A Storage Area Network Analysis and Design Methodology

By

Alicia Mackey

Table of Contents

Table of Figures

Abstract

Although SANs are becoming a more and more popular means to solve storage needs, a standardized methodology to evaluate the needs of companies wishing to implement a storage solution does not exist. This SAN Analysis and Design Methodology creates a SAN analysis and design that could not only be used to create the most cost effective SAN solution that will work best for every unique scenario, but it could also determine if a SAN is indeed the solution to a company's storage needs or if a Network Attached Storage (NAS) or Direct-Attached Storage (DAS) solution might be preferable.
This matrix will enable an organization to determine the following five building blocks that should be assessed to determine "how the supported applications affect the entire data environment," as well as the Interface, Interconnects, and Fabric. [1]

- Application requirements
- Data storage requirements
- Backup and disaster tolerance
- Server connectivity requirements tolerance
- Performance and growth

To keep the length of this document reasonable and I will not elaborate on physical definitions of SAN Interfaces and Interconnects. This is a high level overview of technologies used in building SAN Infrastructures.

Why SANs?

E-business has caused explosive growth in data storage needs through data warehousing, e-commerce, and Enterprise Resource Planning (ERP). These applications demand 100% data availability to customers and employees. Government regulations require "high levels of system and data backup" for healthcare, banking, and financing industries. [6]

Data is now one of the most valuable assets a corporation can have. In assuring its availability, integrity, and disaster recoverability, storage is now the most important IT resource for these mission-critical components. [19] The loss of data availability and integrity can cause severe and sometimes fatal damage to a corporation. Today, data must be readily available to applications as well as departments within corporations. Storage networking is now a strategic component of the IT infrastructure. [19]
Storage Area Networks (SANs) have gained recognition as the potential solution to ever growing data storage needs. "A SAN is a dedicated configuration of multiple servers connected to peripheral storage devices using high-speed fibre and special routers, switches, and hubs" (McData, 2001). SANs enable the consolidation of data from disparate servers to be managed centrally on a storage network. SANs offer key advantages over other storage solutions like NAS (Network Attached Storage) and DAS (Direct Attached Storage). Remote backup and recovery, increased uptime, more efficient data sharing and storage expansion are three key

advantages in implementing a SAN.
A SAN is a specialized network that enables fast, reliable access among servers and external or independent storage resources. [21] SANs use similar interconnect technologies as Local Area Networks (LANs) or Wide Area Networks (WANs) to interconnect its devices: routers, hubs, switches, and gateways. SANs can be used to:

- Connect clustered servers for failover
- Interconnect mainframe and tape resources to distributed network servers
- Connect shared storage arrays
- Create parallel or alternate paths for high performance computing environments.

Enterprise management is now faced with the dilemma of storing critical data that is ever increasing and keeping it secure but available to multiple users, backing it up without driving up cost, and ensuring that the storage hardware and software are interoperable. SANs reduce costs of storage resources, but the data on shared storage is potentially exposed to unauthorized access. [20] There is "heightened concern about security and the ability to prevent unauthorized access to data, although security should be enforced and implemented in a SAN like any other network. [19]

Until recently lack of industry standards posed a challenge to the security of SANs. The Storage Network Industry Association (SNIA) Storage Security Industry Forum (SSIF) is now working towards "increasing the availability of robust storage security solutions" through a collaboration of storage vendors. [16]

Network Attached Storage (NAS)

NAS solutions are typically configured as file-serving appliances accessed by workstations and servers through TCP/IP (Transmission Control Protocol/Internet Protocol) and applications such as NFS (Network File System) or CIFS (Common Interface File System). [3] NAS connections are mostly between workstation clients and the NAS file sharing facility. The connections rely on the existing corporate network infrastructure to operate correctly. NAS solutions are relatively easy to deploy, but have key performance constraints: network congestion of the existing LAN/WAN directly affects NAS performance. Adding additional NAS components is easy, but accessing individual components as one entity is not possible because NAS components have unique identifiers. Data backup in NAS usually cannot be centralized which limits it to direct-attached devices such as tape libraries or dedicated tape drives. Key NAS strengths are it works well in environments where data is transferred over great distances and it works well for organizations that deliver file data to multiple clients over a network.

Direct Attached Storage

Direct Attached Storage is defined simply where each server has dedicated storage. "Storage is seen and accessed by one host system and, should another host system need more storage that host will have to add more physical storage and perhaps I/O interfaces or host bus adapters." [5]
Typically, this setup incorporates different flavors of storage devices and poor storage utilization. Direct attached storage (DAS), in the form of independent drives, RAID arrays or tape libraries, is the most common storage architecture today. [5]
DAS is the classical storage connecting concept, with which a memory unit is assigned to a server directly. DAS solutions are characterized by low costs, high data security and simple handling. DAS solutions offer great investment protection for already existing server environments.

SAN Physical Architecture

SAN Interfaces

SANs are comprised of interfaces, interconnects, and fabrics. Interfaces allow storage to be accessed externally and have the ability to host shared storage configurations for clustering. SAN interfaces come in a number of different flavors: ESCON (Enterprise System Connection), SCSI (Small Computer Systems Interface), SSA (Serial Storage Architecture), HIPPI (High Performance Parallel Interface), Fibre Channel, i-SCSI (Internet Small Computer Systems Interface), Gigabit Ethernet, FCIP (Fibre Channel over Internet Protocol), iFCP (Internet Fibre Channel Protocol), or InfiniBand.

SAN Interconnects

Interconnects are devices such as Hubs, Routers, Gateways, and Switches, and Host Bus Adapters. SAN interconnects join storage interfaces into many network configurations and across great distances. There are three broad classifications of Fibre Channel switches. Directors, Fabric switches, and Arbitrated loop switches.

SAN Fabric

Switched fabric, switched fibre, and switched SSA are common SAN fabrics. SANs can be shared or dedicated, local or remote and are composed of unique externalized and central storage and SAN interconnect components (Vacca, 2002). The following diagram is an example of a SAN Application taken from Business Communications Review (Strechay, 2001).

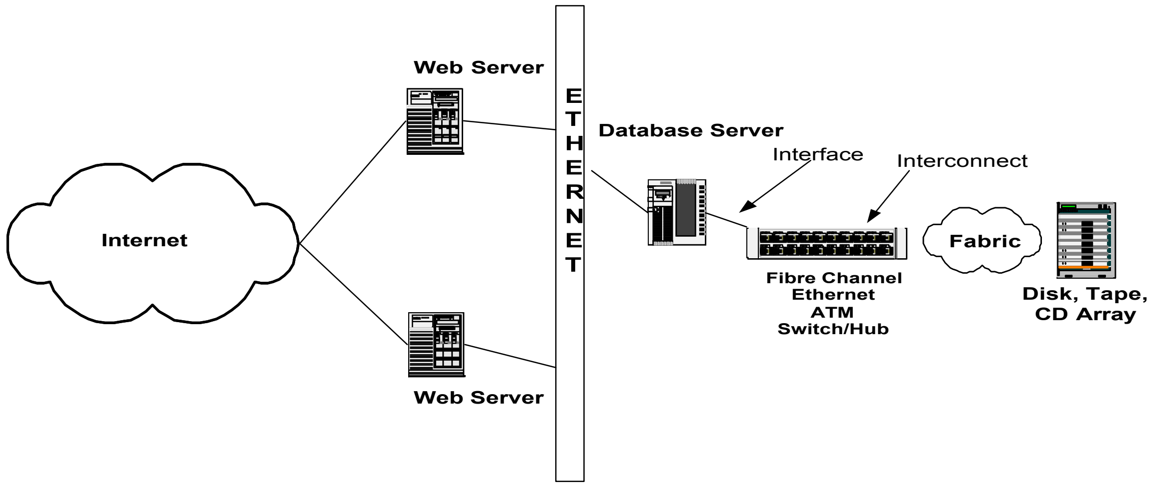

Figure 1 - Storage Area Network

"SANs create a method of attaching storage that is revolutionizing the network because of the improvements in availability and performance". [19] SANs provide better utilization of storage and greater flexibility of a unified storage system
In SANs storage devices are not the exclusive property of any one server. Storage devices are shared among all network servers as peer resources. As recent as the year 2000 SANs were implemented throughout an enterprise as small islands of computing – "mini SANs" is what they were called. These "mini SANs" were incapable of sharing data and resources. This resulted in valuable data being inaccessible or inefficiently deployed. Now SANs are being deployed throughout entire enterprises and what were once "mini SANs," incapable of sharing data, are now entire operations of SANs that result in high data availability, and scalability for the Internet economy. [19] Often referred to as the "network behind the server" SANs can be used to connect server-to-storage, server-to server, and storage-to-storage. [19]
One of its key benefits is externalizing storage from servers and allowing that storage to be shared among multiple host servers without impacting system performance or the primary network (LAN/WAN). [8]

SAN Analysis and Design Methodology Overview

The SAN Analysis and Design Methodology is composed of the following five matrices:

- Server Connectivity Matrix
- SAN Interconnect Matrix
- Data Storage Applications Requirements Matrix

- Backup and Disaster Tolerance Matrix
- Applications Requirements Matrix

Each matrix is outlined as a table of prerequisite elements and questions that must be identified and answered before a correct SAN implementation can take place. Within each section of the matrices, requirements are asked or stated in the form of questions and/or definitions. Each requirement and how it is defined, per matrix, is discussed below.
Using The SAN Analysis and Design Methodology is essentially building a SAN piece by piece. Each matrix is listed below, by title, in the order it must be navigated.

SAN Analysis and Design Matrices

Server Connectivity Requirements

The first step in using the SAN Analysis and Design Methodology is to decipher the storage solution (SAN, NAS, DAS) will be the most effective. This is done by looking at the Server Connectivity matrix and answering these three common key questions:

- How is the data being moved?
- Why is the data being used?
- How much data is being moved?

Answering these three questions will either identify a SAN as the needed storage solution or eliminate it, by choosing NAS or DAS. If the overall goal of a company is to improve the flexibility of storage assets, increase backup speed, storage consolidation, or to implement a dedicated network for storage, then the answers for the above questions will be 'yes' and the user can continue through the SAN Analysis and Design Matrix. If the user answers 'yes' to any of the questions other than the ones associated with SAN, then user should discontinue use of this SAN Analysis and Design Methodology as it was not created to design NAS and DAS storage solutions.
This following matrix serves as the first deciding factor for the use of the SAN Analysis and Design Methodology by guiding the user to the correct overall storage solution. After selecting a storage solution the user must move on to select a SAN Interface and Interconnect(s).

Server Connectivity Matrix

Performance Need	Where is the data being moved?	Why is the data being moved?	How much data being moved?	Recommended Storage Solution

Servers using a dedicated network for data communication between storage devices	• Through a Local Area Network?	• Storage consolidation? Centralized control? • Database shared and heavily used? • Increase speed of backup?	• Several hundred Mb/s? • Large data growth rate forecasted ?	SAN
Internal or directly attached storage server(s).	• Through a Local Area Network?	• Database usage?	Several hundred Mb/day?	DAS
Storage server(s) directly on the storage network.	• Onsite, through a VPN? • FTP?	• Hot backup? • Database usage	Several hundred Mb/day?	NAS

Figure 2 - Server Connectivity Requirements Matrix

SAN Interfaces and Interconnects

This step requires the user to answer questions pertaining to:

- Performance Need
- Scalability
- Distance
- Association
- Location
- Advantage

These five key elements will determine what interface, interconnect, and fabric design will be the best SAN implementation based on the user's needs. The SAN Interface Matrix only uses the Performance Need, Scalability, Distance and Advantage categories. Within this matrix, Performance Need is defined as the speed data will be communicated between storage devices as well as connectivity requirements between storage devices on the SAN.
The Scalability category asks the user about the anticipated number of connected devices. Each Recommended Interface has a maximum number of devices it can accommodate and the user is asked whether the anticipated number of connected devices is less than that number.
The Distance category uses distance metrics in the form of meters or kilometers. Each Recommended Interface's distance is stated as either the maximum or minimum distance it is able to transfer data. That maximum or minimum distance is stated in the form of a question that asks if the user's distance need is equal to or less than the distance metric stated. Finally, the Advantage category gives more insight to advantages of each Interface. I
In this matrix, Performance Need is defined as the summation of the Recommended Interconnects role on a network. The Association category defines the

device/Interconnect most often affiliated with the Recommended SAN Interconnect. Finally, the Location category defines where the Recommended SAN Interconnect is most often located.

SAN Interconnects Matrix

Performance Need	Association	Location	Recommended SAN Interconnect
A central communication point where data comes in from different directions and is transmitted in different directions.	Switches	Gateway	Hub • **Stackable hub**
Part of a network switch, connected to more than one network simultaneously and decides where data needs to be transmitted.	Switches	Gateway	Router • **Edge Router** • **Brouter**
A point in a network that acts as an entrance to another network	Routers	Network entrance	Gateway • **Proxy server** • **Firewall server**
Transmits data from input ports to the output port necessary for it to reach its destination.	Gateway	Backbone and Gateway	Switch • **Circuit switching** • **Packet switching**
Connects storage components to storage resources. Load balancing, storage management, and fail-over.	Hubs or Switches	PCI slot	HBA • **SCSI** • **Fibre Channel**

Figure 3 - SAN Interconnects Matrix

Data Storage Application Requirements

Understanding the location of data, amount of data, and data and access sharing is critical.

- Locations of data – How much data needs to be moved and where will it go? Will it be stored in a storage pool or storage pod? According to O'Brian, the cable distance between storage locations and server locations must be considered. Long distance connectivity will require specific connectivity components that support long distance connections. [15]

- Amount of data - This is a very key factor: How much data will be accessed? This determines necessary network bandwidth and the number of storage connections required. [15]
- Data and Access Sharing – How and how often will the data be accessed and shared? [15]

Data Storage Application Matrix

Performance Need	Differentiating Factors	Data Storage Application
Highly available system composed of multiple storage devices, computers, and redundant interconnections. When one device fails, another is enabled.	• Reduces cost of large installations. • Can be used for load balancing • High Availability • Seamless failover for users	Clustering • IP Cluster • Storage Cluster • Global File System
A centrally managed storage system.	• Reduces administration costs • Easier upgrades	Centralized Storage
A pool of SAN disks that can distribute storage resources on demand and integrate vendor independent storage devices.	• High Availability • Reduces Total Cost of Ownership	SAN Virtualization • Multi-host storage arrays • Host-based LUN masking filters • File System Redirectors • In-band Virtualization Engines • Dedicated Storage Domain Servers

Figure 4 - SAN Data Storage Application Matrix

Backup and Disaster Tolerance

Loss of data due to error in backups "devastates profitability" "A central backup mechanism improves reliability, speeds recovery, and better ensures protection". [13] Assessing how backups are currently accomplished will provide for a "well-founded SAN". [15] This will also dictate what broad connectivity is needed and at what throughput.

In the backup and disaster tolerance matrix the user can select from two categories: Performance Need and Advantage. The Performance Need category is defined as a summation of each Recommended Backup Application and the Advantage category is defined as one or more benefits of the Recommended Backup Application.

Backup and Disaster Tolerance Matrix

Performance Need	Advantage	Recommended Backup Application
Immediate activation of a secondary connection if the primary connection fails without affecting application performance.	• Failover is transparent to the server.	Transparent Failover
If a connection fails the other connection is activated and assumes the workload until the primary connection is re-enabled.	• Utilizes dual or multiple active connections	Dual or Multiple Active Connections
When one connection fails, the other continues to access storage volume. Once the failed connection is enabled load balancing is resumed.	• Utilizes load balancing • Both connections can access storage volume	Storage Path Failover
Data or site replication.	• Rapid Disaster Recovery	Data Mirroring • Symmetric • Asymmetric • Split-Mirror Backup
Backup across the SAN without the use of servers. Data is moved directly between arrays and tape storage.	• Eliminates bottleneck	Serverless Backup
Run data integrity checks without affecting the performance of live data.	• Creates mirror images of active data offline.	Remote Data Replication
Copies of data volume's current state made off line.	• Eliminates performance latency from backups • Enables backups to be started sooner, ending earlier	Snapshot Copy
Grouping of SAN ports into zones that only allow devices in the same zone to see each other.	• Allows servers to share the same storage devices.	Zoning
Minimal performance hits from SAN backup and restore processes.	• Minimal disruption to business processes	Non-disruptive Backups

Figure 5 -SAN Backup and Tolerance Matrix

Application Requirements Matrix

Next, the Application(s) that will be used for integrating and managing the SAN infrastructure will be identified. Application requirements are identified as those applications an organization might need in order to effectively integrate and manage its infrastructure once a SAN is implemented. The seven applications are as follows:

- Enterprise Resource Planning
- Supply Chain Management
- Customer Relationship Management
- Database
- Data warehousing
- Service level management
- Policy-based management

Application Requirement	Application
Integration of departments and functions into one database that serves the entire organization.	**Enterprise Resource Planning**
Collaboration with supply chain members so that management of supply is handled as one (organization and supply chain).	**Supply Chain Management**
Sharing existing data with the necessary personnel throughout the organization. Meeting the needs of existing customers and their demand.	**Customer Relationship Management**
Access a collection of data.	**Database**
Access large, specialized database that extracts and analyzes data using special tools.	**Data Warehousing**
Assurance that services outlined in contracts (SLAs) are being met.	**Service Level Management**
Enforcement of policies established for anticipated issues.	**Policy-based Management**

Figure 6 - SAN Applications Requirements Matrix

Conclusion

Throughout this paper I have identified the elements that I believe are critical in the design of a SAN. Those eight elements are:

- SAN Interconnects
- SAN Fabrics
- SAN Interfaces
- Application Requirements
- Data Storage Application Requirements
- Backup and Disaster Tolerance
- Server Connectivity Requirements
    - NAS vs. DAS vs. SAN

This methodology was created to ease the complexity of analyzing and designing a SAN. It does not give detailed technical information for implementing a SAN. It does however; narrow down possible implementation choices to designs that apply to this methodology, using high level conceptual design. Results from actual SAN implementations using this SAN Analysis and Design Methodology would provide feedback for process improvement of this methodology.

References

[1] "Building high availability enterprise SANs"
URL: www.inrange.com

[2] "Comparing storage area networks and network attached storage"
URL: www.brocade.com/san/white_papers/pdf/SANvsNASWPFINAL3_01_01.pdf

[3] "Compare SANs to Alternate Technologies"
URL: www.brocade.com/san/evaluate/compare_san.jsp

[4] Customer Stories.
URL: www.es.avnet.com/about/customerstories/index.asp

[5] Direct Attached Storage
URL: www.es.avnet.com/solutions/storage/das/

[6] EMC and Cisco Metro Optical Storage Networking Solutions White Paper
URL:
www.cisco.com/warp/public/cc/pd/olpl/metro/on15500/on15540/prodlit/ecnts_wp.htm

[7] Enterprise Storage "Direct Attached Storage"
URL: http://www.mti.com/products/das.asp

[8] "Why Implement a SAN"
URL: www.infrastor.com/tech/reasons.htm

[9] Clark, T. IP SANs. A Guide to iSCSI, iFCP, and FCIP. PROTOCOLS FOR STORAGE AREA NETWORKS MA: Pearson Education, Inc. Publishing co and year

[10] Connor, D. Network World "Cisco Seeks to Star in Storage Nets" August 2002
URL: http://www.nwfusion.com/news/2002/0826cisco.html

[11] Lyford, R. "Business Continuance Through High-Availability SANs"
URL: http://www.mcdata.com/downloads/mkt/wpaper/business_continuance.pdf

[14] Mears, J. "The Ins and Outs of Interconnects" Network World

[15] O'Brian, S. "Building a scalable SAN. Choosing the Right Infrastructure for Your

Environment"
URL: http://www.mcdata.com/downloads/mkt/wpaper/ScalableSANs.pdf

[16] SNIA Storage Security Industry Forum
URL: http://www.snia.org/tech_activities/storage_security

[17] Strechay, R. Business Communications Review "Storage, Apps and Strategy"
URL: http://www.bcr.com/bcrmag/2001/08/p42.asp

[18] Strechay, R. Business Communications Review "Implementing a SAN"
URL: http://www.bcr.com/bcrmag/2002/08/p24.asp

[19] Vacca, J."The Essential Guide to Storage Area Networks" New Jersey MetroVoice Publishing Services Month 2002

[20] Yoshida, H. (1999). "LUN Security Considerations for Storage Area Networks"
URL: http://www.hds.com/pdf/wp91_san_lun_secur.pdf

[21] Network World Fusion SAN (Storage Area Network)
URL: http://www.nwfusion.com/links/Encyclopedia/S/493.html