Open Source XFS™ for Linux®

Providing the World's Most Scalable Journaling-Filesystem Technology to the Open Source™ Community

Features

- Sub-second filesystem recovery after crashes or power failures (never wait for long 'fscks again)
- 64-bit scalability: millions of terabytes, millions of files, and a million files per directory (no more 2GB limits)
- High reliability and performance from journaling and other advanced algorithms

Open Sourcing Key Technologies

SGI® contributed the XFS filesystem to the Open Source Linux community using the copyleft GPL license in March 2000. XFS is one of SGI's core competencies in high-performance computing and is the most scalable, highest performance journaling filesystem available today.

SGI's Open Source Intentions

SGI has helped to scale Linux to run enterprise-class applications by delivering technologies from its core competencies in visualization and high-performance computing; XFS is one of several of these key Open Source contributions. Linux has matured into a world-class, robust, reliable, feature-rich, widely used, and enterprise-ready operating system. Browse the XFS source code, white papers, and design documents at: http://oss.sgi.com/projects/xfs.

XFS: A World-Class Filesystem

XFS combines advanced journaling technology with full 64-bit addressing and scalable structures and algorithms. This combination delivers the most scalable and highest performance filesystem in the world.

Journaling: Quick Recovery

The XFS journaling technology allows it to restart in less than a second after an unexpected interruption, regardless of the number of files it is managing. Traditional filesystems must do special filesystem checks after an interruption, which can take many hours to complete. The XFS journaling avoids these lengthy filesystem checks.

Efficiency: Quick Responses

XFS uses efficient structures for fast searches and rapid space allocation. XFS continues to deliver rapid response times, even for directories with tens of thousands of entries. The performance of most UNIX filesystems significantly degrades as the number of entries per directory grows, not so with XFS. XFS provides extremely sophisticated space allocation techniques.

Optimization: Space Allocation

XFS delivers extremely sophisticated space management through the use of variable sized extents, rather than the simple single-block-at-a-time mechanism of many other filesystems. XFS was the first to implement delayed space allocation for buffered writes, supports direct I/O, has an optional real-time allocator, and is able to align allocations based on the geometry of the underlying storage device. The XFS allocator performs admirably in the presence of multiple, parallel writers, and is renowned for its resistance to space fragmentation under such conditions.

21st Century Scalability

XFS is a full 64-bit filesystem, capable of handling files as large as a million terabytes.

$2^{63}-1 = 9 \times 10^{18} = 9$ exabytes

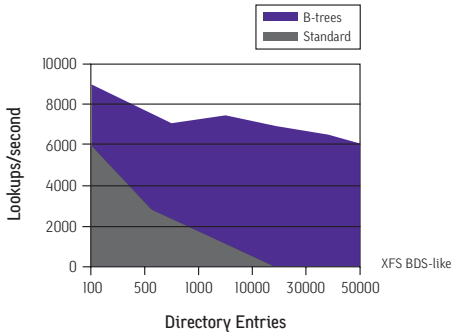
A million terabytes is about a million times larger than most large filesystems in use today. This may seem to be an extremely large address space, but it is actually needed to plan for the exponential disk-density improvements observed in the disk industry in recent years. XFS filesystems with hundreds and thousands of terabytes have been in production use years ahead of any competitors.

Open Source XFS™ for Linux®

With explosive storage growth trends continuing, SGI is actively investing in research and development to expand the capabilities of XFS even further. It is SGI's goal to continue to meet customer needs and remain ahead of the competition through the next decade.

XFS Bandwidth

XFS delivers near-raw I/O performance. This was demonstrated on the largest disk configuration SGI was able to assemble: over 9GB per second (read and write) on 256 disks from multiple parallel processes on an SGI Altix® system.



Scalability

- Scalable file sizes (9 million terabytes)
- Scalable filesystems (18 million terabytes)
- Scalable algorithms for high performance, even with huge filesystems
  - Large numbers of files
  - Large files (including sparse files)
  - Large directories
  - Advanced algorithms for fast performance on huge filesystems
  - Rapid recovery

High Performance

- Extremely fast transaction rates
- Extremely high bandwidths
- Extremely fast directory searches
- Extremely fast space allocation

Advanced Features

- Hierarchical storage (HSMs, DMF)
- Offline storage virtually online

Compatibility

- Backup with popular commercial packages such as Legato NetWorker® for IRIX and VERITAS NetBackup™ or with dump/restore, bru, cpio, or tar
- Support for DMIG-DMAPI compliant HSM packages including SGI DMF.

NFS™ Compatibility

- With NFS version 3 and 4, 64-bit filesystems can be exported to other systems that support the NFS V3 protocol; systems that use NFS V2 protocol may access XFS filesystems within the 32-bit limit imposed by the protocol

Windows NT® Compatibility

- SGI uses the Open Source Samba server to export XFS filesystems to Windows® and Windows NT systems
- Samba speaks the SMB (Server Message Block) and CIFS (Common Internet File System) protocols
- Journaled 64-bit filesystem with guaranteed filesystem consistency
- Algorithms and internal structures that scale to support millions of files per filesystem and millions of files per directory
- XFS uses efficient structures (b-trees) for fast searches and rapid space allocation; XFS continues to deliver rapid response times, even for directories with tens of thousands of entries; the performance of most UNIX filesystems significantly degrades as the number of entries per directory grows not so with XFS

Filesystem Block Size

- 512 bytes to 64KB physical for normal data and up to 1MB for real-time data. Filesystem extents (contiguous data) are configurable up to 4GB in size. Device sector sizes from 512 bytes to 32KB are supported.

Performance

- Throughput in excess of 10GB per second has been demonstrated on a single filesystem using a 24 processor Altix server Single-file reads and writes exceeded 9GB per second

Corporate Office
1200 Crittenden Lane
Mountain View, CA 94043
(650) 960-1980
www.sgi.com

North America +1 800.800.7441
Latin America +55 11.5185.2860
Europe +44 118.912.7500
Japan +81 3.5488.1811
Asia Pacific +1 650.933.3000

2508 [07.01.06]

Jxxxxx