Volume 3, No. 7, July 2012
Journal of Global Research in Computer Science
REVIEW ARTICAL
Available Online at www.jgrcs.info
ISSN-2229-371X

MOVING TOWARDS SAN STORAGE: AN ENTERPRISE PERSPECTIVE

Rabi Prasad Padhy*1, Manas Ranjan Patra2
*1 Senior Software Engineer, Oracle Corporation, Bangalore, India
rabi.padhy@gmail.com
2Associate Professor, PG Department of Computer Science, Berhampur University, Berhampur, India
mrpatra12@gmail.com

Abstract – One of the major components of any enterprise IT infrastructure is data storage. There has been continuous effort to explore new storage technologies to improve performance with respect to input/output operations, data availability, consolidation of storage assets and data protection and archiving. Storage has moved from the traditional Direct Attached Storage (DAS) to Network Attached Storage (NAS) and now to Storage Area Networks (SAN). Traditional storage subsystems like Server Attached Storage (SAS) and Network Attached Storage (NAS) use disks directly attached to servers over a Small Computer Systems Interface (SCSI) bus. However, these have limited capabilities keeping in view the rapidly expanding storage needs. The SAN technology holds great promises to be used for enterprises where issues like data availability, serviceability, consolidation of storage and reliability are major concerns. It can provide high speed data transfer with many other storage services, backboned with a high speed Fibre Channel (FC). Storage-area networks are also more popular and efficient way of building large storage systems both in an enterprise environment and for multi-domain storage service providers. SAN design and implementation requires careful analysis before SAN can be used as an enterprise storage solution.

In this paper we provide an overview of SAN technology from the view point of building large storage systems in an enterprise environment. We also discuss on return on investment (ROI) which can justify the adoption of SAN in an enterprise. Further, it compares two specific technologies, namely, FC SAN and IP SAN. In the sequel different design considerations are made to recommend suitable SAN implementations in order to meet the requirements of any enterprise.

Keywords- SAN, FC-SAN, IP-SAN, Storage Architecture, Enterprise Storage, Storage Consolidation, ROI.

INTRODUCTION

Storage technologies are the backbone of modern business and need constant update and upgrade in order to meet customer demands. Due to the explosion of internet and e-commerce, a tremendous amount of data has been created and made available to users. In addition to this, new type of data such as images, audios and videos have been stored and integrated with applications and databases, further accelerating the demand for storage capacity. In recent years, storage area networks (SANs) have been gaining widespread adoption in enterprise datacenters and are proving effective in supporting a range of applications across a broad spectrum of industries [1]. Majority of IT professionals across a range of corporations, government agencies, and universities have deployed storage area networks in their organizations. Some of the common applications include online transaction processing in finance and e-commerce, digital media production, business data analytics, and high-performance scientific computing.

Storage area networks are popular in enterprise datacenters and are commonly adopted to support the storage needs of data-intensive applications. SAN architecture is quite an attractive choice for parallel clustered applications which demand high-speed concurrent access to a scalable storage backend [2]. Such applications commonly rely on a clustered middleware service to provide a higher-level storage abstraction such as a file system GFS (Global File System), OCFS (Oracle cluster file system), GPFS(General Parallel File System) or a relational database (Oracle RAC) on top of raw disk blocks. SANs aim to provide fully decentralized access to shared application state on disk and in principle, any SAN-attached client node can access any piece of data without routing its requests to a dedicated server.

A SAN is a networked high-speed infrastructure (subnetwork) that establishes direct access by servers to an interconnected group of heterogeneous storage devices such as optical disks, RAID arrays, and tape backups, which are effective for storing large amounts of information and backing up data online in e-commerce, online transaction processing, electronic vaulting, data warehousing, data mining, multimedia Internet/intranet browsing, and enterprise database managing applications [3]. SANs provide additional capabilities (fault tolerance, remote management, clustering, and topological flexibility) to mission-critical, data-intensive applications.

STORAGE AND STORAGE MODELS

There are three storage models commonly used in the storage industry to provide customized storage devices manufactured by various companies [4]. The three common storage models are Direct Attached Storage (DAS), Network Attached Storage (NAS) and Storage Area Networks (SAN). Figure 1 depicts these basic storage models.

a. Direct-attached storage (DAS): DAS is a block device from a disk which is physically (directly) attached to the host machine. User must place a file system upon it before it can be used. Technologies to do this include IDE, SCSI, SATA, etc. This storage could be physically internal or external to the host system and typically not shareable with other host systems. This method does not use any network and thus is unable to share data with other computer (on the storage level).

b. Network attached storage (NAS): Network attached storage uses storage that is connected to a dedicated server and linked with a local area network (LAN) for shared access. An important concept behind NAS is the ability for a client or host system to access storage and data over the network that physically is attached to a host system or network file server. NAS network users centrally share storage, with limited scalability [5]. Technologies to do this include NFS, CIFS, AFS, etc.

Advantages of NAS

a) Storage sharing between multiple host systems
b) Data sharing between multiple hosts and operating systems
c) Overcome distance limitations using networking techniques
d) Simpler management by reducing duplicate data and storage
e) Application based storage access at file level

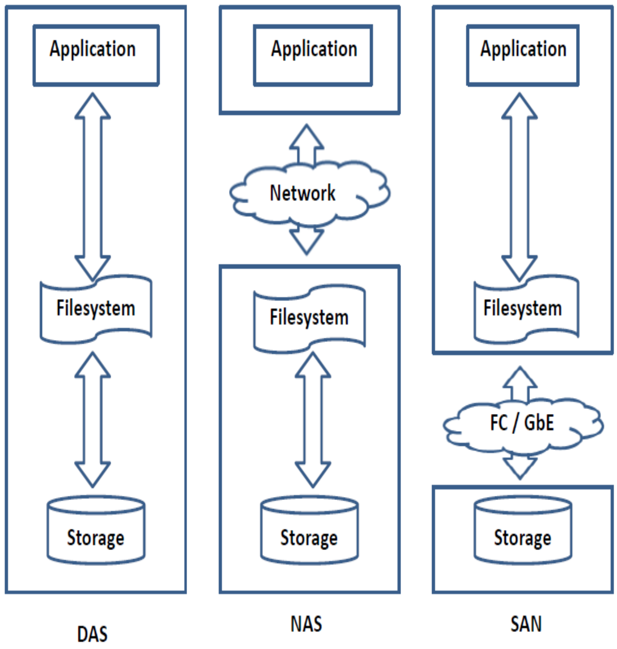

Figure 1. Basic Comparison of DAS, NAS and SAN

c. Storage area network (SAN): Storage area networks use an isolated high speed network, usually interconnected with fibre channel. Heterogeneous storage devices (HDD, Tape, Laser Disks) are connected externally with block level data transfer (Figure 2). SAN normally provides high speed data transfers in three different ways: Server-to-storage, Server-to-server and Storage-to-storage [6]. Storage Area Networks (SAN) is a term generally associated with Fibre and can mean different things to different people. To some a SAN is a Fibre Channel switch, to others a SAN is for backup; others see a SAN as anything with two or more hosts systems using Fibre Channel.

Advantages of SAN

a) Storage sharing between hosts
b) High performance applications
c) LAN-free backup
d) Overcome distance issues
e) Server and storage consolidation

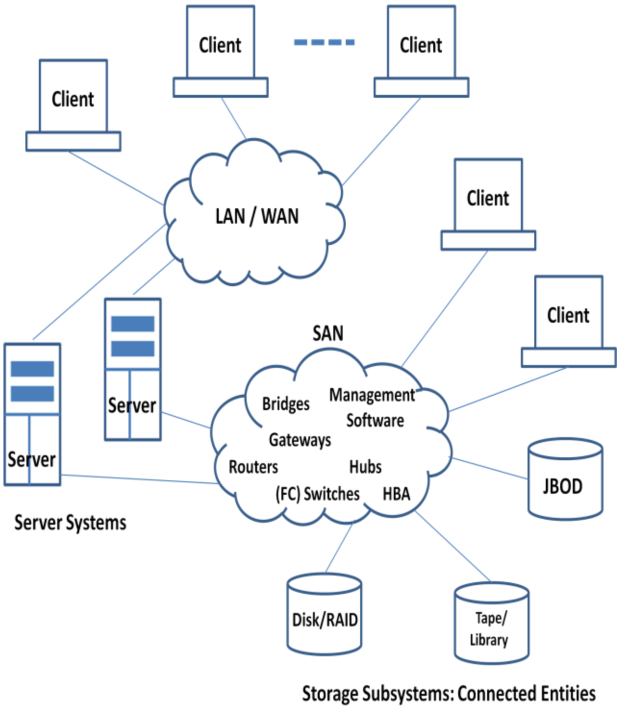

Figure 2. SAN high-level Architecture

SAN STORAGE INFRASTRUCTURE REQUIREMENTS

Today there are two protocols available for building block-based SANs, FC and iSCSI. Both protocols use SCSI commands generated by the file systems of the servers. These SCSI commands are converted by the iSCSI or FC protocol so that those can move through a network to and from centralized disk storage systems where the commands are executed. In case of FC the network equipment is specific to the protocol and in case of iSCSI the network equipment type is anything that will handle IP packets (1GB Ethernet is the most popular one).

SAN Components:

a. FC: FC is a technology designed for very high performance low-latency data transfer among various types of devices. The FC protocol is based on the SCSI protocol and makes use of the common SCSI command set over the FC protocol layer [7]. The FC protocol may be implemented both over optical fibre as well as copper cable.
b. FC Switch: An FC switch provides multiple simultaneous interconnections between pairs of ports with the resultant increase in total bandwidth. FC switches are used to implement FC fabric interconnection.
c. FC Hub: An FC hub is used to implement the FC Arbitrated Loop (FC-AL) protocol. Hubs pass signals arriving from one port to the next port in the loop. It is up to the devices to intercept and process signals addressed to them.
d. FC Arbitrated Loop (FC-AL): FC-AL is an FC topology that provides a solution for attaching multiple communication ports in a loop. In an FC-AL, communication is not broadcast as it is in architectures like Ethernet. Instead it is transmitted from one device

to the next with each device repeating the transmission around the "loop" until the data reaches its destination [8]. The devices arbitrate for access to the loop before sending data.

e. Fabric: A 'fabric' is a network of FC switches providing interconnectivity and scalability. It is used to describe a generic switching environment. With a fabric the bandwidth is not shared.
f. Host Bust Adapters (HBA): An HBA provides the interface between a server and the SANs network. Every HBA has a corresponding device-driver, which handles the I/O and control requests. The HBA connecting a server to a SAN may be an FC HBA (for an FC based SAN) or an iSCSI HBA (for an IP based SAN). Some of the latest HBAs may have the support for both FC and iSCSI.

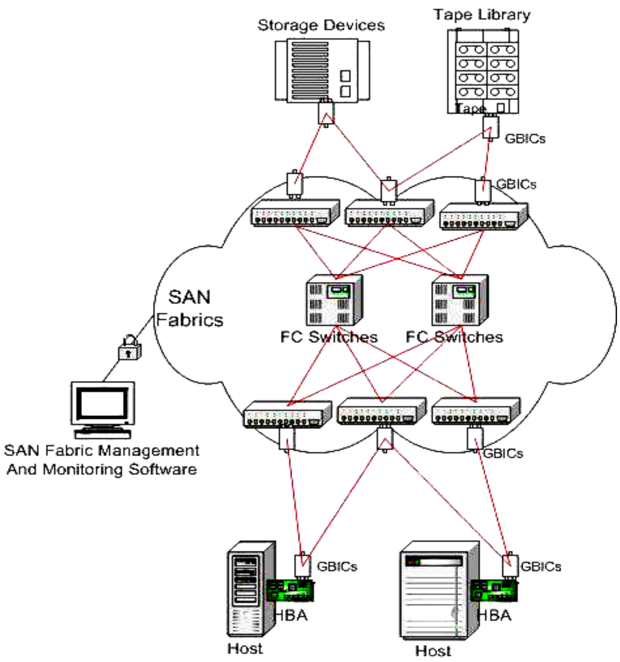

Figure 3. Basic SAN Storage Components

Storage Subsystem:

a. Small Computer Systems Interface (SCSI): This is a parallel interface standard used for attaching peripherals (including disk drives) to a computer. SCSI enables faster data transmission rates compared to other popular interfaces like IDE, serial and parallel ports. In addition to this, many devices can be attached to a single SCSI port. Server grade systems and SAN storage boxes invariably use SCSI disk drives.
b. Redundant Array of Inexpensive Disks (RAID): This is a mechanism for providing disk fault tolerance. Five types of RAID architectures have been defined. They are RAID-1 through RAID-5. Each provides disk fault-tolerance and offers different trade-offs in features and performance. In addition to these five redundant array architectures, it has become popular to refer to a non-redundant array of disk drives as a RAID-0 array. Possible approaches to RAID include hardware RAID and software RAID [9]. Internal hardware RAID solutions involve presence of a RAID controller inside the server. In external hardware RAID solutions, the hardware RAID controller and the disk drives are housed separately from the server in a high-availability external RAID enclosure. The external hardware RAID controller-based storage system may be attached directly to the SAN.
c. Just a Bunch of Disks (JBOD): This refers to a set of disks that has not been configured into a RAID array but can be used as if they were a single volume. This can be used for applications, which require more storage space than that offered by the disks individually.

Gigabit Interface Converters (GBIC):

This is a removable transceiver. It interconverts electrical and optical signals for high-speed networking. GBICs are used in all types of FC devices including switches and HBAs. Initially targeted to support FC data networks, the GBIC standard was quickly adopted for use with Gigabit Ethernet installations as well. By providing hot-swap interchangeability, GBIC modules give net administrators the ability to tailor transceiver costs, link distances, and configure overall network topologies to meet their requirements.

Internet SCSI Protocol (iSCSI):

Protocols like iSCSI enables deployment of a SAN over conventional Ethernet based network. The iSCSI protocol uses TCP/IP as its network transport protocol and is designed to leverage TCP/IP for block storage needs. However, there are a few challenges in the acceptance of iSCSI as the SAN interconnect [10]. TCP/IP has traditionally been tuned to favour short and bursty user transmissions as against large and continuous data transfer requirements of storage. However, several vendors have announced their support for use of iSCSI to reduce the processor overhead. Once this overhead becomes comparable to that of FC, iSCSI would present itself as a competitor of FC.

FC-iSCSI Gateway (SAN Gateway):

In a heterogeneous SAN containing both FC and iSCSI-based devices, an FC-iSCSI gateway provides the internetworking of iSCSI devices with FC devices. The gateway maps selected iSCSI devices into the FC SAN and selected FC devices into the IP SAN.

SAN Management Software:

A SAN management software, as the name suggests, assists in the management of the SAN environment. The tasks of any typical SAN management software include discovery and mapping of storage devices, switches, and servers; monitoring and alerting on discovering devices and logical partitioning and zoning of the SAN. With increase in the number of vendors providing SAN products, the complexity in SAN environments has increased tremendously. This has made the management of a SAN extremely challenging as a good SAN management tool is expected to perform well in a multi-vendor heterogeneous SAN environment.

DESIGN CONSIDERATIONS FORARCHITECHTING A SAN

There are typically two kinds of SAN architectures, which are currently popular and gaining ground. These are: FC-based SAN & IP-based SAN [11].

A SAN can be deployed in two ways: using Fiber Channel or the IP way. The first is called FC-SAN while the second is IP-SAN. IP-SAN can also mean the network built with iSCSI as the protocol.

a. FC-SAN: In an FC-SAN the file I/O is between the application and the file-system, which talks to the volume manager, which in turn makes block I/O to the storage device. The file-system and the volume manager (if any) are on the server and the storage has the RAID and the disks. The Figure 5 shows the high level architect of FC-SAN and IP-SAN.
b. IP-SAN: The basic components of the system are the same, but here the application makes the file I/O to a file-system which is outside the server and the block I/O now happens within the storage between the volume manager and the RAID [12].

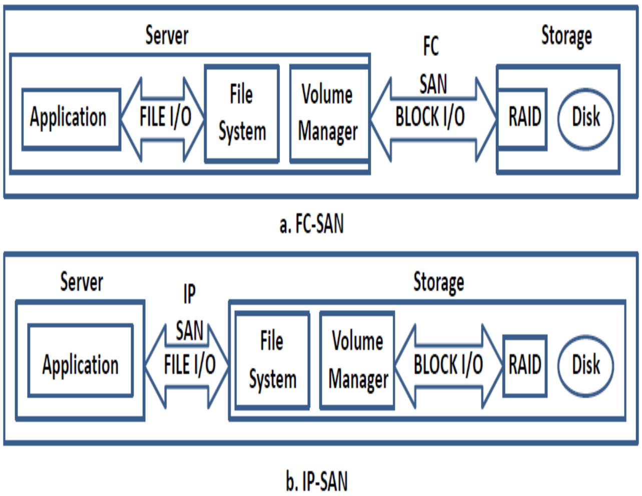

Figure 4. High Level Architecture of FC-SAN and IP-SAN

Traditionally, FC-SAN has been considered to be difficult and IP-SAN easy to deploy. FC-SANs are coming up with management tools, which make deploying FC-SAN relatively simple and vendors on IP-SAN making their devices faster and faster so that there is little distinction on what can run on IP-SAN and on FC [13]. But to date IP-SAN deployments are extremely simple and easy to deploy, manage and provide some exceptional recovery methods.

Traditional SAN is based on Fiber Channel Framework for which IP-SAN is an alternative. IP-SAN is usually over Gigabit Ethernet (GbE) and it represents a new generation of networked storage. It is a combination of the best attributes of Fiber-Channel SANs (FC-SAN) as well as NAS. These include functionality and scalability, ease-of-deployment and ease-of-use. IP-SAN provides virtually all the competence of Fiber-Channel SAN and NAS devices combined at a lower cost.

The issues related to Fiber Channel are expense, complexity, and interoperability, which are absent in IP storage. IP can take advantage of common network, hardware and technologies which makes IP-SAN less complicated to deploy than Fiber Channel. As its name suggests, IP SAN is built on top of IP, making deployment relatively easy and quick.

IP-SAN's iSCSI (Internet SCSI) network is a complete alternative solution to Fiber Channel. There are other alternatives, which offer hybrid solutions and can be used to either extend the existing Fiber Channel frameworks or to migrate from them to an IP storage network, namely iFCP (Internet Fiber Channel Protocol) and FCIP (Fiber Channel over IP). IP-SANs present a good option for SMBs on the storage centralization front, i.e., when they want to consolidate their storage from distributed DAS systems. Combination of six cost-effective and high performance technologies, essentially make the IP-SANs. These are iSCSI, GbE switches, Serial ATA (SATA) hard disk arrays, Linux or similar open source OS kernel, virtual LANs (VLAN), and Cat 6 cables. The IP-SAN storage solutions are convenient and cost effective for an enterprise. It has the capability of switching protocols and share files across multiple application environments without storage infrastructure changes [14]. This solution also facilitates simple, rapid backup, and restoration of damaged or accidentally deleted files and databases.

KEY CONSIDERATION BY ENTERPRISE TO MIGRATE TOWORDS SAN STORAGE

In a business environment, effective information management requires various factors to be considered before the design and deployment of SAN [15]. The decision to adopt SAN is based on a number of factors such as:

a) Size of the organization’s IT infrastructure
b) Data Location and Inter-connectivity
c) Total Cost Ownership and Return on Investment
d) Disaster Management and Backup/Restore
e) Manageability and High Availability of Data
f) Performance and Monitoring
g) SAN Security related issues

Size of the organization’s IT infrastructure”

If the organization is small (20 servers or fewer) or where data management, application performance, or backup is not currently a problem then there is no need to use a SAN. There are certain types of servers that should not be included in a SAN. Such servers are usually better off staying on their internal disk drives; they don’t benefit from SAN-based storage. The following types of servers do not require a SAN [16].

Web servers: Computers set up as Web servers don’t usually have large storage needs; they are usually connected to larger servers that run the databases from which Web pages are automatically built. Although Web servers are good candidates for NAS, database servers can make better use of SAN disks.

Infrastructure servers: Server applications that handle the chores of network infrastructure - such as Domain Name Servers (DNS), Windows Internet Naming Servers (WINS), and Domain Controllers (DC, PDC) - are better left on the server computers’ internal disks. They don’t need a lot of disk space, and their performance requirements are minimal.

All desktop PCs: Personal computers are not good SAN candidates because they usually connect to corporate servers for any applications that require high performance. Those corporate servers, however, could use a SAN.

Servers that require less than 10GB of storage: Internal storage is cheaper than SAN storage. If your server has no performance problems and will never need more than 10GB of storage space, leave it alone.

Servers that don't need fast access to data: If performance is already quite acceptable and you don't mind maintaining the server separately, don't bother hooking it up to a SAN.

Servers that have to share files: Such servers are better off connected to a Network Attached Storage (NAS) server. NAS servers store and transfer data as files, and not blocks of data, so they don't need the high-speed Fibre Channel protocol used in a SAN. NAS devices are best for file-based uses such as user home directories and shared documents.

Data Location and Inter-connectivity:

Storage devices and server deployment should be based on the specific requirements for data location. The important issue in the design of SAN topology is the placement of storage system with respect to servers accessing the storage. The interconnectivity can be done in any of the three ways: one-to-one, many-to-one, and many-to-many. The storage libraries should be selected as per available capacity based on the storage requirement analysis. The devices should have redundancy so that failures can be handled in a better way. As an example, in RAID technology mirroring is done at the disk level to prevent heavy data loss. Interconnectivity is based on the data location and distance from workstations. Fibre channel gives block level access so that data transfer can be high, which is one of the primary factors in adopting SAN as enterprise storage.

Total Cost Ownership and Return on Investment:

In TCO analysis, various costs are calculated as per business requirements. There are two types of costs, hard and soft. Hard costs include total purchase cost, total administration costs, and total utilization costs, and soft costs include storage availability and cost per downtime hour. The installation costs are also included along with storage personnel salary [17]. The ROI is calculated based on the amount spent, calculated from the TCO. In order to get the perfect ROI for SAN deployment, three models can should be analyzed: Payback Period (PP), Internal Rate of Return (IRR), and Net Present Value (NPV). The simplest way to analyze the ROI is given in the following formula:

ROI = (Sum of all Income – Sum of all Investments) / Sum of Investments

Disaster Management and Backup/Restore:

SAN provides centralized backup activity carried out in a secured environment; SAN backup tends to be highly reliable and redundant [18]. The backup strategies need to be developed with customized guidelines which include periodicity of backup, backup method and many more. The disaster recovery management (DRM) system is an important part of safeguarding business data against any unexpected disasters. DRM involves processes such as (1) analysis of the business requirement (2) formulating the DRM team (3) defining the DRM policy and procedure (4) developing the detailed DRM plan and activities (5) implementing and updating the DRM plan.

Manageability and High Availability of Data:

Data availability refers to how reliably data is stored for day-to-day operations. The preventive maintenance of all SAN equipments keeps it alive for long times and reduces the failures of devices. Administrators must carve up the storage space into segments that are only accessible to specific users. SAN management should be done by specialized personnel who are skilled in storage activities; SAN personnel should be involved from the beginning of SAN implementation including installation and configuration [19].

Performance and Monitoring:

In SAN deployment, performance analysis involves the analysis of Bandwidth (storage bandwidth, disk, SCSI or FC), Throughput (the amount of work performed by a component or system over unit time), Delay (wait time induced by contention within a system), Latency –(amount of time between the initiation of an action and the actual start), Response time (time it takes to finish a given storage operation), Scalability (ability of a SAN to grow without adversely affecting its services), Reliability (the degree to which a given computing component produces consistent results), and Availability (the degree to which a SAN keeps running within acceptable performance limits and without any unrecoverable failures). Monitoring records all the activities that are accessing data and the time at which this occurs. Monitoring helps auditors analyze the system from the security perspective.

SAN Security related issues:

SAN security is the group of parameters and settings that makes storage resources available to authorized users and trusted networks and unavailable for others. Storage Area Networks security comprises of many methods and approaches that are linked together to achieve the highest level of security for stored data. Storage Area Networks is one effective technology for organizations that need confidentiality, integrity, scalability, availability, and high performance storage solutions [20]. Since SAN holds mission critical data which is very vulnerable to attackers, it needs multi-factor considerations to address security issues. Effective storage management should implement five basic areas of security in each level of SAN, namely, Storage Array Volume Access Control, Volume Access Control on the HOST, Device Configuration Access Control, Storage Management, Software Access Control and Proactive Detection of Access Violation, and Auditing and Logging. The five broad categories are further individualized into Access Control (Zoning & LUN masking, Intrusion Detection System (IDS), Cryptography (CFS, SFS & EFS), Authentication and Authorization, Fibre Channel Security, and Security by SAN Management Software.

KEY REASONS FOR ADOPTING SAN STORAGE

There could be many benefits in adopting a SAN. Some of the significant ones are enumerated below:

a) Storage consolidation
b) Serverless backup
c) Better utilization of storage facilities
d) Scalable storage
e) Improved fault tolerance and DR capacity
f) Centralized management

g) Interoperability between diverse systems

Storage Consolidation:

With the traditional storage model, administrators have to manage multiple storage devices. Backup for each of these storage devices is also a cumbersome process. Consolidation of these individual storage entities could solve many of such problems. The centralised, consolidated storage environment provided by a SAN provides for easier management of data compared to a DAS environment.

Serverless Backup:

Serverless backup allows disk storage device to copy data across the high-speed links of the SAN directly to a backup device without any intervention of the server. Data is confined to the SAN boundaries and the clients get uninterrupted access to the server resources. As the tape library in a typical Centralised Backup System may be used to backup data from several different servers, the speed of backups can be improved. This is a result of data potentially being inter-leaved on the tape from several sources, sustaining the tape drive at its maximum streaming operating efficiency [21]. The utilisation of tapes may also be improved by this sharing of tapes from many sources, leading to reduced numbers of tapes being needed and also fewer tape drives to backup a given amount of data. A large tape library will essentially operate in unattended mode with the minimum of human intervention. A sufficiently large tape library will enable cycles of backup tapes to be kept in the library for longer periods e.g. a full week or even a month or longer, avoiding the need to physically fetch tapes from the fireproof safes for restores. Clone copies of tapes for safe keeping in fireproof safes should still be made, so some operational effort to remove tapes from the library for storage in a fireproof safe will still be needed. However, considerable operational efficiencies will be achieved and the CBS element alone of any SAN Business Case should produce substantial savings by rationalisation of backup/restore activities.

Better Utilization of Storage Facilities:

With SAN, one may improvise storage as per the requirements, thus considerably increasing the utilization of storage. In the traditional storage model, even though there may have plenty of vacant space on a storage drive attached to a server, it may not be possible to make use of the same for another system, which has run out of storage space. SAN storage is centralized which means that the new hardware procurement for centrally managed systems will require little or no local disk provision. For major corporate systems and large server environments, this achieves a significant cost saving. Furthermore, in those HE/FE establishments with large numbers of servers run by local departmental IT support personnel (outside of the central IT service), further cost savings can be made through server consolidation (reducing the total number of servers needed) across the entire HE/FE establishment. This could be achieved by removing the need for departments to host, operate and manage their own servers for file or email provision. An additional cost saving can be achieved by being able to redeploy some staff effort currently assigned to manage these departmental servers. A SAN would permit central IT providers to offer "managed storage" to users across the campus. Disk space could be made available on the SAN through central high-availability clustered servers. This would maximise the "economies of scale" benefit of deploying a SAN.

Scalable Storage:

A SAN allows for "pay as you grow" storage scalability, pro-active storage planning and non-disruptive growth/reconfiguration. Extra disks can be purchased at any time and added to the total storage provision. Configuration utilities allow for that storage to be 'presented' to any host that requires space. Free space can also be re-assigned as necessary. Typical sites find that up to half of their disk space is unused (not allocated) and in hardware terms alone (i.e. not including maintenance and administration) substantial savings in purchase costs could be made. While individual resources and servers have a restriction on the number of storage and interconnected units they can attach, a SAN is not affected by such constraints, leading to a higher scalability.

Improved fault tolerance and DR capacity:

Integration with a disaster recovery solution or a replication solution also becomes easier, as with a SAN, the challenge is confined to looking for solutions only for an integrated and consolidated storage space rather than a scattered and distributed storage space. The enhanced disaster recovery capability of a SAN would lead to increased availability giving the potential for significant staff productivity savings. It is very difficult to cost accurately the effect of server or system downtime; therefore some sites may not be able to include downtime reduction as a tangible benefit in their cost/benefit analysis. However, it should then be highlighted as a major intangible benefit in the business case if actual costs savings cannot be accurately determined.

Centralized Management:

Centralized management can enhance staff productivity as well as reduce total cost of ownership through intelligent storage management (only one storage dedicated team performing the work of many systems administrators), economies of scale (purchasing terabytes of disk space in one go rather than lots of small disk sub-systems), cheaper hardware (servers no longer need individual disks and expensive RAID controllers), centralised (and shared) backup and restore architecture. Centralised management will also facilitate common standards in storage management according to an institutional storage policy (e.g. standards in disk allocation, levels of fault-tolerance, scalability, utilisation and back up/recovery mechanisms across the institution). This should also assist sites with disaster recovery and security audits.

Interoperability between diverse systems:

Disparate computer platforms can share the SAN as their common mass storage system. The SAN would then provide a totally heterogeneous environment, enabling cross platform functionality (e.g. between UNIX, PC and other operating systems) for activities such as file and print serving, database hosting, high-performance computing needs etc. In fact, new developments allow systems to boot directly from a SAN; servers would then require no local storage whatsoever. This could then possibly allow "spare" servers to be kept that are able to be allocated to any of several operating systems in the event of a server failure.

APPLICATIONS AND FILESYSTEMS DESIGNED FOR A SAN

These are different newer applications which are specially designed for SAN capabilities:

Cluster applications:

Cluster applications are created by tying a group of servers together via a fast network and then allowing those servers to access the same disks' storage where the application is installed. This allows highly scalable and available applications; if one of the servers fails, another server in the cluster can pick up where the first one left off. Some GRID computing applications are examples that need a SAN storage. Common cluster applications are IBM HACMP, Solaris Cluster 3.0 or above, Compaq/HP TruCluster, Oracle Failsafe Cluster, Oracle Real Application Clusters, Microsoft Cluster Server (MSCS), HP MC/Serviceguard Clusters, and Novell Netware Cluster Services [22].

SAN backup/restore applications:

SAN-based backup software is optimized for using SAN hardware. The backup software includes intelligence that takes advantage of what SAN offers [23]. When using SAN backup software, a user can back up data directly over the SAN to a tape drive, which makes backup run much faster. Common backup software that has this capability includes Veritas NetBackup, Tivoli Storage Manager, Veritas Backup, CA ARCserve with the SLO option, Legato NetWorker, and CommVault.

Virtualization software solutions:

Server virtualization hardware and software such as VMware, Virtual Iron, Microsoft Hyper-V, Egenera, and others need the disk-sharing capability of a SAN (especially fail-over for applications between physical servers), and they also gain from the performance benefits that a SAN has to offer.

File systems for SAN:

Many cluster designs require a Shared File System, so cluster members can work with the same data files at the same time. SAN provides Shared Disks, but SAN itself does not provide a Shared File System [24]. Cluster File Systems can be used to build Shared File Systems using SAN. There are many cluster file systems available in the market, such as IBM Global Parallel File system, RedHat's Global File system and VERITAS cluster file system etc. [25]

Table 1. Cluster File System products from various vendors and supported platforms

Cluster File System	Operating system
Sun Cluster 3.0	Oracle Solaris
IBM's Generalized Parallel File System (GPFS)	IBM AIX, Linux
RedHat's Global Filesystem (GFS)	Linux
VERITAS Cluster File System	Sun Solaris, HP/UX
Oracle Cluster File System	Linux
PolyServe	Linux
Tru64 Cluster 5.x	HP Tru64

CONCLUSION AND FUTURE WORK

Information plays a vital role in business continuity. The growth of business is based on the effective management of its valuable information in today's information era. SAN is one promising solution to provide reliability, scalability and high information availability for instant access. To increase productivity, meet customer demand, and to be a leader in industry, any organization must ensure that its data is reliably stored and available as and when access is needed. SAN provides dynamic storage solution for the enterprise business requirements and needs careful analysis before design and its deployment. In order to implement SAN, requirement analysis is to be carried out and the SAN capacity planning is to be done to ensure that the proposed SAN solution never runs out of space for the business requirements under consideration.

In this article, we have discussed on various storage networking solutions, from the traditional direct-attached storage, to NAS to FC-Based SAN to IP-Based SAN. We have provided a comparative view of the capabilities of SAN technology as compared to the previous storage architectures. We have also analyzed the functions of various components that make up a SAN. This study can be useful for enterprises in order to effectively design and deploy SAN in their respective business contexts.

In future we plan to compare the performance of iSCSI with FCIP and iFCP in order to bring out the benefits and drawbacks of the architectures. We also propose to develop an impact analysis framework with various SAN planning tools in order to provide better design and suggest appropriate system policies for any given design requirements. We shall also study how SAN can be judiciously connected to different networks such as WAN, LAN etc. in order to address the ever-growing storage requirements of different enterprises.

REFERENCES

[1]. R. J. Morris, and B. J. Truskowski, The Evolution of Storage Systems. IBM Systems Journal, Vol. 42(2), pp205-217, 2003.

[2]. SAN Essentials: A Complete Guide to Understanding and Implementing SANs – John Wiley & Sons

[3]. Jiwu Shu, Bigang Li, Weimin Zheng, "Design and Implementation of an SAN System Based on the Fiber Channel Protocol," IEEE Transactions on Computers, pp. 439-448, April, 2005.

[4]. NetApp Storage Solutions

http://www.netapp.com/us/products/storage- systems/

[5]. P. Mahalingam, N. Jayaprakash, S. Karthikeyan, "Enhanced Data Security Framework for Storage Area Networks," ICECS, pp.105-110, 2009 Second International Conference on Env. And Computer Science, 2009.

[6]. EMC Storage Solutions:

http://www.emc.com/products/category/storage.htm

[7]. Aziz, M.H.Ong Con Nie, Jesse Chan Mei Yam & Lee Chang Wei. TCO reduction. Communications, 2003.

APCC 2003. The 9th Asia- Pacific Conference. Vol. 3, pp1147- 1151 Vol.3. ISBN: 0-7803-8114-9, 2003.

[8]. D. R. Avresky, V. Shurbanov, R. Horst, W. Watson,L. Young and D. Jewett. Performance Modeling of ServerNet™ SAN Topologies. ISSN. 0920-8542, Springer Netherlands, The Journal of Supercomputing, Computer Science, pp.19-37, December, 2004

[9]. Jiwu Shu, Bigang Li, Weimin Zheng, "Design and Implementation of an SAN System Based on the Fiber Channel Protocol," IEEE Transactions on Computers, pp.439-448, April, 2005

[10]. Computer Data Storage, Wikipedia. Available at http://en.wikipedia.org/wiki/Computer_data_storage

[11]. Paul Ross. Design Considerations in Enterprise Storage Networks.

[12]. Jamie Gruener, Research note: Compliance and digitalcontent growth drive content-storage market, The Yankee Group , May 20, 2003

[13]. Storage Network Infrastructure FC San Components Guide, 2003

http://www4.gartner.com/DisplayDocument?doc_cd=121308

[14]. D. Dunn. “Seagate expands full disk encryption to data center hard drives”, Information Week Magazine, pages 1–2, October 2007.

[15]. The Cost Benefits of a SAN, An Analysis of Total Cost of Ownership of an iSCSI SAN, Fibre Channel SAN, and Direct Attached Storage, Comp. Technology Review. Retrieved on October 20, 2011, from

http://findarticles.com/p/articles/mi_m0BRZ/is_7_24/ai_n6185910.

[16]. J. M. Ng, R.N. Susilo, Controlled Shared Backup Strategy Based on Channel Availability. International Journal of Computer Application Technology, Vol. 24(1), pp25-32, 2005.

[17]. E. Riedel, M. Kallahalla, and R. Swaminathan, A Framework for Evaluating Storage System Security. Proceedings of the 1st USENIX Conference on File and Storage Technologies, January pp28-30, Monterey, CA, 2002.

[18]. P. Mahalingam, N. Jayaprakash, and S. Karthikeyan, Enhanced Data Security Framework for Storage Area Networks. Proceedings of the 2nd International Conference on Environmental and Computer Science (ICECS 2009), December pp28-30, Dubai, UAE, 2009.

[19]. D. R. Avresky, V. Shurbanov, R. Horst, W. Watson, L. Young, and D. Jewett, Performance Modeling of ServerNet™ SAN Topologies. The Journal of Supercomputing, Vol 10(4), pp19-37, 2004.

[20]. O. I. Pentakalos, D. A. Menascé, M. Halem, and Y. Yesha, Analytical Performance Modeling of Hierarchical Mass Storage Systems. IEEE Transactions on Computers, Vol. 46(10), pp1103-1118, 1997.

[21]. G. Schulz. “Decrypting enterprise storage security: Trends and options for securing enterprise data and storage”,Industry Trends and Technology Perspective White paper, pages 1–5, December 2006.

[22]. RedHat Global Filesystem

http://www.redhat.com/software/rha/gfs/

[23]. Sixth NASA Goddard Space Flight Center Conference on Mass Storage and Technologies in cooperation with the Fifteenth IEEE Symposium on Mass Storage Systems March pp23-26, 1998

[24]. Polyserve Matrix Server

http://www.polyserve.com/products.html

[25]. Open GFS http://opengfs.sourceforge.net/

[26]. Lymberopoulos, L., Lupu, E., and Sloman, M. An Adaptive Policy-based Framework for Network Services Management. Journal of Networks and System Management vol (3), 2003.

[27]. hp Storage Solutions :

http://h18006.www1.hp.com/storage/networking/index.html

[28]. IBM SAN Storage Solutions: http://www-03.ibm.com/systems/in/networking/switches/san/

[29]. Hitachi Storage Solutions:

http://www.hds.com/products/storage-networking/

[30]. Dell Storage Solutions: http://supportapj.dell.com/support/edocs/stor-sys/ssan/index.htm

[31]. Y. Lu and D. H. C. Du, “Performance Study of iSCSI-Based Storage Subsystems,” IEEE Communications Magazine, August 2003.

[32]. J. Elmirghani and I. H. White, “Optical storage area networks,” IEEE Communications Magazine, vol. 3(43), 2005, feature section.

[33]. W. W. Hsu and A. J. Smith, “Characteristics of I/O traffic in personal computer and server workloads,” IBM Systems Journal, vol 2 (42), 2003.

[34]. A. Baldwin and S. Shiu. “Encryption and key management and a SAN”, Proceedings of the First International IEEE Security in Storage Workshop (SISW’02), pp.1–10, 2002.

Short Bio Data for the Author

Rabi Prasad Padhy is currently working as a Senior Software Engineer - Oracle India Private Ltd. Bangalore. He has achieved his MCA degree from Berhampur University. He carries 8 years of extensive IT Experience with MNC’s like EDS, Dell, IBM and Oracle. His area of interests include IT Infrastructure Optimization, Virtualization, Enterprise Grid Computing, Cloud Computing and Cloud databases. He has published several research papers in international journals. He is a certified professional for Oracle, Microsoft SQL Server database and also ITIL certified.

Dr. Manas Ranjan Patra holds a Ph.D. Degree in Computer Science from the Central University of Hyderabad, India. Currently he is an Associate Professor in the Post Graduate Department of Computer Science, Berhampur University, India. He has about 24 years of experience in teaching and research in different areas of Computer Science. He had visiting assignment to

International Institute for SoftwareTechnology, Macao as a United Nations Fellow and for sometime worked as assistant professor in the Institute for Development and Research in Banking Technology, Hyderabad. He has about 90 publications to his credit. His research interests include Service Oriented Computing, Software Engineering, Applications of Data mining and E-Governance. He has presented papers, chaired technical sessions and served in the technical committees of many International conferences.