Communications of the Association for Information Systems

Volume 14 Article 27

11-30-2004

Storage Area Networks

Manjari Mehta
University of Houston, Manjari.Mehta@mail.uh.edu

Blake Ives
University of Houston, blake.ives@uh.edu

Communications of the Association for Information Systems

STORAGE AREA NETWORKS

Manjari Mehta
Blake Ives
University of Houston
Blake.Ives@uh.edu

ABSTRACT

This paper is a tutorial on Storage Area Networks (SANs). On the face of it storage network technology might appear to be just another incremental improvement relative to current technology; i.e., faster and cheaper but otherwise nothing new. On the other hand, as with technologies such as the World Wide Web and XML, SAN can also be viewed as an exciting new window into the future applications of technology. The World Wide Web provided a global, if relatively disorganized, file structure that almost immediately transcended organization and national boundaries. XML provides a still finer level of multi-organization data structure, this time at the level of the logical data element. Storage Network technologies provide an even finer level of resolution for global storage; providing powerful tools for organizing vast repositories of data and information within a building, organization, within municipalities, across distributed organizations, and throughout the world.

This paper compares SAN technology with previous storage management solutions with particular attention to promised benefits of scalability, interoperability, and high-speed LAN-free backups. The paper provides an overview of what SANs are, why invest in them, and how SANs can be managed. The paper also discusses a primary management concern, the interoperability of vendor-specific SAN solutions. Bluefin, a storage management interface and interoperability solution is also explained. The paper concludes with discussion of SAN-related trends and implications for practice and research.

Keywords: storage area network, SAN, backup, storage management, interoperability, Bluefin, Fibre Channel

I. INTRODUCTION

THE STORAGE MANAGEMENT CHALLENGE

The world produced around 2 million terabytes of information in 2002 [Anthes, 2002][1]. With emerging media-rich content in the medical and entertainment industries joining such increasingly popular data-intensive applications as ERP, business continuity systems, on-line transaction processing and data warehousing, the demand for storage will escalate [Bird, 2003, Brinkmann et

[1] Anthes quoted data from the School of Information Management and Systems, University of California, Berkeley.

al., 2000, Chernyshov, 1999, Guha, 1999, Molero et al., 2001]. The problem is exacerbated in the United States by new requirements for retaining data because of laws such as the Sarbanes-Oxley act [Moore, 2004] and the ever-burgeoning film and healthcare industries. Sidebar 1 shows some of the many reports and forecast that were made in recent years.

SIDEBAR 1. THE DATA DELUGE

The following are representative of reports in the trade press. Although these reports look at the data from disparate points of view and are hard to correlate with one another, they indicate the nature of the problem.

- For a typical firm in the Fortune 2500, the amount of storage would increase tenfold from 15 to 150 terabytes between 1999 and 2003 [Trellisoft, 2001]
- The nearly insatiable demand for online information resources suggests that storage hardware and software will continue to create a serious dent in the IT budget, already estimated to make up 12-15% of the total IT budget [META Group, 2003]
- The relative amounts spent on hardware and on storage management are expected to shift toward storage management software [Goodwin, 2003] because storage hardware costs are declining at around 30% per year [Derrington, 2002b]
- In 2002, the average cost of a small Storage Area Network[2] with about 10 servers, 0.5 Terabyte of storage and 16-port switches[3] was $100,000, while software costs ranged from $20,000 to $200,000, depending on the network size[4] [Pratt, 2002]
- The global market for storage area networks (SAN) grew to $7.5 billion in 2002 and is expected to reach $84 billion by 2008 [MarketResearch.com, 2003]
- IDC's Worldwide Quarterly Disk Storage Systems Tracker [ByteEnable, 2004] estimated that the networked storage market grew by 17.5% to $1.9 billion just in the first quarter of 2004.

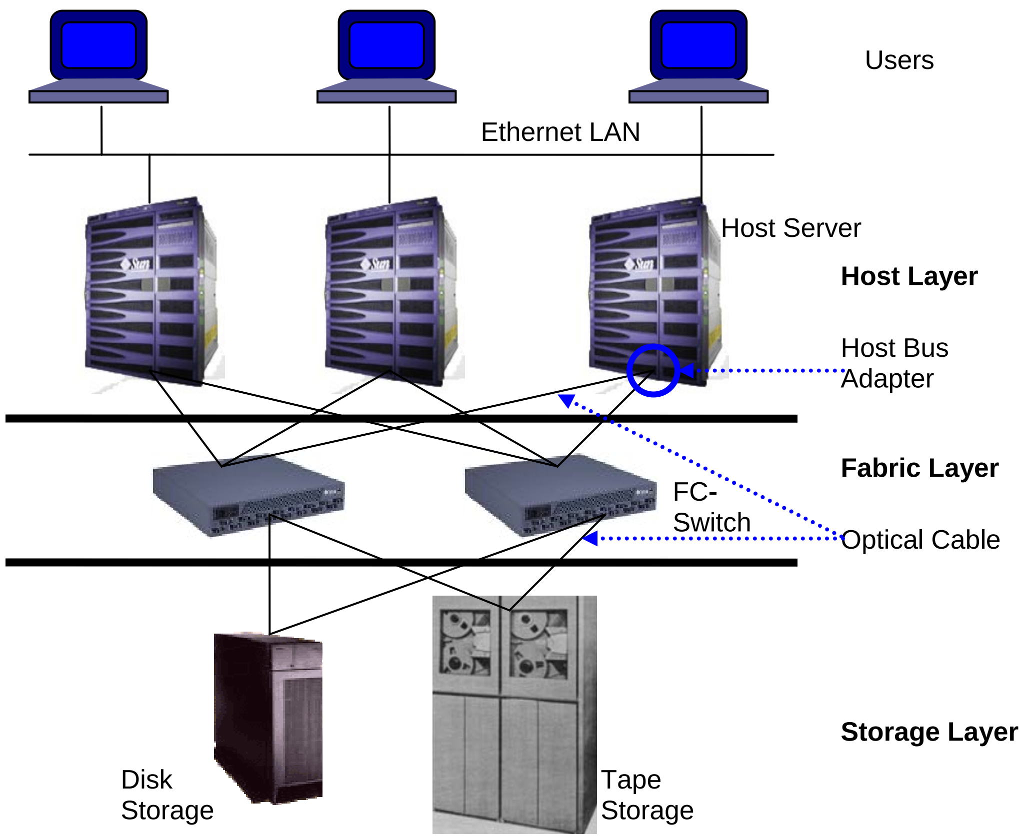
STORAGE AREA NETWORKS

The Storage Area Network (SAN) is a response to the storage management challenge and opportunity. While a relative newcomer in the world of distributed networked servers, the SAN concept was long used in the mainframe environment [Vacca, 2002].

A SAN is a dedicated high-speed network linking storage devices tied to a local area network at the application server [Milanovic and Petrovic, 2001]. SAN is sometimes called the [storage] 'network behind the server' [Vacca, 2002]. A SAN is similar to a LAN in that it establishes direct connections among storage elements, clients, and servers. As explained in Sections II and IV, Storage Area Networks are implemented in a variety of ways, including Fibre-Channel (FC) and IP-based versions. Figure 1 shows a typical Fibre Channel (FC) - based SAN.

The SAN allows client and server devices to view a pool of storage devices, often geographically dispersed, as a single storage unit. Data [file] requests are transmitted only on the SAN, leaving the LAN dedicated for application related requests. This pool of storage is 'virtual', presenting a unified view of all the disparate hardware using a process called virtualization [Preston, 2002, Vacca, 2002]. The storage administrator uses both virtualization hardware and software to ensure that a user simply sees a list of storage options, without knowing where they are located physically.

[2] Explained later in this section
[3] A switch is a network device that provides alternate paths for high-speed data routing
[4] Lower cost implies a smaller, less feature-intensive SAN solution and a lower network size. For example, as of September 2003, HP's OpenView Storage Area Manager 3.1 cost $36,000, but provided minimum functionality [Infoworld, 2003]. The high cost of management software in general is keeping small businesses to migrate to SAN.



Adapted from Poelker and Nikitin, [2003]

Figure 1. FC-Based Storage Area Network

ORGANIZATION OF THIS TUTORIAL

The foregoing section describes a simple SAN. In Section II we discuss the components of SAN, followed in section III by applications and benefits. We discuss the predecessors to SAN and extensions of SAN in section IV. In Section V, we discuss SAN management, including cost implications, problems with SAN including interoperability concerns and SAN management software standards such as Storage Management Initiative Specifications (also known as Bluefin). Section VI highlights major trends related to Storage Area Networks and Bluefin. Section VII examines implications for practice. The eighth section outlines a SAN research program. The final section offers concluding remarks.

II.SAN PHYSICAL COMPONENTS

In this section, we describe a simple SAN architecture and its components. The description covers a typical physical setup. Variations around this configuration define the products of specific vendors.

TYPICAL SAN CONFIGURATION

The main purpose of a SAN is to address storage input/output needs with high speed. It uses optical fiber[5] cables to connect its components. Users are connected to the server via a LAN using Ethernet hubs, routers, and switches. The server is then connected to a pool of storage devices, using interconnecting "fabric" [Adlung, 2002, Heath and Yakutis, 2000, Vacca, 2002]. The term 'fabric' implies a network of switches.

SAN COMPONENTS

SAN components are similar to LAN components, except that the standards they follow are typically Fibre-Channel (FC) rather than IP based[6]. A SAN typically consists of:

- servers (that are connected to the LAN),
- communication media (copper and optical fiber cables),
- switches, hubs, gateways, and host bus adapters at the physical level that interconnect servers with storage devices such as RAID, and software protocols for communication over these physical interconnects.

Additional software is needed to manage the SAN centrally [Poelker and Nikitin, 2003, Vacca, 2002]. The three basic types of SAN components (shown in Figure 1) are:

- **Hosts** – include servers and applications residing on servers, servers' operating systems, host bus adapters, drivers, and routing software. Servers that need large amounts of storage and/or quick access to data (such as file servers, database[7] servers, multimedia servers, and mail servers) are connected to a SAN. Servers that need less storage and/or do not need access to data quickly (e.g., Web servers, Domain Name Servers) and desktops are not typically connected in a SAN.
- **Fabric** – includes FC-based hubs, switches and gateways, related operating systems, and fiber optic cables. In 2002, FC-switches were running at 2Gbps [Thompson et. al., 2003], and were predicted to increase to 10 Gbps in future implementations [Bird, 2002, Thompson et. al., 2003]. By 2004, a survey of 600 Infoworld readers showed that most still preferred to use 2Gbps FC devices. Those who do look at 4Gbps speeds are typically less concerned with cost and more concerned about the response time from their high-end data centers [Apicella, 2004a].
- **Storage** – disk and tape storage, storage software

Protocols/Typologies - Fibre Channel-Arbitrated Loop (FC-AL) and Fabric are two competing FC-based protocols and typologies. In FC-AL8, the servers and storage devices are connected in a loop (similar to Token Ring typology) whereas a Fibre Channel interconnects switches and gateways to create a switched typology [Connor, 1999, Vacca, 2002]. Generally, the term 'fabric' is used colloquially to refer to the switch typology that ties different interconnects together [Brocade, 2001]. Different types of fabric typologies include mesh, star, and ring typologies. The

---

[5]'Fiber' is different from 'Fibre'. 'Fiber' is used in association with optic cables while 'Fibre' is used in association with Channel and implies that Fibre Channel can run over not just fiber optic cables but copper cables as well [Preston, 2002]. Plus Fibre Channel typically implies a set of standards and not just an optic cable [Poelker and Nikitin, 2003]. Hence the standards committee deliberately chose the European spelling of Fibre.

[6] A detailed description of IP-based SANs are beyond the scope of this tutorial because there are several different versions of it, such as Fibre Channel over IP, iSCSI, Fibre Channel Backbone and Nishan's Storage over IP [Vacca, 2002]; However, they are discussed briefly in Section IV.

[7] In a typical data center, each application server is statically assigned its own dedicated database server which then connects to the SAN [Dar et. al., 2004]. But it Is possible for a DBMS to connect to either raw block storage (via block-level-access provided by SAN) or to a file system provided by Network Assigned Storage (NAS , which is discussed in Section IV) [Voruganti et. al., 2004]. For details on how DBMS interacts with application servers, see Dar et. al., []2004]. To find out how DBMS interacts with NAS and SAN, refer to Voruganti et. al. [2004].

[8] A more detailed description of FC-AL is presented in Section III.

server and storage nodes are not connected to each other directly, but to the "network of switches" or fabric [Poelker and Nikitin, 2003].

In Figure 1, the data travels from a host server, via the host bus adapter (HBA) and the optical fiber, to the network of switches. The switch's operating system routes the data to an appropriate RAID or tape device, again over the optical fiber [Apicella, 2004b, Poelker and Nikitin, 2003]. The data can belong to any application. The application simply communicates (i.e., sends data) with its own operating system, the same way it would in the absence of a SAN. SAN works behind the scenes and is not 'seen' by the applications. The operating system then communicates with the HBA using an HBA driver and passes the data on to FC-switches and finally to storage devices[9]. From a user's point of view, a file's datapath (e.g., /Cisco/confirm.doc) would look no different on a SAN [Gibson and van Meter, 2000].

III. SAN APPLICATIONS & BENEFITS

SAN is significantly different from its predecessors such as Direct Attached Storage (DAS) and Redundant Array of Independent Disks (RAID), partly because it creates a shared pool of storage devices that are networked together to create a single view of storage. It is extremely efficient, not just because of its design, but also because it normally uses high-speed media such as fiber optic cables and the Fibre Channel standard that allows for block- level access to storage devices, something that is not offered by TCP/IP. Furthermore, detaching servers from storage (a) protects important organizational data even when a server crashes, because alternate paths to data exist via other servers and switches and (b) allows the server to relinquish its storage responsibilities thus allowing machines with low CPU performance to become a server.

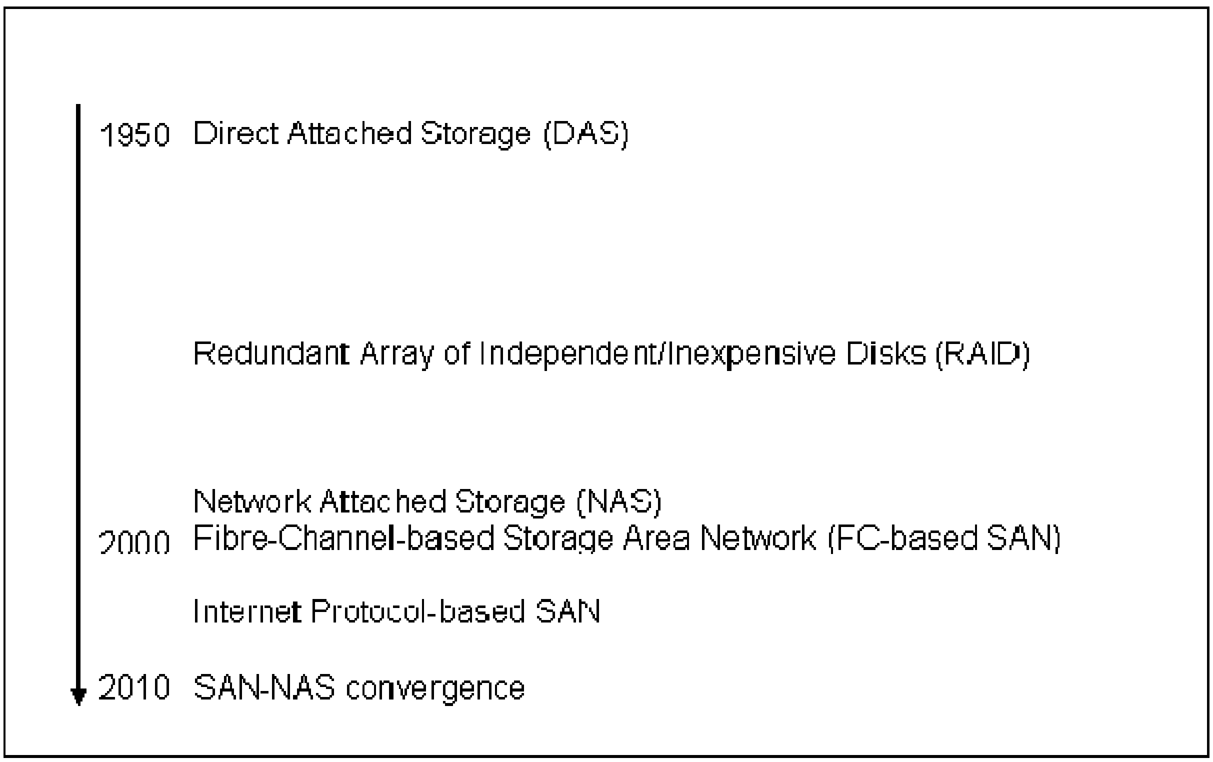


Figure 2. Timeline for Peers and Predecessors of SAN

[9] For a detailed description of how the data passes through different devices and cables, refer to Poelker and Nikitin [2003].

Some of the benefits of SAN include [Bird, 2003, Thompson et. al., 2003, Vacca, 2002]:

- Centralized management – provides a single console with SAN management software that monitors and controls heterogeneous proprietary storage devices and other software. The alternative to this single-console management is quite unattractive. For example, with say, four different vendors' devices and four different software packages that control and monitor those devices, the learning curve for the SAN administrator is four times higher [Scheier, 2003]. In addition, it is more expensive to procure and maintain four different management software packages as opposed to one.
- Network architecture – separates storage resources and servers allow higher application availability and data availability, independent of one another
- Clustering servers – provide server redundancy when servers fail; high availability, and scalability
- Data protection – provides multiple paths to shared storage devices (e.g., RAIDs) and twenty times faster LAN-free backup than conventional LAN backup [Veritas, 1999].
- Data vaulting – transfers archival data or an activity log to a remote site for disaster recovery.

Although some of these features are provided by alternate storage networking solutions, a SAN is both extremely efficient and comprehensive. If large amounts of valuable data is stored or is needed in real-time, SAN can be leveraged successfully. Common application areas are found in the military, immigration services, the media industry, weather forecasting, genetic engineering, drug design, nuclear physics, and others as described below.

ENTERTAINMENT AND HEALTH CARE

The entertainment and healthcare industries demand large amounts of storage for images, video, text, and sound. In the entertainment industry, one-hour[10] of video footage requires 22 Gigabytes for storage [Guha, 1999]! In medicine, a typical eye examination generates 200 megabytes that require approximately 200 seconds to transmit over a 10BaseT Ethernet[11] connection. On a SAN, the same eye exam image could, theoretically, load in 2 seconds [Chernyshov, 1999]. Thus, SANs not only improves the speed of data-delivery, but also allow a less powerful CPU to take on server responsibilities and reduce total cost of ownership. Streaming video over the Internet also use SANs, sometimes in combination with other networking storage solutions [Wu et al., 2001]. Horwitt [2003] describes a New Jersey hospital environment where patient data is sent over a newly implemented SAN to doctors facing critical care decisions. The 300 multi-vendor servers used previously resulted in poor disk utilization.

GRID COMPUTING

SANs also play an important role in grid computing. Grid computing refers to the use of several computational resources connected via a network to solve a single problem which typically requires major computing power and access to large amounts of data (see, e.g., IBM [n.d.]). IBM started testing a SAN virtualization solution – Storage Tank - at the European Organization for Nuclear Research. It was agreed that if a global grid of 8,000 nuclear scientists were able to

---

[10] This number will vary depending on the video standards (e.g., MJPEG, DV) used. For example the DV compressed video requires 13 GB for 1 hour of footage. To put things in perspective, The Lord of the Rings: The Two Towers (the edited version) runs for 3 hours (179 minutes). The RSPB Film Collection on bird footage runs 100 hours and requires one million feet of projection film. Conus maintains over 15,000 hours of footage just on major personalities and events in the U.S. in the past 2 decades. These examples are just a tiny fraction of the entertainment industry's storage requirements.

[11] 10BaseT is an implementation of Ethernet which allows stations to be attached via twisted pair cable [Wikeipedia, 2004].

access the data from their own laboratories, using heterogeneous computing and storage technologies, then the iSCSI-based[12] SAN solution would be deemed successful enough for other commercial ventures [Sayer, 2003]. This experiment, if successful, will demonstrate that SANs are highly scalable for grid computing at a global level.

OTHER EXAMPLES

- Physicists from around the world shared results from experiments with a supercollider using a SAN [Sayer, 2003].
- In March 2004, Greenpeace UK, which relies on e-mails to communicate its campaign, bought Linux-based SAN to revamp its e-mail servers and create reliable access to e-mail data archives which use DB2 databases and Microsoft Exchange servers [Computerworld Staff, 2004].
- Around the same time, *Arizona Republic*, a newspaper company in Phoenix Arizona, created a new disaster recovery plan based on a 10TB SAN [Redding, 2004].
- Johnson Memorial Hospital based in Stafford Springs, Connecticut, expanded its existing FC-based SAN in September 2004 because its data center demands doubled and its user population increased by 30%. The hospital wanted to support paperless medical records, scheduling systems and electronic archiving systems.

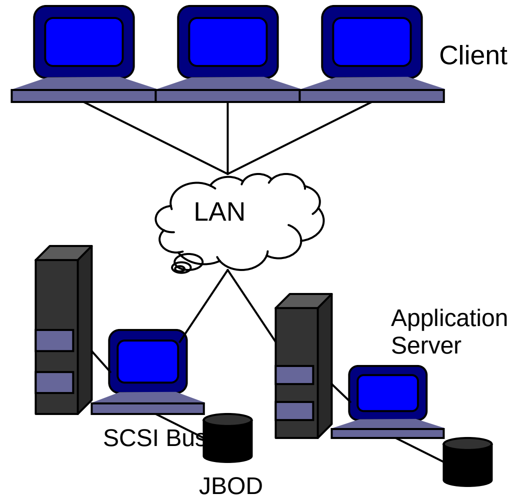
IV. PREDECESSORS AND EXTENSIONS TO SAN

TRADITIONAL DATA STORAGE

Data storage architectures traditionally consisted of a simple computer network with several clients connected to the server. As shown in Figure 3[13], the server is directly connected to storage devices (such as disks and tapes) via a single storage interface [Guha, 1999, Molero et al., 2001] (e.g., SCSI[14] [Small Computer System Interface] or IDE [Integrated Development Environment]). As recently as 2002, as much as 99% of storage remained directly attached to servers via a SCSI or IDE bus [Storage Research Corporation in Vacca, 2002]. In such a scheme, the server is responsible for moving user data and application programs to and from tape and disk storage. An increase in users or user requests can lead to a bottleneck either between the server and the storage device or at the storage device. Adding further storage devices reduces load-leveling[15] and reliability. In addition, if the server goes down, all access to data is lost [Molero et al., 2001, Robinson, 2002]. Application residing on a particular server will access data attached only to that server – other storage attached to other servers, will not be accessible to it, even though they may have usable storage space; each application will have to know what data is accessible from its server [Coates, 2003].

---

[12] iSCSI refers to Internet SCSI. It uses the SCSI protocol over a TCP/IP network. It enables any machine on an IP network (initiator) to contact a remote dedicated server (target) and perform block I/O on it just as it would do with a local hard disk. This definition is from the Wikipedia [2004] SCSI is defined in Footnote 15.
[13] The term JBOD in figures 3 to 7 stands for Just a Bunch Of Disks
[14] SCSI: Small Computer Systems Interface is a parallel-bus architecture and a protocol for transmitting large blocks up to a distance of 25 meters. It is the oldest standard that is still prevalent. One of the main drawbacks of SCSI is its bus length limitation and its data transfer rate (40 MB/sec.)
[15] Depending on the current number of users (load) that are logged on to each server, a new user will be routed to the server with the least load when she logs on. Thus, the users are balanced across many servers in the pool; the users view the server pool as a single unit [Boisvert, 2001].

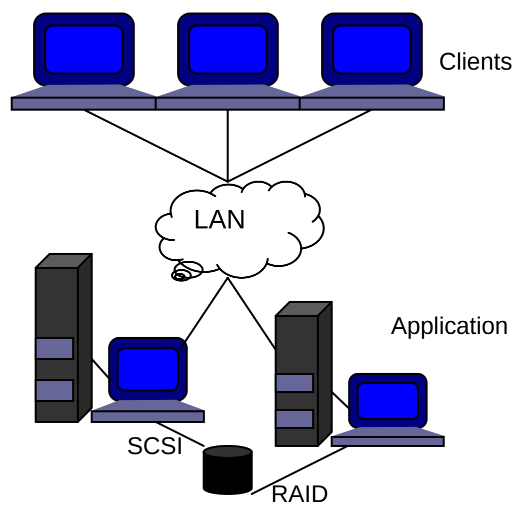


Adapted from Robinson [2002]

Figure 3. Traditional Direct- Attached Storage

RAID

One way to address these problems is to use RAID technology (Figure 4). RAID, (Redundant Array of Independent Drives) uses a dedicated controller[16] that manages load-leveling and provides data redundancy for rebuilding data if a drive fails. That is, several copies of the same data are maintained to ensure high data availability [Boisvert, 2001]. The server-to-storage bottleneck can again only be eliminated by using faster interfaces – faster SCSI. When the server



Adapted from Robinson, 2002

Figure 4. RAID Direct-Attached Storage

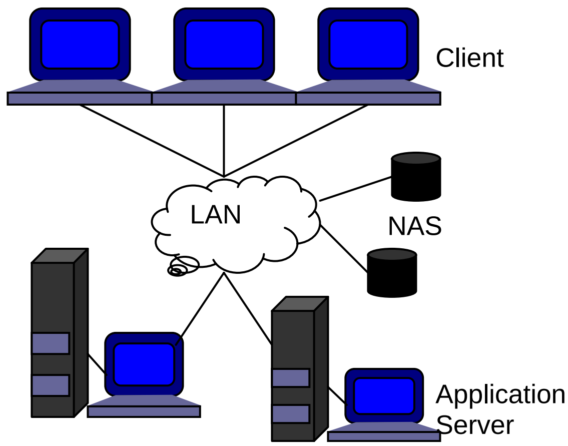
[16] A controller is a program module or a hardware device that interprets signals between a host and a peripheral device so as to control how data is written and retrieved [Vacca, 2002].

itself becomes a bottleneck, more servers need to be added with their own directly attached storage device. This approac h creates problems with management, scalability, and unnecessary data duplication as well as increasing costs [Khurshudov, 2001].

NETWORK ATTACHED STORAGE

NAS (Network Attached Storage) and SAN (Storage Area Network) solve the problems created by attaching servers directly to storage. In both approaches, several servers share a pool of storage devices via a network. A typical server is separated into two specialized servers: one is application-specific and the other is storage-specific. Now, the storage 'device' is no longer just a hardware unit, but loaded with software and protocols. It is a storage server (or appliance) and carries out administrative functions [Khurshudov, 2001].

As shown in Figure 5, a NAS is a dedicated storage server that connects directly to the LAN instead of connecting to the server. The NAS 'server' (also called a NAS appliance) carries LAN interfaces and file-access protocols such as NFS (Network File System for Unix) and CIFS (Common Internet File System for NT) [Gibson and van Meter, 2000].



Adapted from Robinson [2002]

Figure 5. Network-Attached Storage

The application server no longer needs to support traditional storage interfaces (such as SCSI) [Khurshudov, 2001]. The advantage is that now any client or server with any operating system can access NAS storage via an already existing network [Khurshudov, 2001, Robinson, 2002]. The client does not have to go through the server to transfer data, thus conserving CPU cycles previously spent processing storage requests. The drawback of NAS is the lack of a high-speed dedicated connection between CPU and storage units – they still must use the LAN to communicate among one another, thus continuing to create bandwidth bottlenecks[17]. Furthermore, since the LAN uses Internet Protocol (IP) and all client requests for files are processed using file-access protocols (NFS/CIFS), CPU cycle time is required to convert file

[17] Using LAN for both data and application requests is a primary contributor to system bottlenecks, particularly when files are being backed up. If data were to flow on a separate path, then the LAN would be free for applications. The separate data path is frequently a Fibre channel thus permitting extremely high-speed (1Gbps-2Gbps) data backup and recovery – which is usually not possible on a LAN. Therefore, storage administrators look forward to the SAN promise of LAN-free backups.

requests into block-level requests[18] that can directly interact with storage servers [Renato, 2003, Vacca, 2002]. For these and other overhead-related reasons, NAS is normally used only for simple data backup [Khurshudov, 2001].

When bandwidth is critical, and file system functionality is not [Gibson and van Meter, 2000], SAN is a more appropriate solution. SAN is a web of different storage devices sharing a dedicated, high-speed network that, as was shown in Figure 1, is connected to users and servers on a LAN [Adlung, 2002, Brinkmann et al., 2000, Heath and Yakutis, 2000, Thompson et. al., 2003].

TYPES AND EXTENSIONS OF SAN

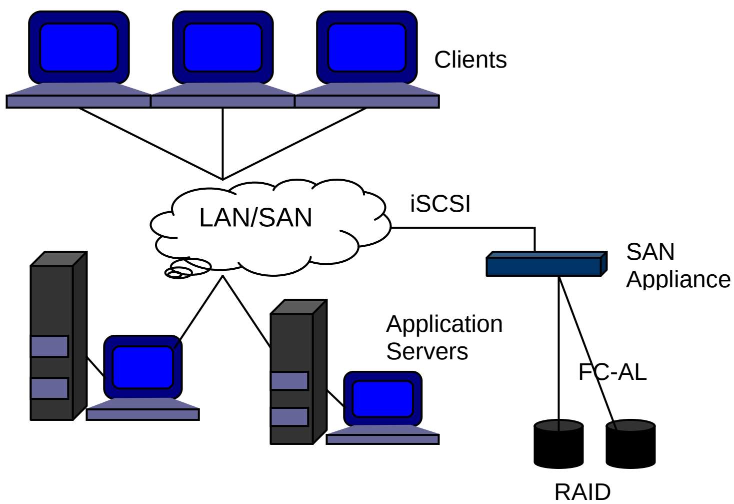
A popular data transport in use for SAN is Fibre-Channel-Arbitrated Loop (FC-AL). FC-AL is an industry-standard, high-speed serial data transfer interface that can be used to connect systems and storage in point-to-point or switched topologies. FC-based switches can support up to 256 ports, a FC-based loop can support 127 devices and FC-based interconnect links can support up to 2Gbps at distances up to 6.25 miles [Heath and Yakutis, 2000, Thompson et. al., 2003].

FC-based SAN's advantages over NAS are numerous. Among the most important are

- scalability,
- availability of a dedicated path between storage and servers and
- the ability of Fibre Channel to allow data transfer at a block-level[19] rather than at the slower file-level.

On the down side, organizations must invest in FC specialists and training.

Further advancements in SAN include the development of IP-based SANs, as shown in Figure 6.



Adapted from Robinson[2002]

Figure 6. IP-based SAN

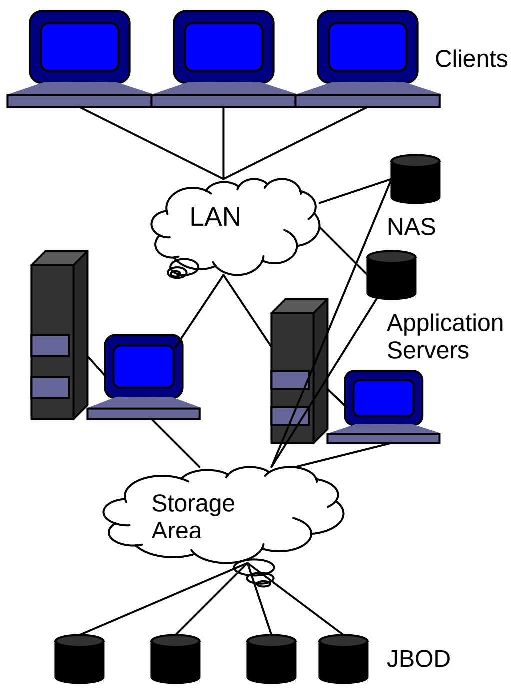
[18] A logical file is actually stored physically in 'blocks' – which may or may not even be on the same disk. Storage devices do not understand file-level requests.

[19] When a NAS device receives a file-level request (e.g., open san.doc) from a user, it searches its file system and translates the logical file name (i.e., san.doc) to a list of the actual physical block addresses where data is located. Converting file-level I/O to block-level I/O requires CPU cycles resulting in considerable overhead. On the other hand, SANs provide direct block-level access to the physical hardware [McConnell, 2001].

IP-based SANs will wrap another layer over the basic IP that will allow block-level data transfer to compete with Fibre Channel capabilities. A major advantage of IP-based SANs is that organizations can use existing IT expertise [Robinson, 2002] and extend the existing IP/LAN infrastructure to build a separate storage area network[20] [Renato, 2003]. iSCSI (Internet SCSI) is also a type of IP-based SAN technology. IP is a mature technology, which now incorporates several advanced security procedures that Fibre Channel technology only just started to address [Renato, 2003, Robinson, 2002].

MERGER OF SAN AND NAS

It is expected in the next few years that NAS and SAN will evolve into a single architecture [David Hitz in Betts, 2002, Gibson and van Meter, 2000, Merhar, n.d., Poelker and Nikitin, 2003, Satchell, 2003], as illustrated in Figure 7. Most of the NAS/SAN differences center around performance and architectural issues such as block-level versus file-level access. The differences in NAS and SAN are slowly fading away in terms of reliability, capacity, speed and other performance measures [Earls, 2003]. In terms of architectural differences, virtualization is proposed as a solution that will kill the block-level versus file-level debate [Benett, 2002]. Other options of merging SAN and NAS include interfacing NAS devices between the clients and the SAN, or packaging SAN and NAS features all in one box [Earls, 2003].



Adapted from Robinson[2002]

Figure 7. Merged NAS-SAN Typology

SWAN

SAN subsystems jointly coordinated over long distances are collectively called a Storage Wide Area Network or SWAN. In addition to providing off-site backup and mirroring, SWANs can also

[20] It must be clarified that IP-based SANs can use a separate network to connect the storage devices – then the SAN components (such as switches) will be IP-compliant rather FC-compliant. For this reason, IP-based SANs also provide the same benefits (such as LAN-free backup) as FC-based SANs. For more information on IP-based SANS, visit http://www.snia.org/ education/ip_storage.pdf

disseminate information globally [Yoshida, 2001] because they can connect FC-based SAN subsystems using the Internet's IP protocol. But it must be noted that SAN's FC protocol differs considerable from typical WAN protocols such as IP. As SANs scale, they will have to deal with address bandwidth congestion and addressing issues that IP had to confront earlier [Falk et al., 2003].

ASSESSMENT

It is indeed difficult to choose among the variety of storage networking configurations discussed in this section. Lei and Rawles [2003] assessed many of these solutions on different Linux and Windows environments, for NAS, SAN and DAS typologies. Their assessment is based on real options theory using measures such as affordability, scalability, manageability, adaptability and usability. For example, in a networking laboratory setting, they found that NAS design, with Windows hosts and virtual machine technology was the most appropriate choice with respect to affordability and manageability from the real options grid.

Table A1 in Appendix I outlines the components and potential benefits and weaknesses for storage architectures discussed in this section, starting from the earliest DAS, and ending with IP-based SANs.

V. SAN MANAGEMENT

JUSTIFYING THE SAN INVESTMENT

Return on Investment for SAN is estimated to range from 65% to close to 300% [Zamer, 2001]. SANs' major benefits include reliability, scalability, and flexibility [Beck et al., 2002, Guha, 1999, Milanovic and Petrovic, 2001]. As discussed in Section III, SAN's proponents also promise a single-console view and management of all the heterogeneous distributed proprietary SAN hardware and software components [Bird, 2003, Milanovic and Petrovic, 2001] through interoperability[21]. Without interoperability, most servers are typically not tested to work together, and if complex customized server configurations are added to the mix, it can result in more server crashes [Gibson and van Meter, 2000].

SAN makes it possible to add more storage capacity without adding servers and upgrading servers without adding new storage and simultaneously allowing operations to continue 24x7 without any downtime [Satchell, 2003]. Thus, one of its biggest promises is scalability [Beck et al., 2002, Brinkmann et al., 2000, Clark, 2002, Guha, 1999, Satchell, 2003]. SANs also offer 'LAN-Free Backup' i.e., a number of backup, mirroring[22], and snapshot copy features that do not add a load on the LAN. With SAN, it is now possible for a user to work late and access the LAN quickly, because it is not being used to backup data overnight anymore [Riedel, 2003]. International Data Corporation (IDC) estimates that 50% or more of direct-attached disk space is currently unused. In a SAN, all servers share all the storage devices. Thus, unused disk space belonging to one

[21] Interoperability allows the storage administrator to identify, monitor, and control storage devices belonging to different vendors using a single management software environment (instead of different vendors' software) to identify and control their respective storage devices [Scheier, 2003]. This approach reduces the cost of learning to use the software and the manual labor involved in using different consoles (software) for different devices. Interoperability also prevents vendors from locking-in the customer into their solution. If all SAN solutions are interoperable, then a customer can choose the best (i.e., the most suitable) servers, the best HBAs, the best switches, the best cables, the best storage arrays, the best tape libraries, the best identifying, monitoring, routing, backup and recovery software – all from different vendors, and integrate them into one SAN.

[22] Mirroring allows data to be written simultaneously to two hard disks. Data will be lost only if both hard disks fail at the same time [Vacca, 2002].

server that was previously inaccessible to other servers, can now be shared, thereby increasing disk space utilization by at least 25% [Yoshida and Dolcini, 2000].

POTENTIAL PROBLEMS WITH SANS

Although SANs offer numerous advantages over previous storage network typologies, the technology is still in the early stages of its evolution. Among the reasons for delaying adoption are:

1. SANs are still quite expensive, although prices are dropping [Goodwins, 2002, Vacca, 2002]
2. SANs are still not completely interoperable. The user is forced to stay with one vendor to manage the SAN centrally [Goodwins, 2002, Vacca, 2002]
3. Standards that support interoperability are evolving at a slow pace. For example, Bluefin (see below) is not expected to mature until 2006 [Reich, 2002]
4. SAN vendors are suing each other over patent infringements [Goodwins, 2002]
5. SAN management software that provides central SAN administration is sometimes proprietary and difficult to integrate with current IT infrastructure [Goodwins, 2002]
6. Like other networking storage options, SAN solutions are also more vulnerable to data integrity and privacy attacks [Gibson and van Meter, 2000].

In the next two sub-sections, we explain SAN management software and interoperability further.

SAN MANAGEMENT SOFTWARE

Traditionally, network management involves reliable data transfer from its source to its destination. Reliable transfer requires attention to server uptime, bandwidth utilization, alternate data path guarantees, multiple protocol support, and error-free delivery. But reliable storage management also involves the organization and placement of data once it arrives at its destination. Concerns here include RAID levels, backup, and disk utilization (i.e., ensure that no disks remain under-utilized, a typical problem with the traditional DAS). Because SAN is a network of servers and storage, SAN management requires a pooled approach that includes both traditional network management and traditional storage management [Dot Hill Systems, n.d.]. The SAN management software may reside either on the servers or on the SAN appliances.

Examples of network management functions [Kerns, 2000] include:

- Monitoring the network
- Automatic discovery of devices
- Logging changes
- Managing events and alerts
- Setting thresholds and rules
- Managing security
- Managing service level agreements
- Managing chargeback
- Managing cluster of servers
- Managing policies

Examples of storage management functions [Kerns, 2000] include:

- Installing and configuring drives and related software
- Adding and upgrading storage
- Managing data capacity (used vs. available)
- Analyzing usage trends
- Migrating data
- Retiring devices
- Managing backup/restore (server less/LAN-Free)
- Managing file and data sharing
- Balancing server load
- Maintaining SAN file systems

Managers need to assess whether the features of a given SAN works only in certain environments (e.g., NT but not UNIX) [Dot Hill Systems, n.d.]. Managers must also know which OS platforms the SAN supports, as well as the vendor's interpretation of 'interoperability'.

INTEROPERABILITY

It is often difficult to recognize a proprietary solution, one that ensures a customer is locked-in to a specific vendor. Deciding whether a solution is proprietary is often relative to time and scope and location of the solution (Sidebar 2).

SIDEBAR 2. CHARACTERISTICS OF PROPRIETARY SAN SOLUTIONS

A solution is proprietary when:

1. An operational change required for a particular type of (hardware) device eliminates the possibility of replacing that device with one offered by a competitor.
2. A SAN's scalability is limited to particular devices.
3. Proprietary hardware and control software are incorporated with storage devices and switches to provide a complete SAN.
4. The SAN is limited in the number of servers and the storage systems involved. [Kerns, 2000]

Vacca [2002] claims that the future success of SAN depends on meeting "the key assumption that standards will be developed and incorporated into SAN products". Products that comply with the standards will be interoperable, thus allowing for truly open SAN solutions.

THE BLUEFIN INITIATIVE

Because vendors' preferred definitions vary, SAN components differ in the criteria that identify them as being open. One way of avoiding multiple interpretations of SAN management interoperability is to arrive at industry standards. In 2002, the SNIA[23] pioneered a project to define a network storage management API (application programming interface) that can be incorporated into both the hardware devices and the management applications. This effort is intended eventually to allow vendor-specific products to inter-operate so that

- they can be integrated and then managed centrally.
- customers are free to select the product that best suits their needs without being locked-in to a vendor.

This Storage Management Initiative Specification (SMIS), commonly referred to as 'Bluefin', is based on the Common Information Model (CIM) - Web Based Enterprise Management (WBEM)[24]. Wide acceptance of this standard is expected because it relies on open-source code that supports CIM/WBEM [Reich, 2002, Reich, 2003, Storage Networking Industry, Association].

Bluefin will allow SAN vendors to decrease their products' time-to-market while reducing the tedious effort of integrating incompatible management interfaces; vendors can instead focus resources on building more functional management engines [Reich, 2002, Reich, 2003, Storage Networking Industry Association, n.d.].

---

[23] The Storage Networking Industry Association (SNIA) is a not-for-profit organization, consisting of over 300 organization and individual members (www.snia.org). The objectives of the Storage Management Initiative are available at http://www.snia.org/smi/home.

[24] The WBEM pyramid of standards has three layers – the bottom-most is HTML, the middle layer is XML that is coded specifically for the top-most layer or the CIM [Webster, 2004]. Common Information Model is a management structure enabling disparate resources to be managed by a common application [Vacca, 2002]. All storage devices and applications are types of the CIM Object with different attributes. Thus, the CIM schema is object-oriented. As long as different objects (marketed by different vendors) are defined using the CIM schema, they be able to interoperate [Webster, 2004]. For more information, visit http://www.dmtf.org/ standards/index.php

Bluefin/SMIS provides the following features needed for its success [Storage Network Industry Association]:

- The "Common Information Model - XML over HTTP" standard allows vendors to extend the features and functions of their products dynamically without starting from the beginning.
- One single object model (classes, properties, methods) allows SAN developers to understand and implement SAN-management components. It also specifies how to build these components.
- A discovery system automatically announces the presence and capabilities of a newly plugged-in component.

In early 2003, the first Bluefin specifications were released, but the industry still needs to implement Bluefin in products. SNIA's goal is that all storage products are SMIS compliant by 2005.

VI. CURRENT AND FUTURE TRENDS

Where is SAN headed? What does the SAN market look like? In this section, we show how the SAN market is progressing in mid-2004, the struggle between FC-based and IP-based SANs, progress on Bluefin specifications and Bluefin-compliant products.

FIBRE CHANNEL VERSUS INTERNET PROTOCOL

In 2002 several technologies were competing for standardization to carry SCSI traffic over IP networks (iSCSI, FC over IP, and iFCP – Internet FC Protocol). In early 2003, iSCSI met with industry approval to become a protocol standard. Following that decision, Hewlett Packard (HP) announced that it would market iSCSI-compliant routers and Microsoft announced its would introduce iSCSI-compliant drivers in Windows XP. Introduction of iSCSI-compliant storage in the market implies that current LAN capabilities can be extended to build a SAN, without the need for large-scale Fibre Channel deployment. However limitations such as network management, interoperability, performance, and cost continue to restrict the use of iSCSI through 2004 [Derrington, 2002b]; by 2005 it is expected to be the de facto standard for IP-based SANs and might be a popular choice primarily for branch offices - not located within a corporate data center [Derrington, 2002a] where FC-based SANs may prevail. IDC reported that iSCSI totaled 1% of the networked storage market in the first quarter 2004, but that small amount was still a 40% increase from the previous quarter [SystemsWorld, 2004]. Although FC-based SANs are predicted to dominate through 2005-2006 [Derrington, 2002b], IP-based SAN equipment sales will exceed FC-based equipment sales by 2007 [MarketResearch.com, 2003].

As of 2004, Fibre Channel devices support both 2 Gbps and 4Gbps products though the 2Gbps ($1000 per port) devices are more popular because most organizations are concerned about the high cost of SANs. The 4Gbps devices are selected only by those organizations that require higher performance and are less concerned about costs [Apicella, 2004a]. Standards are being developed for 8Gbps products that are expected to arrive in the market in 2007. All of these units are backward compatible with their slower versions, except 10Gbps devices, which are available even today, but at a steep price of $5000 per port [Mearian, 2004].

THE FUTURE OF BLUEFIN / SMIS

SNIA released the first version of the Bluefin standard (Section V) in April 2003. But SNIA officials predict that Bluefin may take up to seven years [Kerns, 2000] to become an accepted industry standard with stable specifications. By June 2004, thirteen vendors, including Brocade, EMC, Hitachi and HP promised to incorporate Bluefin into their storage devices. CISCO made its MDS 9000 switches Bluefin compliant [Byte and Switch, 2004]. In addition, AppIQ Inc. is incorporating Bluefin into its management suite. Bluefin-compliant management policies are not expected to emerge in end-user products before 2005-2006. More Bluefin-compliant features (e.g., local snapshot copy) will not be available before 2006-2007 [Goodwin, 2002]. For some time to come, vendors are expected to continue to offer proprietary management features and/or centralized proprietary management software for managing a specific constellation of heterogeneous vendor hardware (e.g., EMC WideSky[25], HDS TrueNorth, IBM Storage Tank, and HP OpenView) [Goodwin, 2002]. As demonstrated regularly at SNIA sponsored conferences featuring live prototypes of Bluefin-compliant SAN components, the storage industry progressed significantly; however, complete interoperability and plug-and-play product installations remain in the future. Some skeptics predict that progress will be slow as vendors try to retain their market share by locking in customers with their own proprietary solutions; more optimistic analysts believe that vendors are more serious this time since they too want some interoperability and consequent reductions in R&D costs [Byte and Switch, 2004].

VII. IMPLICATIONS FOR PRACTICE

New storage network technologies provide finer level of resolution for global storage; providing powerful tools for organizing vast repositories of data and information within a building, organization, within municipalities, across distributed organizations, and throughout the world. SAN is a strong contender in realizing this vision, because it is not simply a more efficient solution, it is a more effective one. It is not based on simply adding more existing components (e.g. RAID); instead it is based on a new architecture and some new technology.

If storage is to be a key corporate differentiator, an organization can take several measures to harness SAN technology quickly:

1. Ensure that the organization does not become involved in competing standards initiatives, especially those being fought among FC, IP-based protocols (such as iSCSI, FC over IP), and iFCP. Explore whether vendors provide Bluefin-compliance in their products.
2. As discussed in Section V, prospective buyers should be certain that vendors are correct about their interpretation of interoperable systems before buying an 'open' SAN solution. Find out whether specific SAN features work only in certain environments.
3. SAN implementation can begin at a departmental level; for example, by investing in fewer FC-based switches and connecting only the most critical servers [Apicella, 2004b]. New servers and switches can be added once a comfort-level is reached with the current functionality of the SAN.
4. Select the SAN typology (arbitrated loop, fabric[26]). The choice will depend on the number of servers, types of switches, and cost.
5. Although SANs can be scaled to span an entire campus or city by creating SWANs (Section IV), security trade-offs must be considered before making a choice.
6. Participate in the Storage Management Initiative by enrolling as members of SNIA (www.snia.org/ tech_activities /SMI/).
7. Since standards are still emerging, and the interpretations of these standards vary considerably from vendor to vendor, an initial SAN experiment might be limited to a single vendor [Apicella, 2004b].
8. If the organization is in a high storage demand industry such as entertainment, healthcare, or energy, consider SAN as a viable option.

A detailed hands-on approach to installing and managing SANs can be found in Poelker and Nitikin [2003] and Vacca [2002].

---

[25] EMC dropped its WideSky initiative in September 2003 in favor of the SMIS (Bluefin).
[26] Within the fabric typology, firms can select from a star, a ring, or a mesh typology. A detailed discussion can be found in Poelker and Nikitin [2003].

VIII. WHAT WE COULD STILL LEARN ABOUT SANS

In this section, we briefly summarize what we know about SANs and what we still must learn in a systematic study of SAN development and management. Advances on the technical side are many, including different types of SAN such as FC-based SANs, IP-based SANs, combinations of SAN and NAS, and extended SANs or SWANs. However, security, scalability, interoperability, adoption of standards, and trade-offs between cost and manageability remain on-going concerns. We discuss some of these issues in this Section.

SAN SECURITY AND SCALABILITY

As SANs are scaled to become campus or metropolitan area networks (i.e. SWANs), they become more vulnerable to security breaches. How security is implemented in such SWANs will depend on whether they are FC-based or IP-based SANs [Beck et al., 2002]. To ensure a secure network, private IP-based SWANs may end up sacrificing the scalability provided by the Internet [Beck et al., 2002]. Also, SANs assume that connectivity will not fail unless the entire system fails completely; this assumption creates problems when SANs are extended to SWANs [Beck et al., 2002]. Thus, we still need to know:

- How to scale up a SAN, while continuing to improve connectivity, data availability, reliability, and security?.

BUSINESS VALUE OF SANS

Currently, performance of SANs is assessed by technical metrics, such as instructions per second, sequential reads and writes per second, and CPU utilization. Translating these metrics into improved organizational performance is a far greater challenge. Some work was already done in this area by Lei and Rawles [2003] who use real options theory (i.e., trade-offs between affordability, manageability, scalability, usability and adaptability) to evaluate different combinations of SAN infrastructure and host operating systems. Their work can be used as a starting point to answer the following question:

- What evaluation measures may be used to justify and monitor SAN expenditure and return on investment?

SAN OUTSOURCING

SAN and SWAN implementation and management may be outsourced to Storage Service Providers (SSPs) that provide storage on demand [Yoshida, 2001]. Thus, the role of SANs in providing SSP services needs to be considered. For example,

- How does one develop better service level agreements for data availability, backup protection, and coordination between SSPs, SAN vendors, SAN technology, and end-users?

LEVERAGING TECHNOLOOGIES

The line between different types of SANs and NAS and their costs and benefits is blurring. Thus it is increasingly difficult to decide which technology is better for solving a given managerial problem. Therefore, further areas for leveraging the potential of SAN and storage networking technologies need to be explored. For example:

- How can SANs be deployed in an environment where data is increasingly shared between organizations (e.g., to provide high levels of customer service among the airlines serving the same customers and airports)?

SAN – ORGANIZATIONAL STRUCTURE ALIGNMENT

Another challenge is the choice of the appropriate typology for a particular organizational structure. Molero et. al. [2001] describe several typologies that could span a departmental SAN within one building. They discuss various performance and cost issues related to different typologies. But that research assumes that the required network size and target environment are known. Such is often not the case, particularly in rapidly changing organizational environments resulting from mergers, acquisitions, and divestitures. As a result of acquisitions, for example, data centers may be consolidated, directly affecting SAN infrastructure.

- How does data center consolidation change the current SAN typology and at what cost?

SAN INTEROPERABILITY

SAN interoperability issues are of prime importance today. All standards (including SMIS) contain a technical component. But, arriving at an industry standard also requires consideration of the socio-political factors that force organizations to prefer one standard to another. For example, the SMIS is being developed by active members of the SNIA, but Microsoft is not participating. While the vendors indulge in this tug-of-war, what happens to the customers? We still require answers to the following questions:

- How will political issues affect acceptance of SMIS as an industry standard, even though the standard may address the technical aspects of SAN interoperability?
- How should/do interoperability standards develop? How much time should elapse before a technology demands a standard?
- How should/do customers deal with different standards supported by competing groups of vendors?

DEVELOPING TECHNICAL SAN EXPERTISE

How can IT staff be trained for hands-on SAN expertise (e.g., FC-based SANs, FC-based and IP-based SAN hybrids)? New skills are required to handle trouble-shooting, deployment, and maintenance of SANs. SNIA is currently providing certification for IT professionals in this realm. But is SNIA's certification sufficient?

- What framework should be used to ensure effective, efficient, and timely SAN training?

IX.CONCLUSION

Data is an essential asset for all organizations. Data is related to customers, suppliers, employees, products, inventories, equipment, policies, intellectual property, financial results, business processes and more. New image-oriented applications, particularly in healthcare and entertainment, are increasing demand for storage significantly. Organizations now manage escalating storage demands, instant data access, and backup and recovery that cannot interrupt normal operations. Even though storage hardware prices continue to fall [Guha, 1999], it is the cost of storage software and management that will become the factors that both constrain and energize organizational success.

Ushered in by a cacophony of new technologies and vendor hype, the Storage Area Network is an evolving architecture that appears to offer considerable promise in meeting these challenges. Among the promises of SAN are increased reliability, unlimited scalability, lower management costs, central management of disparate heterogeneous proprietary hardware, and software (viz., interoperability), automatic resource (volume and file) management, network management, typology choices, and security and business continuity. Still, largely unexplored, but risky to ignore are the strategic, organizational and societal implications that will accompany these massive and highly integrated storage systems.

This tutorial summarized the various stages of storage networking solutions, from the traditional direct-attached storage, to NAS to FC-Based SAN to IP-Based SAN. It answers basic questions about what the capabilities of SAN technology are over those of previous storage architectures and describes various components that make up a SAN. The tutorial also examines SAN interoperability and management, and development of related industry standards (Bluefin). Finally, this article looks into the uncertain future in terms of general storage growth, pricing trends, FC versus IP struggles and Bluefin, followed by implications for MIS practice and research.

Editor's Note: The initial draft of this article was received on April 23, 2003. The article was with the authors for 8 months for 4 revisions. It was published on November 30, 2004

REFERENCES

Editor's Note: The following reference list contains hyperlinks to World Wide Web pages. Readers who have the ability to access the Web directly from their word processor or are reading the paper on the Web, can gain direct access to these linked references. Readers are warned, however, that

1. these links existed as of the date of publication but are not guaranteed to be working thereafter.
2. the contents of Web pages may change over time. Where version information is provided in the References, different versions may not contain the information or the conclusions referenced.
3. the author(s) of the Web pages, not AIS, is (are) responsible for the accuracy of their content.
4. the author(s) of this article, not AIS, is (are) responsible for the accuracy of the URL and version information

Adlung, I. (2002) "FCP for IBM eServer zSeries Systems: Access to Distributed Storage," *IBM Journal of Research and Development* (46) 4/5, pp. 487-502.

Alexander, S. (2004) "Hospital Gets Faster Access to Storage With iSCSI Technology," Computerworld, http://www.computerworld.com/hardwaretopics/storage/story/0,10801,95832,00.html (September 29, 2004).

Anthes, G. H. (2002) "Information Overload," Computerworld, http://www.computerworld.com/databasetopics/storage/story/0,10801,72500,00.html (September 2, 2003).

Apicella, M. (2004a) "Research Report: Enterprise storage," Infoworld, http://archive.infoworld.com/article/04/05/28/22FErrstorageT_1.html?STORAGE%20AREA%20NETWORKS%20-%20SAN (current June 30, 2004).

Apicella, M. (2004b) "SAN breaks out of the datacenter," Infoworld, http://archive.infoworld.com/article/04/05/28/22FErrstoragesan_1.html?STORAGE%20AREA%20NETWORKS%20-%20SAN (current June 30, 2004).

Beck, M., T. Moore, and J. Plank. (2002) An End-to-End Approach to Globally Scalable Network Storage. *Proceedings of the 2002 conference on Applications, Technologies, Architectures, and Protocols for Computer Communications, Pittsburgh, Pennsylvania, 2002*, pp. 339-346.

Bennett, C. (2002) "From Storage Virtualization to NAS SAN Convergence," Infostor, http://www.storage-virtualization.com/sannas.htm (current September 30, 2004).

Betts, M. (2002) "Predictions on the Future of Data Storage," Computerworld, http://www.computerworld.com/databasetopics/storage/story/0,10801,75252,00.html (January 18, 2003).

Bird, D. (2002) "Storage Basics - Storage Area Networks," http://www.enterprisestorageforum.com/technology/features/article/0,10564_981191,00.html (September 2, 2003).

Bird, E. (2003) "Connectivity," *Strategic Finance* pp. 47-49.

Boisvert, J. (2001) "Application Delivery Models for GIRO's Products: Workstations, Application Servers, and Web-Enabling with Windows Terminal Server," GIRO, http://www.giro.ca/English/Publications/GIRO/PDF/App_delivery.pdf (September 1, 2003).

Brinkmann, A., K. Salzwedel, and C. Scheideler. (2000) Efficient, Distributed Data Placement Strategies for Storage Area Networks. *Proceedings of the 12th ACM Symposium on Parallel Algorithms and Architectures (SPAA), Bar Harbor, Maine, 2000*, pp. 119-128.

Brocade (2001) "Comparing Storage Area Network and Network Attached Storage," http://storage.ittoolbox.com/browse.asp?c=StoragePeerPublishing&r=http%3A%2F%2Fwww%2Ebrocade%2Ecom%2Fsan%2Fwhite%5Fpapers%2Fpdf%2FSANvsNASWPFINAL3%5F01%5F01%2Epdf (current April 1, 2004).

Byte and Switch (2004) "SMI-S Slogs Along," http://www.byteandswitch.com/document.asp?doc_id=54696 (current July 7, 2004).

Chernyshov, D. (1999) What to do if Your LAN Isn't Fast Enough. *Proceedings of IEE Eurpoean Workshop on Distributed Imaging, Savoy Place, London, 1999*, pp. 7/1-7/5.

Clark, T. (2002) *IP SANs: A Guide to iSCSI*. iFCP and FCIP, Protocols for Storage Area Networks, Boston, MA: Addison-Wesley.

Coates, J. (2003) Departments: Opinion: Big Storage: Make or Buy? in *Queue*, vol. 1.

Connor, D. (1999) "How to Cook Up a Storage Area Network," NetworkWorldFusion, http://www.nwfusion.com/archive/1999/73611_08-30-1999.html (current April 1, 2004).

Dar, S., G. Hecht, and E. Shochat. (2004) Industrial Sessions: Database Applications: dbSwitch(tm): Towards a Database Utility. *Proceedings of the 2004 ACM SIGMOD International Conference on Management of Data,* Paris, France*:* pp. 892-896.

Derrington, S. (2002a) "Storage Network Evolution," ZDNet, http://techupdate. zdnet.com/techupdate/stories/main/0,14179,2874939,00.html (current January 18, 2003).

Derrington, S. (2002b) What's the Right Storage Infrastructure Decision? A 3-Year Look into the Future of Storage Technology, in *Server Infrastructure Strategies-META Group Inc.*

Dot Hill Systems, C. "SAN Tutorial," http://www.dothill.co.jp/tutorial/topic07.htm (current January 18, 2003).

Earls, A. (2003) "NAS and SAN converging but debates remain," InfoStor, http://is.pennnet.com/Articles/Article_Display.cfm?ARTICLE_ID=170693 (current July 7, 2004).

Falk, A., T. Faber, J. Bannister, A. Chien et al. (2003) "Blueprint for the Future of High-Performance Networking: Transport Protocols for High Performance," *Communications of the ACM* (46) 11, pp. 42-49.

Gibson, G. and R. van Meter (2000) "Network Attached Storage Architecture," *Communications of the ACM* (43) 11, pp. 37-45.

Goodwin, P. (2002) "Bluefin Storage - a Wide Net, or Full of Holes?" ZDNet, http://techupdate. zdnet.com/techupdate/stories/main/0,14179,2882017,00.html, 2002 (current September 2, 2003).

Goodwin, P. (2003) "Storage: Cheap Hardware, Pricey Software," ZDNET, http://techupdate. zdnet.com/techupdate/stories/main/0,14179,2907886,00.html (current January 18, 2003).

Goodwins, R. (2002) "The FAQs on SANs," ZDNet, http://techupdate.zdnet.com/techupdate/stories/main/0,14179,2901466,00.html (current, September 2, 2003).

Guha, A. (1999) The Evolution to Network Storage Architectures for Multimedia Applications. *IEEE International Conference on Multimedia Computing and Systems, Florence, Italy, 1999*, pp. 68-73 1.

Heath, J. R. and P. J. Yakutis (2000) "High Speed Storage Area Networks Using a Fibre Channel Arbitrated Loop Interconnect," *IEEE Network* (14) 2, pp. 51-56.

Horwitt, E. (2003) "Hospital consolidates far-flung multi-vendor environment," Storage Networking World, http://www.snwonline.com/case_studies/ hackensack_medical_center_02-24-03.asp?article_id=206 (current September 2, 2003).

IBM (n.d.) IBM Grid Computing. http://www-1.ibm.com/grid/ (Current November 18, 2004

Infoworld (2003) "Stepping Up SAN Synchronization," http://www.infoworld.com/pdf/special_report/SANmgmt_2003.pdf (current June 30, 2004).

Japan Storage Systems (2003) IDC, http://www.idc.com/getdoc.jsp?containerId=IDC_P4924 (current September 22, 2004).

Kerns, R. (2000) "Open SANs: An In-depth Brief," Evaluator Group, http://www.storage.ibm.com/ibmsan/whitepaper/open_san.html (current September 2, 2003).

Khurshudov, A. (2001) *The Essential Guide to Computer Data Storage*. Upper Saddle River, NJ: Prentice Hall.

Lei, K. and P. T. Rawles. (2003) Technology to Enable Learning: Strategic Decisions on technology Selections for Facilitating a network/systems laboratory using real options & total cost of ownership theories. *Proceeding of the 4th Conference on Information Technology Education, Lafayette, Indiana, 2003*, pp. 76-92.

MarketResearch.com (2003) "Storage Area Network (SAN) - Revolutionary Data storage Technology," http://www.marketresearch.com/product/display.asp?productid=891988&SID=87623368-293853472-347791336 (current July 7, 2004).

McConnell, D. (2001) "IP Storage: A Technology Overview," Dell Power Solutions, (September 2, 2003)

Mearian, L. (2004) "SANs Come Up to Speed Outlook: Cost and backward-compatibility issues could delay acceptance of 10Gbit Fibre Channel in favor of two interim standards," Computerworld,http://www.computerworld.com/hardwaretopics/storage/story/0,10801,91470,00.html?nas=ST-91470 (current April 1, 2004).

Merhar, M. "IP and the Convergence of NAS and SAN," http://www.wwpi.com/sms/IPConvergence.htm (current January 18, 2003).

Meta, G. (2003) "Storage Constitutes 12%-15% of Total IT Budget," http://domino.metagroup.com/pressHome.nsf/(webPressRelease)/48CCE303238E56D085256CD2006DD209 (current September 2, 2003).

Milanovic, S. and Z. Petrovic. (2001) Building the Enterprise-wide Storage Area Network. *Proceedings of the IEEE Region 8 EUROCON 2001, Bratislava, Slovakia, 2001*, pp. 136-139 1.

Molero, X., F. Silla, V. Santonja, and J. Duato. (2001) On the Interconnection Topology for Storage Area Networks. *Proceedings of the 15th International Parallel & Distributed Processing Symposium (IPDPS-01), San Francisco, CA, 2001*, pp. 162-163.

Poelker, C. and A. Nikitin (2003) *Storage Area Networks for Dummies*. New York, NY: Wiley Publishing.

Pratt, M. K. (2002) "SAN Costs," Computerworld, http://www.computerworld.com/managementtopics/roi/story/0,10801,72289,00.html (current September 2, 2003).

Preston, W. C. (2002) *Using SANs and NAS*. Sebastopol, CA: O'Reilly & Associates.

Redding, A. (2004) "Newspaper's SANs Copy Data Between Sites," Computerworld, http://www.computerworld.com/hardwaretopics/storage/story/0,10801,91530,00.html?Nas=ST-91530 (current April 1, 2004).

Reich, R. (2002) "Bluefin: The First SAN Management Industry Standard," WWPI, http://www.wwpi.com/Archive/show_article.asp?ArticleID=703 (current September 2, 2003).

Reich, R. (2003) "Storage Management Initiative in 2003," Veritas, http://www.snia.org/news/newsletter/2003/01/smi/ (current September 2, 2003).

Renato, J. R. (2003) Promises and Reality: Server I/O Networks Past, Present, and Future. *Proceedings of the ACM SIGCOMM workshop on Network-I/O convergence, Karlsruhe, Germany, 2003*, pp. 163-178.

Riedel, E. (2003) "Features: Storage Systems: Not Just a Bunch of Disks Anymore," *Queue* (1) 4, pp. 33-41.

Robinson, S. (2002) "Exploring the Benefits of Storage Consolidation," Datalink, http://www.appiq.com/webinar (January 18, 2003).

Satchell, S. (2003) "Minding the store," *netWorker* (7) 1, pp. 27-32.

Sayer, P. (2003) "IBM to Build Storage Tank for CERN Supercollider," IDG News Service, http://www.nwfusion.com/news/2003/0402ibmtobu.html (current September 2, 2003).

Scheier, R. L. (2003) "SMI-S: What's in it for Storage Managers?" SearchStorage.com, http://searchstorage.techtarget.com/originalContent/ 0,289142,sid5_gci922603,00.html (current July 7, 2004).

Staff, C. (2004) "Linux-based SAN helps Greenpeace UK Keep E-Mail Flowing," Computerworld, http://www.computerworld.com/hardwaretopics/storage/story/0,10801,91778,00.html?nas=ST-91778 (current April 1, 2004).

Storage Networking Industry Association "Bluefin Backgrounder," http://www.snia.org/tech_activities/SMI/bluefin/bluefinbackgrounderv3.doc (current September 27, 2004).

SystemsWorld (2004) "External Disk Storage Systems Posts Third Consecutive Quarter of Year-on-Year Growth," http://www.systems-world.de/id/ 6556/CMEntries_ID/50804 (current July 7, 2004).

Thompson et. al. (2003) Considerations and Performance Evaluations of Shared Storage Area Networks. *Proceedings of the 20th IEEE/11th NASA Goddard Conference of Mass Storage Systems and Technologies (MSST), San Diego, CA, 2003*, pp. 135-144.

Trellisoft (2001) "The Storage Resource Management Challenge: Taming Explosive Data Growth in a 24x7 World," http://www.whatstorage.co.uk/pdf/whitepapers/Trellisoft%20Storage%20Challenge.pdf (current January 18, 2003).

Vacca, J. (2002) *The Essential Guide to Storage Area Networks*. Upper Saddle River, NJ: Prentice Hall.

Veritas (1999) "West Chester University Leverages SAN Scalability to Grow LAN Free Backup Solution Implemented in 1998," http://www.veritas.com/us/ aboutus/pressroom/1999/99-10-27-2.html (current July 24, 2004).

Voruganti, K. and P. Sarkar (2001) "An analysis of three gigabit networking protocols for storage area networks," *Performance, Computing, and Communications: IEEE International Conference* pp. 259-265.

Webster, J. (2004) "Everything You Need to Know About SMI-S," InfoStor, http://www.snia.org/apps/group_public/download.php/6264/InfoStor-EverythingyouneedtoknowaboutSMI-S.pdf (current July 7, 2004).

Wikipedia (n.d.) http://en.wikipedia.org

Wu, D. P., Y. W. T. Hou, and W. W. Zhu (2001) "Streaming Video Over the Internet: Approaches and Tirections," *IEEE Transactions on Circuits and Systems for Video Technology* (11) 3, pp. 282-300.

Yoshida, H. (2001) "SWAN: Storage Wide Area Network - An Overview of Today's Capabilities and Future Directions," Hitachi Digital Systems, http://www.hds.com/pdf/wp98_swan.pdf (current January 18, 2003).

Yoshida, H. and M. Dolcini (2000) "The Hitachi Freedom Storage Lightning 9900 for Optimized Storage Area Networks," Hitachi Data Systems, http://www.hds.co.uk/pdf/9900 optSANS.pdf (current January 18, 2003).

Zamer, A. (2001) "Clearing the Confusion: A Primer on Internet Protocol Storage," Storage Networking Industry Association, http://www.snia.org/education/tutorials/ip_storage.pdf (current January 18, 2003).

BIBLIOGRAPHY

Andrews, P., T. Sherwin, and B. Banister. (2003) A Centralized Data Access Model for Grid Computing. *Proceedings of the 20th IEEE/11th NASA Goddard Conference, San Diego, CA, 2003*, pp. 280-289.

Baldwin, A. and S. Shiu. (2002) Encryption and Key Management in a SAN. *Proceedings of the First International IEEE on Security in Storage Workshop, Greenbelt, Maryland, 2002*, pp. 35-44.

Berenbrink, P., A. Brinkmann, and C. Scheideler. (2001) SIMLAB - A simulation Environment for Storage Area Networks, Parallel and Distributed Processing. *Proceedings of the Ninth Euromicro Workshop on Parallel and Distributed Processing, Mantova, Italy, 2001*, pp. 227-234.

Burns, R. and W. Hineman. (2001) A Bit-Parallel Search Algorithm for Allocating Free Space. *Proceedings of the 9th International Symposium on Modeling, Analysis, and Simulation in Computer and Telecommunication Systems (MASCOTS), Cincinnati, Ohio, 2001*, pp. 302-310.

Chiu, S. C., W. K. Liao, and A. N. Choudhary (2004) "Processor-embedded Distributed Smart Disks for I/O-Intensive Workloads: Architectures, Performance Models and Evaluation," *Journal of Parallel and Distributed Computing* (64) 3, pp. 427-445.

DeCusatis, C. (2002) "Storage AREA NETWORK APPLICATIONS," *Optical Fibre Communication Conference and Exhibit* pp. 443-444.

Duncan, T. S. and M. S. Sichler. (2003) High bandwidth Scientific Data Management Using Storage Area Networking: Lessons Learned At The Starfire Optical Range. *Proceedings of the 20th IEEE/11th NASA Goddard Conference on Mass Storage Systems and Technologies, San Diego, CA, 2003*, pp. 105-109.

Meth, K. Z. and J. Satran (2003) Storage Area Networking - Features of the iSCSI protocol, in *Communications Magazine, IEEE*, vol. 41, pp. 72 -75.

Molero, X., F. Silla, V. Santonja, and J. Duato. (2000) Modeling and Simulation of Storage Area Networks. *Modeling, Analysis and Simulation of Computer and Telecommunication Systems: Proceedings of the 8th International Symposium, Washington, D.C., 2000*, pp. 307-314.

Molero, X., F. Silla, V. Santonja, and J. Duato. (2000) Performance Analysis of Storage Area Networks using High-Speed LAN Interconnects. *Proceedings of the IEEE Networks International Conference, Singapore, 2000*, pp. 474-478.

Molero, X. S., F. Santonja, V. Duato, J. (2001) On the Scalability of Topologies for Storage Area Networks in Building Environments. *Network Computing and Applications: IEEE International Symposium, Cambridge, Massachusetts, 2001*, pp. 332-335.

Voruganti, K. and P. Sarkar (2001) "An Analysis of Three Gigabit Networking Protocols for Storage Area Networks," *Performance, Computing, and Communications: IEEE International Conference* pp. 259-265.

Yingping, L. and D. H. C. Du (2003) "Storage Area Networking - Performance Study of iSCSI-Based Storage Subsystems," *Communications Magazine, IEEE* (41) 8, pp. 76 storage area networks," *Communications Magazine IEEE* (41) 8, pp. 92 -99.

APPENDIX I. STORAGE ARCHITECTURES

Table A1. A Comparison of Evolving Storage Solutions

Storage Networking Architecture	Components/Environment	Benefits	**Weaknesses**
Traditional Direct Attached Storage **(Figure 1)**	▪ Storage directly attached to individual servers ▪ Adding servers requires adding storage ▪ Adding storage may require adding servers	▪ Easy to use ▪ Low cost ▪ Widely available industry standard ▪ Efficient for low storage demands	▪ Low scalability ▪ Potential bottlenecks at the SCSI interconnect or at the storage device ▪ Data inaccessible when server fails
Enterprise RAID **(Figure 3)**	▪ Enterprise RAID system ▪ Point-to-point connectivity ▪ High-speed Fibre channel interface	▪ Reduce cost through centralized management ▪ Increase data availability ▪ Increase disk utilization ▪ Simplify Scaling ▪ Provide foundation to add SAN infrastructure later	▪ Unnecessary data duplication ▪ Does not address server-to-storage bottleneck ▪ Adding servers increases data duplicity problems and costs involved
Network Accessed Storage (NAS) **(Figure 4)**	▪ NAS Unit ▪ NT and/or UNIX ▪ Typical Applications: email servers, search engines, web hosting, libraries	▪ Connect any user connected to any server ▪ Leverage existing network infrastructure & IT knowledge base ▪ Allow users to directly access storage without accessing servers ▪ Provide software for system's snapshot[27] and data replication	▪ Lack of a high-speed dedicated connection between CPU and storage units ▪ Overhead in converting file-level requests into block-level requests ▪ Low scalability ▪ Good for only small segments of data transfer
Scalable FC Storage Area Networks (SAN) **(Figure 5)**	▪ FC Host Bus Adapters[28] (HBA) in each server ▪ FC switches ▪ RAID system(s) ▪ SAN management software	▪ Support heterogeneous servers and storage devices ▪ Reduces cost through centralized management reduces costs ▪ Increase data availability and data	▪ Requires special FC- expertise ▪ Still addressing security concerns ▪ Still addressing interoperability concerns ▪ Cannot be used long distance, unless

[27] Snapshot copy keeps a log of a storage disk's current state in terms of tracks and sectors and volume used so far, in preparation of upcoming backups [Vacca 2002].
[28] Host Bus Adapter is an interface between a server and the Fibre Channel network [Vacca, 2002].

	▪ Separation of servers and storage ▪ Any-to-any connectivity ▪ High speed Fibre channel interface ▪ Separate storage network ▪ Typical Applications: Performance critical client/server applications, databases and transaction-processing systems, graphics and real-time video	utilization ▪ Good for both large and small segments of data transferred ▪ Use block-level protocols unlike IP-Based SANs ▪ Provide very high connectivity – dedicated bandwidth ▪ Scale easily ▪ Allow for addition of servers based on application needs (not on storage needs) ▪ Allow LAN free backup	coupled with IP-based devices and interconnects
IP SAN (Figure 6)	▪ Ethernet Switch ▪ Storage Router ▪ RAID system ▪ Any-to-any connectivity ▪ Low to medium data speed ▪ Separate or shared storage network	▪ Leverage existing IT expertise ▪ Build of Ethernet technologies ▪ Offer more mature security features within the TCP/IP layer ▪ Extend the benefits of storage networking to mid-range servers ▪ Reduce cost through centralized management reduces costs ▪ Support interoperability ▪ Allow for addition of servers based on application needs (not on storage needs)	▪ Low to medium data speed ▪ Still evolving

Based On Robinson [2002] and Storage Network Industry Association[n.d.]

LIST OF ACRONYMS

CIFS – Common Internet File System	NFS - Network File System
CIM – Common Information Model	RAID – Redundant Array of Independent Drives
FC-AL – Fibre Channel – Arbitrated Loop	SAN – Storage Area Networks
HTTP – Hyper Text Transfer Protocol	SMIS – Storage Management Initiative Specifications
iSCSI – Internet SCSI	SCSI – Small Computer System Interface
iFCP – Internet Fibre Channel Protocol	SNIA – Storage Networking Industry Association
IDC – International Data Corporation	SSP – Storage Service Provider
IDE – Integrated Development Environment	SWAN – Storage Wide Area Networks
IP – Internet Protocol	WBEM – Web-Based Enterprise Management
iSCSI – Internet SCSI	
LAN – Local Area Network	
NAS – Network Attached Storage	

ABOUT THE AUTHORS

Manjari Mehta, **who** will join the University of Melbourne in July 2005 isa doctoral candidate in MIS in the Decisions and Information Sciences department, C. T. Bauer College of Business, University of Houston. She is currently working on her dissertation on “IT Integration Decisions during Mergers and Acquisitions.” Her other research interests include strategic alignment, power relations within information systems, IT evaluation, and knowledge management.

Blake Ives holds the C. T. Bauer Chair in Business Leadership at the University of Houston where he is also the Director of Information Systems Research Center. He is the founder of *MISQ Discovery* and ISWorld Net and is a past President and Fellow of the Association of Information Systems (AIS). A two-time co-Chair of the International Conference on Information Systems, Blake is also past Editor-in-Chief of *MIS Quarterly* and currently chairs the *MIS Quarterly*’s Policy Committee. A frequent contributor to *Communications of the AIS*, Blake’s research is also published in *Sloan Management Review, Management Information Systems Quarterly, Management Science, Decision Sciences, Information Systems Research, and Communications of the ACM*, among other outlets. His research interests lie in virtual organizations, IT driven sustainable advantage, customer service and electronic scholarship. He is also the author of numerous case studies, three of which have been awarded prizes by the Society for Information Management International.

Communications of the Association for Information Systems

ISSN: 1529-3181