Planning & Implementation of SAN in Heterogeneous Environments

By Dan Carson
VP Marketing and Business Development

Open Systems Solutions, Inc.
710 Floral Vale Boulevard
Yardley, PA 19067

Tel: (215) 579-8111
Fax: (215) 579-8113
E-mail:danc@ossi.net
Web: www.ossi.net

The advent of the Storage Area Network (SAN) has captured the attention and imagination of nearly everyone associated with the creation, implementation and management of data storage devices and data center storage infrastructure. From small graphics houses to the largest multi-national corporations, the promise of simplified administration, improved performance and reliability, lower capital costs, and greater scalability are all compelling reasons to investigate this new storage architecture.

Storage Requirements Expanding Across the Board

Driven by the veritable explosion of data fueled by the Internet, e-commerce and e-business, not only are data centers under pressure to store massive amounts of data, they are also responsible for managing its availability, securing its accessibility, and protecting it from disaster. Information is the lifeblood of almost every organization. The effective, efficient storage and retrieval of this collective corporate knowledge is an absolute business necessity.

The need for increased storage efficiency cuts across virtually all industry segments. Pharmaceutical, financial, insurance, healthcare, telecommunications, government, aerospace, manufacturing, retail, broadcasting and video and Internet services are each experiencing rapid expansion of their information capacity requirements.

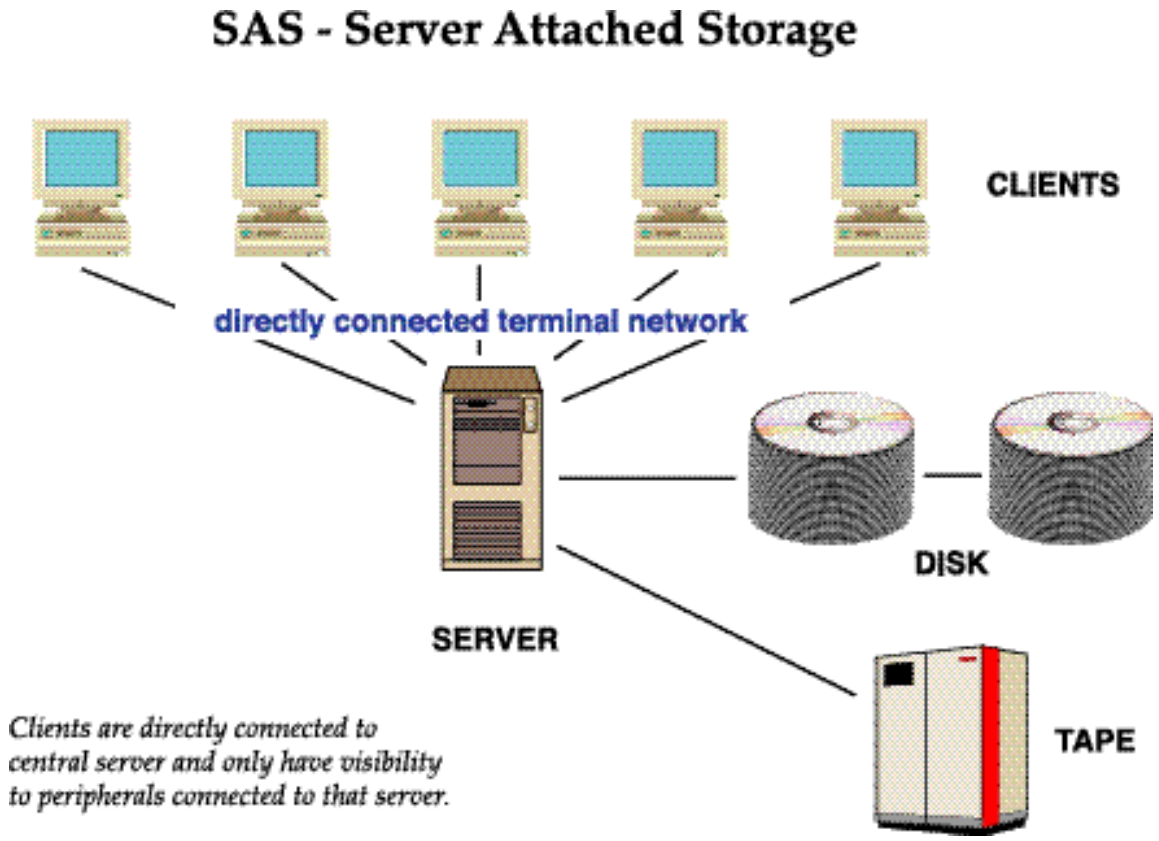
Quick Overview of Storage Options

Before taking a closer look at the planning and implementation of SAN, a brief review of historical storage trends and environments will help set the stage for understanding the issues involved in building a SAN infrastructure.

In general, companies implement storage infrastructure in one of three ways.

SAS (Server Attached Storage)
Historically, this is the oldest storage infrastructure arrangement, but it is still in use, particularly in mainframe data center environments. Essentially, disk drives and tape libraries are attached to the central server. All file transfer and read/write activity require processor cycles on the central server, as do tape operations. Visibility to storage is limited to those workstations directly attached to the central server.

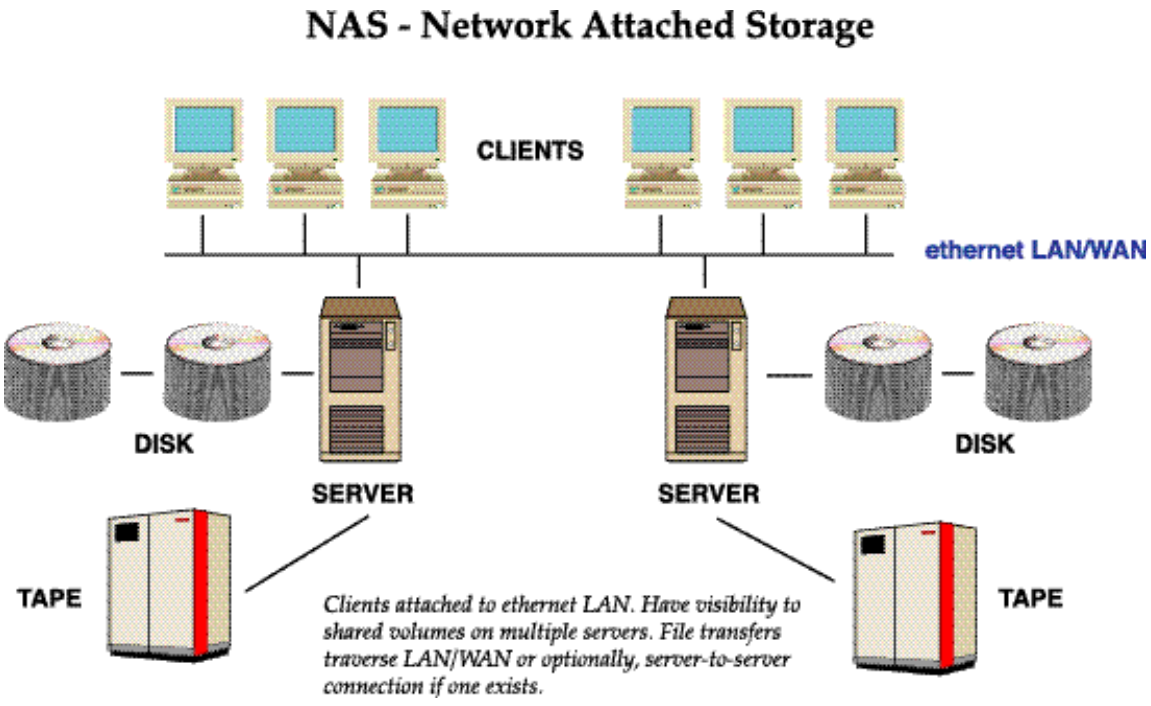
Another form of server attached storage, familiar to many consumers, is the storage associated with a standalone, non-networked personal computer. While the average computer user would not even think to categorize their disk and tape units in such a fashion, nonetheless, they do constitute a SAS infrastructure.



NAS (Network Attached Storage)

This storage infrastructure is the most common implementation in business today. The rapid development of distributed computing environments, enabled by substantial advancements in networking technology and desktop computing allows storage devices to be distributed across servers and desktops throughout an enterprise. Interconnected by LANs and WANS, desktop workers can choose to share all or parts of their disk drives with others on the network. They also have visibility to a wide variety of devices whose physical locations can be virtually anywhere. File transfers and read/write activities can occur from storage attached to any desktop or server to any other connected to the network.

The limiting factor in NAS environments is the speed and bandwidth of the network connections. File manipulation and I/O activities between storage devices attached to different CPU's must traverse the intervening network. These activities also incur processor overhead on a server or desktop computer (and often, on both).

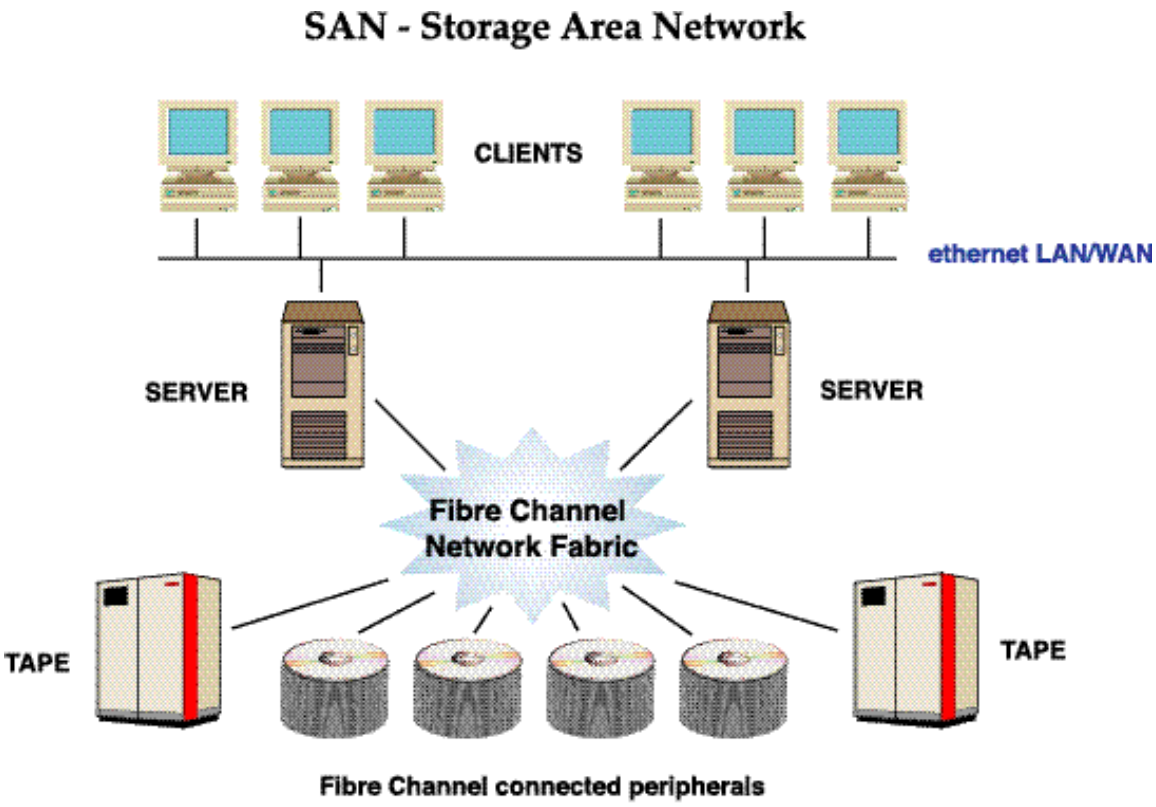


SAN (Storage Area Network)

The difficulty and cost of managing highly distributed computing environments has triggered a significant trend toward centralized disk and tape storage. Accompanying this trend has been the evolution of a storage infrastructure that takes advantage of new, high speed interconnect technologies as well as the increasing ability to add intelligence to storage devices themselves.

SAN provides the ability to create a centralized storage pool, where the storage devices themselves reside on their own network. This network is distinct and separate from the corporate LANs and WANs used by employees and sits 'behind' the servers. The primary objective of SAN is to remove the 1:1 server-to-storage paradigm that exists in both SAS and NAS infrastructures and allow any server to see and use any storage device on the storage network.

Another major difference is that file transfer and manipulation activity can occur between storage devices without traversing the 'employee' LAN or WAN. In fact, other than processing the command to execute a file copy, for example, the server(s) involved are not required to use processor cycles.



Clients have same visibility to shared storage as in NAS. File movement can traverse fibre channel network "behind" the servers without using the client LAN/WAN.

Why SAN is the Future of Storage

Given that central storage is a rapidly developing corporate strategy, SAN infrastructure provides a number of performance and cost-saving advantages:

Increased Configuration Flexibility
Storage (both disk and tape) is shared among numerous servers and can be dynamically or manually allocated as needed to address changes in storage requirements. Addition of storage capacity is not necessarily accompanied by the requirement to add new servers, and vice versa.

High Availability
Combine the configuration flexibility noted above with RAID disk technology and automated tape libraries and new levels of data availability can be achieved.

Throughput Performance
Fibre Channel connectivity, running at 100MB/s, rivals all but the latest Ultra3 SCSI protocols. New Fibre Channel standards will push transmission speeds further in the near future.

Protection
Fibre Channel is a key factor in storage protection due to its ability to support significant distances between nodes – up to 10km. The high bandwidth of Fibre Channel also promotes quick duplication of data. This capability can either supplement or replace RAID mirroring data protection.

Reduced Management
Because storage can be consolidated physically, building and data center costs can be minimized. Also, with storage in a central location, administration, configuration management, and maintenance can be conducted from one place, with fewer staff resources. Studies have shown that in a distributed storage environment, fully 55% of overall IT budget costs are related to ongoing storage management. In a centralized storage deployment, such as SAN, these costs drop to 15% of the overall budget, a 40% savings.

Lower Total Cost of Ownership
A number of factors contribute to overall lower storage costs when a SAN infrastructure is implemented. In addition to the lower management and administration costs, elimination of the 1:1 server-to-storage-device limitation means that additional server and storage components can be added independently of each other, thereby minimizing capital costs. Finally, the ability of SAN to operate across heterogeneous hardware and software platforms means optimal pricing for companies investing in SAN infrastructure.

Key Components of SAN Architecture

SAN is comprised of the following elements:

Network Fabric
The fundamental principle of SAN is that storage elements can reside on their own network. This network is made up of components similar to those in a traditional LAN or WAN; hubs, switches, and host adapters. The difference between an ethernet LAN and a SAN is the Fibre Channel connectivity protocol. It is the key to SAN infrastructure. In fact, one can't talk about one without including the other. Each component of the network fabric must be Fibre Channel capable.

Disk Storage
Many SAN implementations will also take advantage of RAID disk, configuring it either for performance (striping) or protection (mirroring). However, RAID is not required for SAN implementation. Some companies may choose to satisfy their disk capacity requirements using the old standby JBOD configuration (just a bunch of disks). SAN can also work with other types of disk storage media, such as CDROM, Magneto-Optical, and DVD.

Tape Storage
Like disk storage, tape backup and archiving can come in several forms. Depending on the company's needs for long-term storage capacity and quick recoverability, they may implement individual tape drives or turn to automated tape libraries.

Planning for SAN is Critical
Building a Total Solution

Interoperability – The big question
Before delving into the key elements of planning for a SAN infrastructure, it should be noted that this paper assumes that open standards solutions that support heterogeneous hardware and software is where the SAN market is headed. It is not the intent of this presentation to explore the issues involved in the sometimes contentious discussion of open standards solutions versus single-vendor proprietary storage solutions. Three general industry observations lead to the position taken by this paper:

1) Companies want choices in storage hardware and software
2) Several standards groups are working diligently to create specifications that promote interoperability among heterogeneous SAN hardware and software
3) There are already heterogeneous SAN solutions available, and both vendors and integrators seem committed to continued efforts to expand the interoperability of these solutions.

The Big Picture – Phased Approach versus Complete Conversion
Companies considering a SAN solution take one of two general paths to SAN implementation. Depending on budget, staff, risk assessment and competitive pressures, some organizations elect to implement SAN in smaller increments, perhaps converting a specific department or particular service, like e-mail, to a SAN environment.

Clearly, there are benefits to the phased approach – it affords the company a 'test case' in which to review such factors as hardware and software interoperability, performance, ease of management and administration, vendor support and other related functions. At the same time, capital costs can be minimized. The downside to the phased approach is that the company now has two separate storage environments to manage, which may cause confusion and will not result in any immediate administrative savings. Nor will the enterprise realize the overall performance and availability improvements of SAN. As such, the company is reducing their return-on-investment potential. The other risk is that competitors may gain an advantage by completely migrating to SAN.

Moving all enterprise storage into a SAN infrastructure is the other implementation alternative. The risk may be perceived to be greater, depending on numerous business factors, but the rewards of quickly deriving the benefits of increased operating efficiency, lower costs and improved competitive profile may outweigh these risks. For companies choosing this route, perhaps more effort is required in both the planning and implementation activities to ensure an effective transition without disruption.

In either implementation case, the move to a SAN solution represents a substantial change in the way companies structure their data storage. Therefore, it also represents an opportunity to conduct a thorough audit of corporate storage strategies, policies and practices. At the same time, the company can compare their parameters against industry Best Practices.

Best Practices for storage solutions encompass the following areas:

- Performance
- Availability
- Backup
- Archiving
- Disaster Recovery
- Scalability

The assistance of a third-party storage solution specialist can be invaluable in the assessment of a company's current storage infrastructure vis-à-vis Best Practices. Storage specialists can also be critical partners in the planning and implementation process, particularly given some of the complexities and issues surrounding SAN, such as interoperability of heterogeneous components. A good independent storage specialist brings experience across industries and platforms and therefore has a finger on the pulse of what the best, most successful companies are doing in the storage arena.

Storage Audit – The Key to Successful Planning

Storage consolidation is not an easy task. Before any SAN design or implementation steps are developed, the first job is to assemble a comprehensive picture of current storage capacity and infrastructure, as well as a thorough review of anticipated storage needs based on business plans and projections.

The following areas should be closely examined and analyzed with an eye toward how a SAN infrastructure will integrate with or replace related elements:

Storage Capacity
Both disk and tape storage capacities need to be reviewed. Use data from other departments of the company to assess storage growth rates. Determine whether any special initiatives will require a spike in storage capacity needs.

SAN considerations: Unlike previous storage environments, storage capacity can and should be viewed as a large pool, accessible by any server, as configured and allocated by the storage administration staff. Review and assess capabilities of SAN management tools.

Servers
Inventory current servers for information such as OS and release level, file system, file structures, data formats, attachment protocols (SCSI, SCSI 2, UltraSCSI, etc.). Document usage (email server, web server, DB server, application server etc.). Also catalog LAN/WAN connectivity, note usage patterns and document who uses the server.

SAN considerations: Review Fibre Channel connectivity issues, think about how specific data or groups of data currently assigned to specific servers might be better reconfigured within a storage pool.

Network
Review current network topology and protocols. Analyze current performance levels to help set baselines against which SAN performance can be measured. Document current geographic distribution and network connectivity.

SAN considerations: Look at how separate storage network will be implemented and assess impact to existing network infrastructure. Determine whether existing SCSI-based equipment will be bridged or routed into the new Fibre Channel SAN. Assess best Fibre Channel topology (point-to-point, arbitrated loop, switched, with or without cascading, or a mixture of these topologies). Depending on choice, assess equipment requirements, such as hubs, host adapters, switches etc.).

Application/DB
Review which server(s) host which applications and databases. Review and document current performance versus required performance.

SAN considerations: Comprehensive assessment helps to accurately determine the overall number of users accessing applications and database and helps to define usage patterns. This information is an important criterion in determining the type of SAN implementation architecture that will be implemented – arbitrated loop or switched. This, in turn, directly affects the cost of the SAN implementation (switched environments are more expensive, but provide better use of Fibre Channel bandwidth).

Security
Review current server, application and database environment to document user access rights.

SAN considerations: Assess SAN management and administration software and determine how security will be migrated from current system and subsequently administered.

Backup/Archiving
Hand-in-hand with disaster recovery, backup is a critical element that is sometimes overlooked in the overall storage environment. With many companies moving into 24x7 environments, backup windows are shrinking therefore requiring improved performance. Speed of restore in the event of a failure is also critical. A thorough review of current backup procedures, equipment and capacity is an absolute must.

SAN considerations: Determine interoperability and performance characteristics of existing tape devices in a networked storage environment to guide equipment requirements. Like disk, SAN makes it possible to share tape units among multiple servers. Review backup administration software for flexibility and interoperability. One of SAN's big advantages is the promise of serverless backup, which can improve backup performance dramatically. Will the SAN solution being considered support this capability.

Disaster Recovery
Review current procedures and plans and re-assess disaster risks, if necessary. Is an off-site backup maintained?

SAN considerations: Fibre Channel, with its 100Mbs speed and 10km distance capabilities offers new flexibility in disaster recovery planning. Remote disk mirroring may become a viable disaster recovery alternative or supplement to off-site tape archiving.

Data Migration
After deciding to implement a SAN, data will need to be moved from the current production storage environment to the new SAN storage configuration. Planning for the orderly migration of this data so as to minimize disruption of the production environment is crucial. Test cases should also be established to ensure that applications and databases are accessible and functioning properly following data migration.

SAN considerations: Develop mapping of logical and physical data location in the SAN for each specific set of data involved. With SAN, data no longer needs to be attached to the server to support application(s) or database(s), which is a new way of thinking. Determine timing and hierarchy of data migration from old storage system.

Implementation Issues

Pre-Install Testing
Because an array of new hardware devices and software required by SAN, the optimal situation for a company is to have their vendor(s) or storage solution specialist independently model the implementation in a lab environment and pre-test all components for interoperability. The capability to do this should be factored in to the company's choice of SAN solution vendors.

Parallel Operation
As with any significant IT infrastructure change, it is advisable to run the old and the new system in parallel for a period of time. This gives operations and administrative staff a chance to become familiar with new equipment and procedures. At the same time the data migration strategy and tactics can be tested. IT staff can also monitor storage performance characteristics prior to moving into full production.

The following implementation concerns can all be monitored and refined as part of pre-install testing and parallel operation phases:

Physical Location
Depending on the size, geographic distribution and disaster recovery policies of a company, decisions must be made about whether to consolidate storage to a single facility or to multiple (but fewer) facilities. These decisions directly affect implementation tactics. In general, multi-facility implementations require more coordination and testing due to data migration and inter-facility networking issues.

Networking and Device Connectivity
With Fibre Channel as the connectivity protocol, a new set of adapters, hubs, routers and bridges are required. Interoperability is critical in a heterogeneous environment, so equipment should be pre-tested, if possible, particularly where existing devices are being adapted to operate via Fibre Channel. Network performance should be tested and verified during a parallel operation phase.

Device Allocation & Configuration
Properly establishing physical and logical device partitions, assigning access rights, configuring RAID devices and validating performance criteria are all critical elements of the SAN implementation. During parallel operation, administrative staff can familiarize themselves with any new software components and administrative procedures required by SAN.

Training
The best time for vendors and storage specialists to effect knowledge transfer is during a parallel operation phase. Not only can IT staff be brought up to speed on the new system, but with the old system still in place, comparison between old and new processes and procedures is easier. Additionally, it provides the company's IT staff a final opportunity to confirm that no critical elements of the old system have been left out in the migration.

The Bottom Line

A SAN storage infrastructure offers substantial short and long-term benefits for companies engaged in storage consolidation efforts. However, the creation of a true storage network, one that sits 'behind' the companies servers separate from the end-user LAN, and all the attendant considerations of storage consolidation can be complex. Environments with a mix of storage devices, servers and operating systems require extra attention to ensure a smooth transition to a SAN architecture. Thorough planning and testing of all storage related issues is essential.

Without question, planning for SAN is similar to planning for any major system upgrade. The key difference is that the way storage is managed and used in a SAN architecture is different than the way storage has been handled since the inception of computers. It represents a whole new paradigm and requires a new way of thinking about storage. For those who grasp this difference, the rewards are many.