



Managing the information that drives the enterprise

[SUBSCRIBE](#) | [CUSTOMER SERVICE](#)[HOME](#)[CURRENT ISSUE](#)[BACK ISSUES](#)[BROWSE CONTENT](#)[BROWSE BY TOPIC](#)[ABOUT US](#)[CONTACT US](#)Search For: [ALL FEATURES](#) [MAY '03 TABLE OF CONTENTS](#)

Additional Features

- [Distance: Meeting the new mandate for disaster recovery](#)
- [Optimize your storage for fixed content](#)
- [The shape of the new data center](#)
- [Getting real about iSCSi](#)

[>> SEE ALL FEATURES](#)

Columns

- [Editorial: Finally, a storage user group](#)
- [Behind the Firewall: EMC goes dancing, without a partner ... Softek going it alone? ... CA picks up where BMC left off ... Prices stay put](#)
- [Hot Spots: You've selected your tools for implementation. Now get ready to deploy them.](#)
- [Best Practices: Keeping your eyes on all of your devices might require more than just two eyes.](#)

Shared file systems: a mixed blessing

by: [Benjamin Kuo](#)Issue: [May 2003](#)

printer friendly

e-mail a friend

request a reprint

[PREV PAGE](#) | [1](#) [2](#) | [NEXT PAGE](#)

Seamless, consolidated and efficient management of storage through an advanced shared file system has long been a promise of storage area networks (SANs). Shared file systems--which let hosts share files on a SAN--promise to simplify the management of storage and save money by consolidating storage resources. So, how close are we to that dream? The short answer: There's some progress to report, but to date, not many companies have employed shared file systems.

If you have multiple hosts that need to access a common set of files on a SAN, a shared file system is necessary to coordinate between those hosts. Otherwise, if two systems try to read and write from the same file, it's likely that data corruption will occur. A shared file system coordinates access to a file, and ensures that reads and writes are consistent between the hosts. And if the two hosts use different operating systems, you'll also need a shared file system to handle normalizing file operations between multiple operating systems.

File sharing product roundup

PRODUCT	OPERATING SYSTEMS
ADIC StorNext File System	Linux, Solaris, WinNT, Win2000, IRIX
EMC Celerra HighRoad	Win 2000, WinNT, HPUNIX, Solaris, AIX, IRIX
IBM Tivoli SANergy	AIX, Solaris, HPUNIX, Linux, WinNT, Win2000, Mac
IBM StorageTank	AIX, Solaris, HPUNIX, Linux, Win2000, Win XP
IBM General Parallel File System (GPFS)	AIX, Linux
Oracle Cluster File System (OCFS)	Linux

▣ [Integration: Is HSM ready for open systems or has it had its day?](#)

>> SEE ALL COLUMNS

Trends and Analysis

- ▣ [Storage Firsts](#)
- ▣ [ATA Drives Move Up the Ranks](#)
- ▣ [EMC/Accenture Partnership Bears Few Fruit](#)
- ▣ [DDS Lives as Replacement Fails to Materialize](#)
- ▣ [Outsourced Backup: Pricey But Worth It](#)

>> SEE ALL TRENDS AND ANALYSIS

Our Sponsors

Find out more about this month's *Storage* advertisers by clicking on the company names below to contact them and request more information on their products and services.

▣ [VERITAS](#)

▣ [Quantum](#)

▣ [Fujitsu Softek](#)

▣ [Computer Associates](#)

▣ [HP](#)

PolyServe Matrix Server	Linux, Windows (2Q 2003)
SGI CXFS IRIX, Win NT	Solaris
Sistina Global File System (GFS)	Linux
Sanbolic MelioFS	Windows
Sun QFS	Solaris
Veritas SANPoint Direct	Windows
Veritas SANPoint Foundation Suite/Cluster File System	Solaris
Veritas Database Edition/Advanced Cluster for Oracle 9iRAC	Solaris

Users are also looking to shared file systems to help solve issues with the speed of accessing data over Ethernet. If you directly connect clients to the SAN through a shared file system, you eliminate the overhead and bottlenecks of transmitting that data over an Ethernet network. That's a technique that works well with large files, where the throughput more than offsets the overhead of the shared file system. In speed-dependent applications such as scientific computing, database clusters or multimedia handling, the additional speed is directly linked to increased performance of those applications.

The use of shared file systems can also significantly ease the amount of storage and handling required for data, particularly when there's a large amount of data which would need to be moved or duplicated, such as in multimedia applications. A shared file system is also a requirement for many high-availability systems, providing a shared storage pool for a failover pair or shared access for scaling an application cluster. Finally, by using a shared file system you can optimize use of your storage, and allocate storage on a finer granularity than disks or LUNs.

Shared file systems aren't a new technology. Systems such as OpenVMS have had clustered filesystem support for years in mainframe and midrange environments. Now with the advent of widely available storage networking equipment, shared and cluster file systems for Unix and Windows server environments are starting to gain acceptance, especially in such data intensive areas of video editing, oil and gas exploration and genomic research applications.

There are several types of shared file systems in use today on SANs, says Philippe Nicolas, SNIA data sharing tutorial manager and SNIA France chairman. Shared file systems can be broadly grouped

Future directions

[Network Appliance](#)

[Overland Storage](#)

[Sony SAIT PetaSite](#)

[MaXXan](#)

[StorageTek](#)

>> SEE ALL ADVERTISERS

into three categories. First, there are SAN file systems where access to files on a device is shared, but not the file system itself. The second type is clustered file systems where all nodes understand the file system structure. The third type is shared file systems that are integrated within an application engine, such as Oracle 9i Real Application Clusters (RAC) (see "[Shared file systems types](#)").

Accelerating NAS

Typical of a shared SAN file system is IBM's SANergy, which targets multimedia and small- to medium-size workgroups. The solution-- which was purchased from Mercury a few years ago--uses a metadata server and presents a network-attached storage (NAS)-like access to systems, using the SAN for large block transfers. "SANergy is an accelerator of network file systems. For someone with a NAS box, SANergy takes advantage of Fibre Channel and splits the control and data path. Control data goes over an IP network; information is shared back to client and actual data I/O goes over the Fibre Channel SAN," says Greg Tevis, one of IBM's software architects for its Tivoli Storage Area Network Manager software.

EMC's HighRoad solution also provides a combined NAS/SAN approach to shared file systems. Paul Ross, director of storage network marketing at EMC says, "Two years ago, we released a product called HighRoad. It enables file sharing between a bunch of servers, but they don't have to access the file system through the NAS device." Using the EMC's network-attached storage heads in its NS600 servers which contain HighRoad drivers and a host bus adapter (HBA), the EMC servers can access a volume via NAS over Ethernet, using the SAN for

How will shared file systems evolve? Bill Yaman, VP of software at ADIC, says, "We're taking a Switzerland approach to platform coverage. Our primary areas for future development are in platform support, heterogeneity in operating systems and better support for back end storage."

EMC is looking to scale shared file systems into the enterprise. According to Paul Ross, director of storage network marketing at EMC, the next step in EMC's development is a single name space. Because of file system corruption issues, he says, "Everyone is afraid of creating a terabyte file system, because the amount of time required to do a fsck (fsck is a suite of programs for RISC OS that are able to analyze and fix E-Format discs. The suite is made up of three programs: one (fsck) for everyday disc scanning, one (eliminate) for removing broken files and one (hardfix) for fixing heavily corrupted discs) is exponential to the size of a file system." However, he says this is one of the areas EMC is tackling next, in order to allow companies to further consolidate their storage.

IBM's StorageTank project is keyed to global name spaces, which is the concept of managing storage in containers, such as grouping files by business. IBM's Greg Tevis, a software architect for IBM's Tivoli Storage Area Network Manager software, says integrating global files in a SAN with policy-based storage management will be important direction. Tevis also says IBM is moving toward a future where "administrators can set up a policy that sets up when and how often to back up, what performance characteristics of storage data should be placed on and allow data to move according to policy."

PolyServe, Beaverton, OR, is focusing in on the Oracle RAC market. Steve Norall, director of product marketing at PolyServe, says, "We are the only company to license and implement the Oracle Disk Manager (ODM) interface on Linux. As a consequence, our product provides the same performance as a raw data partition, but also provides major file system and cluster manageability benefits. In addition, with Matrix Server, we enable the use of all the Oracle9i features [parallel ETL, OMF, external tables] that are unavailable when using raw partitions or OCFS."

high speed, direct access for large block transfers.

Clustered file systems

Unlike SAN file systems, clustered file systems mount an entire volume on the nodes in the cluster. Clustered file systems work by joining a set of servers together in tight coordination, allowing them to share and access common files over a SAN. When a client requests to read or write a file, the file system drivers determine if another user is currently reading or writing a block of data through a locked server. If not, the client locks the file, directly accesses the data through the SAN and holds that lock until a read or write is completed. This coordination ensures what is written to disk in the SAN is always consistent.

Advanced Digital Information Corp.'s (ADIC) StorNext file system is one of the original shared file systems to run on a SAN. Bill Yaman, VP of software at ADIC, says StorNext is a heterogeneous file system designed for data-intensive SAN environments.

IBM's StorageTank is also a clustered file system. Unlike SANergy, StorageTank is focused on providing strategic, enterprise-level reliability and features in a clustered file system. Tevis describes the difference, saying SANergy isn't an enterprise-level generic global SAN file system. It's a department or area file sharing solution with file system limitations in terms of performance and scalability. SANergy can only support hundreds of clients. According to Tevis, StorageTank can support "tens of thousands of clients."

Start-up Sanbolic Inc., Watertown, MA, also offers a fully clustered file system, with initial availability of Windows support. The company also says that its architecture will support Unix in the future.

PREV PAGE | **1** 2 |> NEXT PAGE



printer friendly



e-mail a friend



request a reprint

SUBSCRIBE | **CUSTOMER SERVICE**

HOME | **CURRENT ISSUE** | **BACK ISSUES** | **BROWSE BY TOPIC** | **BROWSE CONTENT** | **ABOUT US** | **CONTACT US** | **ADVERTISE**

Storage magazine is part of the [TechTarget](#) portfolio of enterprise IT-focused media.

© 2002-2003 TechTarget. All Rights Reserved.