restrictions (such as operating system and scaling support) to IBM System i servers to match their IBM System p counterparts:

Table 8.

IBM System i	IBM System p
i-595	p5-595
i-570	p5-570, p6-570
i-550	p5-550
i-520	p5-520

No service updates are required for this additional support.

---

1. GPFS for Linux on Itanium Servers is available only through a special Programming Request for Price Quotation (PRPQ). The install image is not generally available code. It must be requested by an IBM client representative through the RPQ system and approved before order fulfillment. If interested in obtaining this PRPQ, reference PRPQ # P91232 or Product ID 5799-GPS.

**Q3.3: What machine models has GPFS for Linux been tested with?**

**A3.3:** GPFS has been tested with:

- IBM x86 xSeries machine models:
  - 330
  - 335
  - 336
  - 340
  - 342
  - 345
  - 346
  - 360
  - 365
  - 440
  - x3550
  - x3650
  - x3655
- IBM BladeCenter x86 blade servers:
  - HS20
  - HS21
  - HS40
  - LS20
  - LS21
- IBM POWER processor-based blade servers:
  - JS20
  - JS21
  - JS22
- IBM BladeCenter Cell/B.E.<sup>™</sup> blade servers
  - QS21
- IBM AMD processor-based servers:
  - 325
  - 326
- IBM eServer p5 :

**For both the p5-590 and the p5-595:** See the question *What is the current service information for GPFS?*

  - 510
  - 520
  - 550
  - 570
  - 575
  - 590
  - 595
- IBM Power POWER6<sup>™</sup>
  - 570
  - 575
  - 595
- IBM eServer OpenPower servers:
  - 710
  - 720
- IBM eServer pSeries<sup>®</sup> machines models that support Linux
- The IBM eServer Cluster 1300
- The IBM System Cluster 1350

For hardware and software certification, please see the IBM ServerProven site at <http://www.ibm.com/servers/eserver/serverproven/compat/us/>

**Q3.4: Is GPFS for Linux supported on all IBM ServerProven servers?**

**A3.4:** GPFS for Linux is supported on all IBM ServerProven servers:

1. with the distributions and kernel levels as listed in the question *What are the latest distributions and kernel levels that GPFS has been tested with?*
2. that meet the minimum hardware model requirements as listed in the question *What are the minimum hardware requirements for a GPFS cluster?*

Please see the IBM ServerProven site at <http://www.ibm.com/servers/eserver/serverproven/compat/us/>

**Q3.5: What interconnects are supported for GPFS daemon-to-daemon communication in a GPFS cluster?**

**A3.5:** The interconnect for GPFS daemon-to-daemon communication depends upon the types of nodes in your cluster.

*Table 9. GPFS daemon -to-daemon communication interconnects*

Nodes in your cluster	Supported interconnect	Supported environments
Linux	Ethernet	all supported GPFS environments
	10-Gigabit Ethernet	all supported GPFS environments
	Myrinet	IP only
	InfiniBand	GPFS Multiplatform V3.2 for Linux: <ul style="list-style-type: none"><li>• IP and optionally VERBS RDMA</li></ul> <b>Note:</b> See the question <i>Are there any considerations when utilizing the Remote Direct Memory Access (RDMA) offered by InfiniBand?</i> GPFS for Linux on POWER V3.2: <ul style="list-style-type: none"><li>• IP only</li></ul> GPFS V2.3 or 3.1: <ul style="list-style-type: none"><li>• IP only</li><li>• SLES 9 or Red Hat EL 4.0, on System x servers</li><li>• SLES 9 SP 3 on System p5 servers with GPFS V3.1.0-4 or later</li></ul>
Linux/AIX/Windows	Ethernet	all supported GPFS environments
	10-Gigabit Ethernet	<ul style="list-style-type: none"><li>• all supported GPFS Linux environments</li><li>• AIX V5.3</li><li>• AIX V6.1</li></ul>
AIX	Ethernet	all supported GPFS environments
	10-Gigabit Ethernet	AIX V5.3 AIX V6.1
	Myrinet	AIX V5.2 and V5.3 64-bit kernel BladeCenter JS20 and p5 POWER5 servers IP only
	InfiniBand	AIX V5.3 GPFS V3.1 or V3.2 IP only
	eServer HPS	homogenous clusters of either AIX V5.2 or V5.3
	SP™ Switch2	GPFS V2.3

**Q3.6: Does GPFS support exploitation of the Virtual I/O Server (VIO) features of POWER5 processors?**

**A3.6:** Yes, GPFS allows exploitation of POWER5 VIOS configurations. Both the virtual SCSI (VSCSI) and the shared Ethernet adapter (SEA) are supported in single and multiple central electronics complex (CEC) configurations. This support is limited to GPFS nodes that are using either the AIX V6.1 or V5.3 operating system.

All LPARs in a CEC that are GPFS cluster members must have all the VIO disks mapped to each LPAR using virtual SCSI. This creates to GPFS a SAN environment where each node has access to disk on a local path without requiring network access. All of the NSD's in these configurations must NOT be coded with any NSD server associated with them.

Minimum required code levels:

- VIOS Release 1.3.0.0 Fix Pack 8
- AIX 5L V5.3 Service Pack 5300-05-01

There is no GPFS fix level requirement for this support, but it is recommended that you be at the latest GPFS level available. For information on the latest levels of GPFS go to <https://www14.software.ibm.com/webapp/set2/sas/f/gpfs/home.html>

For further information on POWER5 VIOS go to [techsupport.services.ibm.com/server/vios/documentation/faq.html](http://techsupport.services.ibm.com/server/vios/documentation/faq.html)

For VIOS documentation, go to [techsupport.services.ibm.com/server/vios/documentation/home.html](http://techsupport.services.ibm.com/server/vios/documentation/home.html)

---

## Disk questions

### Q4.1: What disk hardware has GPFS been tested with?

A4.1: This set of tables displays the set of disk hardware which has been tested by IBM and known to work with GPFS. GPFS is not limited to only using this set of disk devices. Other disk devices may work with GPFS but they have not been tested by IBM. The GPFS support team will help customers who are using devices outside of this list of tested devices, to solve problems directly related to GPFS, but not problems deemed to be issues with the underlying device's behavior including any performance issues exhibited on untested hardware.

It is important to note that:

- Each individual disk subsystem requires a specific set of device drivers for proper operation while attached to a host running GPFS or IBM Recoverable Virtual Shared Disk. The prerequisite levels of device drivers are not documented in this GPFS-specific FAQ. Refer to the disk subsystem's web page to determine the currency of the device driver stack for the host's operating system level and attachment configuration.

For information on IBM disk storage subsystems and their related device drivers levels and Operating System support guidelines, go to [www.ibm.com/servers/storage/support/disk/index.html](http://www.ibm.com/servers/storage/support/disk/index.html)

- Microcode levels should be at the latest levels available for your specific disk drive.

For the IBM System Storage™, go to [www.ibm.com/servers/storage/support/allproducts/downloading.html](http://www.ibm.com/servers/storage/support/allproducts/downloading.html)

- GPFS for Windows can only operate as an NSD client at this time, and as such does not support direct attached disks.

**DS4000 customers:** Please also see

- The IBM TotalStorage DS4000 *Best Practices and Performance Tuning Guide* at [publib-b.boulder.ibm.com/abstracts/sg246363.html?Open](http://publib-b.boulder.ibm.com/abstracts/sg246363.html?Open)
- For the latest firmware and device driver support for DS4100 and DS4100 Express Midrange Disk System, go to <http://www.ibm.com/systems/support/supportsite.wss/selectproduct?brandind=5000028&familyind=5329597&osind=0&oldbrand=5000028&oldfamily=5345919&oldtype=0&taskind=2&matrix=Y&psid=dm>
- For the latest storage subsystem controller firmware support for DS4200, DS4700, DS4800, go to:
  - <https://www.ibm.com/systems/support/supportsite.wss/docdisplay?lnocid=MIGR-5075581&brandind=5000028>
  - <https://www.ibm.com/systems/support/supportsite.wss/docdisplay?lnocid=MIGR-5073716&brandind=5000028>

Table 10. Disk hardware tested with GPFS for AIX on POWER

GPFS for AIX on POWER:	
	IBM System Storage DS6000™ using either Subsystem Device Driver (SDD) or Subsystem Device Driver Path Control Module (SDDPCM)  <b>Configuration considerations:</b> GPFS clusters up to 32 nodes are supported and require a firmware level of R9a.5b050318a or greater. See further requirements below.
	IBM System Storage DS8000™ using either SDD or SDDPCM  <b>Configuration considerations:</b> GPFS clusters up to 32 nodes are supported and require a firmware level of R10k.9b050406 or greater. See further requirements below.

Table 10. Disk hardware tested with GPFS for AIX on POWER (continued)

	<p><b>DS6000 and DS8000 service requirements:</b></p> <ul style="list-style-type: none"> <li>• AIX 5L V5.2 maintenance level 05 (5200-05) - APAR # IY68906, APAR # IY70905</li> <li>• AIX 5L V5.3 maintenance level 02 (5300-02) - APAR # IY68966, APAR # IY71085</li> <li>• GPFS for AIX 5L V2.3 - APAR # IY66584, APAR # IY70396, APAR # IY71901</li> </ul> <p><b>For the Disk Leasing model install the latest supported version of the SDD fileset supported on your operating system.</b></p> <p><b>For the Persistent Reserve model install the latest supported version of SDDPCM fileset supported for your operating system.</b></p>
	IBM TotalStorage DS4100 (Formerly FAStT 100) with DS4000 EXP100 Storage Expansion Unit with Serial Advanced Technology Attachment (SATA) drives
	IBM TotalStorage FAStT500
	IBM System Storage DS4200 Express all supported expansion drawer and disk types
	IBM System Storage DS4300 (Formerly FAStT 600) with DS4000 EXP710 Fibre Channel (FC) Storage Expansion Unit, DS4000 EXP700 FC Storage Expansion Unit, or EXP100
	IBM System Storage DS4300 Turbo with EXP710, EXP700, or EXP100
	IBM System Storage DS4400 (Formerly FAStT 700) with EXP710 or EXP700
	IBM System Storage DS4500 (Formerly FAStT 900) with EXP710, EXP700, or EXP100
	IBM System Storage DS4700 Express all supported expansion drawer and disk types
	IBM System Storage DS4800 with EXP710, EXP100 or EXP810
	IBM System Storage DS3400 (1726-HC4)
	IBM TotalStorage ESS (2105-F20 or 2105-800 with SDD)
	IBM TotalStorage ESS (2105-F20 or 2105-800 using AIX 5L Multi-Path I/O (MPIO) and SDDPCM))
	<p>IBM System Storage Storage Area Network (SAN) Volume Controller (SVC) V2.1 and V3.1</p> <p>The following APAR numbers are suggested:</p> <ul style="list-style-type: none"> <li>• IY64709 - Applies to all GPFS clusters</li> <li>• IY64259 - Applies only when running GPFS in an AIX V5.2 or V5.3 environment with RVSD 4.1</li> <li>• IY42355 - Applies only when running GPFS in a PSSP V3.5 environment</li> <li>• SVC V2.1.0.1 is supported with AIX 5L V5.2 (Maintenance Level 05) and AIX 5L V5.3 (Maintenance Level 01).</li> </ul> <p>See <a href="http://www.ibm.com/support/docview.wss?rs=591&amp;uid=ssg1S1002471">www.ibm.com/support/docview.wss?rs=591&amp;uid=ssg1S1002471</a> for specific advice on SAN Volume Controller recommended software levels.</p>
	IBM 7133 Serial Disk System (all disk sizes)



Table 10. Disk hardware tested with GPFS for AIX on POWER (continued)

	<p>Hitachi Lightning 9900™/9910, 9960, 9970 and 9980 Hitachi Universal Storage Platform 100/600/1100</p> <p><b>Notes:</b></p> <ol style="list-style-type: none"> <li>1. In all cases Hitachi Dynamic Link Manager™ (HDLM) (multipath software) or MPIO (default PCM - failover only) is required</li> <li>2. AIX ODM objects supplied by Hitachi Data Systems (HDS) are required for all above devices.</li> <li>3. Customers should consult with HDS to verify that their proposed combination of the above components is supported by HDS.</li> </ol>
	<p>EMC Symmetrix DMX Storage Subsystems (FC attach only)</p> <p>Selected models of CX/CX-3 family including CX300, CX400, CX500 CX600, CX700 and CX3-20, CX3-40 and CX3-80</p> <p>Device driver support for Symmetrix includes both MPIO and PowerPath. <b>Note:</b> CX/CX-3 requires PowerPath.</p> <p>Customers should consult with EMC to verify that their proposed combination of the above components is supported by EMC.</p>
	<p>HP XP 128/1024 XP10000/12000 HP StorageWorks Enterprise Virtual Arrays (EVA) 4000/6000/8000 and 3000/5000 models that have been upgraded to active-active configurations <b>Note:</b> HDLM multipath software is required</p>
	<p>IBM DCS9550 (either FC or SATA drives) FC attach only minimum firmware 3.08b must use IBM supplied ODM objects at level 1.7 or greater</p> <p>For more information on the DCS9550 go to <a href="http://www.datadirectnet.com/dcs9550/">http://www.datadirectnet.com/dcs9550/</a></p>

Table 11. Disk hardware tested with GPFS for Linux on x86 xSeries servers

GPFS for Linux on xSeries servers:	
	IBM TotalStorage FASTT 200 Storage Server
	IBM TotalStorage FASTT 500
	IBM TotalStorage DS4100 (Formerly FASTT 100) with EXP100
	IBM System Storage DS4200 Express all supported expansion drawer and disk types
	IBM System Storage DS4300 (Formerly FASTT 600) with EXP710, EXP700, or EXP100
	IBM System Storage DS4300 Turbo with EXP710, EXP700, or EXP100
	IBM System Storage DS4400 (Formerly FASTT 700) with EXP710 or EXP700
	IBM System Storage DS4500 (Formerly FASTT 900) with EXP710, EXP700, or EXP100
	IBM System Storage DS4700 Express all supported expansion drawer and disk types
	IBM System Storage DS4800 with EXP710, EXP100 or EXP810
	IBM System Storage DS3400 (1726-HC4)
	IBM TotalStorage Enterprise Storage Server® (ESS) models 2105-F20 and 2105-800, with Subsystem Device Driver (SDD)

Table 11. Disk hardware tested with GPFS for Linux on x86 xSeries servers (continued)

<b>GPFS for Linux on xSeries servers:</b>	
	EMC Symmetrix Direct Matrix Architecture (DMX) Storage Subsystems 1000 with PowerPath v 3.06 and v 3.07
	IBM System Storage Storage Area Network (SAN) Volume Controller (SVC) V2.1 and V3.1  See <a href="http://www.ibm.com/support/docview.wss?rs=591&amp;uid=ssg1S1002471">www.ibm.com/support/docview.wss?rs=591&amp;uid=ssg1S1002471</a> for specific advice on SAN Volume Controller recommended software levels.
	IBM DCS9550 (either FC or SATA drives) FC attach only minimum firmware 3.08b QLogic drivers at 8.01.07 or newer and IBM SAN Surfer V5.0.0 or newer <a href="http://support.qlogic.com/support/oem_detail_all.asp?oemid=376">http://support.qlogic.com/support/oem_detail_all.asp?oemid=376</a>  For more information on the DCS9550 go to <a href="http://www.datadirectnet.com/dcs9550/">http://www.datadirectnet.com/dcs9550/</a>
<b>Restrictions:</b>	IBM ServeRAID™ adapters are not supported.

Table 12. Disk hardware tested with GPFS for Linux on POWER

<b>GPFS for Linux on POWER:</b>	
	IBM System Storage DS4200 Express all supported expansion drawer and disk types
	IBM System Storage DS4300 (Formerly FAStT 600) all supported drawer and disk types
	IBM System Storage DS4500 (Formerly FAStT 900) all supported expansion drawer and disk types
	IBM System Storage DS4700 Express all supported expansion drawer and disk types
	IBM System Storage DS4800 all supported expansion drawer and disk types
	IBM System Storage DS8000 using SDD

Table 13. Disk hardware tested with GPFS for Linux on AMD processor-based servers

<b>GPFS for Linux on eServer AMD processor-based servers:</b>	No devices tested specially in this environment.
---	--

- Q4.2: What Fibre Channel Switches are qualified for GPFS usage and is there a FC Switch support chart available?**
- A4.2:** There are no special requirements for FC switches used by GPFS other than the switch must be supported by AIX or Linux. For further information see [www.storage.ibm.com/ibmsan/index.html](http://www.storage.ibm.com/ibmsan/index.html)
- Q4.3: Can I concurrently access SAN-attached disks from both AIX and Linux nodes in my GPFS cluster?**
- A4.3:** The architecture of GPFS allows both AIX and Linux hosts to concurrently access the same set of LUNs. However, before this is implemented in a GPFS cluster you must ensure that the disk subsystem being used supports both AIX and Linux concurrently accessing LUNs. While the GPFS architecture allows this, the underlying disk subsystem may not, and in that case, a configuration attempting it would not be supported.

**Q4.4: What disk support failover models does GPFS support for the IBM System Storage DS4000 family of storage controllers with the Linux operating system?**

**A4.4:** GPFS has been tested with both the Host Bus Adapter Failover and Redundant Dual Active Controller (RDAC) device drivers.

To download the current device drivers for your disk subsystem, please go to <http://www.ibm.com/servers/storage/support/>

**Q4.5: 4.5 What devices have been tested with SCSI-3 Persistent Reservations?**

**A4.5:** The following devices have been tested with SCSI-3 Persistent Reservations:

- DS8000 (all 2105 and 2107 models) using SDDPCM or the default MPIO PCM on AIX.
- DS4000 subsystems using the IBM RDAC driver on AIX. (devices.fcp.disk.array.rte)

The most recent versions of the device drivers are always recommended to avoid problems that have been addressed.

**Q4.6: Are there any special considerations when my cluster consists of two nodes?**

**A4.6:** Customers who previously used single-node quorum and are migrating to a supported level of GPFS, must be aware that the single-node quorum function has been replaced with *node quorum with tiebreaker disks*. The new node quorum with tiebreaker disks support does *not* depend upon the availability of SCSI-3 persistent reserve. All disks tested with GPFS can now utilize node quorum with tiebreaker disks as opposed to GPFS node quorum (one plus half of the explicitly defined quorum nodes in the GPFS cluster). For further information, see the *GPFS: Concepts, Planning, and Installation Guide* for your level of GPFS.

---

## Scaling questions

**Q5.1: What are the GPFS cluster size limits?**

**A5.1:** The current maximum tested GPFS cluster size limits are:

*Table 14. GPFS maximum tested cluster sizes*

GPFS Multiplatform for Linux	2441 nodes
GPFS on POWER for AIX	1530 nodes
GPFS Multiplatform for Windows	64 nodes
<b>Contact <a href="mailto:gpfs@us.ibm.com">gpfs@us.ibm.com</a> if you intend to exceed:</b> 1. Configurations with Linux larger than 512 nodes 2. Configurations with AIX larger than 128 nodes 3. Configurations with Windows larger than 32 nodes	

Although GPFS is typically targeted for a cluster with multiple nodes, it can also provide high performance benefit for a single node so there is no lower limit. However, there are two points to consider:

- GPFS is a well-proven, scalable cluster file system. For a given I/O configuration, typically multiple nodes are required to saturate the aggregate file system performance capability. If the aggregate performance of the I/O subsystem is the bottleneck, then GPFS can help achieve the aggregate performance even on a single node.
- GPFS is a highly available file system. Therefore, customers who are interested in single-node GPFS often end up deploying a multi-node GPFS cluster to ensure availability.<sup>2</sup>

**Q5.2: What are the current file system size limits?**

**A5.2:** The current file system size limits are:

*Table 15. Current file system size limits*

GPFS 2.3 or later, file system architectural limit	2 <sup>99</sup> bytes
GPFS 2.2 file system architectural limit	2 <sup>51</sup> bytes (2 Petabytes)
Current tested limit	approximately 2 PB
<b>Contact <a href="mailto:gpfs@us.ibm.com">gpfs@us.ibm.com</a> if you intend to exceed 200 Terabytes</b>	

**Q5.3: What is the current limit on the number of mounted file systems in a GPFS cluster?**

**A5.3:** The total number of mounted file systems within a GPFS cluster depends upon your service level of GPFS:

*Table 16.*

GPFS Service Level	Number of mounted file systems
GPFS V3.2.0.1 or later	256
GPFS V3.1.0.5 or later	64
GPFS V3.1.0.1 thru V3.1.0.4	32
GPFS v2.3 all service levels	32

**Q5.4: What is the architectural limit of the number of files in a file system?**

**A5.4:** The architectural limit of the number of files in a file system is determined by the file system format. For file systems created prior to GPFS V2.3, the limit is 268,435,456. For file systems created with GPFS V2.3 or later, the limit is 2,147,483,648. Please note that the effective limit on the number of files in a file system is usually lower than the architectural limit, and could be adjusted using the **-F** option of the **mmchfs** command.

---

2. *GPFS Sequential Input/Output Performance on IBM pSeries 690*, Gautam Shah, James Wang available at <http://www.redbooks.ibm.com/redpapers/pdfs/redp3945.pdf>

**Q5.5: What is the current limit on the number of nodes that may concurrently join a cluster?**

**A5.5:** The total number of nodes that may concurrently join a cluster depends upon the level of GPFS which you are running:

- GPFS V3.2 is limited to a maximum of 8192 nodes.
- GPFS V3.1 and V2.3 are limited to a maximum of 4096 nodes.

A node joins a given cluster if it is:

- A member of the local GPFS cluster (the **mmlscluster** command output displays the local cluster nodes).
- A node in a different GPFS cluster that is mounting a file system from the local cluster.

For example:

- GPFS **clusterA** has 2100 member nodes as listed in the **mmlscluster** command.
- 500 nodes from **clusterB** are mounting a file system owned by **clusterA**.

**clusterA** therefore has 2600 concurrent nodes.

**Q5.6: What are the limitations on GPFS disk size?**

**A5.6:** The maximum disk size supported by GPFS depends on the file system format and the underlying device support. For file systems created prior to GPFS version 2.3, the maximum disk size is 1 TB due to internal GPFS file system format limitations. For file systems created with GPFS 2.3 or later, these limitations have been removed, and the maximum disk size is only limited by the OS kernel and device driver support:

Table 17.

OS kernel	Maximum supported GPFS disk size
AIX, 64-bit kernel	>2TB, up to the device driver limit
AIX, 32-bit kernel	1TB
Linux 2.6 64-bit kernels	>2TB, up to the device driver limit
Linux 2.6 32-bit kernels, Linux 2.4	2TB

**Notes:**

1. The above limits are only applicable to nodes that access disk devices through a local block device interface, as opposed to NSD protocol. For NSD clients, the maximum disk size is only limited by the NSD server large disk support capability, irrespective of the kernel running on an NSD client node.
2. The basic reason for the significance of the 2TB disk size barrier is that this is the maximum disk size that can be addressed using 32-bit sector numbers and 512-byte sector size. A larger disk can be addressed either by using 64-bit sector numbers or by using larger sector size. GPFS uses 64-bit sector numbers to implement large disk support. Disk sector sizes other than 512 bytes are unsupported.
3. GPFS for Windows can only operate as an NSD client at this time, and as such does not support direct attached disks.

---

## Configuration and tuning questions

Please also see the question *What is the current service information for GPFS?*

**Q6.1: What specific configuration and performance tuning suggestions are there?**

**A6.1:** In addition to the configuration and performance tuning suggestions in the *GPFS: Concepts, Planning, and Installation Guide* for your version of GPFS:

- If your GPFS cluster is configured to use SSH/SCP, it is suggested that you increase the value of **MaxStartups** in **sshd\_config** to at least 1024.
- You must ensure that when you are designating nodes for use by GPFS you specify a non-aliased interface. Utilization of aliased interfaces may produce undesired results. When creating or adding nodes to your cluster, the specified hostname or IP address must refer to the communications adapter over which the GPFS daemons communicate. When specifying servers for your NSDs, the output of the **mmlscluster** command lists the hostname and IP address combinations recognized by GPFS. Utilizing an aliased hostname not listed in the **mmlscluster** command output may produce undesired results.
- If your system consists of the eServer pSeries High Performance Switch, it is suggested that you configure GPFS over the **ml0** IP network interface.
- On systems running with the Linux 2.6 kernel, it is recommended you adjust the **vm.min\_free\_kbytes** kernel tunable. This tunable controls the amount of free memory that Linux kernel keeps available (i.e. not used in any kernel caches). When **vm.min\_free\_kbytes** is set to its default value, on some configurations it is possible to encounter memory exhaustion symptoms when free memory should in fact be available. Setting **vm.min\_free\_kbytes** to a higher value (Linux **sysctl** utility could be used for this purpose), on the order of magnitude of 5-6% of the total amount of physical memory, should help to avoid such a situation.

Also, please see the GPFS Redpapers:

- *GPFS Sequential Input/Output Performance on IBM pSeries 690* at [www.redbooks.ibm.com/redpapers/pdfs/redp3945.pdf](http://www.redbooks.ibm.com/redpapers/pdfs/redp3945.pdf)
- *Native GPFS Benchmarks in an Integrated p690/AIX and x335/Linux Environment* at [www.redbooks.ibm.com/redpapers/pdfs/redp3962.pdf](http://www.redbooks.ibm.com/redpapers/pdfs/redp3962.pdf)
- *Sequential I/O performance of GPFS on HS20 blades and IBM System Storage™ DS4800* at [ftp://ftp.software.ibm.com/common/ssi/rep\\_wh/n/CLW03002USEN/CLW03002USEN.PDF](http://ftp.software.ibm.com/common/ssi/rep_wh/n/CLW03002USEN/CLW03002USEN.PDF)

**Q6.2: What configuration and performance tuning suggestions are there for GPFS when used primarily for Oracle databases?**

**A6.2:**

**Note:** Only a subset of GPFS releases are certified for use in Oracle environments. For the latest status of GPFS certification:

- For AIX go to, [http://www.oracle.com/technology/products/database/clustering/certify/tech\\_generic\\_unix\\_new.html](http://www.oracle.com/technology/products/database/clustering/certify/tech_generic_unix_new.html)
- For Linux go to, [http://www.oracle.com/technology/products/database/clustering/certify/tech\\_generic\\_linux\\_new.html](http://www.oracle.com/technology/products/database/clustering/certify/tech_generic_linux_new.html)

In addition to the performance tuning suggestions in the *GPFS: Concepts, Planning, and Installation Guide* for your version of GPFS:

- When running Oracle RAC 10g, it is suggested you increase the value for **OPROCD\_DEFAULT\_MARGIN** to at least 500 to avoid possible random reboots of nodes.

In the control script for the Oracle CSS daemon, located in **/etc/init.cssd** the value for **OPROCD\_DEFAULT\_MARGIN** is set to 500 (milliseconds) on all UNIX derivatives except for AIX. For AIX this value is set to 100. From a GPFS perspective, even 500 milliseconds maybe too low in situations where node failover may take up to a minute or two to resolve. However, if during node failure the surviving node is already doing direct IO to the **oprocd** control file, it should have the necessary tokens and indirect block cached and should therefore not have to wait during failover.

- Using the IBM General Parallel File System is attractive for RAC environments because executables, trace files and archive log files are accessible on all nodes. However, care must be taken to properly configure the system in order to prevent false node evictions, and to maintain the ability to perform rolling upgrades of the Oracle software. Without proper configuration GPFS recovery from a node failure can interfere with cluster management operations resulting in additional node failures.

If you are running GPFS and Oracle RAC 10gR2 and encounter false node evictions:

- Upgrade the CRS to 10.2.0.3 or newer.

The Oracle 10g Clusterware (CRS) executables or logs (the CRS\_HOME) should be placed on a local JFS2 filesystem. Using GPFS for the CRS\_HOME can inhibit CRS functionality on the surviving nodes while GPFS is recovering from a failed node for the following reasons:

- In Oracle 10gR2, up to and including 10.2.0.3, critical CRS daemon executables are not pinned in memory. Oracle and IBM are working to improve this in future releases of 10gR2.
- Delays in updating the CRS log and authorization files while GPFS is recovering can interfere with CRS operations.
- Due to an Oracle 10g limitation rolling upgrades of the CRS are not possible when the CRS\_HOME is on a shared filesystem.
- CSS voting disks and the Oracle Clusterware Registry (OCR) should not be placed on GPFS as the IO freeze during GPFS reconfiguration can lead to node eviction, and the inability of CRS to function. Place the OCR and Voting disk on shared raw devices (hdisks).
- Oracle Database 10g (RDBMS) executables are supported on GPFS for Oracle RAC 10g. However, the system should be configured to support multiple ORACLE\_HOME's so as to maintain the ability to perform rolling patch application. Rolling patch application is supported for the ORACLE\_HOME starting in Oracle RAC 10.2.0.3.
- Oracle Database 10g data files, trace files, and archive log files are supported on GPFS.

See also:

- *Configuring IBM General Parallel File System (GPFS) with Oracle RAC on IBM pSeries with AIX 5L and Linux on POWER* at <http://www.ibm.com/servers/enable/site/peducation/wp/zb726/zb726.pdf>
- *Deploying Oracle9i RAC on eServer Cluster 1600 with GPFS* at <http://www.redbooks.ibm.com/abstracts/sg246954.html?Open>
- *An Oracle 9i RAC Implementation over GPFS* at <http://www.redbooks.ibm.com/abstracts/tips0263.html?Open>

**Q6.3: 6.3 Are there any considerations when utilizing the Remote Direct Memory Access (RDMA) offered by InfiniBand?**

**A6.3:** GPFS Multiplatform V3.2 for Linux supports Infiniband RDMA in the following configurations:

**Notes:**

1. Ensure you are at the latest firmware level for both your switch and adapter.
2. See the question *What are the current GPFS advisories?*

- SLES 10 or RHEL 5, x86\_64
- OFED Infiniband Stack VERBS API – GEN 2
  - OFED 1.2, OFED 1.2.5, OFED 1.3
  - OFED 1.1 – Voltaire Gridstack only
- Mellanox based adapters
  - RDMA over multiple HCAs/Ports/QPs
  - For multiple ports - GPFS balances load across ports
- Single IB subnet
  - QPs connected via GPFS RPC
- RDMA support for Mellanox memfree adapters requires GPFS V3.2.0.2, or later, to operate correctly



**Q6.4: What Linux configuration settings are required when NFS exporting a GPFS filesystem?**

**A6.4:** If you are running at SLES 9 SP 1, the kernel defines the `sysctl` variable `fs.nfs.use_underlying_lock_ops` that determines if the NFS `lockd` is to consult the file system when granting advisory byte-range locks. For distributed file systems like GPFS, this must be set to **true** (the default is **false**).

You can query the current setting by issuing the command:

```
sysctl fs.nfs.use_underlying_lock_ops
```

Alternatively, the record `fs.nfs.use_underlying_lock_ops = 1` may be added to `/etc/sysctl.conf`. This record must be applied after initially booting the node and after each reboot by issuing the command:

```
sysctl -p
```

As the `fs.nfs.use_underlying_lock_ops` variable is currently not available in SLES 9 SP2 or later, when NFS exporting a GPFS file system ensure your NFS server nodes are at the SP1 level (until such time the variable is made available in later service packs).

For additional considerations when NFS exporting your GPFS file system, see the:

- *GPFS: Administration Guide* chapter on *Managing GPFS access control lists and NFS export*
- *GPFS: Concepts, Planning, and Installation Guide* chapter *Planning for GPFS on File system creation considerations*.

**Q6.5: Sometimes GPFS appears to be handling a heavy I/O load, for no apparent reason. What could be causing this?**

**A6.5:** On some Linux distributions the system is configured by default to run the file system indexing utility **updatedb** through the **cron** daemon on a periodic basis (usually daily). This utility traverses the file hierarchy and generates a rather extensive amount of I/O load. For this reason, it is configured by default to skip certain file system types and nonessential file systems. However, the default configuration does not prevent **updatedb** from traversing GPFS file systems.

In a cluster this results in multiple instances of **updatedb** traversing the same GPFS file system simultaneously. This causes general file system activity and lock contention in proportion to the number of nodes in the cluster. On smaller clusters, this may result in a relatively short-lived spike of activity, while on larger clusters, depending on the overall system throughput capability, the period of heavy load may last longer. Usually the file system manager node will be the busiest, and GPFS would appear sluggish on all nodes. Re-configuring the system to either make **updatedb** skip all GPFS file systems or only index GPFS files on one node in the cluster is necessary to avoid this problem.

**Q6.6: What considerations are there when using IBM Tivoli Storage Manager with GPFS?**

**A6.6:** Consideration when using Tivoli Storage Manager (TSM) with GPFS include:

- When using TSM with GPFS, please verify the supported environments:
  - IBM Tivoli Storage Manager Requirements for IBM AIX Client at [www.ibm.com/support/docview.wss?rs=663&context=SSGSG7&uid=swg21052226&loc=en\\_US&cs=utf-8&lang=en](http://www.ibm.com/support/docview.wss?rs=663&context=SSGSG7&uid=swg21052226&loc=en_US&cs=utf-8&lang=en)
  - IBM Tivoli Storage Manager Linux x86 Client Requirements at [www.ibm.com/support/docview.wss?rs=663&context=SSGSG7&uid=swg21052223&loc=en\\_US&cs=utf-8&lang=en](http://www.ibm.com/support/docview.wss?rs=663&context=SSGSG7&uid=swg21052223&loc=en_US&cs=utf-8&lang=en)
  - To search TSM support information go to [www.ibm.com/software/sysmgmt/products/support/IBMTivoliStorageManager.html](http://www.ibm.com/software/sysmgmt/products/support/IBMTivoliStorageManager.html) and enter **GPFS** as the search term
- Quota limits are not enforced when files are recalled from the backup using TSM. This is because **dsmrecall** is invoked by the root user who has no allocation restrictions according to the UNIX semantics.



**Q6.7: How do I get OpenSSL to work on AIX and SLES8/ppc64?**

**A6.7:** To help enhance the security of mounts using Secure Sockets Layer (SSL) a working version of OpenSSL must be installed. This version must be compiled with support for the Secure Hash Algorithm (SHA).

- On AIX V5.2, V5.3 and V6.1, the supported versions of OpenSSL libraries are available at <https://www14.software.ibm.com/webapp/iwm/web/preLogin.do?source=aixbp>. The minimum supported versions of openssl.base are 0.9.8.411 and 0.9.8.601 (versions 0.9.8.40 and 0.9.8.41 are known not to work with GPFS). Additionally,

1. GPFS APAR IZ21177 is required.
2. GPFS configuration needs to be changed to point at the right set of libraries:

- On 64-bit kernel:

```
mmchconfig openssllibname="/usr/lib/libssl.a(libssl64.so.0.9.8)" -N AffectedNodes
```

- On 32-bit kernel:

```
mmchconfig openssllibname="/usr/lib/libssl.a(libssl.so.0.9.8)" -N AffectedNodes
```

On AIX V5.1, OpenSSL 0.9.7d-2, or later, as distributed by IBM in the AIX Toolbox for Linux Applications, is supported. To download OpenSSL from the AIX Toolbox for Linux Applications:

1. Go to <http://www-03.ibm.com/systems/p/os/aix/linux/toolbox/download.html>
2. Under **Sorted download**, click on **AIX Toolbox Cryptographic Content**.
3. Either register for an IBM ID or sign-in.
4. To view the license agreement, click on **View license**.
5. By clicking **I agree** you agree that you have had the opportunity to review the terms and conditions and that such terms and conditions govern this transaction.
6. Scroll down to **OpenSSL -- SSL Cryptographic Libraries**
7. Ensure you download 0.9.7d-2 or later

- For the supported versions of Linux:
  - For the Red Hat EL 3, Red Hat EL 4, Red Hat EL 5, SUSE Linux ES 9 and SUSE Linux ES 10 distributions, GPFS supports the version that comes with your distribution.
  - For the SUSE Linux ES 8 distribution on x86, this is currently OpenSSL 0.9.6, as included with your distribution.
  - For SUSE Linux ES 8 for PowerPC64 you must compile and install OpenSSL version 0.9.7f, according to these directions, before mounting any GPFS file systems that belong to other GPFS clusters (If you are running GPFS V2.3, ensure you are at least at the minimum service level. See the question *What is the current service information for GPFS?*):

1. Download the file **openssl-0.9.7f.tar.gz**, or later, from [www.openssl.org](http://www.openssl.org).

2. Unpack the file **openssl-0.9.7f.tar.gz**:

```
tar xzf openssl-0.9.7f.tar.gz
cd openssl-0.9.7f
```

3. Edit the script **Configure** , changing **gcc** to **/opt/cross/bin/powerpc64-linux-gcc**:

```
398c398
< "linux-ppc64", "gcc:-bpowerpc64-linux -DB_ENDIAN -DTERMIO -O3 -fomit-frame-pointer
-Wall:-D_REENTRANT:-ld1:SIXTY_FOUR_BIT_LONG RC4_CHAR RC4_CHUNK DES_RISC1
DES_UNROLL:asm/linux_ppc64.o:::::::::dlfcn:linux-shared:-fPIC:-bpowerpc64-linux:.so.
\$(SHLIB_MAJOR).\$(SHLIB_MINOR)",
---
> "linux-ppc64", "/opt/cross/bin/powerpc64-linux-gcc:-bpowerpc64-linux -DB_ENDIAN
-DTERMIO -O3 -fomit-frame-pointer -Wall:-D_REENTRANT:-ld1:SIXTY_FOUR_BIT_LONG
RC4_CHAR RC4_CHUNK DES_RISC1
DES_UNROLL:asm/linux_ppc64.o:::::::::dlfcn:linux-shared:-fPIC:-bpowerpc64-linux:.so.
\$(SHLIB_MAJOR).\$(SHLIB_MINOR)",
```

4. Run this script:

```
./Configure --prefix=/usr/local/ linux-ppc64
```

5. Build and install the OpenSSL library:

```
make  
make install
```

6. Update the library cache:

```
ldconfig
```

7. Configure all of the PowerPC64 nodes in the GPFS cluster, listed in the file **PPC64nodes**, to use the edited library:

```
mmchconfig openssllibname=/usr/local/lib/libssl.so.0.9.7 -N PPC64nodes
```

**Q6.8: What ciphers are supported for use by GPFS?**

**A6.8:** You can specify any of the RSA based ciphers that are supported by the OpenSSL version installed on the node. Refer to the `ciphers(1)` man page for a list of the valid cipher strings and their meaning. Use the `openssl ciphers` command to display the list of available ciphers:

```
openssl ciphers RSA
```

In addition, GPFS supports the keyword **AUTHONLY**. When **AUTHONLY** is specified in place of a cipher list, GPFS checks network connection authorization. However, data sent over the connection is not protected

**Note:** When different versions of OpenSSL are used within a cluster or in a multi-cluster setup, ensure that the ciphers are supported by all versions.

**Q6.9: When I allow other clusters to mount my file systems, is there a way to restrict access permissions for the root user?**

**A6.9:** Yes. Starting with GPFS 2.3.0.4, a *root squash* option is available when making a file system available for mounting by other clusters using the **mmauth** command. This option is similar in spirit to the NFS root squash option. When enabled, it causes GPFS to *squash* superuser authority on accesses to the affected file system on nodes in remote clusters.

This is accomplished by remapping the credentials: user id (UID) and group id (GID) of the root user, to a UID and GID specified by the system administrator on the home cluster, for example, the UID and GID of the user **nobody**. In effect, root squashing makes the root user on remote nodes access the file system as a non-privileged user.

Although enabling root squash is similar in spirit to setting up UID remapping (see [www.ibm.com/servers/eserver/clusters/whitepapers/uid\\_gpfs.html](http://www.ibm.com/servers/eserver/clusters/whitepapers/uid_gpfs.html)), there are two important differences:

1. While enabling UID remapping on remote nodes is an option available to the remote system administrator, root squashing need only be enabled on the local cluster, and it will be enforced on remote nodes.
2. While UID remapping requires having an external infrastructure for mapping between local names and globally unique names, no such infrastructure is necessary for enabling root squashing.

When both UID remapping and root squashing are enabled, root squashing overrides the normal UID remapping mechanism for the root user. See the **mmauth** command man page for further details.

---

## Service questions

### Q7.1: What support services are available for GPFS?

#### A7.1: Support services for GPFS include:

- GPFS support page at [www14.software.ibm.com/webapp/set2/sas/f/gpfs/home.html](http://www14.software.ibm.com/webapp/set2/sas/f/gpfs/home.html)
- IBM Global Services - Support Line for Linux

A 24x7 enterprise-level remote support for problem resolution and defect support for major distributions of the Linux operating system. Go to [www.ibm.com/services/us/index.wss/so/its/a1000030](http://www.ibm.com/services/us/index.wss/so/its/a1000030).

- IBM Systems and Technology Group Lab Services

IBM Systems and Technology Group (STG) Lab Services can help you optimize the utilization of your data center and system solutions.

STG Lab Services has the knowledge and deep skills to support you through the entire information technology race. Focused on the delivery of new technologies and niche offerings, STG Lab Services collaborates with IBM Global Services and IBM Business Partners to provide complementary services that will help lead through the turns and curves to keep your business running at top speed.

Go to <http://www.ibm.com/systems/services/labservices/>.

- Subscription service for pSeries, p5, and OpenPower

This service provides technical information for IT professionals who maintain pSeries, p5 and OpenPower servers. Subscribe at [www14.software.ibm.com/webapp/set2/subscriptions/pqvcmj](http://www14.software.ibm.com/webapp/set2/subscriptions/pqvcmj)

- GPFS software maintenance

GPFS defect resolution for current holders of IBM software maintenance contracts:

- In the United States contact us toll free at 1-800-IBM-SERV (1-800-426-7378)
- In other countries, contact your local IBM Service Center

- GPFS technical discussion forum

General questions can use the GPFS technical discussion forum at [www.ibm.com/developerworks/forums/dw\\_forum.jsp?forum=479&cat=13](http://www.ibm.com/developerworks/forums/dw_forum.jsp?forum=479&cat=13).

Contact [gpfs@us.ibm.com](mailto:gpfs@us.ibm.com) for all other services or consultation on what service is best for your situation.

### Q7.2: What is the current service information for GPFS?

#### A7.2: The current GPFS service information includes:

- For GPFS v3.1, if there are foreign characters in file or directory names, the **mmapplypolicy** command may fail

GPFS: 6027-902 Error parsing work file /tmp/tsmigrate.  
inodeslist.<pid>

The workaround for this problem is to:

- Upgrade to GPFS v3.2 where this problem no longer exists.
- If you need to stay on GPFS v3.1:
  1. Install GNU sort contained in the GNU **coreutils** from the AIX Toolbox for Linux Applications at <http://www-03.ibm.com/systems/p/os/aix/linux/toolbox/download.html>

2. Set the environment variables

```
MM_SORT_CMD = "LC_ALL=C  
/local-or-opts-whenever-gnu-binaries-happen-to-be/sort -z"  
MM_SORT_EOR = "" #empty string
```

- For GPFS V3.2 use with AIX V6.1:

- GPFS is supported in a Ethernet/10-Gigabit Ethernet environment, see the question *What interconnects are supported for GPFS daemon-to-daemon communication in my GPFS cluster?*

- The versions of OpenSSL shipped as part of the AIX Expansion Pack, 0.9.8.4 and 0.9.8.41, *ARE NOT* compatible with GPFS due to the way the OpenSSL libraries are built. To obtain the level of OpenSSL which will work with GPFS, see the question *How do I get OpenSSL to work on AIX and SLES8/ppc64?*
- Role Based Access Control (RBAC) is not supported by GPFS and is disabled by default.
- Workload Partitions (WPAr) or storage protection keys are not exploited by GPFS.
- If you get errors on RHEL5 when trying to run GPFS self-extractor archive from the installation media, please run **export \_POSIX2\_VERSION=199209** first.
- When installing or migrating GPFS, the minimum levels of service you must have applied are:
  - GPFS V3.2 you must apply APAR IY99639 (GPFS V3.2.0-1)
  - GPFS V3.1 you must apply APAR IY82778
  - GPFS V2.3 you must apply APAR IY63969

If you do not apply these levels of service and you attempt to start GPFS, you will receive an error message similar to:

```
mmstartup: Required service not applied. Install GPFS 3.2.0.1 or later
mmstartup: Command failed Examine previous error messages to determine cause
```

### Upgrading GPFS to a new major release on Linux:

When migrating to a new major release of GPFS (for example, GPFS 3.1 to GPFS 3.2), the supported migration path is to install the GPFS base images for the new release, then apply any required service updates. GPFS will not work correctly if you use **rpm -U** command to upgrade directly to a service level of a new major release without installing the base images first. If this should happen you must uninstall and then reinstall the **gpfs.base** package.

**Note:** Upgrading to the GPFS 3.2.1.0 level from a pre-3.2 level of GPFS does not work correctly, and the same workaround is required.

- GPFS V3.1 maintenance levels 10 (GPFS-3.1.0.10) thru 12 (GPFS-3.1.0.12) do not coexist with other maintenance levels

All nodes in the cluster must conform to one of these maintenance level compatibility restrictions:

- all nodes must be at maintenance levels 1-9 or 13 and later (GPFS-3.1.0.1 thru GPFS-3.1.0.9 or GPFS-3.1.0.13 and later)
- all nodes must be at maintenance levels 10-12 (GPFS-3.1.0.10 - GPFS-3.1.0.12)
- Required service for support of SLES 10 includes:
  1. If running GPFS V3.1, service update 3.1.0-8 available at <https://www14.software.ibm.com/webapp/set2/sas/f/gpfs/download/home.html>
  2. The GPFS required level of Korn shell for SLES 10 support is version ksh-93r-12.16 and can be obtained using one of these architecture-specific links:
    - x86** at <https://you.novell.com/update/i386/update/SUSE-SLES/10/PTF/43ed798d45b1ce66790327fe89fb3ca6/20061201>
    - POWER** at <https://you.novell.com/update/ppc/update/SUSE-SLES/10/PTF/43ed798d45b1ce66790327fe89fb3ca6/20061201>
    - x86\_64** at [https://you.novell.com/update/x86\\_64/update/SUSE-SLES/10/PTF/43ed798d45b1ce66790327fe89fb3ca6/20061201](https://you.novell.com/update/x86_64/update/SUSE-SLES/10/PTF/43ed798d45b1ce66790327fe89fb3ca6/20061201)
  3. For SLES 10 on POWER:
    - a. The gpfs.base 3.1.0-0 rpm must be installed using the rpm **--nopro** flag BEFORE any updates can be applied.
    - b. **/etc/init.d/running-kernel** shipped prior to the availability of the SLES 10 SP1 kernel source rpm contains a bug that results in the wrong set of files being copied to the kernel source tree. Until SP1 is generally available, the following change should also address the problem:

```

--- running-kernel.orig 2006-10-06 14:54:36.000000000 -0500
+++ /etc/init.d/running-kernel 2006-10-06 14:59:58.000000000 -0500
@@ -53,6 +53,7 @@
 arm*|sa110)    arch=arm ;;
 s390x)    arch=s390 ;;
 parisc64)    arch=parisc ;;
+ ppc64)    arch=powerpc ;;
 esac
 # FIXME: How to handle uml?

```

- GPFS V2.3 file systems may mount file systems that belong to other GPFS V2.3 clusters. This function is available with APARs IY64709 and IY66584 (see the question, *How do I download fixes for GPFS?*):  
Please see the updated HTML files at [publib.boulder.ibm.com/infocenter/clresctr/index.jsp?topic=/com.ibm.cluster.gpfs.doc/gpfsbooks.html](http://publib.boulder.ibm.com/infocenter/clresctr/index.jsp?topic=/com.ibm.cluster.gpfs.doc/gpfsbooks.html) or the documentation errata file at [publib.boulder.ibm.com/infocenter/clresctr/topic/com.ibm.cluster.gpfs.doc/gpfs23\\_doc\\_updates/docerrata.html](http://publib.boulder.ibm.com/infocenter/clresctr/topic/com.ibm.cluster.gpfs.doc/gpfs23_doc_updates/docerrata.html) for:
  - A replacement of chapter 3 *Accessing GPFS file systems from other GPFS clusters* from the *GPFS: Administration and Programming Reference*.
  - Updates to the **mmauth** and **mmremoteccluster** commands.
- Included with GPFS 2.3.0-6 is a performance feature for very large files. After this maintenance level is applied, issue the **mmchfs device -V** command to enable this feature for existing file systems.

Prior to issuing the **mmchfs -V** command on existing file systems or creating new files systems, it is suggested that all nodes in the cluster are at this maintenance level. If for any reason there are nodes in the cluster which are not at the 2.3.0-6 maintenance level, they should not be designated as nodes available to become the File System Manager. Should a node at a level lower than 2.3.0-6 become the File System Manager, conflicting messages will be sent to nodes at the 2.3.0-6 level causing them to go down.

- If you applied GPFS fix level 2.3.0-4 to your Linux system, please move to 2.3.0-5. See the question *How do I download fixes for GPFS?*
- When running GPFS on either a p5-590 or a p5-595:
  - The minimum GFW (system firmware) level required is SF222\_081 (GA3 SP2), or later.  
For the latest firmware versions, see the IBM Technical Support at [www14.software.ibm.com/webapp/set2/firmware/gjsn](http://www14.software.ibm.com/webapp/set2/firmware/gjsn)
  - The supported Linux distribution is SUSE Linux ES 9.
  - Scaling is limited to 16 total processors.
- IBM testing has revealed that some customers using the Gigabit Ethernet PCI-X adapters with the **jumbo frames** option enabled may be exposed to a potential data error. While receiving packet data, the Gigabit Ethernet PCI-X adapter may generate an erroneous DMA address when crossing a 64 KB boundary, causing a portion of the current packet and the previously received packet to be corrupted.

These Gigabit Ethernet PCI-X adapters and integrated Gigabit Ethernet PCI-X controllers could potentially experience this issue:

- Type 5700, Gigabit Ethernet-SX PCI-X adapter (Feature Code 5700)
- Type 5701, 10/100/1000 Base-TX Ethernet PCI-X Adapter (Feature code 5701)
- Type 5706, Dual Port 10/100/1000 Base-TX Ethernet PCI-X Adapter (Feature code 5706)
- Type 5707, Dual Port Gigabit Ethernet-SX PCI-X Adapter (Feature code 5707)
- Integrated 10/100/1000 Base-TX Ethernet PCI-X controller on machine type 7029-6C3 and 6E3 (p615)
- Integrated Dual Port 10/100/1000 Base-TX Ethernet PCI-X controller on machine type 9111-520 (p520)
- Integrated Dual Port 10/100/1000 Base-TX Ethernet PCI-X controller on machine type 9113-550 (p550)
- Integrated Dual Port 10/100/1000 Base-TX Ethernet PCI-X controller on machine type 9117-570 (p570)

This problem is fixed with:



- For AIX 5L 5.2, APAR IY64531
- For AIX 5L 5.3, APAR IY64393
- IBM testing has revealed that some customers with the General Parallel File System who install AIX 5L Version 5.2 with the 5200-04 Recommended Maintenance package (**bos.mp64** at the 5.2.0.40 or 5.2.0.41 levels) and execute programs which reside in GPFS storage may experience a system wide hang due to a change in the AIX 5L loader. This hang is characterized by an inability to login to the system and an inability to complete some GPFS operations on other nodes. This problem is fixed with the AIX 5L APAR IY60609. It is suggested that all customers installing the **bos.mp64** fileset at the 5.2.0.40 or 5.2.0.41 level, who run GPFS, immediately install this APAR.
- Service bulletins for pSeries, p5, and OpenPower servers at [www14.software.ibm.com/webapp/set2/subscriptions/pqvcmj](http://www14.software.ibm.com/webapp/set2/subscriptions/pqvcmj)
  1. Sign in with your IBM ID.
  2. Under the **Bulletins** tab:
    - For the **Select a heading** option, choose **Cluster on POWER**.
    - For the **Select a topic** option, choose **General Parallel File System**.
    - For the **Select a month** option, select a particular month or choose to **All months**.

**Q7.3: How do I download fixes for GPFS?**

**A7.3:** To download fixes for GPFS, go to <https://www14.software.ibm.com/webapp/set2/sas/f/gpfs/home.html>

**Q7.4: What are the current GPFS advisories?**

**A7.4:** The current GPFS advisories are:

- Currently with GPFS Multiplatform for Linux V3.2.1-4 and lower, with Infiniband RDMA enabled, an issue exists which under certain conditions may cause data corruption. This is fixed in GPFS 3.2.1-6. Please apply 3.2.1-6 or turn RDMA off.
- **GPFS 2.3.0.x not compatible with AIX 5.3 TL6**

Currently GPFS 2.3.0.x on AIX TL6 has a known private heap memory leak.

**USER'S AFFECTED:** All customers using GPFS 2.3 and AIX 5.3

**DESCRIPTION:** GPFS 2.3.0.0 through 2.3.0.23 do not work with AIX 5.3 TL6 due to the changes that AIX made in the threading library. GPFS 2.3 PTF 24 and up do have the necessary code changes to work with TL6 but they produce a private heap memory leak due to AIX APAR IZ04791. The AIX fix for this problem is scheduled for AIX TL6 SP4. A workaround that can be used until obtaining AIX TL6 SP4 is to change the GPFS configuration to not use the sigwait library call (**mmchconfig asyncSocketNotify=no**). Therefore, until the issue is resolved please be advised not to use GPFS 2.3.0.0 through 2.3.0.23 and AIX 5.3 TL6 in a production environment. AIX 5.3 TL1 through 5 are known to work with all GPFS 2.3 PTFs.

**EFIX AVAILABLE:** There are no fixes at this time. Once one is available notice will be given. Please see <https://www14.software.ibm.com/webapp/set2/sas/f/gpfs/download/aix.html>

- In certain GPFS 2.3 and 3.1 PTF levels there is a subtle GPFS issue in truncate, where if multiple nodes are accessing the same file against which a truncate is issued on one of the nodes, a time window existed during which incorrect size information could be communicated to some nodes, which may cause GPFS to mishandle the last fragment of the file. This could lead to various failed internal consistency checks, manifested by the GPFS daemon shutting down abnormally.

The affected GPFS PTF levels are:

- GPFS 3.1.0-6
- GPFS 3.1.0-5
- GPFS 2.3.0-17
- GPFS 2.3.0-16
- GPFS 2.3.0-15

**Recommended action:**

- For customers running GPFS 3.1.0.x PTF 7 contains a fix and is available at [www14.software.ibm.com/webapp/set2/sas/f/gpfs/download/home.html](http://www14.software.ibm.com/webapp/set2/sas/f/gpfs/download/home.html)
- For all other versions, please contact support.

- Customers running IBM Virtual Shared Disk V4.1 using a communications adapter other than the IBM eServer pSeries High Performance Switch, who have configured IBM Virtual Shared Disk with an IP packet size greater than the Max Transfer Unit (MTU) of the network, may experience packet corruption.

IP must fragment packets that are greater than the MTU size of the network. On faster interconnects such as Gigabit Ethernet, the IP fragmentation buffer can be overrun and end up incorrectly assembling the fragments. This is an inherent limitation of the IP protocol, which can occur when the number of packets transferred exceeds the counter size, which then rolls over, potentially resulting in a duplicate packet number.

If a duplicate packet number occurs, and the checksum matches that of the expected packet, corruption of the IBM Virtual Shared Disk packets can result in GPFS file system corruption. IBM Virtual Shared Disk will attempt to validate the incoming packets and discard misformed packets, but it can not identify them every time (since checksums for different data patterns may be the same).

The level of IBM Virtual Shared Disk affected (shipped in AIX 5.2.x and later releases) has been available since October 2003, and the problem has only been confirmed as having occurred in an internal IBM test environment.

IP fragmentation can be prevented by configuring the IBM Virtual Shared Disk IP packet size less than or equal to the MTU size of the network. This will move the fragmentation into the IBM Virtual Shared Disk layer, which can correctly process the fragmentation.

The current IBM Virtual Shared Disk infrastructure allows for 160 packets per request which will limit the maximum buddy buffer size that can be used. For example:

- for an MTU of 1500, you need to set the IBM Virtual Shared Disk IP packet size to 1024 effectively limiting the maximum buddy buffer size to 128 KB.
- for an MTU of 9000, you need to set the IBM Virtual Shared Disk IP packet size to 8192 effectively limiting the maximum buddy buffer size to 1 MB.

You can check the IBM Virtual Shared Disk IP packet size with these two commands:

**vsdata1st -n**

Shows you the value that will take affect at the next reboot.

**statvsd**

Show you the current value that the IBM Virtual Shared Disk device driver is using.

Here is an example of how to set the IP packet size when using jumbo Ethernet frames (MTU = 9000):

```
updatevsdnode -n ALL -M 8192
dsh -a ct1vsd -M 8192
```

For more information see the *RSCT for AIX 5L Managing Shared Disks* manual at [publib.boulder.ibm.com/infocenter/clresctr/index.jsp?topic=/com.ibm.cluster.rsct.doc/rsctbooks.html](http://publib.boulder.ibm.com/infocenter/clresctr/index.jsp?topic=/com.ibm.cluster.rsct.doc/rsctbooks.html) and search on the commands **vsdnode**, **updatevsdnode**, and **ct1vsd**.

APAR IY66940 will completely prevent IP fragmentation and will enforce the IBM Virtual Shared Disk IP packet size being less than the MTU size. This will also remove the restrictions relating to the maximum IBM Virtual Shared Disk buddy buffer size.

Anyone who cannot take the preventive action, for whatever reason, or is unsure whether their environment may be affected, should contact IBM service to discuss their situation:

- In the United States contact us toll free at 1-800-IBM-SERV (1-800-426-7378)
- In other countries, contact your local IBM Service Center

**Q7.5: What Linux kernel patches are provided for clustered file systems such as GPFS?**

**A7.5:** The Linux kernel patches provided for clustered file systems are expected to correct problems that may be encountered when using GPFS with the Linux operating system. The supplied patches are currently being submitted to the Linux development community but may not be available in particular kernels. It is therefore suggested that they be appropriately applied based on your kernel version and distribution.

A listing of the latest patches, along with a more complete description of these patches, can be found at the General Parallel File System project on SourceForge®.net at [sourceforge.net/tracker/?atid=719124&group\\_id=130828&func=browse](http://sourceforge.net/tracker/?atid=719124&group_id=130828&func=browse):

1. Click on the **Summary** description for the desired patch.
2. Scroll down to the **Summary** section on the patch page for a description of and the status of the patch.
3. To download a patch:
  - a. Scroll down to the **Attached Files** section.
  - b. Click on the **Download** link for your distribution and kernel level.

**site.mcr consideration:** Patches listing a **site.mcr** define have additional steps to perform:

1. Apply the patch to the Linux kernel, recompile, and install this kernel.
2. In **site.mcr** either **#define** the option or uncomment the option if already present. Consult **/usr/lpp/mmfs/src/README** for more information.
3. Recompile and reinstall the GPFS portability layer.

**Q7.6: What Windows hotfix updates are required for GPFS?**

**A7.6:** The current Windows hotfix updates required for GPFS consist of :

- KB article 956548 at <http://support.microsoft.com/kb/956548/en-us> ;only the hotfix for Windows Server 2003 (Fix243497) is required.
- KB article 950098 at <http://support.microsoft.com/kb/950098/en-us>

**Q7.7: Where can I find licensing and ordering information for GPFS?**

**A7.7:** The Cluster Software Ordering Guide provides the following information:

- Licensing information  
Licenses can also be viewed at <http://www.ibm.com/software/sla/sladb.nsf>
- Ordering information
- Software Maintenance Agreement information
- Product End of Market/Service dates  
Software support lifecycle information can also be viewed at [http://www-306.ibm.com/software/support/lifecycle/index\\_a\\_z.html](http://www-306.ibm.com/software/support/lifecycle/index_a_z.html)
- Hardware and Software requirements

To view the Guide please go to [http://www.ibm.com/systems/clusters/software/reports/order\\_guide.html](http://www.ibm.com/systems/clusters/software/reports/order_guide.html)



## Notices

This information was developed for products and services offered in the U.S.A.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. 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. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

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

IBM Director of Licensing  
IBM Corporation  
North Castle Drive  
Armonk, NY 10594-1785  
USA

For license inquiries regarding double-byte (DBCS) information, contact the IBM Intellectual Property Department in your country or send inquiries, in writing, to:

IBM World Trade Asia Corporation  
Licensing  
2-31 Roppongi 3-chome, Minato-ku  
Tokyo 106-0032, Japan

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law:

INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not

allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.

IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Licensees of this program 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  
Intellectual Property Law  
2455 South Road, P386  
Poughkeepsie, NY 12601-5400  
USA

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

The licensed program described in this document and all licensed material available for it are provided by IBM under terms of the IBM Customer Agreement, IBM International Program License Agreement or any equivalent agreement between us.

Any performance data contained herein was determined in a controlled environment. Therefore, the results obtained in other operating environments may vary significantly. Some

measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurements may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

#### COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrates programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs.

If you are viewing this information softcopy, the photographs and color illustrations may not appear.

---

## Trademarks

IBM, the IBM logo, and [ibm.com](http://ibm.com) are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. If these and other IBM trademarked terms are marked on their first

occurrence in this information with a trademark symbol (® or ™), these symbols indicate U.S. registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the Web at *Copyright and trademark information* at [www.ibm.com/legal/copytrade.shtml](http://www.ibm.com/legal/copytrade.shtml)

Cell Broadband Engine is a trademark of Sony Computer Entertainment, Inc. in the United States, other countries, or both and is used under license therefrom

Intel, Intel logo, Intel Inside, Intel Inside logo, Intel Centrino, Intel Centrino logo, Celeron, Intel Xeon, Intel SpeedStep, Itanium, and Pentium are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

Java and all Java-based trademarks and logos are trademarks or registered trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.

Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.

Red Hat, the Red Hat "Shadow Man" logo, and all Red Hat-based trademarks and logos are trademarks or registered trademarks of Red Hat, Inc., in the United States and other countries.

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

Microsoft, Windows, Windows NT, and the Windows logo are registered trademarks of Microsoft Corporation in the United States, other countries, or both.

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

---

December 2008

© Copyright International Business Machines Corporation 2004,2008.

US Government Users Restricted Rights – Use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM Corp.