

Enterprise File Management with GPFS 3.2



Highlights

- **Scalable high-performance shared disk file system for AIX® and Linux® systems.**
- **Capable of supporting multi-petabytes of storage and thousands of disks within a single file system**
- **High reliability/availability through redundant paths and automatic recovery from node and disk failures**
- **Information lifecycle management (ILM) tools simplify data management and enhance administrative control**
- **Powers many of the world's largest supercomputers**

The IBM General Parallel File System™ (GPFS™) is a high-performance shared-disk file management solution that provides fast, reliable access to a common set of file data from two computers to hundreds of systems. GPFS integrates into your environment by bringing together mixed server and storage components to provide a common view to enterprise file data. GPFS provides online storage management, scalable access and integrated information lifecycle tools capable of managing petabytes of data and billions of files.

The proven GPFS file management infrastructure provides the foundation for optimizing the use of your computing resources.

Optimizes storage utilization by centralizing management

GPFS allows multiple systems and applications to share common pools of storage. This allows dynamic allocation of storage space and the ability to transparently administer the infrastructure without disrupting applications.

GPFS supports multi-vendor storage and server hardware, enabling striping, replication, and snapshots across heterogeneous systems.

Provides a flexible scalable alternative to a growing number of NAS appliances

Reduce management overhead by using GPFS to share a common set of files concurrently across multiple systems. Even a subset of GPFS systems can serve the data over NFS while others concurrently run applications that require high performance direct SAN access to the files. GPFS includes tools for highly scalable NFS serving and supports SAMBA file serving.

Highly available grid computing infrastructure

GPFS is an essential component to implementing a high throughput grid environment. High-availability features including data replication and automatic failure handling provide a solid base on which to build an analytics grid or digital media management system.

Scalable information lifecycle tools to manage growing data volumes

You can use storage pools and the highly scalable policy engine to introduce structure to your unstructured data. GPFS can manage *billions* of files using the integrated SQL based policies for file placement and migration over time. The scalable policy engine can be run on one system or concurrently on all systems depending on processing needs

The GPFS difference

GPFS takes enterprise file management beyond a single computer by providing scalable access from multiple systems to a single file system. GPFS interacts with applications in the same manner as a local file system but is designed to deliver much higher performance, scalability and failure recovery by allowing access to the data from multiple systems directly and in parallel.

High-performance data access is achieved by automatically spreading the data across multiple storage devices and the ability to read and write data in parallel. In addition, for high-bandwidth environments like digital media, GPFS can read or write large blocks of data in a single operation minimizing the overhead of I/O operations.

For optimal reliability, GPFS can be configured to eliminate single points-of-failure. Availability is further improved by automatic logging and data replication. You can mirror data within a site or across multiple locations. The file system can be configured to remain available automatically in the event of a disk or server failure.

High performance and scalability

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 analytics, seismic data processing, scientific research and scalable file serving.

GPFS has been proven in applications with aggregate data and bandwidth requirements that exceed the capacity of typical shared file systems.

GPFS allows multiple applications or users to share access to a single file simultaneously while ensuring file data integrity. For example, multiple animators or editors can work on different parts of a single video file simultaneously from multiple work stations. This reduces storage and management overhead typically required

to maintain multiple copies of the source file. To ensure data consistency during concurrent access, GPFS uses fine grained locking based on a sophisticated scalable token (lock) management system. This prevents multiple applications or users from updating the same portion of a file at the same time.

In addition to nodes that are directly attached to the storage, a single GPFS file system can be accessed by thousands of nodes using a LAN connection like Ethernet or InfiniBand™. GPFS can outperform network based sharing technologies like Network File System (NFS) and even local file systems like EXT3 or JFS.

Flexible high-performance configurations

GPFS has a very flexible cluster architecture providing many options to develop a solution including: direct attached, network block I/O, a combination of the two and multi-site operations.

Applications needing the highest per node throughput often deploy the *Direct-Attached* model, where disks are physically attached to all nodes in the cluster which access the file system. This direct attached model requires an operating system device interface to the storage for example using a Fibre Channel SAN or InfiniBand connection.

For applications running on a large number of systems, for applications that need to make the most of available LAN bandwidth or require guaranteed data consistency, the network block I/O (also called network shared disk or NSD) model can be used. Network block I/O is a software layer that forwards block I/O requests from an NSD client application node to the LAN, which then passes the request to an NSD storage node to perform the disk I/O and pass data back to the client. GPFS makes the LAN-based I/O operation transparent to the application. Using a Network Block I/O configuration can be more cost-effective than a full-access SAN and can be used to tie together systems across a WAN.

How the systems in the cluster are attached to the storage is an implementation question and can be tailored to fit the needs of multiple applications in a single cluster. For example, you can use a combination of the direct-attached and network block I/O models. This allows you to run applications with the greatest throughput needs on systems with direct attachment to the storage and integrate that with other systems using LAN connectivity. This flexibility allows you to centralize data management and use available storage most effectively across multiple applications.

GPFS supports petabytes of storage and billions of files.

Scalable NFS file service

GPFS provides a Clustered Network File Service (NFS) feature. Clustered NFS gives you the ability, on a Linux cluster, to designate one or all of the nodes to service a common data set using the NFS protocol. GPFS provides the tools to administer and monitor a highly available, parallel NFS server cluster including IP address failover and load balancing.

Availability and recoverability

A GPFS file system can survive many system and infrastructure failures. It is designed to transparently failover token (lock) operations and other GPFS cluster services, which can be distributed throughout the entire cluster eliminating dedicated token/metadata points of failure. GPFS can be configured to automatically recover from node, storage and other infrastructure failures.

GPFS Feature Summary	
Feature	Benefit
Highly Scalable File System Designed for a cluster environment, GPFS can support thousands of nodes and thousands of disks comprising petabytes.	GPFS enables the scale out expansion of the file system using a building block approach with commodity components as service demands and data volumes increase.
Data Protection Both user data and metadata (file system transaction data) can be replicated (mirrored) in the GPFS installation.	Ensures reliable access to data and metadata.
Parallel File System (Data Striping) Divides individual files into multiple blocks and stores these blocks across multiple disks or disk arrays in parallel.	Higher performance and reliability by eliminating the bottlenecks that typically arise when an entire file resides on a single node.
Storage Pools User-defined grouping of storage resources.	Storage can be grouped or tiered based on factors such as location, performance and reliability.
External Storage Pools Policy driven interactions with external applications for near-line data storage.	External pools leverage the high performance policy engine to expand your data storage options. For example, an external pool data can be stored on a tape library.
Policy-Driven Automation Automatically place file data in a specific storage pool based on attributes of the file. Once files exist file data can be moved between pools or deleted automatically. File movement retains logical path, so is transparent to users and applications.	Simplifies data management and allows matching the cost of storage to the value/importance of the data. For example, critical data can be stored on the fastest storage.
Concurrent File Access and Block-Level Locking Multiple applications on different systems can access a single file simultaneously; sophisticated lock management prevents collisions or conflicts ensuring data consistency.	Enables multiple users, programs or nodes to access a single file simultaneously – accelerates processing for parallel applications that share data and eliminates the need for multiple copies of data, reducing data storage and management costs.
High Availability Automatically recovers from events that would normally interrupt data availability.	Applications and users can continue without interruption in the event of a node, path or storage system failure.
Clustered NFS Tools to deploy and manage a clustered NFS file serving solution.	Provides an integrated solution enabling scalable NFS file serving from a cluster of Linux servers.
Large Blocksize Options Data blocks of 16K, to 4MB can be used to optimize throughput for different applications.	Large blocks allow more data to be written or read in a single operation. The wide range of block sizes enables more efficient use of disk bandwidth and available storage space.

GPFS provides this functionality by supporting data replication to increase availability in the event of a storage media failure, multiple data paths in the event of a communications or server failure, and file system activity logging, enabling consistent fast recovery after system failures. Additionally, GPFS uniquely provides a 'Multi-Cluster' function, enabling multi-site operations for disaster recovery solutions or cross-site joint operations.

Central administration

GPFS simplifies administration by providing control of the entire system from any node in the cluster. Administration functions are based on existing UNIX® and Linux administrative file system commands.

Heterogeneous servers and storage systems can be added to and removed from a GPFS cluster while the file system remains online. When storage is added or removed the data can be dynamically rebalanced to maintain optimal performance.

GPFS supports file system snapshots, providing a space efficient point in time image of a file system at a specified time. This provides an online backup to protect from user errors or a frozen view from which to take a backup.

GPFS supports mixed clusters of AIX and Linux systems in a single cluster. In addition multiple versions of GPFS can exist within a single cluster supporting rolling upgrades and providing operational flexibility.

ILM tools

GPFS provides information lifecycle management (ILM) tools that allow you to simplify data management by enabling more control over the placement of data in a file system. These tools include storage pooling and a scalable rule based policy engine.

Storage pools enable you to transparently manage multiple tiers of the storage based on factors such as performance or reliability. For example, you can use storage pools to

transparently provide the appropriate type of storage to multiple applications or different portions of a single application within the same directory tree. For example you can automatically provide low latency disks for index operations and high capacity disks for data operations of a relational database. This can be done even if all of the database files are created in the same directory. In addition storage pools can simplify storage management by allowing you to transparently re-task older storage systems moving them from a tier 1 role to tier 2.

As files are created, policy-driven automation can place file data in the appropriate physical storage pool based on file attributes, regardless where the file is placed in the directory structure. Once files exist, the policy engine can move files between pools based on powerful SQL-based policies, or delete them after a specified time period.

The GPFS policy engine leverages the parallel metadata infrastructure to provide highly scalable policy processing. This allows you to actively manage *billions* of files across multiple storage tiers.

Summary

GPFS provides world-class performance, scalability and availability of your file data. It complies with UNIX file system standards and interfaces for both end-user applications and administrative extensions for seamless management and tuning. GPFS scales with the needs of your business providing application infrastructure and NFS file serving capabilities. GPFS is designed to optimize the use of storage, support scale-out applications and provide a high availability platform for data intensive applications..

More information

To learn more about IBM GPFS, contact your IBM representative, IBM Business Partner or visit the GPFS Web site:

ibm.com/systems/clusters/software/gpfs.html



© IBM Corporation 2007

IBM Corporation
Marketing Communications
Systems Group
Route 100
Somers, New York 10589

Produced in the United States of America
August 2007
All Rights Reserved

This document was developed for products and/or services offered in the United States. IBM may not offer the products, features, or services discussed in this document in other countries. The information may be subject to change without notice. Consult your local IBM business contact for information on the products, features and services available in your area.

All offers subject to availability. IBM reserves the right to alter product offerings and specifications at any time without notice. IBM is not responsible for photographic or typographic errors.

All statements regarding IBM future directions and intent are subject to change or withdrawal without notice and represent goals and objectives only.

IBM, the IBM logo, AIX, General parallel File System, GPFS are trademarks or registered trademarks of International Business Machines Corporation in the United States or other countries or both. A full list of U.S. trademarks owned by IBM may be found at <http://www.ibm.com/legal/copytrade.shtml>.

UNIX is a registered trademark of The Open Group in the United States, other countries, or both.

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

InfiniBand is a trademark and/or service mark of the InfiniBand Trade Association.

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

Copying or downloading the images contained in this document is expressly prohibited without the written consent of IBM.

Information about non-IBM products was obtained from suppliers of those products. Questions concerning those products should be directed to those suppliers.

The IBM home page on the Internet can be found at <http://www.ibm.com>.

CLF03001-USEN-00