CXFS™ 7 **SGI**® InfiniteStorage Shared Filesystem

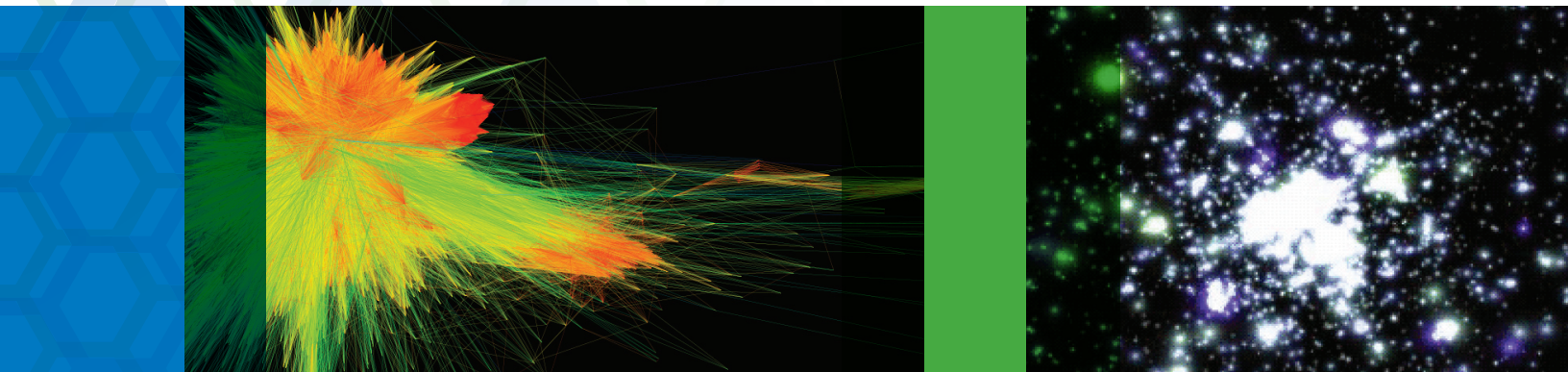## High Performance Shared Filesystem

Key Features

Delivers on the promise of SANs

Provides full cache coherency across heterogeneous nodes

Instant, multiOS, no-copy data sharing

Time tested, proven solution

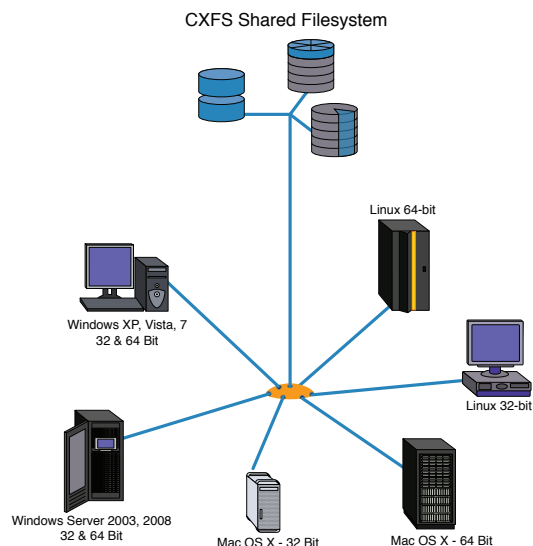## Delivering on the Promise of Storage Area Networks (SANs)

A SAN provides direct, high-speed physical connections between multiple hosts and disk storage. But in a typical data-intensive environment, 20–80% of the time it takes to produce results is wasted by slow, resourceintensive data access methods. The mission-critical ability to share data among computer systems is often accomplished by using a network filesystem such as NFS or manually copying files using FTP, a cumbersome and often unacceptably slow process.

SGI InfiniteStorage Shared Filesystem CXFS provides instant data sharing with no network mounts and no data copies. CXFS combines unparalleled 64-bit scalability and performance with SAN technology, enabling computer systems to efficiently share data and do more work in less time. This allows customers to realize the benefits of SAN consolidation with a significant increase in overall productivity for real return on investment (ROI).
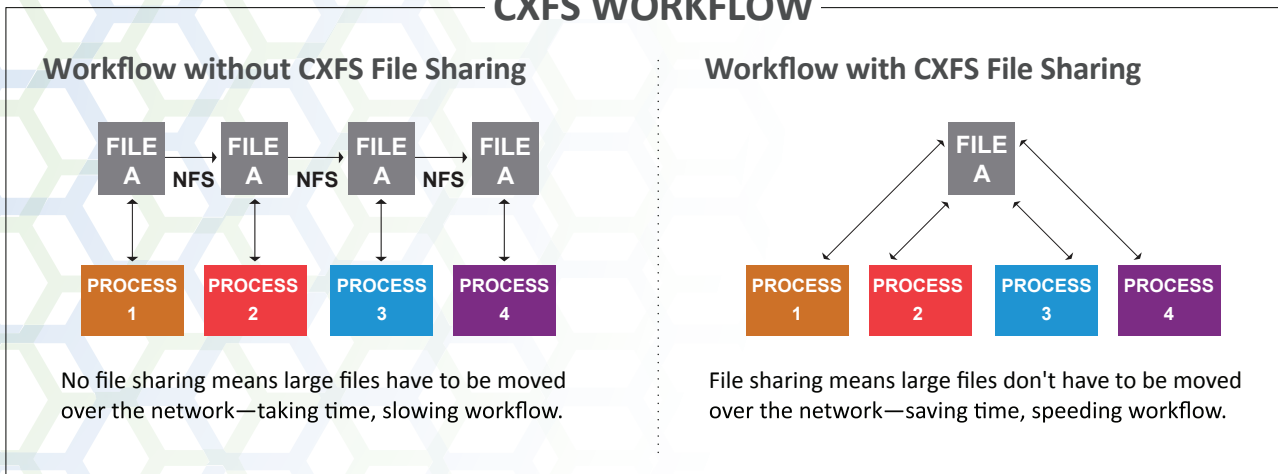
Instant, No-copy Data Sharing Among Major Operating Systems

CXFS provides the software infrastructure to allow simultaneous shared access to the storage—all systems have direct access to all data and are no longer bottlenecked by slow, congested networks or overloaded fileservers. There is a huge reduction in the time required to access data plus the added benefit of no extra data copies to store and manage.

CXFS 7 supports metadata servers running SUSE Linux Enterprise Server (SLES) and Red Hat Enterprise Linux (RHEL). Client-only nodes running Apple Mac OS X, SLES, RHEL, and Microsoft Windows.

CXFS Shared Filesystem

Linux 64-bit

Windows XP, Vista, 7
32 & 64 Bit

Linux 32-bit

Windows Server 2003, 2008
32 & 64 Bit

Mac OS X - 32 Bit

Mac OS X - 64 Bit

CXFS WORKFLOW

Workflow without CXFS File Sharing

FILE A → NFS → FILE A → NFS → FILE A → NFS → FILE A

PROCESS 1 | PROCESS 2 | PROCESS 3 | PROCESS 4

No file sharing means large files have to be moved over the network—taking time, slowing workflow.

Workflow with CXFS File Sharing

FILE A

PROCESS 1 | PROCESS 2 | PROCESS 3 | PROCESS 4

File sharing means large files don't have to be moved over the network—saving time, speeding workflow.

Time-tested, Proven Solution

CXFS was designed as a logical extension to the award-winning SGI 64-bit XFS® filesystem, widely recognized as the most scalable, highest performance filesystem available. CXFS technology is field proven, running reliably at hundreds of customer sites. CXFS provides:

- Industry-leading performance

- Journaling for reliability and fast recovery

- 64-bit scalability, to support extremely large files and filesystems

- Dynamically allocated metadata space

Scalability and Performance Ensure Investment Protection

CXFS delivers the same performance as XFS®. All systems in a SAN equipped environment with CXFS have access to the same file at the same time, at local or near-local filesystem speeds. CXFS is architected to address single files as large as 9 million terabytes and filesystems as large as 18 million terabytes, guaranteeing that CXFS can exceed user data requirements far into the future. Dynamic allocation algorithms ensure that a filesystem can store millions of files without wasting disk space. A single directory can contain millions of files without degrading performance. With the advanced XVM volume manager, disk volumes can be configured across thousands of disks, ensuring that CXFS will grow to meet future storage needs.

Data sharing using CXFS is more efficient than traditional methods, reducing the total cost of ownership for storage. CXFS is designed to reduce costs by centralizing and consolidating storage, reducing data duplication, lowering administration costs, removing LAN bottlenecks, and reducing time wasted waiting for data. CXFS configurations scale easily through the addition of disks for more storage capacity or Fibre Channel connections for more bandwidth. CXFS performance far exceeds that of most filesystems, for typical workloads such as:

- Reading and writing to a file opened by a single process

- Reading and writing to a file where all processes with that file open reside on the same host

- Multiple processes on multiple hosts reading the same file

- Multiple processes on multiple hosts reading and writing the same file, using direct I/O

- Excellent low latency real-time streaming performance

- Resolution independent, multiple real-times to a single filesystem

These capabilities make CXFS ideal for large file-based applications such as film and video post-production, weather forecasting, and geospatial imaging. CXFS performance is superior to that of network based filesystems such as NFS. The speed of the network, the size of the server, and the protocol itself can bottleneck NFS performance. NFS protocol is synchronous—each block of data requires two network I/Os. CXFS requires a few metadata I/Os, after which the data I/O is direct to disk. CXFS can use multiple Host Bus Adapters to scale a single system's I/O path. CXFS and NAS combined provide both high speed IO to SAN connected CXFS clients while providing the ability to connect to network-based clients to provide a single namespace filesystem environment.

sgi

Solid, Standard Data Integrity

Streamlined metadata management is the key to CXFS performance. Metadata includes information about files (size, ownership, timestamps, location, etc.) as well as information about the filesystem itself. While all file data in CXFS flows directly between clients and disk, transactions that alter metadata are brokered by a metadata server for each CXFS filesystem, which coordinates access and ensures data integrity.

Metadata transactions are routed over a TCP/IP network to the metadata server. Because metadata transactions are typically small and infrequent relative to data transactions, a fast Ethernet connection is typically adequate; however, faster connections such as Gigabit Ethernet are also supported and are recommended for high-availability environments.

Several underlying XFS® features enhance metadata performance:

• Fast metadata algorithms with excellent buffering

• Sophisticated structures and algorithms for fast lookups

• Ability to allocate large extents, minimizing metadata

A unique RPC design enables high-speed processing of metadata transactions. CXFS utilizes:

• Specialized thin RPCs

• Excellent buffering of both data and metadata on clients

• Multiple metadata servers (one per filesystem)

• Ability to bypass the CXFS layers on the metadata server

• Ability to dedicate small configuration servers as metadata servers

CXFS uses a fast and efficient token mechanism to control file access, guaranteeing transaction accuracy that far exceeds that of typical network filesystems. Because file locking is critical, CXFS takes great care in the implementation of POSIX®, BSD, and SVR4 file locks. No application changes are required to use file locking with CXFS.

Guaranteed Rate I/O (GRIO) Ensures Application Bandwidth Priorities

Applications like satellite stream acquisition, film postproduction and media broadcast frequently have the highest quality of service requirements. GRIO V2 is a unique solution which guarantees that a required level of bandwidth is available for a specific user or application within the CXFS SAN.

With GRIO V2, any system in the SAN can be guaranteed a dynamic or static stream for a specified period of time. Other systems can be given priorities for their claim on the remaining available bandwidth. GRIO V2 central service throttles non-guaranteed system available bandwidth based on the total available SAN bandwidth. This capability scales with the SAN environment and does not restrict configuration flexibility.

GRIO servers on SLES and RHEL server-capable nodes.

High Availability with Complete Data Access Failover

CXFS is designed to detect and automatically recover from singlesystem failure, failure of the active metadata server, failure of SAN or TCP/IP network components, and failure of the network or SAN partition.

In a CXFS configuration, one system is the metadata server, and all other server-capable nodes may be designated as backup metadata servers. A backup metadata server becomes the metadata server on detection of the failure of the initial metadata server; this failover is designed to take place in seconds, ensuring continual access to data regardless of the state of any single system.

Integrated with the Most Robust Storage Resource Management Available

CXFS is configured, monitored, and managed using an intuitive user interface based on Java™. CXFS can be managed from any Web browser compliant with Java, for the utmost in flexibility. A command-line interface is also available with functionality that can easily be incorporated in automated scripts.

True LAN-free Backup and Restore

LAN-based backups are becoming increasingly time consuming and often put an excessive burden on the LAN. CXFS allows existing backup applications to be easily migrated from the LAN to the SAN. A SAN-based backup server simply mounts the CXFS filesystems and backs up data as if it were local. All data travels across the SAN, the LAN is free for user traffic, and no new software is required.

Complete SAN / NAS Gateway

Systems not participating in the CXFS membership still have access to all data. Nonparticipating systems simply remain in their current configurations and access data in the CXFS volumes through a member system via traditional data sharing methods such as NFS, FTP, and CIFS.

Experienced Consultants

When you're ready to deploy, rely on the experts in SGI Global Services for implementation, training and support services. SGI Professional Services consultants have helped hundreds of organizations realize the benefits of CXFS. We offer both packaged and custom implementation services to install, configure, tune and test the CXFS environment.

Investment Protection

SGI Customer Education offers courses in CXFS administration and flexible training options at www.hpctraining.com. SGI Customer Services provides a wide range of Support programs to protect your investment. Contact your SGI Sales Rep for more details and pricing.

Technology

Journaled, shared, 64-bit with guaranteed filesystem consistency

Supported Platforms

CXFS 7.0 Supports:
    Server:
    SUSE Linux Enterprise Server (SLES) and Red Hat Enterprise Linux (RHEL) x86_64
    Client-Only:
    RHEL ia64 & x86_64
    SLES ia64 & x86_64
    Mac OS X Snow Leopard (10.6.2 or later) for Intel
    Mac OS X Lion (10.7.4 or later) for Intel
    Mac OS X Mountain Lion (10.8.1 or later) for Intel
    Windows XP including x64
    Windows Server 2003 including x64
    Windows Vista including x64
    Windows Server 2008 including x64
    Windows 7, including x64
    Windows 8

Storage Connectivity Technologies

All platforms are supported in a SAN configuration by Fibre Channel, InfiniBand, and SAS.

CXFS is also supported in a WAN environment via the use of appropriate WAN extension devices.

Scalability

Maximum File Size
• 9 million terabytes (or the system drive limits)

Maximum Filesystem Size
• 18 million terabytes (or the system drive limits)

Filesystem Block Size
• Selectable at filesystem creation time
• 512 bytes to 64KB for normal data and up to 1MB for real-time data
• Filesystem extents (contiguous data) are configurable at file creation time using fcntl and are multiples of the filesystem block size; single extent can be up to 2GiB in size

Partitioning
• Up to 64,000 volume elements, 64,000 wide stripes and dynamic configurations

Physical Disk Sector Size Supported
• 512 bytes

Backup/Restore

• Dump/restore, bru, cpio, tar; Atempo Time Navigator,EMC Legato Networker, and many popular commercial backup packages

• Dumps of active CXFS filesystems are supported

Support for Data Lifecycle Management Policy Automation

The Data Management API (DMIG-DMAPI) allows implementation of data lifecycle management policy automation software (such as SGI Data Migration Facility - DMF) without any kernel modifications as well as high-performance dump programs without requiring raw access to the disk and knowledge of filesystem structures.

Global Sales and Support: **sgi.com/global**