GVSAN

Technical Overview

GVSAN and XSAN
Technology Overview (featuring point by point comparison)
February 2007

390 Fremont Street, San Francisco, CA 94105 ● www.gvs9000.com ● Ph. 415-777-0320 ● Fax 415-777-9544

Table of Contents

Diagrams

Introduction

Storage Area Network (**SAN**) technology has been a common storage solution since 1995 allowing storage consolidation, high throughput, highly available and scalable access to storage resources for many computers. GVS Inc.'s GVSAN SAN solution has a legacy going back to 1995, the beginning of Fibre Channel based SAN solutions. Apple Computer, Inc.'s recent offerings into this space include XSAN, a SAN file system built from ADIC's StorNext.
Although both GVSAN and XSAN can utilize the same underlying SAN hardware, exactly how they do that is fundamentally different resulting in drastic differences in cost, ease-of-use, performance, scalability, and availability.

Most SAN architectures deliver these important benefits:

• **Storage and server consolidation**. Servers have the ability connect and reconnect to any storage resource through software without needing to deal with physical connections.

• **Higher performance**. Fibre Channel switch fabric technology supports 200 megabytes per second per port, with 400 megabytes and 1024 megabytes per second per port widely available soon. Small to large networks can easily be built with 8, 16, 32, 64 and 128 port switches being very common.

• **Lower cost**. Total cost of ownership is lowered by more efficient use of existing storage, reducing unutilized storage, meaning less storage purchases. Also, centralized management of the storage and how the servers use that storage reduces time to reconfigure applications, and allows fewer IT professionals to manage a larger amount of storage.

• **Easy scalability**. The networks can grow without disrupting users work by simply adding or reconfiguring switches, storage, and servers as required.

• **High availability**. Redundancy in Fibre Channel paths, switches, storage, and servers can be built into the SAN in order to reduce unplanned downtime.
In the work together spirit, GVS Inc., the developers of the GVSAN® SAN solution, will provide readers with a technology overview of these systems outlining the true capabilities of each. Armed with this information, the reader will then be able to ask intelligent questions about each SAN solution and ultimately decide which solution provides the capabilities they require.
The true story of cost, ease-of-use, performance, flexibility, and scalability is not what all the hype and marketing would make one think. So, sit back and relax, as the true stories behind the different SAN solutions are unveiled.

GVSAN SAN direct file access
GVSAN architecture utilizes all existing disk formats, file systems, and utilities including all third-party disk utilities. GVSAN uses existing non-proprietary file systems, such as the very high-speed journal protected HFS+.

XSAN Metadata Controller per SAN
XSAN architecture utilizes a metadata controller (**MDC**) for each SAN. XSAN groups storage by LUNs to form Volumes where the proprietary XSAN (StorNext) file system stores its data. Once

committed, the entire storage LUN must always belong to that Volume. All requests routed via Ethernet to MDC as in standard NAS, but data portion of I/O goes through Fibre Channel.

SAN Background

Each decade more and more information is converted to digital format causing an exponential growth in storage requirements in almost every industry. Recently, for the first time in history, the cost of storage in the IT infrastructure has surpassed the cost of the servers. As the amount of data increases, so has the need to manage and protect that data. Over the years different strategies have been developed to deal with storage needs. Three architectures are described herein: direct-attached storage (**DAS**), network-attached storage (**NAS**), and storage area network (**SAN**). Before SAN – DAS and NAS Architectures

DAS is one of the most common storage architectures in use today because the storage directly attaches to the computer either internally or externally. NAS is common, especially in businesses, as it provides an easy way to centrally locate, administer, and access common storage. Each of these architectures however has important limitations.

Direct-attached storage (DAS)

The first storage strategy developed early in the digital revolution was direct attached storage. This storage architecture provides for the storage device to be directly attached to the computer/server. Internal drives of your computer are directly attached. External storage devices are often directly attached using the American National Standards Institute (**ANSI)** Small Computer Systems Interface (**SCSI**) or ANSI Fibre Channel (**FC**) standard protocols. The storage devices are prepared for use by the computer by writing information on them that is in a recognized format of the software installed on the computer. This format then allows the computer to utilize a file system where data and executable code is stored in files. The computer can then share these files with others often via Ethernet using a protocol such as Apple File Protocol (**AFP**), Server Message Block / Common Internet File System (**SMB/CIFS**), or Network File System (**NFS**).

Limitations of DAS and NAS systems

DAS and NAS architectures have some important limitations, including:

• **Single point of failure**. The single computer attached to the storage in DAS (illustrated as file server) and the NAS appliance can fail cutting off all I/O from client systems.

• **Performance bottlenecks**. Because all I/O traffic must flow from storage into a computer (DAS file server or NAS appliance) and be copied in memory and packetized for sending to client computers over bandwidth limited Ethernet (known as the I/O store and forward operation), the maximum performance levels are easily reached.

• **Scaling limits**. Besides quickly exhausting the potential maximum bandwidth to a few clients, scalability in terms of total clients, and total storage, are severely limited by both DAS and NAS architectures in terms of scalable bandwidth, capacity, and manageability (the ability to centrally and easily manage a growing number of resources in terms of clients, bandwidth, and capacity).

Network-attached storage (NAS)

In NAS architecture, a NAS appliance connects to client computers via a local area network. The NAS appliance is when user access to the data over the network interface over to the storage device, the storage device in this case is Nomadic Firewire drive which are formatted with software RAID 1 (mirror) for data security in Nomadic Firewire doc, it provides simple removal of drive for transport and remote storage for safety. File-sharing is done over the network at very high speed for DV playback applications up to 10 users or 100 of users for data sharing.

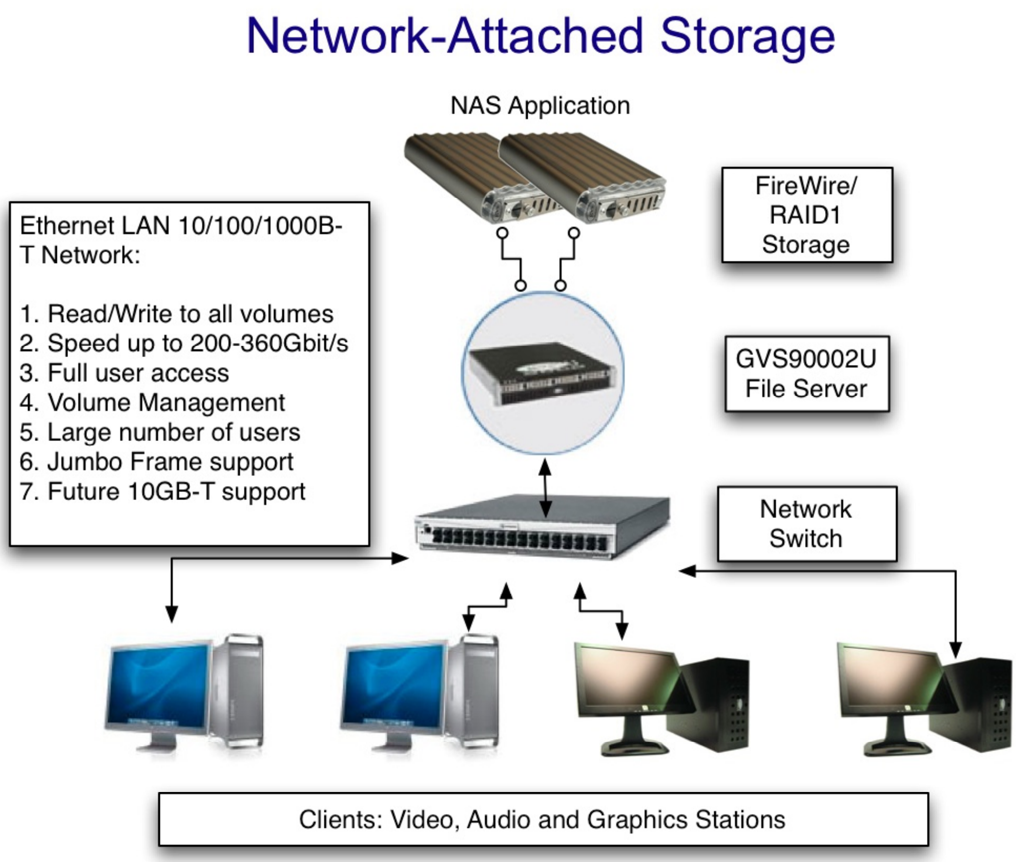


Diagram 1: Network Attached Storage

DAS and NAS Sharing

The act of sharing a file system to Ethernet clients is basically the same for DAS and NAS, save for NAS being optimized for just this purpose. Basically every file system request consists of a metadata portion and the actual data portion, which both transfer over Ethernet **and** SCSI.

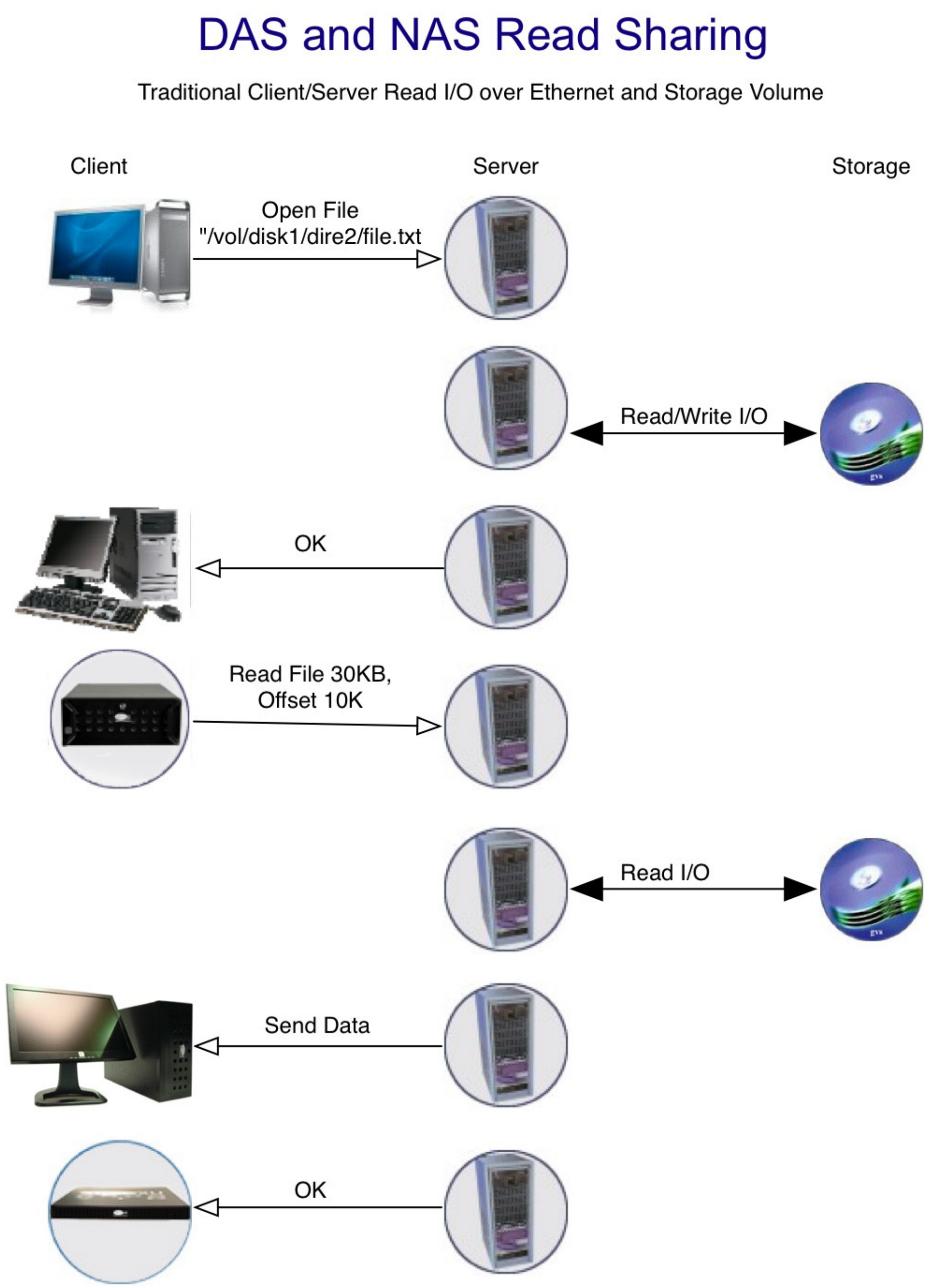


Diagram 2: DAS and NAS Read Sharing

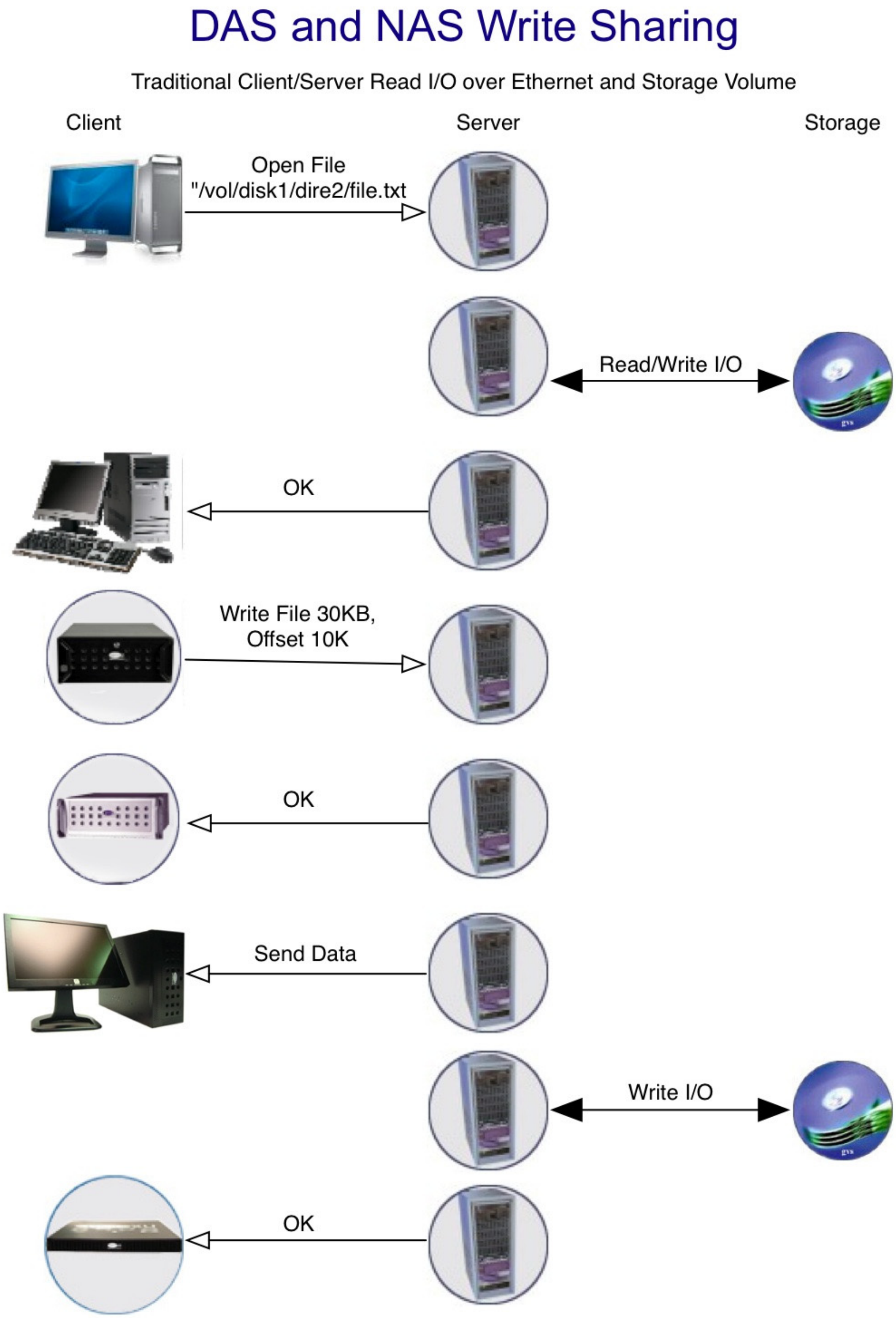


Diagram 3: DAS and NAS Write Sharing

DAS and NAS sharing bottlenecks

When DAS and NAS architectures share file systems to clients over Ethernet the data must transfer twice: first from the storage to the server usually through SCSI and then from the server to the clients over Ethernet. This operation is known as I/O store and forward and involves packetizing the information to be sent over Ethernet. This is usually a CPU intensive operation as it often involves moving memory from one location to another in preparation before DMA engines can offload the CPU. This bottleneck is a major limiting factor to network file sharing performance and scaling limitations and also why it is a single point of failure for all clients in the network.

Everything Connected in SAN

SANs allow computers and storage to connect in a many-to-many fashion through the high-speed, low-latency interconnect technology of Fibre Channel. This paradigm shift from one computer - many storage to many computers – many storage creates several challenges for Operating System (**OS**) software which were designed thinking all the storage the computer can access is fair game. This shift also provides new benefits by consolidating storage and computer resources together.

SAN Implementation Types

Because simply connecting computers into a SAN will result in data corruption as each computer tries to write to the storage in an uncoordinated manner several approaches have been developed to manage this issue. Each approach has different advantages and disadvantages.

Solving SAN multiple-writer corruption

With all computers connected to the same storage, corruption must be prevented. Several SAN architectures have addressed this issue, including:

• **Shared Nothing Zoning**. By zoning storage to only one computer the SAN mimics the DAS architecture and its limitations.

• **Shared Everything - single-writer per file system**. This architecture is an improvement allowing actual sharing among computers of the same storage at the same time with the only drawback that there can only be one writer at a time to the same file system. GVSAN uses this architecture.

• **Shared Everything – multiple writers per file system**. This architecture is an improvement in sharing allowing multiple writers to the same file system at the same time at the cost of getting back some of the DAS and NAS performance, scalability, availability, and manageability limits. This approach sometimes introduces nonstandard proprietary file system formats that are not compatible with standard disk partitioning or repair utilities and severely restricts how the storage is utilized. XSAN uses this architecture.

Shared Nothing – Storage Islands

An early method to control access to the shared storage in a SAN was to defeat the sharing of everything and instead share nothing, as in DAS, through a technique known as zoning. Zoning can occur in software and hardware including the FC Host Bus Adapter (**HBA**) and the FC switch depending on the implementation. In this approach everything is physically connected, but the zoning of the SAN restricts access. This allows each computer to only see the storage it has been given access, and can be configured such that each storage unit is seen by only one computer and not shared among others. In essence this recreates the situation with DAS although reconfiguration is easier since the SAN can be rezoned to make changes.

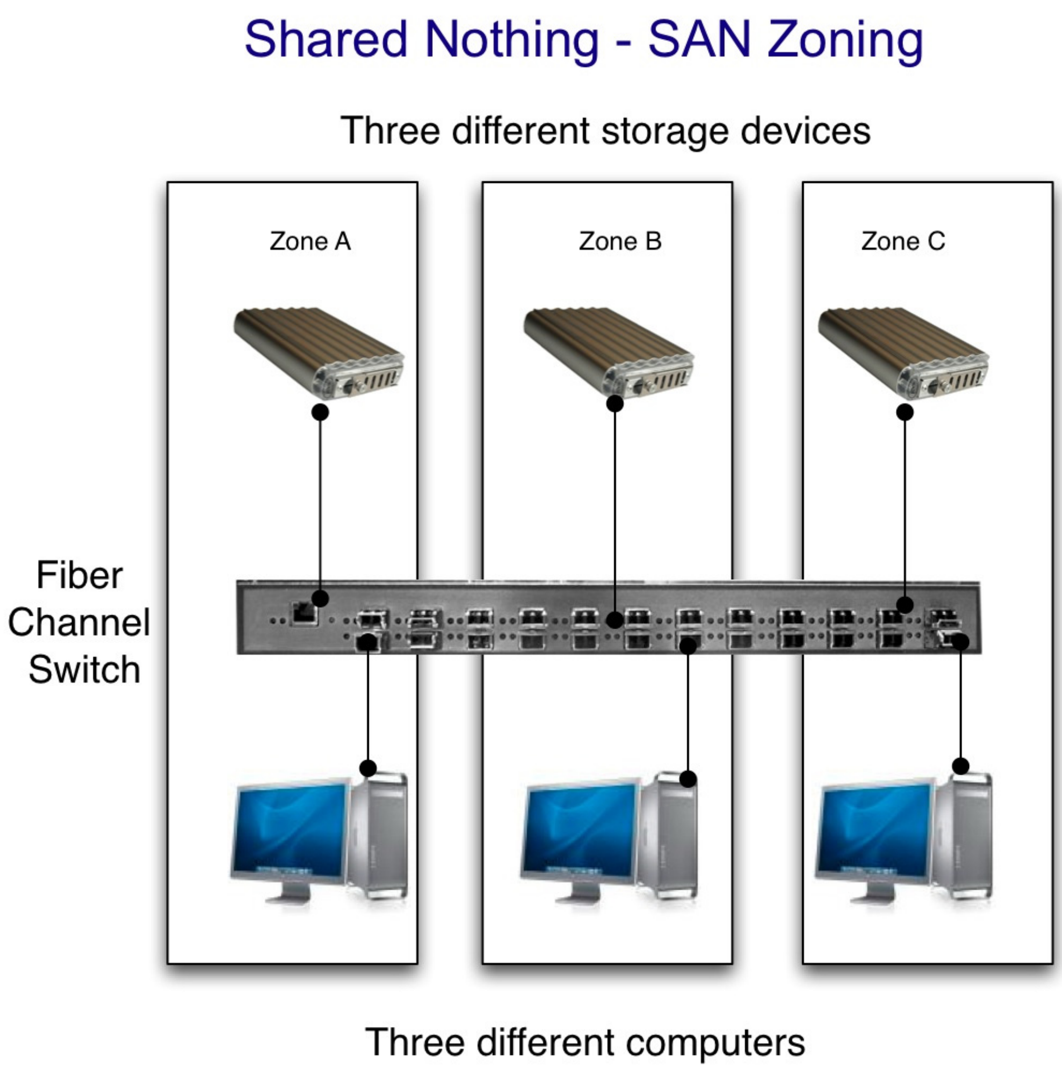


Diagram 4: Shared Nothing – SAN Zoning

Although this approach prevents the multiple-writer data corruption problem, there are several important disadvantages to shared-nothing zoning, including:

• It defeats actual sharing of storage with others at the same time

• Making changes are not dynamic, requiring restarting one or more computers disrupting work
• Tedious management problem requiring special training to make changes and prone to human error
• Granularity of control typically doesn't match with real world use of storage at the individual file system level, but is usually limited to all storage attached to a particular switch port, or a storage device logical unit (**LUN**), or a storage device world wide name (**WWN**) known as port-level zoning, LUN-level zoning and WWN-level zoning respectively. NOTE: Not all switches support all zoning types.

Shared Everything – Shared Storage

An improvement over strict DAS-like zoning is to allow each computer in the SAN to dynamically access storage on demand while preventing multi-writers from corrupting the file systems. Even in a small SAN it is desirable to have some access control that allows a subset of storage to be available to particular computers on demand. Over the past two decades many products have been developed to address sharing storage in the SAN environment. In general these products fall into two architectures featuring different tradeoffs.

High-performance, concurrent file sharing. Unlike DAS or NAS, which use network file-sharing protocols to deliver files to clients, a Fibre Channel SAN allows faster file access and more efficient sharing allowing many users in a workflow to work with the same data at the same time, facilitating collaboration and increasing productivity.

Network-based storage management. By consolidating storage resources in a single accessible SAN all available storage can be utilized more efficiently—and with greater flexibility. Any computer can dynamically use any of the storage on demand. Management is also streamlined making it easier to control user access.

Eliminating single points of failure. Because SANs can be provisioned with multiple paths to the storage, redundant HBAs, switches, controllers, and power supplies, Shared Nothing – SAN Zoning

Benefits of consolidated storage

A SAN file system provides important benefits over DAS and NAS architectures:
• Fast, concurrent file sharing for streamlined workflows
• Increased uptime through the elimination of single points of failure
• Simplified administration and access controls using directory-based management
• Reduced cost through more efficient disk use
• Flexible deployment and easy scalability without interrupting operations single points of failure can be eliminated. Even the computers can become redundant as each can be made to access the same storage at the same time.

Flexible SAN topology. With a SAN file system, such as GVSAN or XSAN, it's easy to expand the capacity of your SAN as storage needs grow. Simply add more storage, and you can expand existing volumes—or create new ones—that can be shared among the attached computers. To increase available bandwidth or processing power to your network services, you can add more servers running GVSAN or XSAN. These new systems can have immediate block-level access to the same storage volumes and host network services, such as web serving, file serving, video editing, or media streaming, to additional network or Internet clients.

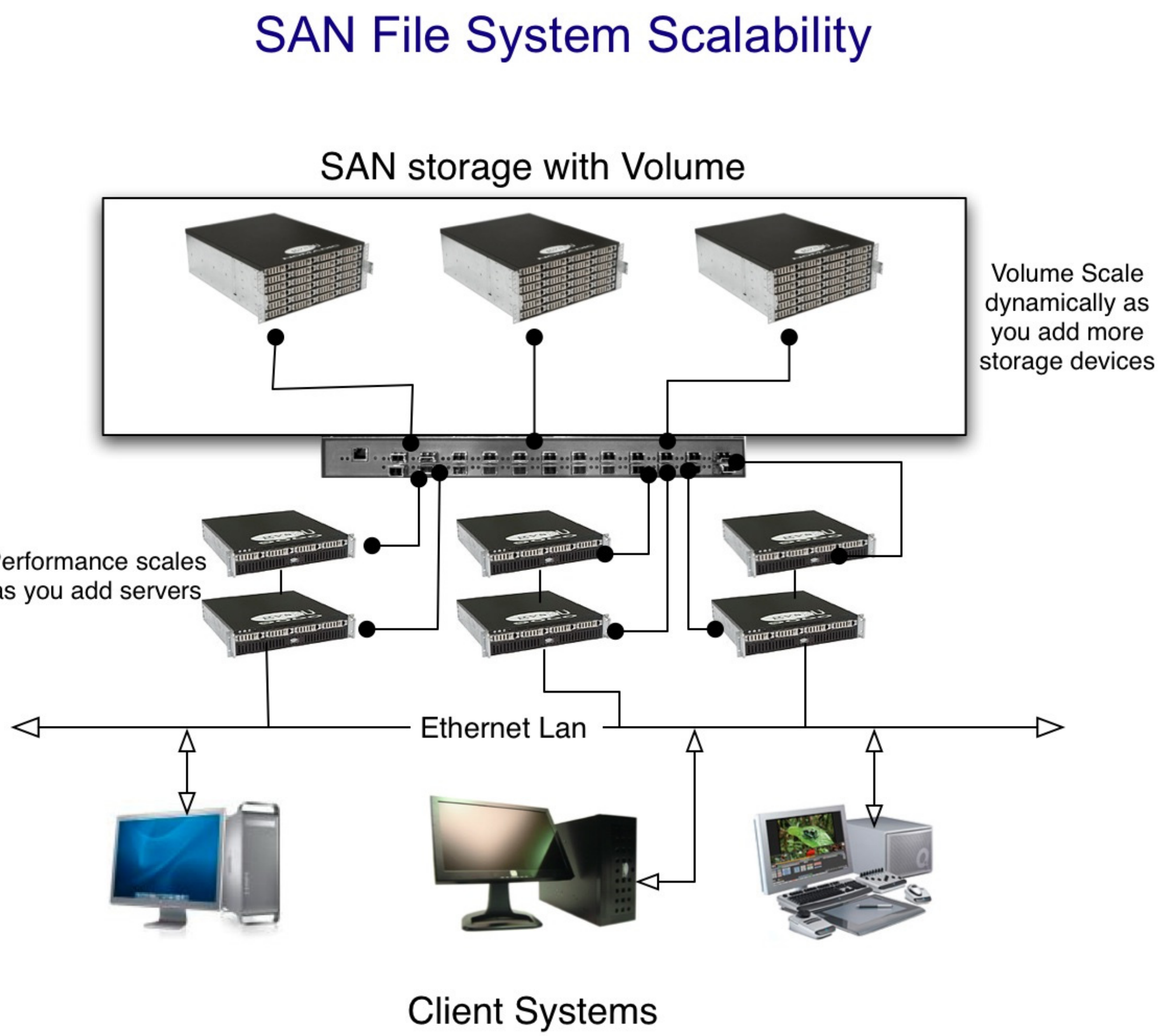


Diagram 5: SAN File System Scalability

SAN File Systems – – Serverless Architectures allowing single writer to file system at a time
GVSAN utilizes an architecture that operates similar to DAS. The customer can partition, format and utilize the LUNs in the SAN without restriction. All existing third-party disk utilities function as normal and can be used on the SAN storage. Because the Mac OS X native high-speed journaled HFS+ is supported all existing repair utilities can also be used. Because of this, LUNs can be used and reused at will anytime.

For example, the GVSAN customer may software-stripe between portions of LUNs to create any number of volumes with a particular performance characteristic, such as for high-definition video. The customer can then stripe between partitions of the same LUNs to create any number of additional volumes with a different performance characteristic, such as for standard-definition video. As time goes by, the customer can reconfigure just those portions of LUNs as needed for new purposes without disrupting any other existing SAN file systems in use. This results in a highly efficient utilization of SAN storage.

Although this architecture limits the number of computers that can write to the same file system at the same time to one, it does not limit to number of computers that can write to the same LUN at the same time. This means that many file systems may exist in the SAN, even on the same LUN, and all can be written to at the same time by different computers. Additionally collaboration is eased because there are no restrictions to the number of readers of the same file system at the same time.

Two SAN file system sharing approaches

Two architectures with different tradeoffs:

• **No MDC (serverless)**. This architecture, utilized by GVSAN, requires no additional computers or Ethernet network. Each computer accesses the storage and file systems as if they were DAS. This allows the SAN to scale in terms of capacity and performance virtually unlimited. Only one writer at a time to a particular file system is allowed. There is no single point of failure and no restrictions on how LUNs are utilized.

• **MDC (server) required**. This architecture, utilized by XSAN, requires two additional computers, one for failover. The MDC serves the metadata portion of each I/O. As a result, clients can write to the same file system at the same time. The cost of this ability is that many of the scaling and performance limitations of DAS and NAS exist. The entire SAN storage is controlled by the MDC making it a single point of failure. (A failover MDC can be used with limited success). Usage of all LUNs is strictly controlled.

Ethernet LAN Client Systems

Volumes scale dynamically as you add more storage devices. As a result much more file system I/O can be conducted than on the equivalent amount of storage in an MDC based architecture, such as XSAN. The SAN storage is partitioned into file systems to work with the customer's workflow so that having a single writer to a file system at a time does not limit how the customer gets work done in any manner.

SAN File Systems – Proprietary Client / Server Architecture allowing multiple-writers to file system at a time

XSAN utilizes a metadata controller (MDC) architecture that operates similar to traditional NAS save for the fact that it splits the I/O into a metadata portion and a data portion. Some examples of metadata traffic include open, close, get/set date, get/set attributes, delete, directory listing, and extent mapping. Extent mapping describes a portion of the file, as it is actually located on the storage. The data portion of the I/O can be conducted directly by the client computer (which is SAN attached in this case) over Fibre Channel to the storage while the metadata portion goes over Ethernet between the MDC and the client. This is commonly referred to as third-party transfer and is described as Distributed I/O Store. By unburdening the MDC of the data portion of the I/O, the client can transfer the actual data through the high speed FC connection with the benefit of giving more CPU cycles for the MDC to dedicate to the metadata traffic over Ethernet. The net effect is that the MDC can handle more client traffic than the equivalent NAS-appliance. MDC Optimized SAN I/O

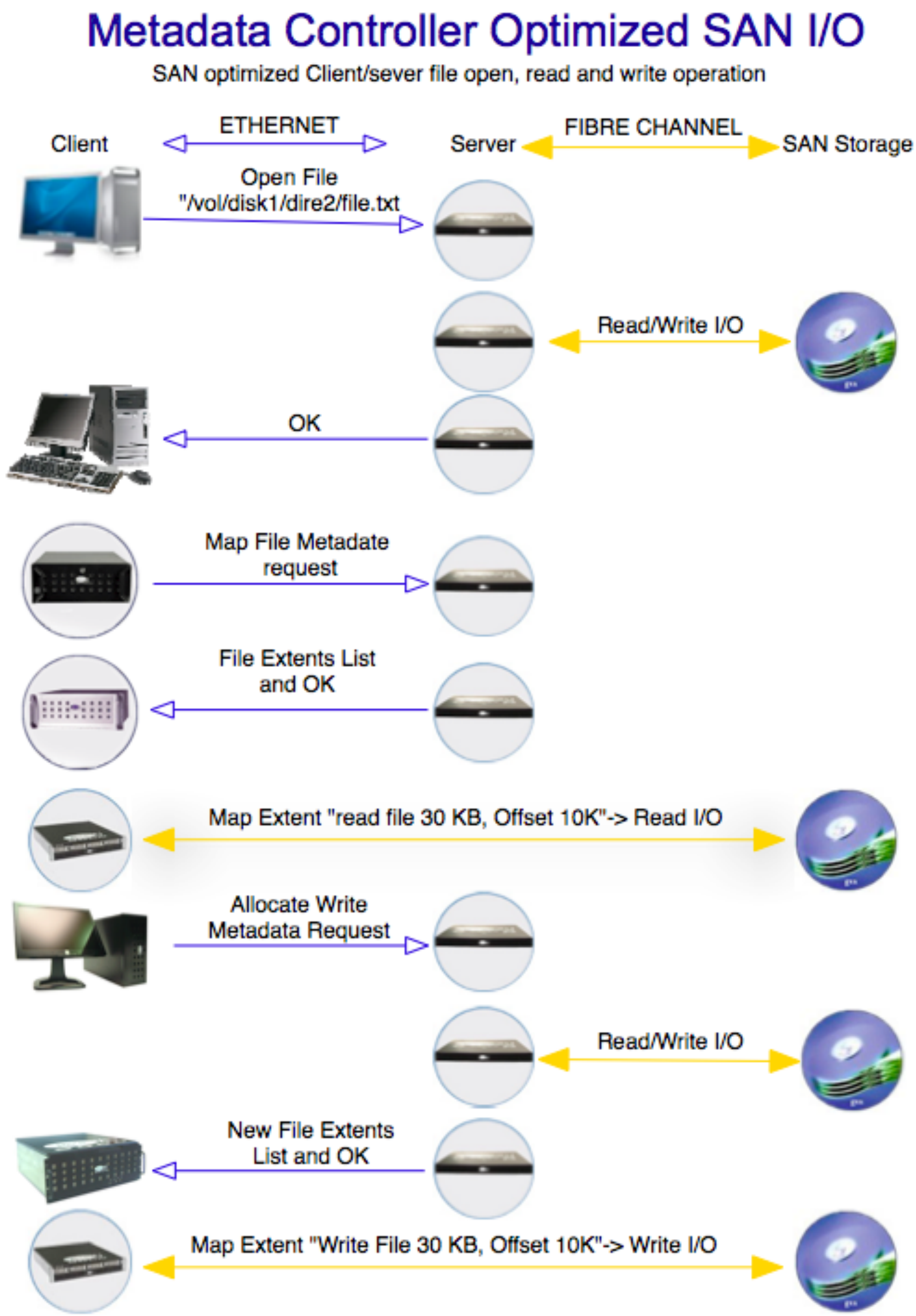


Diagram 6: Metadata Controller Optimized SAN IO

SAN optimized client/server file open, read and write operation

In the case of XSAN, the resulting SAN File System is based on ADIC's proprietary StorNext file system (Fully support/implemented by GVS in large number user site). As a result, the XSAN customer is limited how LUNs are utilized to form volumes. No third-party disk partitioning, formatting, or repair utilities will function in an XSAN system. Additionally, the high-speed journaled HFS+ file system native to Mac OS X cannot be used with XSAN. For example, an entire LUN must belong to either a data-pool or a metadata-pool of storage. These pools are then formed to create a single large volume resulting in a single large file system. Once a LUN belongs to a volume it must always belong to that volume and never by used for any other purpose. The grouping of LUNs into pools defines the maximum performance characteristics as a group of LUNs in a pool are striped together for greater performance. A volume must consist of at least one pool for storage just for the metadata portion and one or more pools for the data portion. This results in inefficient utilization of SAN storage. As a result, multiple computers can write to the same file system at the same time at the cost of getting less done on the same amount of storage as compared to the serverless architecture.

The grate advantage of ADIC's software (XSAN) in mixed user environment, with different Unix, MAC and Windows needing to share data at high speed over fibre channel network SAN. This workflow is much more costly but provide user with centralized work flow at a higher cost.

GVS Metropolis server product delver just that at fraction of the cost, provide centralized storage at hart of production, delivering the user with centralized storage with easy of management, at high speed over TCP/IP to sustain the date rate with utilization of jumbo framing and trucking MCR technology that is been in every Metropolis server line.

Multi CPU Routing (MCR)

MCR is a simple shared disk and GVSAN (storage area network) configurations of up to 8 gigabit can allow users on your network a direct "pipeline" connection to the Metropolis MCR without the necessity of multitude expensive switch and routers. Users appear as network device (as an independent zone or grouped by department) just like output devices and storage volumes; and each pipeline can be easily reassigned or truncated, for greater throughput. Cascading and multi-directional application failover is supported and application services can also be manually migrated to alternative servers for maintenance purposes. A Java-based management GUI provides a single point of administration for multiple Metropolis MCR servers. Metropolis MCR provides a comprehensive availability management solution designed to minimize both planned and unplanned down time.

Implementing cluster storage management solutions is done with a single goal in mind: to reduce the impact of downtime on business operations. To maximize availability, a comprehensive approach that addresses potential failures and potential administrative causes of downtime should be utilized. GVS provides an end-to-end storage management solution that protects availability from the point of data creation, through storage and use, to backup and offsite archiving. Integration points (with the Nomadic Cluster Storage) all along this "data life cycle" for all of GVS' Nomadic family of products.

SAN Performance

Where SANs Bottleneck

SANs connect computers and storage together in a high-speed, low-latency network. The path data travels from its location on the storage devices to the computer have many intervening components, many of which can become bottlenecks in the right circumstance.

• Disk spindle, rotation RPM, seek time, data transfer rate, maximum **IOPS** (I/O's per second)
• Disk drive controller cache memory, bus and memory speed, maximum IOPS
• RAID controller internal paths to storage elements (PCI bus, FC switch paths, cache memory, bus and memory speed, maximum IOPS)
• Storage device enclosure or storage controller external connections to SAN such as the number of FC ports that can connect to a switch and what parts of the storage are visible through each connection
• The Fibre Channel fabric, which consists of one or more FC switches and has a maximum number of ports, and each port has a maximum link speed.
• XSAN architecture suffers an additional important bottleneck factor involving the Ethernet network and the MDC, which is the traffic-cop for the entire SAN, and a single concentration point for I/O traffic.
• The FC HBA(s) in the computer which may have more than one port that can be connected to the switch, the PCI card type of the HBA
• The speed of the computer, its bus, memory, and PCI All of these resources can reach their maximum performance capacity. When this happens, any attempt to access the resource will be delayed. For example, if enough computers access the same storage at the same time then either the maximum performance of the storage will be reached, or the maximum performance of the connecting FC ports will be reached, resulting in delays in accessing the storage.

Scaling Performance, Capacity, and Availability

Fibre Channel ports on switches, HBAs, and storage have a maximum link speed. Today's most common link speed is 2 Gbps, which roughly translates to about 200 megabytes per second. The prior generation of link speeds was 1 Gbps. Speeds of 4 Gbps and 10 Gbps are becoming more and more common. FC switches can be bought with a variety of port connections, including 8, 16, 32, 64, and 128 being common. Additionally, they can be linked together to form a larger fabric of switches that act together as one larger FC network. Increasing the size of the fabric can scale the raw network bandwidth. SAN storage has different performance characteristics in terms of IOPS and throughput. As more computers use the storage at the same time the IOPS and throughput to each computer can change in unexpected ways. Simply adding additional storage or computers to the network increases storage or computer capacity. The trick to scaling storage performance however is not as easy and involves many factors including determining the number of concurrent accesses to the same LUNs at the same time, the number of paths to the same LUNs, and the maximum IOPS and throughput the LUN can achieve, and the typical I/O profile (streaming video, reads of a certain size, writes of a certain size, etc…). Another complicating factor is the formation of the file system where data will be read and written, and how it uses the underlying LUNs. For example, a file system may be software-striped across many LUNs. These LUNs may be located on different or same switch ports. The overall performance picture is affected by all these factors. Generally speaking, it is desirable to make a file system striped across multiple LUNs and multiple switch ports, however it turns out that is not the way to achieve maximum users in all cases. Experience has shown going across too many LUNs or switch ports actually can reduce the maximum performance that can otherwise be achieved, so the right balance for the right situation must be learned, usually through experimentation. Matching the stripe segment size to a particular I/O profile (such as uncompressed video) is also an important consideration when striping across LUNs. An important aspect of availability is the ability of the SAN to function when failures happen. Multipathing is a technique allowing multiple paths to the storage. There can be multiple paths from the HBA or another HBA in the computer to multiple ports in a switch or different switches, and finally to multiple ports to a storage device. To achieve protection from a single point of failure such as an HBA, a cable to the fabric, a bad switch port or bad switch, or a bad storage port or controller, multiple paths can be used. NOTE that not all storage devices provide protection against single points of failure. For example, Apple's Xserve RAID system provides no multipathing protection against the storage device, and if the storage port or controller fails, all the storage fails on that port or controller. What's worse is because XSAN encourages all storage to be placed into a single large volume, all data in the entire SAN would be effected by a Xserve RAID single-point-of failure controller. Redundant paths can sometimes be used to achieve load balancing by alternating switch path is used for each I/O.

GVSAN Product Overview

How GVSAN Works

GVSAN provides an affordable end-to-end SAN solution for demanding storage environments. GVSAN is compatible with most all Fibre Channel HBAs, switches, storage, and computers. Here's how the GVSAN solution works.

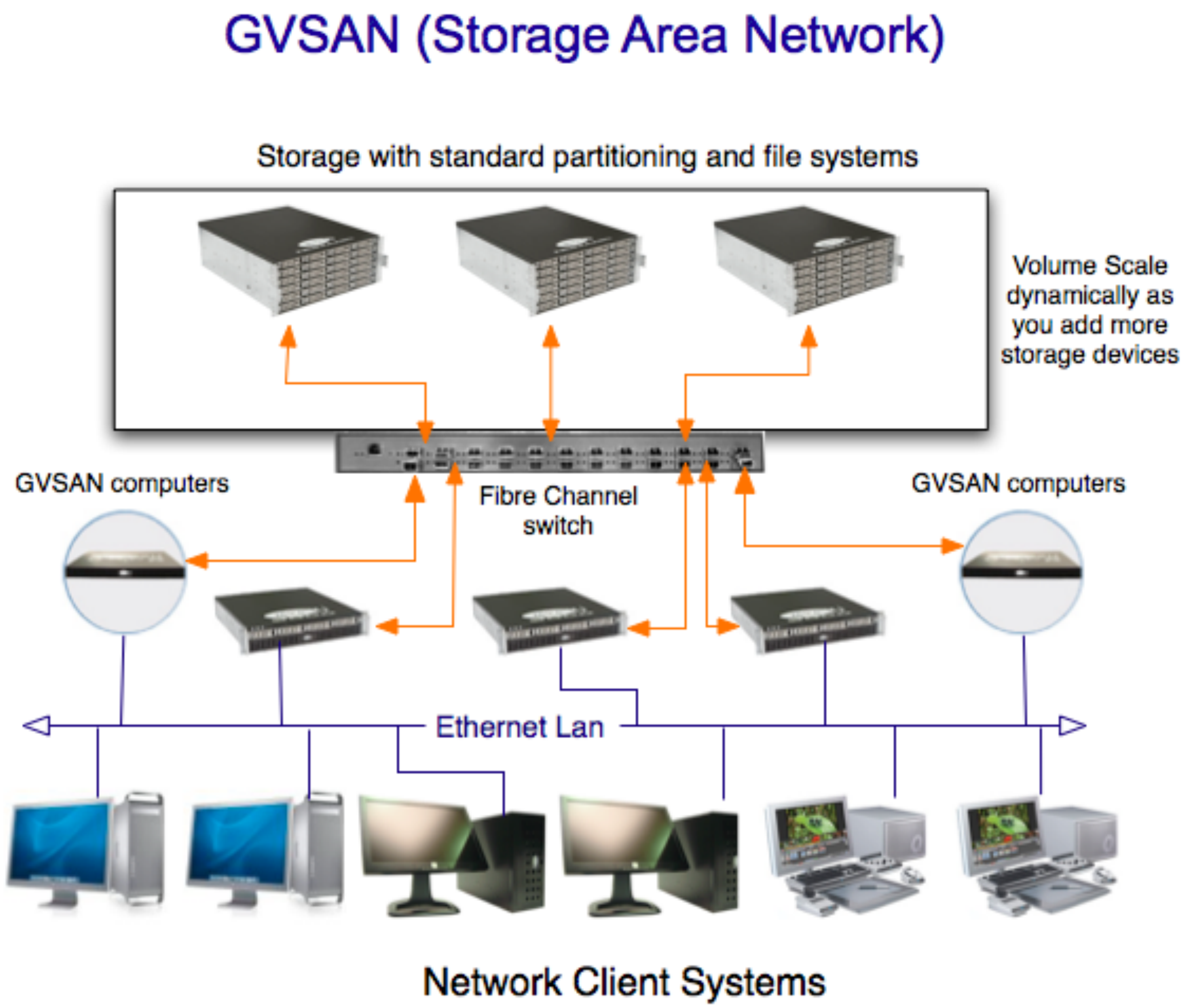


Diagram 7: GVSAN Storage Area Network

SAN file systems. GVSAN allows you to consolidate and use the SAN storage as you would any other storage without restriction. This means you can partition it any way you wish with standard utilities and create any number of file system volumes you desire, all accessible to all systems on the SAN. Adding capacity is as easy as attaching more storage systems to your Fibre Channel network.

Fibre Channel network. The SAN storage connects to the SAN computers through a high-speed Fibre Channel switch. GVSAN has qualified many popular third-party switches for use with GVSAN.

Built-in Project Management. SAN clients can directly control portions of the storage they access as grouped into projects. On small SANs this could be as simple as accessing all the storage all the time. For larger SANs, storage can be organized into more manageable chunks that groups of users collaborate.

GVSAN clients. Computers and servers running GVSAN have direct block-level access to files stored on the SAN file system volumes and read/write capability. As performance needs grow, GVSAN allows you to add servers and computers to the SAN, only limited by the underlying Fibre Channel network you build.

Network clients. SAN computers can share data from the SAN volume with an unlimited number of networked computers using network file-sharing protocols, such as AFP, SMB/CIFS, and NFS.

Key Features

GVSAN is a professional SAN management solution that enables you to manage terabytes of consolidated storage and provides high-availability, high-performance data access across your network.

Standard disk and file system formats. GVSAN supports standard file systems, including journal protected HFS+, the file system native to Mac OS X. As a result, you can use your storage, as you are already familiar, including standard disk partitioning, repair, and striping and mirroring utilities. This allows the most efficient utilization of the SAN storage pool, and is also a lot faster than the proprietary XSAN file system from ADIC's StorNext system, which does not use any standard utilities, and does not work with any of the repair utilities. Multiple users can access even very large files simultaneously—improving efficiency of post-production and other data-intensive workflows, because everyone can be working with the same files on the same network. GVSAN also reduces transfer time, thanks to the high-performance Fibre Channel.

High availability. GVSAN has a high-availability design that allows users to access mission-critical data even in the event of a system or Fibre Channel network failure. Because there is no server, there is no MDC that can fail. Standard HFS+ file system journaling tracks modifications to metadata, enabling quick recovery of the file system in case of unexpected interruptions in service. Multipathing allows file system clients to automatically use an alternate data path in the event of a failure.

Volume management. Volume management and easy file system scalability maximize the flexibility of your storage deployments. Using standard disk partitioning utilities, GVSAN allows you to create storage pools of identical sets of LUNs and stripe them together for fastest-possible performance. Different pools offering special storage characteristics can be combined into file system volumes allowing you to place data on specific file system volumes depending on performance and protection requirements.

Data access control. GVSAN simplifies administration with flexible data management features. Using password-protected Projects, you can control which SAN file system volumes are visible to specific users of the SAN, and thereby control resource utilization on a per-user, per-group, or per-application basis depending on how the Projects are configured. With GVSAN users collaborate on groups of storage that can be organized into projects. Projects consist of any number of file systems, each of which can be configured for a different performance characteristic, such as uncompressed HD video. The users workflow is designed so that not more than one SAN computer at a time can write to the same file system at the same time, but many SAN computers can read from the same file system at the same time. In this way you can have many users writing to the same storage at the same time, just not the same file system at the same time. Users can pass the ability to write to a file system on demand amongst themselves or programmatically depending on the need and the workflow. The benefit of this is that much more can be done on the same amount of storage as compared to other products, such as XSAN.

Remote administration tools. Built-in GVSAN administration software provides volume management; SAN file system configuration, and remote monitoring in one integrated application. GVSAN contains a log of events built-in. These easy-to-use administration tools can even be controlled programmatically with AppleScript and send you notifications via email or pager.

Compatibility and interoperability. GVSAN works with all utilities and equipment.

Components of a GVSAN solution
GVSAN works with these components to create a managed SAN solution:

• **Any PCI computer**. GVSAN will work with any computer with a PCI slot for the FC HBA, and works cross platform with various operating systems.

• **Any storage**. GVSAN works with any storage that can be attached to a SAN FC switch including JBODs and RAID.

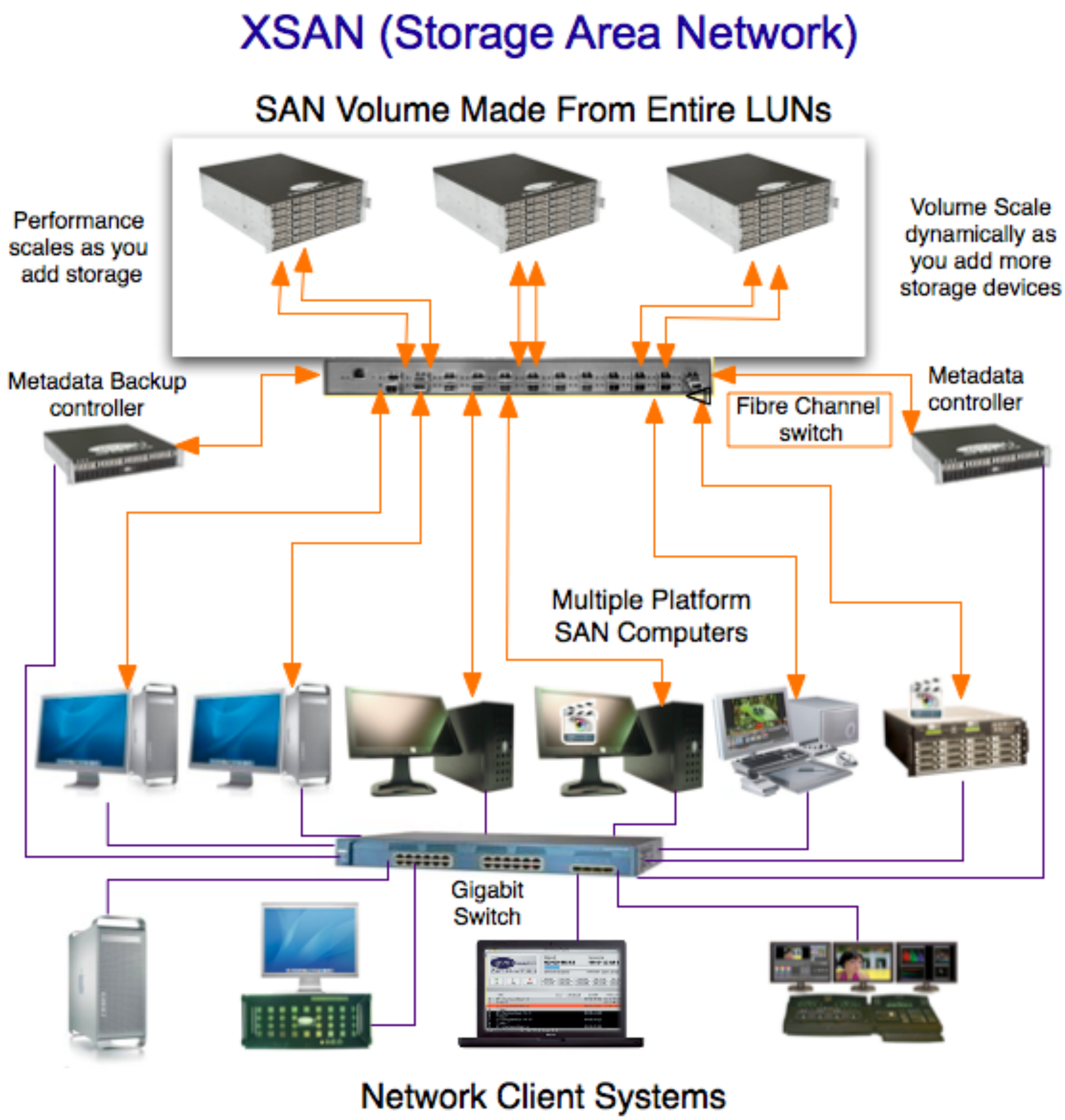
• **Any FC switches**. GVSAN works with all FC switches.

• **Any FC HBA**. GVSAN works with all FC HBAs.

XSAN Network Diagram

Components of a XSAN SAN

XSAN can use some of the same SAN components that GVSAN has used for years, however a XSAN deployment requires considerable extra equipment including a dedicated Gigabit Ethernet switch, connections to each SAN computer (including Gigabit Ethernet card if you want to connect to another network with your SAN computers), two dedicated Xserves for the MDC, and failover MDC (which are the traffic cops for all I/O on the entire SAN), each equipped with HBA and Gigabit Ethernet. Besides all this extra equipment amounting to more moving parts that can break, it also represents considerable additional cost, as well as several additional performance and scaling bottleneck.



.

Diagram 8: XSAN Storage Area Network

Contrasting Product Overviews

At first glance GVSAN and XSAN may appear to offer many of the same benefits, however going deeper will uncover important differences between these products.

Cost differences for a small 6-seat SAN

XSAN TOTAL EXTRA COST FOR 6 SEAT SAN FOR 3 YEARS:

XSAN per seat 3 years support and updates. XSAN retails for x amount. To receive major XSAN software upgrades for three years is x amount. AppleCare XSAN support for one year is x amount, and is 24/7 telephone support covering XSAN GUI.

Additional SAN Cost

GVSAN. It should be noted that Apple offers a separate AppleCare Protection Plan for Mac, AppleCare Premium Service and Support Plan for Xserve and Xerve RAID, AppleCare Professional Video Support, and Max OS X Server Software Support to be purchased separately and at considerable cost if your issue falls outside the XSAN GUI. This additional cost is not reflected in the x number above.

XSAN requires a separate dedicated Gigabit Ethernet network for the metadata traffic. A 24-port Gigabit Ethernet switch cost. Each SAN attached computer must be connected to this separate network requiring an additional cable and Gigabit Ethernet card for each computer, which is an additional total for the 6 computers in our small SAN.

XSAN requires a computer to act as the Metadata Controller, and if you want fail-over capability an additional computer. It is recommended that these computers be dedicated to this purpose to ensure smooth operation. Each MDC costs and includes an Xserve, Xserve service and support, PCI-X Gigabit Ethernet card, Apple FC HBA, Apple remote desktop, and a 50 foot CAT-6 Gigabit Ethernet cable.

Storage Utilization

Today's computers have Operation Systems and native File System that have been developed, optimized, and refined over a long time to have maximum performance. Computers are shipped with the native high-performance file systems. For the Windows world this is NTFS. For the Mac OS X world, this is journal protected HFS+.

GVSAN uses much faster standard file systems and storage formats

GVSAN uses standard file systems, formats, partitions and utilities that have been developed, optimized, and refined to have maximum performance. For Mac OS X, this means using the native journal protected HFS+ file system. The Performance section (see *Performance Comparison* starting on page 38) documents that journal protected HFS+ has been tested to be over 500% faster as compared to the XSAN StorNext file system from ADIC. This can literally translate to copying more than 5 times as much data as compared to XSAN in the same amount of time, or, in another example, streaming an additional 13 uncompressed streams over XSAN. Using a standard file system also means that GVSAN can utilize standard disk utilities to perform important tasks such as defragmentation and repair. GVSAN works with all third-party disk utilities in this regard. Underneath the file system is the storage partition formats that determine how the file system maps to the actual storage devices.

Here GVSAN also uses all standard partition formats and utilities. For example if you are used to using Disk Utility, ATTO ExpressStripe, etc., to partition, mirror, or stripe your storage, then these utilities work the same with GVSAN. Even more important is that use of these standard utilities benefits from years of storage utilization technology that allows extremely efficient use and reuse of the SAN storage pool. Take the case of striping storage to increase performance. Striping is configured differently depending on the I/O profile, for example high-definition video vs. standard-definition. With standard utilities you can create some file systems striped one way on the storage, then on the same storage create more file systems that are striped another way. Later, if you need to reconfigure, a single file system can be unstriped freeing up space that can be restriped another way without disturbing the rest of the storage. With GVSAN, this is no problem, because all these standard techniques can be used.

XSAN uses much slower, proprietary, non-standard file systems and storage formats

XSAN uses proprietary file systems, formats, and partitions that have been developed by ADIC for the StorNext file system. Besides being much slower than journal protected HFS+, the implications of a proprietary file system go much deeper. XSAN's proprietary file system cannot use standard disk utilities to perform tasks such as partitioning, defragmentation, or repair. XSAN does not work with any 3rd party disk utilities.

As a result storage utilization is not efficient and reuse of storage is almost impossible. With XSAN entire LUNs much be dedicated to a single storage pool. A storage pool has a single striped performance characteristic. Once a storage pool belongs to a Volume, is must belong to the volume forever and never be removed. Since XSAN requires whole LUNs to be used as the smallest chunk, rather than a partition slice of a LUN, the user is locked into a very limited striping choice which must be made at the very beginning of life of the SAN and persist through the entire life of the SAN Volume. XSAN prefers to have one large single file system volume, which locks up all your storage. In order to have multiple XSAN file system volumes requires a choice of purchasing much more storage or cutting the performance of your SAN in two since a file system volume exclusively owns all the LUNs that make it up. Reuse of storage is not an option with XSAN since it would require destroying the entire XSAN file system volume. Efficient storage utilization over the SAN life cycle is non-existent since XSAN locks the storage into a single purpose for the entire life of the XSAN file system. Failure points GVSAN and XSAN can suffer from some of the same SAN hardware failures that can happen, if the same equipment is used, but you will see that XSAN has additional glitches that are non-existent in the GVSAN solution.

GVSAN has no single point of failure

GVSAN coordinates a cluster of SAN attached computers usage of storage without any server, leaving each computer completely independent and un-reliant upon another. It therefore, has no single point of failure.

Additionally, because GVSAN is compatible and supported on a wide variety of storage, switch, and HBA equipment, further network-based single points of failure can be eliminated by not using an Xserve RAID for example, and instead using a RAID that has redundant controllers.

XSAN has multiple failure points

XSAN requires a separate independent Gigabit Ethernet network; therefore additional cabling and cards are necessary. If anything goes wrong with the Ethernet switch, or a port or cable problem occurs at the MDC, the entire SAN and every client will stop working.

XSAN requires a MDC computer and for some fail-over capability, another MDC computer. The MDC is a single point of failure in the tests conducted because crashing the MDC, or pulling the MDC Ethernet cable causes the entire SAN and all clients to stop working completely requiring a full reboot. The fail-over mechanism only worked as advertised when the MDC was gracefully shutdown. Because XSAN is only certified and supported to work with limited hardware, the entire SAN volume and all its clients suffer another dangerous single point of failure in the RAID controller, port, and cable to the storage. Because XSAN likes to use all the storage for a single large volume, if any of these fail, the entire SAN and its data will be inaccessible as all the LUNs on that controller will no longer be accessible. Furthermore, there have been cases reported in that replacing a failed RAID Controller does not bring back the data on those drives and needs to be reformatted, meaning the entire SAN and all its data would be completely lost.

Contrasting File System Capabilities

GVSAN uses the HFS+ file system, which is the native file system of Mac OS X. XSAN uses the StorNext file system developed by ADIC. This section compares how Enterprise-class SAN capabilities that meet requirements for data consolidation and fast, shared-access to storage volumes are done with each file system.

Cluster File System

Neither HFS+ nor XSAN's StorNext file system from ADIC are cluster file systems. The definition of a true cluster file system utilizes high-speed low latency interconnect technology, such as infiniband, to connect a group of computers together and implement shared-memory file system data structures for the file system. In this definition, XSAN is not a true cluster file system, just as NAS is not a cluster file system. XSAN is actually just a SAN-optimized version of NAS as its architecture is a traditional client/server model. The optimization simply allows the clients themselves to access the file system data directly from the SAN storage rather than transferring through a server. The metadata traffic in NAS is still done in the XSAN MDC.

Support for 16 TB files and volumes

HFS+ supports 63 bits (2 to the 63 power) for the maximum file size (billions of billions of bytes, much greater than 16 TB) and 32 bits of allocation blocks per volume (over 4 billion). StorNext file system volumes can be as large as 16 TB, and support billions of files per volume. Therefore, on the surface both file systems appear to offer similar capabilities and allow users to share multiple file and volumes, including more than 24 hours of uncompressed 1080i high definition (HD) video. Under the surface however are important differences.

Total number and type of volumes per SAN

Given an amount of SAN storage, GVSAN can easily create virtually an unlimited number of file system volumes on the SAN, each with different striping performance characteristics. XSAN however is strictly limited to a very few number of volumes given the same amount of SAN storage. This XSAN limitation is due to the fact that a minimum XSAN file system volume must consist of at least one LUN for the metadata portion and one LUN for the data portion of a file system volume. With GVSAN, a LUN can be partitioned into thousands of volumes if the user desires. Or, the user might stripe across many LUNs and create thousands of volumes tuned to different I/O profiles (such as HD video). What's more, reusing the storage in a different way is simply a matter of reclaiming the space used by that particular volume, without disturbing any other SAN file system volume.

With XSAN, your choices are strictly limited. To create an XSAN volume the user chooses entire whole LUN to belong to storage pools that are striped for one particular I/O profile. Once a LUN belongs to a storage pool, it cannot be used for any other purpose, ever, without destroying the file system volume to which it belongs. Next the user decides which storage pool will be for the metadata portion of a volume and which storage pool(s) will be used for the data portion of a volume. Once this file system is formed the storage pools are committed to that single volume and cannot be used for any other purpose without destroying all data in that volume. To reconfigure the storage for any other purpose down the road requires destroying the file system volume and all associated whole LUNs that belong to that volume.

Total number of concurrent users, or maximum file system speed

The amount of work that the SAN can conduct on a given amount of storage is much greater with GVSAN as compared to XSAN with the same amount of storage. In one test GVSAN sustained 24 streams of uncompressed video as compared to XSAN's 14 uncompressed streams. In another test, GVSAN copied files and folders over 500% faster than XSAN. The additional performance with GVSAN means more users and more work can be done at the same time as compared to XSAN.

File sharing over Fibre Channel

Fibre Channel's per connection port speed combined with multipathing allow for greater aggregate throughput, allowing transfer of dense formats such as HD video. The switching power of the Fibre Channel fabric allows storage and bandwidth to scale, which is needed for multiple editors working on a video project, or a compute cluster that needs fast data access for maximum utilization of processing power. GVSAN does not get in the way of the maximum performances capabilities of a Fibre Channel network, which is why it is much faster than XSAN. XSAN has reduced, but not eliminated the bottlenecks of traditionally network file servers that use Gigabit Ethernet. As a result, XSAN gets in the way of the maximum potential of Fibre Channel.

File-level locking

Both HFS+ and StorNext file system support fine-grained file-level locking. Although GVSAN architecture uses volume-level locking to control multiple writers, all clients can access all the files on the volume read-only. With XSAN, file-level locking allows multiple users to write to the same volume at the same time, and is very dangerous in post-production workflows where multiple editors are using a single large file, or the same files of any size. With GVSAN in a Final Cut Pro workgroup, each editor has read-only access to all the file system volumes they require and has write access to the area they are working on, including a project segment that they are editing and writing. Because of this, no one can accidentally make changes to a FCP project that someone else is working on.

GVSAN is safe and provides enormous productivity advantages in post-production workflows where multiple editors are using a single large file. In the same situation with XSAN multiple editors will lose all their work. With Final Cut Pro HD v4.5 and XSAN multiple editors can open the same FCP project and start working inside the project, without any warning or error. As each editor saves their work thinking it is safe, it actually completely overwrites the FCP project that was saved by another editor earlier, losing all their work. In this fashion, only the editor who saved the FCP project last will have their video editing changes saved! Clearly, XSAN was not intended to function in a multiple editor video workflow environment.
To avoid such an easy disaster requires strict human workflow planning which is prone to human error. This is not necessary with GVSAN, as GVSAN will not let you lose your work because someone else opened the same FCP project and saved the work they did later than you. XSAN is not practical for true multi-user sharing environment such as digital video editing networks with AVID systems. This is because AVID systems require media databases at the root of the volume and this is unlikely to work in a multi-user environment where everyone has the same "root" volume.

Additionally, an AVID Protools audio system requires throughput that is not practical in metadata controller architectures such as XSAN. GVSAN on the other hand allows complete control of the storage lifecycle and the user can have multiple different stripe characteristics on the same storage.

Bandwidth reservation

GVSAN includes a simple easy-to-use GUI interface method allowing the user to exclusively own access to the file system volume, during which time no other user can access the volume or any of its bandwidth. This is commonly used for important real-time operations that cannot fail, such as ingesting or outputting uncompressed HD video.
In contrast, XSAN offers APIs (Application Programmer Interfaces) that developers and administrators (e.g. not users!) can program for only one system on the SAN for the same purpose. Which would you rather use, GVSAN's user-control through an easy-to-use GUI or XSAN programmer/administration control through a programmatic API?

High-Availability Features

Both GVSAN and XSAN offer Fibre Channel multipathing to eliminate a single point of failure. And although both GVSAN and XSAN provide a means for bandwidth reservation, you learned earlier that GVSAN is much easier and more versatile in this respect. Beyond that, XSAN has many single points of failure that do not even exist as points of failure in the GVSAN architecture.

Metadata controller failover

With XSAN, a second MDC is supposed to take over if the first MDC fails for any reason. In reality the only time failover was successful during testing was when the first MDC was gracefully shutdown. In the cases where the first MDC was crashed, or the Ethernet cable was pulled, the entire SAN and all its clients completely hung requiring a complete restart. Once restarted from the crash, the second MDC did take over the SAN. Failback is supposed to occur when the first MDC comes back online. In testing, the failback failed to occur, leaving the first MDC useless. The only way to get the XSAN back to normal operation was to shut down all the computers, bring the first MDC online, then the second MDC online, and then the remaining computers. XSAN network will also completely fail if there are any problems with the Gigabit Ethernet switch. Perhaps these shortcomings will be fixed over time. With GVSAN, this nightmare will never happen, because its architecture doesn't require an Ethernet network, nor any primary MDC or secondary MDC computer.

Fibre Channel multipathing and storage

Single points of failure in the host-side cabling layer can be eliminated by using a variety of multi-ported FC HBA cards, or multiple FC HBAs in a computer. GVSAN supports just about all FC HBAs, including those from QLogic, Apple, and ATTO Technology. Multipathing, built into Mac OS X 10.3.5 and later, also provides for load-balancing capability. XSAN is only supported on Apple's dual-ported card, which is based on LSI's card. This is catastrophic in the XSAN case as you not only loose access to all the LUNs on that controller, but because XSAN typically uses all the LUNs to form its large file system volume it will cease to function bringing down the entire SAN and all the data on the SAN.

With GVSAN, using the same storage, you would only loose those volumes that you chose to put on those controllers LUN. Higher-end RAID devices can fail-over the storage to surviving controllers, but they are not supported in the XSAN case.

File system journaling

File system journaling protects against file system corruption, not data loss. This means that the user can be working on a file, and the computer can crash and the data in the file, especially if it was cached and not yet written out to disk, will be lost. The journal protects the integrity of the data structures on disk that describe the file system. In the past, without journaling to protect file system transactions, it would be possible to crash a computer at just the right time when important file system structures were in an intermediate state, causing the file system to be corrupted. HFS+ with journaling (used by GVSAN) and ADIC's StorNext file system (used by XSAN) are both examples of file systems that are journal protected. Journal protection also eliminates the need for integrity checks after a crash.

Volume Management

It is important to maximize the efficiency of your storage resources throughout their lifetime. Some examples of this include having storage pools with different performance and protection characteristics for storage files and expanding storage resources on demand, all without any downtime or interruption in service. GVSAN and XSAN differ widely on this front, especially in the area of managing the storage lifecycle—reusing portions of the storage over time without disrupting the SAN or the data in it. Xserve RAID is a low-end RAID system that has some failure points to be aware before choosing it as the basis for your SAN storage.

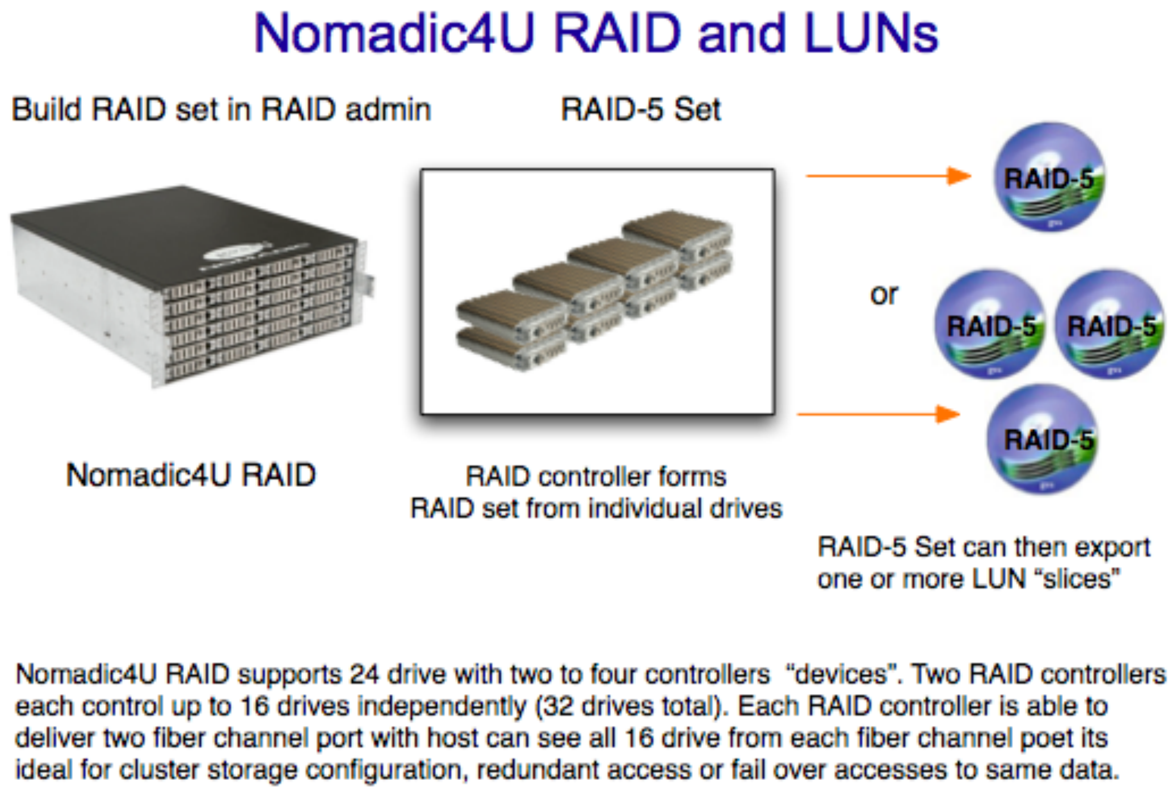


Diagram 9: Nomadic4U RAID and LUNs

Xserve RAID is actually two totally separate RAID "devices". Two RAID controllers each control up to 7 drives independently (14 drives total). Each RAID controller is a single point of failure since one RAID controller has no access to the others drives. If it fails you can no longer access the data on those 7 drives.

Storage pools and SAN volumes

Grouping storage into pools with different performance and protection characteristics allows for efficient management and utilization of the SAN. With GVSAN, file system volumes can be created from any portions of the storage pools, through the use of portions of LUNs or entire whole LUNs. In contrast, with XSAN, file system volumes can only contain entire whole LUNs as different types of storage pools are combined into the same volume. This makes XSAN less efficient and less flexible than GVSAN.

For example with GVSAN, you can create one storage pool of RAID 0 LUNs, another with RAID 5 LUNs and another pool of RAID 1 LUNs. The RAID 0 LUNs provide highest performance, but no protection, the RAID 5 LUNs provide less performance but are parity-protected against a single drive failure. The RAID 1 pool is the least performance, but provides a high level of protection through completely redundant mirroring of the drives. By creating any number of SAN volumes from these storage pools, you can achieve a flexible balance between performance and safety. This provides the administrator with a transparent means to determine the performance and safety of user-stored data. By creating projects that contain the different types of file system volumes (created from the desired pools), the administrator automatically creates an environment in which the user automatically, without user involvement, places their work on the storage to which they have access.

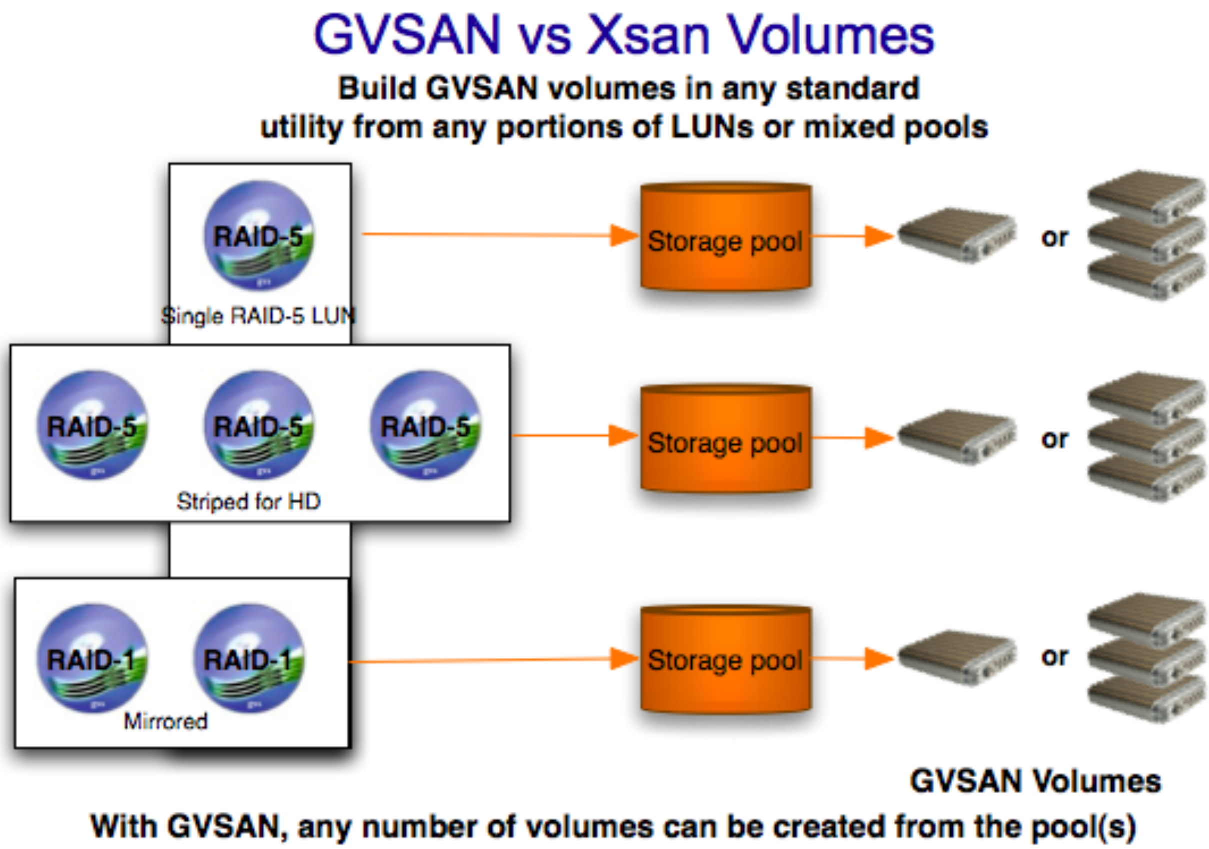


Diagram 10: GVSAN vs. XSAN Volumes

In contrast, with XSAN, you first create a single large file system volume that contains both the RAID 5 and RAID 1 storage pool. The administrator then must configure the single large volume with different file folder directories. XSAN is then configured so that files placed in one of the folders will be directed to the RAID 5 pool, while files placed into another of the folders will be directed to the RAID 1 volumes. This can be confusing to users, as it requires that the administrator involve the SAN users in the process by instructing them specifically where to place their files for the right effect. An additional XSAN restriction is that storage pools are limited to 32 LUNs, and a volume 512 LUNs total.

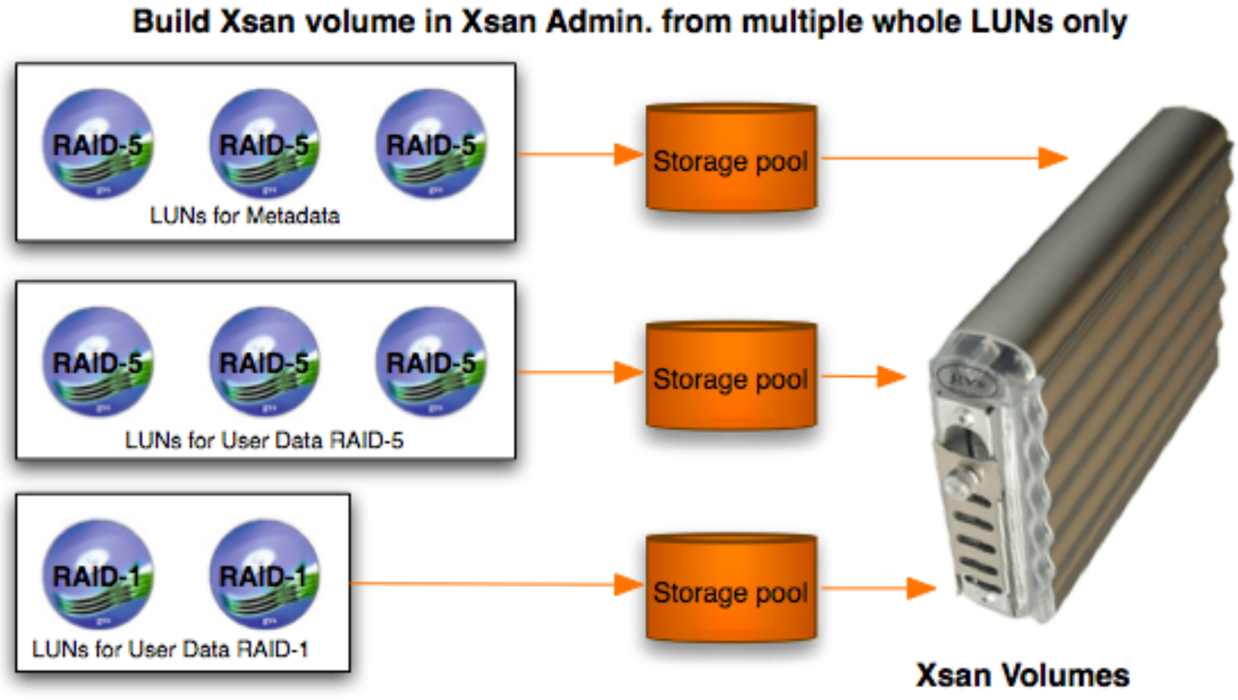


Diagram 11: Build XSAN volume in XSAN Admin from multiple whole LUNs only

Affinities for automated data placement

GVSAN allows you to allocate different classes of storage to different purposes, without forcing users to adhere to annoying rules. Using GVSAN administration tools, you can set up projects, which are groups of file system volumes. By creating projects with file systems created from a storage pool that has a particular characteristic, such as fastest, or most protected, users who place files on that storage will ensure its located only on that pool. Seamless to the end user, projects ensure that an application or task that requires speed or extra protection always stores it files in a suitably fast or protected pool.

In this example, users have access to one project that contains two file system volumes: One is named “Fastest”, associated with the pool of RAID 5 sets; and the other is named “Most Protected”, associated with the pool of RAID 1 sets. During development, working files can be maintained in the Fastest volume, which automatically places them in a RAID 5 pool. Finished work is saved to the Most Protected volume and maintained in a RAID 1, or mirrored, pool. To control what storage pool a user stores files in without giving them the choice between “Fastest” or “Most Protected” the project could be made to only give access to one or the other.

In contrast, XSAN typically has a single large volume because of its inefficient use of whole LUNs. With XSAN you can set up an affinity between a particular folder and a specific storage pool. Instead of volumes named “Fastest” and “Most Protected” you would create folders in the XSAN volume with these names and assign an affinity.

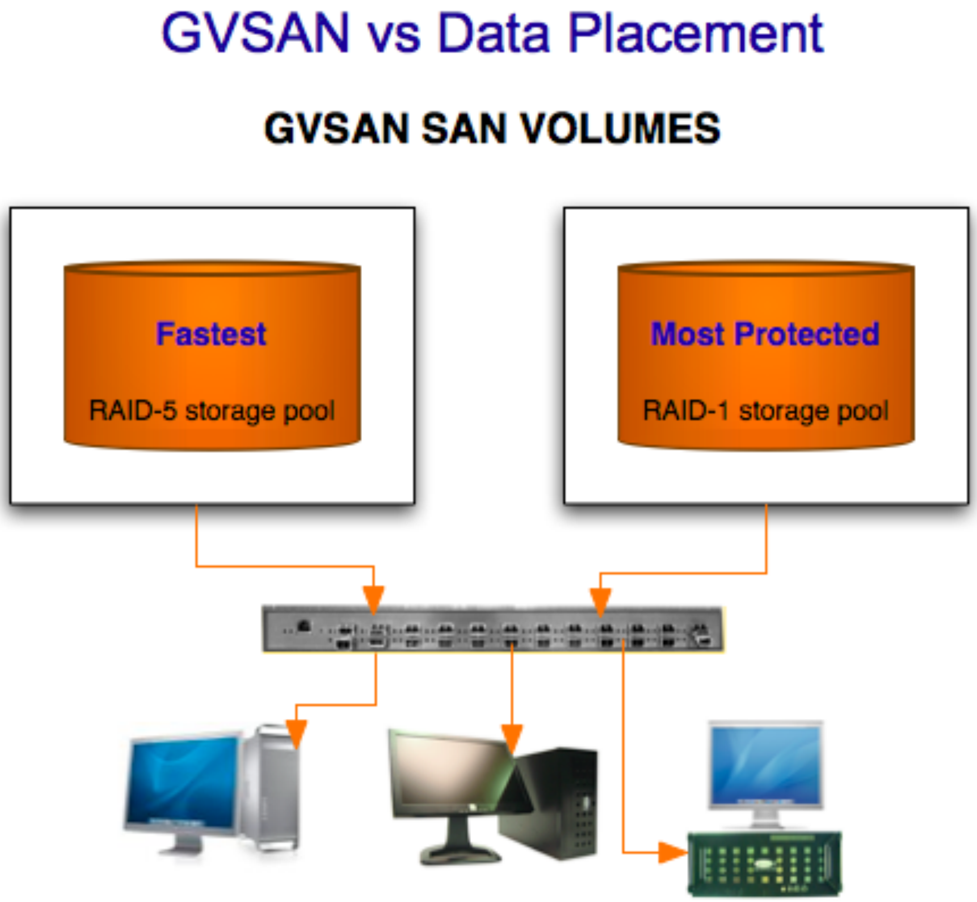


Diagram 12: GVSAN vs. Data Placement

Data Access Control

Before SANs were available, users wishing to share large amounts of data on disks used a method commonly referred to as Sneaker-Net. This is because removable disk drives were hauled from one room to another for use on different computers.

Striping at a higher level

GVSAN uses standard utilities to perform standard partitioning, striping, and mirroring. These utilities all allow custom specification of the stripe segment size, also known as interleave. This is important for tuning for specific purposes, such as HD, SD, or DV video media.

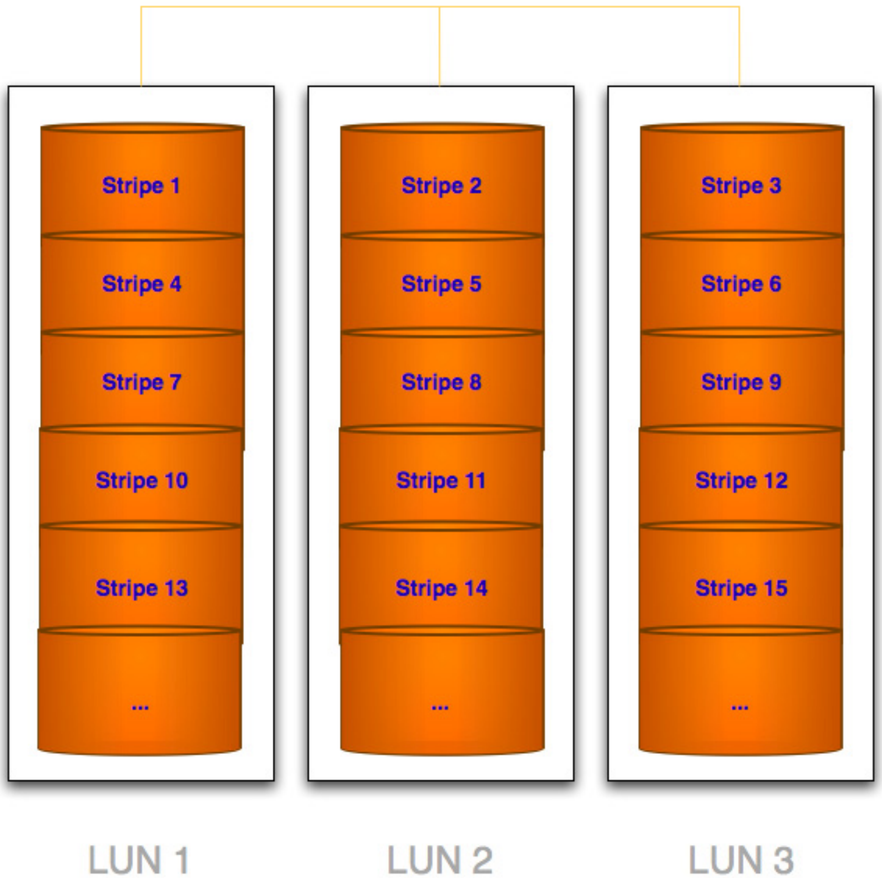


Diagram 13: Striping at higher level

Stripe size is important as it relates to the I/O profile expected on the file system. It directly corresponds to how parallel a particular file I/O operation can be. If a file I/O is of a particular size, large enough to at least equal the size of all the stripes together, then the resulting I/O will work in parallel with all the LUNs accessing at once. Many factors affect this, including boundary conditions on where the file I/O starts and ends with regard to the stripe segments.

XSAN can tune stripe performance on a per storage pool basis only. However, because XSAN requires use of whole LUNs to construct a storage pool, you commit to a single I/O profile characteristic for the entire storage pool, which limits your tuning ability. Further, since the metadata requires a separate storage pool, you eliminate more whole LUNs that would otherwise be used in parallel. Therefore to achieve more than one I/O profile for an entire large XSAN volume you need a huge amount of additional LUNs to create other storage pools with different characteristics. Then you must establish affinity rules of where users must store their data for a particular kind of performance. As a result XSAN is much more costly, very inefficient, and inflexible how it uses the SAN storage as compared to GVSAN.

Computers, including Unix computers like Mac OS X have no problem with this method of data sharing. In this method, someone disconnected storage from one machine and walked it over to another machine for reconnection. With SANs, this same process of moving storage from one computer to another is replaced with a simple software command. Therefore, a SAN can be thought of essentially as Sneaker-Net save for the physical act of hauling a removable disk from one room to another being replaced by a SAN software command to instantly reassign storage to another computer. Mac OS X allows multiple users per computer. Each of these users has access to their own files. Each file is associated with an owner, group, and other permission settings that dictate who and how the files may be accessed (read, write or none). Users and groups are standard, and numbered the same on a machine. For example, when you install the OS, the first user created is always numbered 501. This 501 number is what is recorded on the disk. Therefore, if another computer accesses this storage (thinking it is directly attached), the first user on the computer (501) would appear to be the owner. Also when the OS is installed some standard users and groups are also created, such as system, wheel, root, etc. These also have the same numbers across computers. In a shared storage SAN users typically want to share storage. If they don't, the administrator can configure the SAN so that only particular computers access particular storage. Because a file system's numbers for users and groups will be translated to the local meaning on a computer, a SAN user and administrator simply needs to be aware of this fact. Under Mac OS X, the user has the option whether or not permissions are to be obeyed by the OS on a per file system basis. With GVSAN, if an Administrator wants to still use permissions in a shared-storage SAN, and is not comfortable with the local translations, then he can set up each SAN computer to have the same users, and groups, created in the same order, so that they have the same numbers, to enforce a SAN-global mapping of permissions settings. Normally this is not at all necessary as everyone in a shared-storage SAN that has been given access to the shared-storage by the Administrator will typically want to share their creations. If not, they have the option of password protecting and encrypting their data, or implementing a SAN-global mapping of permissions. Normally the user wants others to read and write to the files they create, so the permissions for "others" are set to read/write. A simple SAN-global group is an easy method for ensuring read/write access as well. By default however, the first user, 501, will appear as the owner on each local machine, so as long as the owner has read/write access everything will be able to be shared with ease.

Volume mapping

GVSAN offers an added layer of control and security through the use of password protected projects, which can be configured to require login each time they are used, and can also be made invisible to the users GUI if desired. In this way, an Administrator can control which users have access to which storage by creating password protected projects that only particular users know how to access. This built-in workflow management mechanism is nice because it also puts some of the control of which storage the user can access at any particular time in the users own hand as along as they know about the different projects and how to access them. In contrast, with XSAN, only an administrator has the ability to assign a particular XSAN volume to a particular SAN computer with a particular permission. The user never has this control and if a change is needed the user must get an administrator to make the change.

LDAP integration

With GVSAN administration tools and SAN global permission mapping, you can also

set user and group permissions, as well as access privileges, at several levels:
• Restrict user access to folders on a volume by specifying owner, group, and general access permissions.
• Create projects with password protection and access restrictions, such as read-only
• Users can unmount, mount exclusive, read/write or read-only depending on permissions.
• Remove a client from a SAN.

Disk quotas
GVSAN supports hard quotas that can be assigned on a user or project basis by defining which whole file system volumes that user/project can have access.

Remote Administration Tools

GVSAN supports SAN managing and monitoring directly from any computer in the SAN, and remotely with a tool such as Apple Remote Desktop. In addition to the easy-to-use graphical user interface, GVSAN can be controlled using AppleScript.

Real-time monitoring

GVSAN features real-time monitoring of the SAN, providing the following information:
• Total, free and used space of a file system volume
• Log file
• Event list
• Connected clients and states
• SAN communication status
• List of BSD names, and word-wide-names of physical devices making up file systems

Event notification

Using AppleScript to control GVSAN, you can have automatic notification via email or pager when an event happens such as a backup is completing.

Compatibility and Interoperability

GVSAN lets you create a SAN that is compatible with your existing infrastructure— giving you the flexibility to build a system that's right for your organization, maximizing the return on your storage investment. In contrast, XSAN is not supported on any existing FC SAN hardware in your organization, unless you are already using Apple HBAs and Apple storage.

Support for native Mac OS X applications

GVSAN supports native Mac OS X and UNIX applications and delivers incredible performance on GVS9000, Power Mac, and Xserve systems. This is because it is based on the native HFS+ file system used by Mac OS X. HFS+ is a multi-fork file system that is not case sensitive, although the HFSX extension supports case-sensitivity as an option.

Fibre Channel infrastructure integration

GVSAN supports most all Fibre Channel HBAs, switches and storage. In contrast, XSAN is only supported on Apple's storage and HBAs, and qualified switches.

ADIC integration

Since XSAN is based on ADIC's StorNext File System it means that it also compatible with ADIC's other products (which must be purchased separately at considerable cost), including its heterogeneous support for other OS. XSAN however does not work with Apple's own Mac OS X prior to 10.3 and does not work with Mac OS 8 or 9! In contrast, GVSAN works with Mac OS 8 and 9, and also Mac OS X 10.2 forward, with support for heterogeneous Windows in development.

Sharing data over Ethernet

With GVSAN you can use standard file sharing to enable network-wide access to the SAN storage.

Performance Comparison

In an effort to better understand the different characteristics of GVSAN and XSAN, each product features and benefits, and which customers would benefit best from each product, they were tested at GVS Labs.

The Test Bed

A single SAN was used for testing GVSAN and XSAN. This was done so that results comparing identical equipment would be more meaningful.

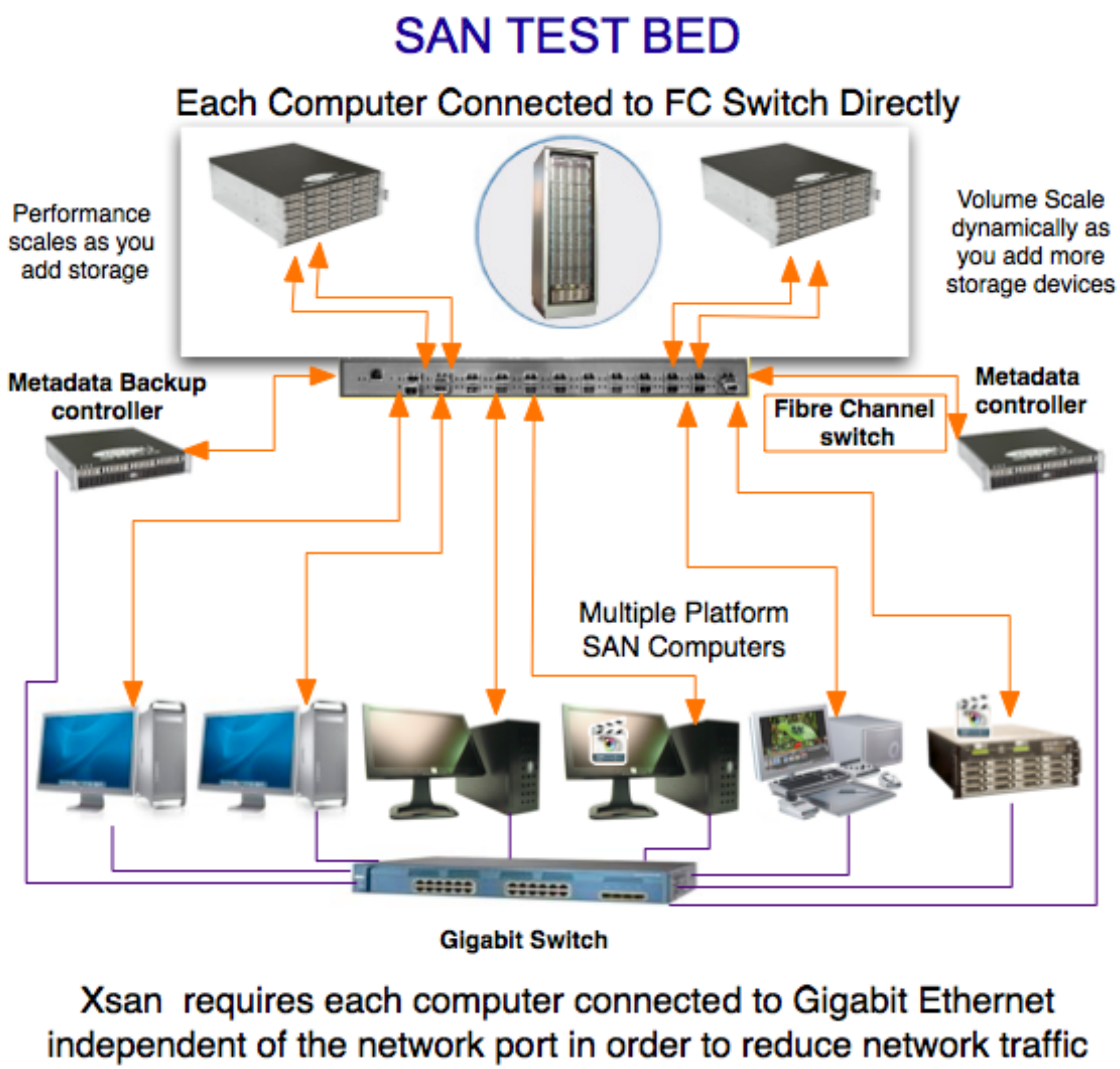


Diagram 14: SAN Test Bed

The Primary Tests

Primarily 3 tests were conducted at each stage. Test 1 was uncompressed Standard Definition (SD) streaming. Test 2 was copying a single 12.14 GB video file. Test 3 was copying a folder of 5,648 files of various sizes totaling 1.09 GB of data. With both GVSAN and XSAN, the optimal storage configuration resulting in the best test results was discovered and used for each result. Xserve RAID storage was tested separately from Seagate JBOD testing. Streaming tests consisted of starting up additional streams on each computer until the storage bandwidth was exhausted resulting in dropped frames on multiple machines. When dropped frames were observed the total number of streams was reduced, all were restarted, until a stable number of streams was run for an extended period of time (sometimes overnight) without any frame drops. In the uncompressed 10-bit SD streaming test GVSAN was able to sustain a total of 13 additional streams. **GVSAN is 146.4% the performance of XSAN**. In other words, GVSAN on identical equipment and storage was able to sustain about 351 megabytes per second more aggregate throughput than XSAN for this test.

The next test compares the relative speed of a simple Finder-level file copy of a single large 12.14 GB file occurring simultaneously on 6 machines. On the XRAID storage, GVSAN and XSAN finished copying the single large file in virtually the same amount of time, on the JBOD storage however GVSAN was 140% the performance of XSAN. This test is interesting because it illustrates an important fact about XSAN, which will become extremely apparent in the next test—metadata traffic effects performance. With a Finder-level copy of a single large 12.14 GB file, there is relatively little metadata traffic to the Metadata Controller machine. Because the 12.14 GB file is not fragmented, it can be described by a small amount of metadata. When XSAN clients begin copying the file, it asks the MDC for a description of where the data of the file exists on the actual storage. Additionally, the MDC controller is asked to allocate chucks of space on the storage to place the copy of the file.

Since there is little work for the MDC controller to perform with a simple copy of a single large file that is not fragmented, the bulk of the work can be performed by the XSAN client itself. Nonetheless, this test shows that **GVSAN is 140% the performance of XSAN on the JBOD storage**. This next test illustrates how metadata traffic can bring your SAN to a crawl. Instead of a single 12.14 GB file, this test copies only 1.09 GB of files, consisting of 5,648 files of various sizes, ranging from as small as 23 KB, to as large as 7 MB. This test, showing **GVSAN to be about 500% performance of XSAN** demonstrates that anytime XSAN has to deal with a fair amount of metadata traffic, it quickly deteriorates the performance of the entire SAN and all its clients.

Watching the XSAN metadata network traffic is very revealing. During file copying, such as the test above, there is a tremendous flurry of traffic for the first several seconds as everyone starts the file copying, then the traffic settles down into a distinct pattern. For the test above about 10,000 packets were coming out of the MDC while 6,000 packets were coming in. Then it settled down to about 6,500 packets out and 4,200 packets in. This results in many megabytes of traffic coming into and out of the MDC, scaling as more clients simultaneously access the file system. This is why XSAN behaves as NAS with its same limitations. Because XSAN's MDC doesn't have to actually transfer data over the Gigabit Ethernet, it can handle more metadata traffic than the equivalent NAS. What's important here to understand is that a result of the metadata traffic in XSAN slowing the SAN clients causes GVSAN to be about 500% faster than XSAN. Another way to look at this is that **5 times the amount of work can be done with GVSAN** in the same time that XSAN took.

Other XSAN behavior and miscellaneous notes

Because XSAN is a MDC architecture, it naturally has more things that can go wrong, including more single points of failure that do not even exist in GVSAN. As a result, we tested the following capabilities:

XSAN metadata traffic slowdown

Anytime XSAN has to deal with metadata traffic to clients, the overall SAN performance will decrease, as the MDC is a single traffic-cop for the entire SAN. We tried working with XSAN clients, and performing a Finder file search on an XSAN volume. All clients video workflow can be disrupted. This is because, as a single traffic-copy for the entire SAN, a tremendous additional amount of metadata traffic is produced when searching for files. A find that took less than **5 seconds with GVSAN took about 4 minutes and 30 seconds with XSAN**. This doesn't happen with GVSAN as its distributed architecture allows each SAN computer to handle its own metadata traffic directly over Fibre Channel. With GVSAN, the SAN will only slow when the performance level of the actual storage or network has been saturated, not simply because there are more computers in the SAN.

XSAN Fail over/back test

Failing over and back are important aspects of high availability architecture. With XSAN, the MDC is not only the single traffic-cop for the entire SAN, it also represent a single point of failure. To mitigate this fact, XSAN is supposed to be designed so that the MDC can fail over and back in case of failure and recovery. Unfortunately were only able to get the XSAN metadata failover to work under a single unrealistic circumstance. If we were to gracefully shutdown the metadata controller the fail-over function seemed to work. In the more realistic failure cases where the metadata controller crashes or its Ethernet cable is removed the entire SAN froze and every system was brought down. We also tested fail-back and found that also not to function with XSAN. The metadata controller therefore adds another single point of failure reducing the otherwise high availability of the entire SAN.

Xserve RAID pull-drive test

Although not specifically a GVSAN or XSAN issue, we thought it important to point out the facts surrounding high availability and Apple Xserve RAID. Pulling a drive while I/O is occurring will cause the I/O in progress to fail. Copying files and removing a Xserve RAID drive can illustrate this. The copy will fail with an error. If the copy is restarted it will then function. True drive fail-over should be transparent to the I/O's in progress. Since the Xserve RAID controllers cannot access the others drives, if a Xserve RAID controller fails, all SAN data located on that controller will be inaccessible as another single point of failure.

Contrasting Deployments

GVSAN relies on a proven architecture, backed by over 10 years and over 20,000 users, ideal for film, TV, video, and effects post production workflows. XSAN architecture, attempted by many other products, including StorNext, SANergy, Unity and so on has proven time and again that NAS extension to the SAN problem is not scalable to very demanding video workflows. Data centers and high-performance computing are also categories ideal for SAN implementations. GVSAN is available at a fraction of the cost of an XSAN solution.

Post-Production Workflows

The film, broadcast, and video post-production customers were early adopters of SAN technology because of its ability to move huge amounts of data to multiple workstations. Gone are the dreaded days of Sneaker-netting to collaborate on projects. The SAN workflow simplifies everything:

• **Concurrent access to source files**. Multiple editors can access the same media simultaneously from the same storage. Storage file systems are shared by directly attaching to each computer via the SAN. With GVSAN, volume-locking ensures that only one stations at a time can modify the contents of a particular file system. With XSAN, file-level locking is supposed to ensure that only one station at a time can modify the contents of the file, but this doesn't work—take the Final Cut Pro case: With XSAN and Final Cut Pro, multiple editors can all open the same project file and do work. Without any warning XSAN allows this, resulting in all editors losing all their Final Cut Pro editing work in the project file, save for whoever saves their file last! This can never happen with GVSAN; therefore with GVSAN the editors would not lose their work. Since GVSAN uses disks and file systems in a truly standard way, it works in environments that XSAN cannot, including AVID audio and video systems!

• **Consistent high-bandwidth performance**. SANs are designed for high-bandwidth real-time applications including, audio and video where frame drops are not acceptable. To control bandwidth to a single resource amount multiple potential users, GVSAN lets the user exclusively control all the bandwidth to a particular file system to enable all the bandwidth available to that storage resource. XSAN on the other hand provides complicated APIs to reserve bandwidth for a particular SAN computer.

• **Storage and server consolidation**. Server and storage are consolidated in a SAN via the FC fabric where they have the ability to connect to each other without need to deal with physical connections. This reduces overall cost by allow efficient use of storage and reducing unutilized storage (less storage purchases). Centralized management of server and storage use reduces time to reconfigure applications, projects, access control and allows IT professional to manage more for the same effort.

Nomadic Storage

This high-performance, high-availability storage system delivers data protection and enormous capacity—up to 48 terabytes— at a groundbreaking price.

Standard Bandwidth requirements

When designing a SAN it is useful to translate what you are doing into hard numbers that can be utilized to create a SAN design supporting the flow of data. Here are some common standard data requires for different streams of media:

• **Audio:**
- Less than 15 MBps for hundreds of tracks,
dual-channel 44-192 KHz 24-bit audio.
The requirement here is seek time and
low-latency for small I/O, which is why
15,000 RPM FC drives are ideal.

• **Standard definition:**
- 3.6 MBps for MiniDV, DVCAM, and DVCPRO
- 7.7 MBps for DVCPRO 50
- 20 MBps for uncompressed 8-bit
- 27 MBps for uncompressed 10-bit

• **Compressed high definition:**
- 5.8-14 MBps for DVCPRO HD

• **Uncompressed high definition:**
- 46 MBps for 720p 24 fps
- 50 MBps for 720p 30 fps
- 100 MBps for 720p 60 fps
- 98 MBps for 1080 24p (8-bit)
- 120 MBps for 1080i (8-bit)
- 110 MBps for 1080 24p (10-bit)
- 165 MBps for 1080i (10-bit)

• **Tremendous capacity**. With SANs it is easy to work with and share huge amounts of SAN data from terabytes to petabytes making media storage a breeze.

• **Dynamic scalability**. Fibre Channel SAN architecture is designed to scale dynamically. Attaching more servers and storage systems to the Fibre Channel network scale capacity and performance. When expanding the number of computers, simply install each with the GVSAN SAN software and attach them to the SAN.

Misconception between XSAN and GVSAN workflow

There have been many misunderstandings about the difference between a file-level and volume-level architecture and how workflows might be different. For example, in a Final Cut Pro editing group using XSAN a customer will probably create a single large volume and have different folders for each users work area, and perhaps a common folder for some shared areas. With GVSAN, instead of folders, the different users can have access to shared volumes in addition to their own volumes (controlled by built-in project management software). In an XSAN world with multiple project editors, there is a belief that multiple editors can work on the same project file at the same time. **This**

is not true. If multiple editors open the same Final Cut Pro project file and begin working in that project, nothing that each editor does will be seen by any of the other editors. In fact, when each editor saves or auto-saves the project file, only the last editor who saves the project file will retain any work that was done. The work of all the other editors will be lost forever without even a warning! This is because XSAN allows multiple users to open and use a Final Cut Pro project file, and when an editor saves what they are doing only that data gets written to disk, overwriting the old file. The next editor to do a save will overwrite the last editors file. So when everyone quits and reopens the same project, only the project from the last editor that saved will actually exist! Therefore, to accomplish multi-user editing with XSAN each user must have a separate Final Cut Pro project that the other user will not accidentally overwrite. **This is prone to human error** as there is no warning from the software preventing this from happening. To finish a project the users have to very carefully coordinate their actions to allow their projects to be combined a piece at a time. In contrast, with GVSAN, this situation is impossible to happen. GVSAN prevents multiple users from opening a Final Cut Pro file with write access at the same time. GVSAN does allow Final Cut Pro to safely open projects read-only by multiple editors at the same time. GVSAN allows easy and safe automatic coordination to allow multiple editors projects to be combined easily without fear of data loss.

Now back to the workflow misconceptions—because XSAN requires multiple users and editors to have their own working directories and projects (in order not to lose data by accident), there is in fact no workflow difference between XSAN's multiwriter file system workflow and GVSAN's multi-volume workflow. Backup, project consolidation, and media consolidation is identical. In fact, GVSAN prevents human error, such as that described above, in addition to providing carefully designed Project Management built into the system designed specifically for video/audio editing professionals. XSAN has no such ability.

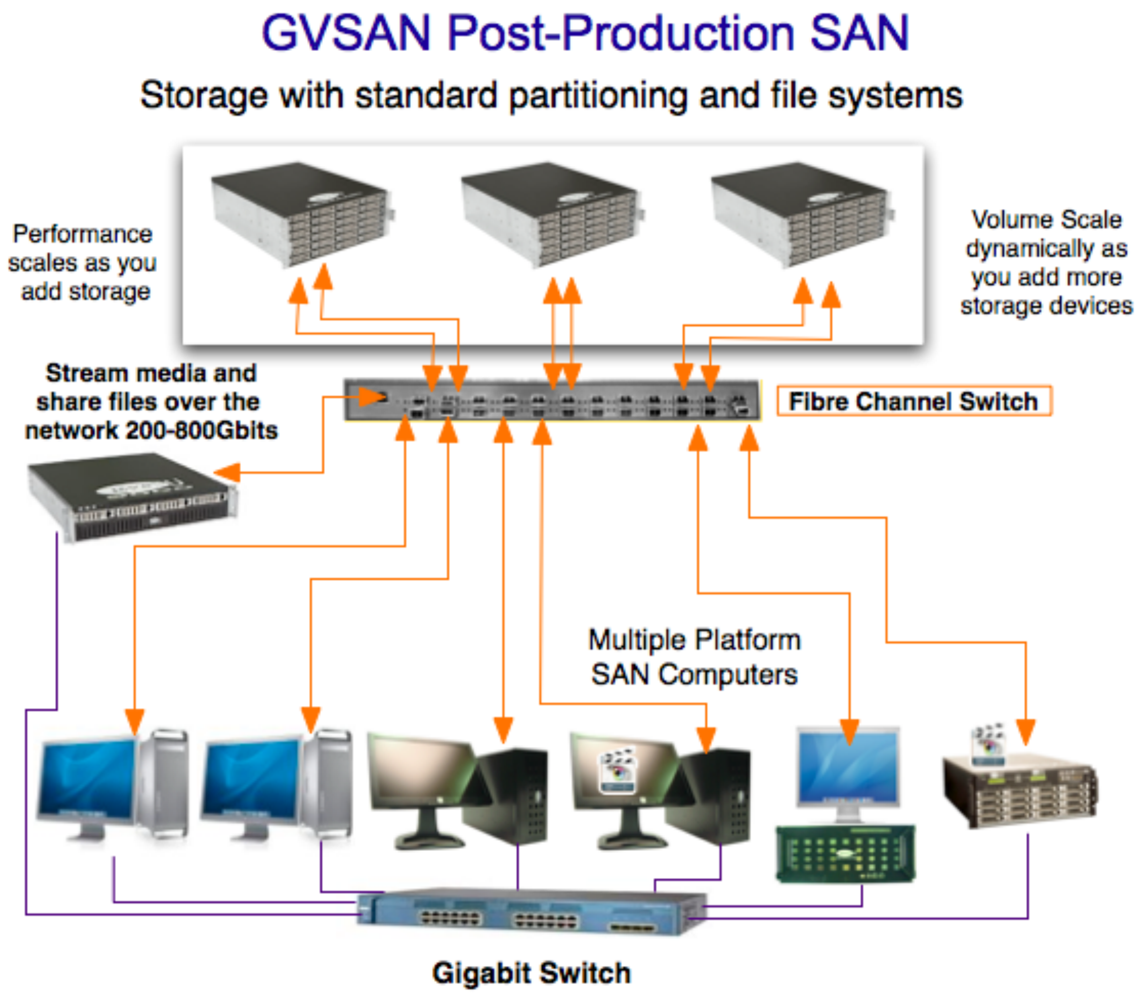


Diagram 15: GVSAN Post-Production SAN

Data Center SANs

IT departments can benefit greatly by deploying a SAN. Consolidating server and storage resources reduces unutilized storage and allows IT personnel to manage much more resources with the same effort. Sharing single copies of data, and allowing on-the-fly reconfiguration of the storage resources allow IT personnel to react quickly to changing demands. These time and money saving benefits include:

• **Consolidated storage resources**. Consolidation of server and storage assets under a high performing and scalable SAN architecture allows centralized management, elimination of unutilized storage, and dynamic redeployment of storage assets among servers.

• **Simultaneous data delivery**. Some ideal data center applications for SAN include home directory services, mail services, web services, and certain database services. For example, the ability to share this data among multiple servers is ideal for scaling distributed delivery of fixed-content html pages from a website among many servers.

• **Increased uptime**. Using SANs with multipathing, redundant Fibre Channel fabrics and RAID protected storage with truly redundant components will help keep the storage infrastructure and applications that use it up more. Upgrading the SAN with more capacity or more computers can all be done without taking systems offline.

• **Cross-platform integration**. Using NAS sharing, your storage can be shared with others on the Ethernet network, including Windows, Linux and other UNIX systems.

Cluster Computing

Computation clusters are ideally suited for the SAN environment. This is because the clusters often share a data repository, which can be placed on the SAN and shared among all the computers in the cluster for processor-intensive work. Each computer accesses the shared data sets and writes its results back to the storage where it is combined into the final result.

Using GVSAN there are no bottlenecks to the computational cluster, as exist when it is done traditionally over NAS. In contrast, XSAN can provide some of this as well, but because its architecture is only an extension to NAS it will eventually be limited by the same network file system techniques, thus limiting overall scalability.

Deploying small clusters over Fibre Channel

With a small cluster in which all cluster computer nodes are directly attached to the Fibre Channel fabric, each node can read from the same data set, perform its calculations, and write its results back out to the SAN. With GVSAN the data-set can be on a single (or multiple) file system volume, accessed read-only by the inputting cluster nodes, the results can be written to a file system owned by each of the cluster nodes.
With GVSAN your cluster can scale in this fashion easily to hundreds of nodes, actually only limited by how large a Fibre Channel network can be built. Also with GVSAN, performance will not decrease as nodes are added, as will happen with XSAN and NAS.

Deploying larger clusters with Fibre Channel and Ethernet

In larger clusters, where it is impractical or too costly to put all the nodes directly on Fibre Channel, the SAN attached computers can share the data sets to a larger number of compute nodes located over Ethernet. All the nodes then perform just as before; computing and writing their results back to the SAN (either directly or as a result of file sharing over Ethernet).

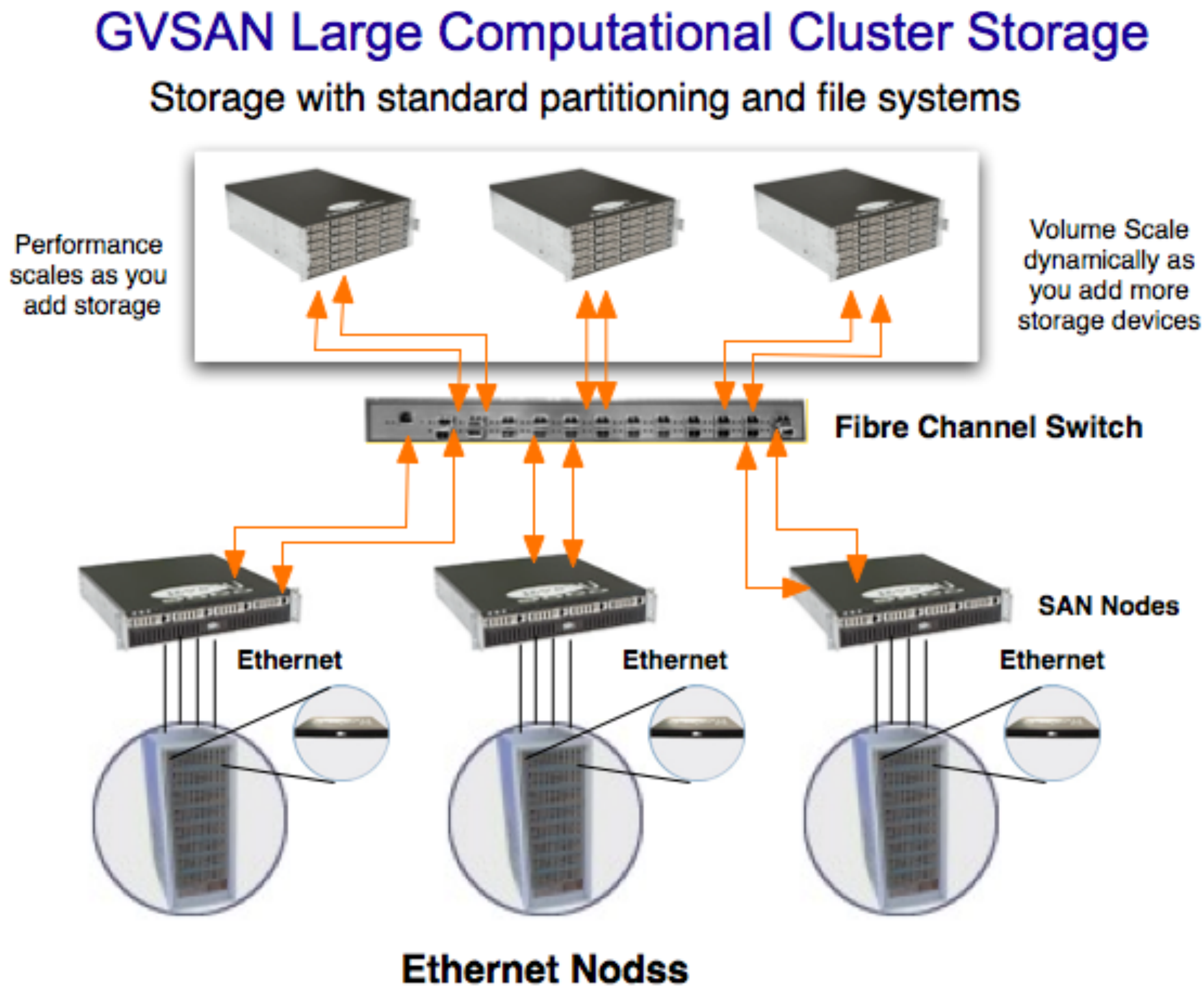


Diagram 16: GVSAN Large Computational Cluster Storage

Contrasting Support and Training

Both GVSAN and XSAN come with 90 days of toll-free telephone support for installation, launch, and basic troubleshooting.

Full support per seat

XSAN major software upgrades for three years a. AppleCare XSAN support for one year offering 24/7 telephone support covering the XSAN GUI only. Therefore 3 years full **XSAN support**

Additional Hardware Requirement

a. **2nd** Gigabit card pre XSAN users
b. Additional Gigabit switch non-manage
c. Metadata Controller 2nd Xserve w/ software support
d. **2nd** Metadata backup controller
e. Independent RAID volume on SAN for metadata file
f. 4xaddition fiber port for metadata controllers
g. 4xFiber cables and GIBIC for Metadata controllers

Total Apple hardware cost for proper configuration of XSAN:

Total GVS hardware cost using GVS90001U for Metadata controller:

The total cost saving of using GVS hardware

Training and Certification

GVS provides an instructor-led SAN courses, as well as related certification programs at GVS San Francisco Training Centers or, by special arrangement, onsite at your business or institution.

Conclusions

GVSAN offers many advantages over XSAN.

Cost

XSAN costs more than GVSAN for a small SAN. **See *Contrasting Product Overviews* starting on page 24.**

XSAN also requires the additional cost of a GB Ethernet network. It is recommended that this be isolated and separate from any other existing network and dedicated to the metadata traffic.

XSAN requires 2 machines for the function of Metadata controller (one is for failover). It is best these machines are dedicated to just this purpose as it has been shown any interruption to the metadata controller will hang every system on the SAN stopping and possibly losing everyone's work in progress. This is unacceptable in a professional video-editing network where large amounts of money are being spent in post-production.

Performance

GVSAN performance is drastically superior to XSAN on many fronts. As the tests illustrate GVSAN outperforms many times what XSAN can do on the same storage, sometimes over 500% the performance of XSAN! If you want to do more work on the same amount of storage, then GVSAN is clearly the choice. This is especially true for video editing professionals where we demonstrated the ability to stream 13 additional uncompressed SD streams (or about 351 megabytes per second) more than XSAN. **See *Performance Comparison* starting on page 38**

Ease-of-use and workflow

GVSAN is safer and easier to use than XSAN, especially in a video editing environments with multiple editors. XSAN makes it is very easy to overwrite data from multiple users and lose work. **See *Contrasting Deployments/Misconception between XSAN and GVSAN workflow* starting on page 36**.

GVSAN works with third party equipment including HBAs and storage that is not supported with XSAN. GVSAN also uses standard partitioning and formatting utilities, making storage more usable, maintainable, and repairable. XSAN's awkward use of entire LUNs to form volumes greatly reduces the customer's choice as to how the storage will be utilized throughout its lifecycle. **See *GVSAN Product Overview/Key Features* starting on page 20, *Contrasting Product Overviews* starting on page 24 and *Contrasting File System Capabilities* starting on page 27.**

Availability and Scalability

XSAN introduces more single points of failure into the SAN as compared to GVSAN. **See *Contrasting Product Overviews page 24 /Failure Points* starting on page 20.**

GVSAN can scale in computer, capacity and performance linearly. As XSAN scales the performance steadily decreases, until it drastically slows down. **See *Contrasting File System Capabilities* starting on page 21 and *Performance Comparison* starting on page 38.**

Since XSAN relies on a single metadata controller as traffic-cop to control the entire SAN, it does not scale very well and is a single point of failure to the entire SAN. This will limit its use to small SANs and customers that do not require a growth path, do not need much performance, and do not mind that they are only using a fraction of their SAN storage performance potential requiring the purchase of much more storage to accomplish the same task (more cost). These users must also not mind the human-error that can easily lead to data-loss in the XSAN environment. **See *SAN Background/SAN File systems – Proprietary Client / Server Architecture allowing multiple-writers to file system at a time* starting on page 11, *Contrasting Product Overviews* starting on page 24, *Contrasting File System Capabilities* starting on page 21, *Performance Comparison* starting on page 38 and *Contrasting Deployments/Post-Product workflows* starting on page 35.**

Legal Notices, Credit and References