FOR PUBLIC RELEASE

Part 1

Overview of SANs Technology

1 SANs Fundamentals

Network and server downtime is costing companies hundreds of millions of dollars in business and productivity losses. At the same time, the amount of information to be managed and stored is increasing dramatically every year.

A new concept called the Storage Area Network (SAN) could offer an answer to the increasing amount of data that needs to be stored in an enterprise network environment. By implementing a SAN, users can offload storage traffic from daily network operations while establishing a direct connection between storage elements and servers.

Basically, a SAN is a specialized network that enables fast, reliable access among servers and external or independent storage resources. In a SAN, a storage device is not the exclusive property of any one server. Rather, storage devices are shared among all networked servers as peer resources. Just as a Local Area Network (LAN) can be used to connect clients to servers, a SAN can be used to connect servers to storage, servers to each other, and storage to storage.

A SAN does not need to be a physically separate network, either. It can be a dedicated subnetwork, carrying only the business-critical I/O traffic between servers and storage devices. A SAN, for example, would not carry general-purpose traffic such as email or other end-user applications. This type of net avoids the unacceptable trade-offs inherent in a single network for all applications, such as dedicating storage devices for each server and burdening a LAN with storage and archival activity.

Furthermore, as distributed networks are re-engineered to achieve continuous operations and to host mission critical applications, a common data-center technology is being applied to them. Data centers use a network storage interface called Enterprise System Connection (ESCON) to connect mainframes to multiple storage systems and distributed networks. This type of network is also called a SAN. In other words, SANs are employed by mainframe data centers and account for approximately 58% of all network traffic. What is new is that SAN architectures are now being adopted in distributed networks out of low cost SAN technologies such as Small Computer System Interface (SCSI), Serial Storage Architecture (SSA), and Fibre Channel.

BUT WHAT IS A SAN—REALLY?

As previously mentioned, a SAN is a high speed network, similar to a LAN, that establishes a direct connection between storage elements and servers or clients. The SAN is an extended storage bus which can be interconnected using similar interconnect technologies as LANs or Wide Area Networks (WANs): routers, hubs, switches, and gateways. A SAN can be local or remote; shared or dedicated; and includes unique externalized and central storage and SAN interconnect components. SAN interfaces are generally ESCON, SCSI, SSA, High-Performance Parallel Interface (HIPPI), or Fibre Channel, rather than Ethernet. It doesn't matter whether a SAN is called a Storage Area Network or System Area Network, the architecture is the same in either case.

SANs create a method of attaching storage that is revolutionizing the network because of the improvements in availability and performance. SANs are currently used to connect shared storage arrays; clustered servers for failover; interconnect mainframe disk or tape resources to distributed network servers and clients; and to create parallel or alternate data paths for high performance computing environments. In essence, a SAN is nothing more than another network, like a subnet, but constructed from storage interfaces.

SANs enable storage to be externalized from the server and in doing so, allow storage to be shared among multiple *host* servers without impacting system performance or the primary network. The benefits are well proven as this architecture emerges from mainframe Direct Access Storage Device (DASD). It is nothing new. In fact, the DEC VMS network environment is based on SAN architectures and clustered servers. For example, EMC already has a large installed base of SAN attached disk arrays (for example, EMC's disk array product: Symmetrix) and has achieved such a high level of customer confidence that they are the standard of comparison [1]. So, what's new? This important technology is moving into the mainstream in distributed networking and is now the normal, adopted way of attaching and sharing storage.

Often referred to as the *network behind the server*, SANs represents a new model that has evolved with the advent of shared, multi-host connected enterprise storage. A SAN bypasses traditional network bottlenecks and supports direct, high-speed data transfer in three different ways:

- Server-to-storage
- Server-to-server
- Storage-to-storage [2]

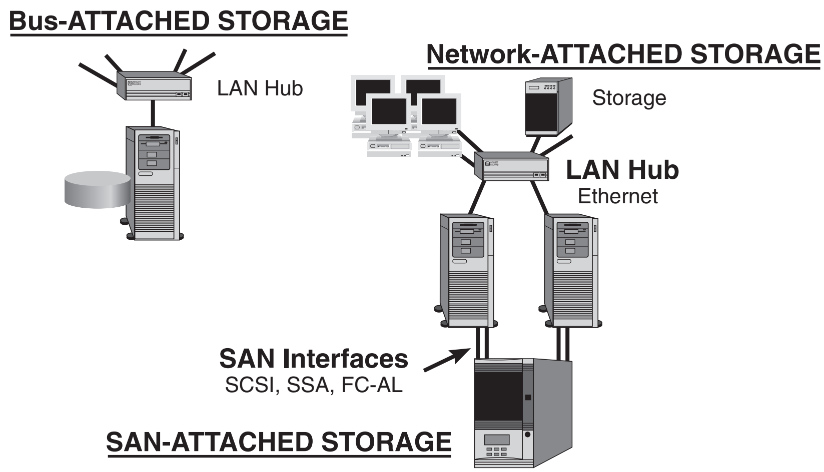

Figure 1.1
Storage attachments.

SAN architecture and terminology is getting confused as each product camp goes about praising the merits of their solutions. The following discussion is aimed at providing a simple set of definitions and terms that should be adopted by the industry. To begin, storage can be attached to the network in one of three ways. According to Strategic Research Corporation [2], ninety-nine percent of today's server storage connections are *bus-attached* via a form of SCSI or Integrated Development Environment (IDE) as shown in Figure 1.1 [2]. Bus attached storage operates through the server. Availability and performance are limited to the server's capabilities and loading. Storage is externalized from the server via Network Attached Storage (NAS) or SAN Attached Storage (SAS). NAS and SAS are very similar from an engineering standpoint, but it is essential to differentiate them to help the customer understand the differences in implementations.

Network Attached Storage (NAS)

NAS is a disk array that connects directly to the messaging network via a LAN interface such as Ethernet using common communications protocols (see sidebar, "Storage Sorting"). It functions as a server in a client/server relationship, has a processor, an OS or microkernel, and processes file I/O protocols such as Server Message Block (SMB) and Network File System (NFS).

Storage Sorting

Storage area networks are fast becoming part of the IT lexicon, but the abundance of other storage management acronyms is making things a bit confusing. As previously explained, a SAN is a collection of networked storage devices that can automatically communicate with each other.

Note: *A SAN doesn't have to use Fibre Channel as its underpinnings. For example, the mainframe environment's Enterprise Systems Connection channels could form the SAN interface.*

The key to understanding what makes a SAN is understanding that the goal is to divorce all users and network administrators from storage management. Storage, retrieval, and file transfers are automatically managed in a true SAN.

OK, so what is network attached storage? SANs may include NAS-enabled devices, but they aren't the same thing.

A NAS system is connected to application servers via the network. But unlike a SAN, with NAS users can directly access stored data without server intervention. A SAN will automate management of storage systems, while NAS devices don't have this capability.

And when it comes to automated management in a SAN, there's yet another definition floating around: hierarchical storage management (HSM). HSM is simply managing data movement from online to offline storage, such as to tape devices.

SAN Attached Storage (SAS)

A shared storage repository attached to multiple host servers via a storage interface such as SCSI, Fibre Channel-Arbitrated Loop (FC-AL), or ESCON. The SAN is an extended, shared storage bus which can be interconnected using similar interconnect technologies as LANs or WANs, routers, switches, and gateways.

Note

FC-AL is not the only fibre channel interconnect used in "SAN." More and more fibre channel users use switched fibre channel as an interconnect. Fibre channel can be used by itself, and *arbitrated loop* can be removed.

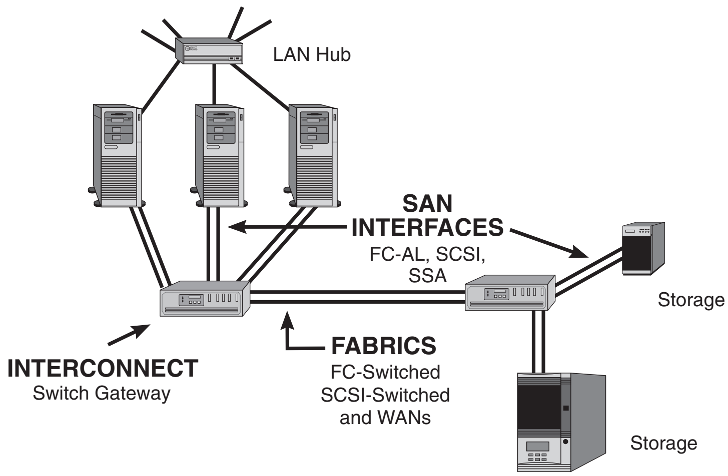

Figure 1.2
SAN network terminology.

The next key terminology point deals directly with SAN architectures. The three SAN components are the SAN interfaces, the SAN interconnects, and the SAN fabric as shown in Figure 1.2 [2]. These are often mixed together, but really are distinct elements of the SAN. Think of the relationship between these three fitting together in a chained sequence, server-to-interface-to-interconnect-to-fabric-to-interconnect-to-interface-to-storage array.

SAN Interfaces

SCSI, FC-AL, SSA, ESCON, bus-and-tag, and HIPPI are common SAN interfaces. All allow storage to be externalized from the server and can host shared storage configurations for clustering. Multiple channels can be installed or loops built to provide increased performance and redundancy. It is incorrect to say that SCSI cannot be extended, multiplexed, switched, and connected via gateways to WANs like serial interfaces.

SAN Interconnects

Extenders, Multiplexors (Mux), Hubs, Routers, Gateways, Switches, and Directors are the SAN interconnects. Sounds just like a LAN or WAN, and it is. SAN interconnects tie storage interfaces together into many network configurations and across large distances. Interconnects also link SAN interfaces to SAN Fabrics. One common misconception is that FC-AL, a SAN interface, is a SAN Fabric, Fibre Channel Switched (FCS). It is not (see sidebar, “SAN Myths”).

SAN MYTHS

Anyone considering a storage area network quickly encounters a number of myths. Like most technology myths, SAN myths contain a grain of truth, but the reality is often quite different. The following are the top five common SAN myths:

The Fibre Channel Myth

When first conceived, SAN technology was specified on Fibre Channel as the preferred communications link. Fibre Channel was able to provide the speed and the distance SAN required. Today, the majority of SANs are being implemented with Fibre Channel, either arbitrated loop or switched topologies. However, the SAN is not locked into Fibre Channel. Rapid developments are occurring with SCSI over Internet Protocol (IP) for use in SANs. The Fibre Channel SAN will continue to have a place in the enterprise data center, but protocols using SCSI commands are emerging, and vendors will introduce SAN products using the SCSI protocol over the next few quarters.

The Interoperability Myth

Early SANs indeed suffered from a lack of interoperability among components from different vendors. However, interoperability is improving, especially within the switch, hub, and host bus adapter market. Through a number of interoperability events, dubbed Plug Fests, competing vendors come together to iron out many of the interoperability issues. Observers expect any remaining interoperability issues to completely fade away within 12 months.

The Skills Barrier Myth

Certainly SANs introduce new technologies into the enterprise storage world, particularly fibre and networking, which require new skills. Information technology (IT) storage experts groomed for directly attached SCSI storage now must learn new protocols and new configurations. Switched fibre SANs, in particular, require advanced networking skills. There is no getting around it; SANs require an understanding of networking. However, new tools, new products, and new service offerings from the vendors are making SANs easier. And, service providers increasingly have a solid stable of trained people who can help any organization implement a SAN.

The Management Myth

Early SAN adopters complained that SANs were hard to manage and administer. And they were, due mainly to a lack of tools. Today, however, SAN administrators are finding a growing selection of tools to manage the SAN, perform backup, create virtual storage pools, monitor resources, manage the topology, and more. Storage vendors are responding with tools to manage the various SAN components, and more and better tools are in the pipeline.

The Cost Myth

SANs entail a large, initial capital outlay, but the long-term benefits are significant. While it is cheaper initially to attach low-cost disk storage to a server, the cost of administering storage attached to multiple servers and the inefficiency that results from underutilized pools of storage, shift to the overall total cost of ownership advantage clearly to the SAN. Recent studies suggest that half of all server-attached storage goes unused because it can't be shared. With a SAN, storage utilization increases to 70% and, ultimately, can hit 90%. And with a SAN, each administrator can manage far more storage.

The trouble with technology myths is that technology keeps changing. Even if a SAN myth was true once, it probably isn't today.

SAN Fabrics

Switched SCSI, FCS, and Switched SSA form the most common SAN fabrics. With gateways, SANs can be extended across WAN networks as well. Switches allow many advantages in building centralized, centrally managed, consolidated storage repositories shared across a number of applications.

Building A SAN

Building a SAN requires network technologies with high scalability, performance, and reliability in order to marry the robustness and speed of a traditional storage environment with the connectivity of a network. As the SAN concept has developed, it has grown beyond identification with any one technology. In fact, just as LANs use a diverse mix of technologies, so can SANs. This mix can include Fiber Distributed Data Interface (FDDI), Asynchronous Transfer Mode (ATM), and IBM's Serial Storage Architecture, as well as Fibre Channel. SAN architectures also allow for the use of a number of underlying protocols, including Transmission Control Protocol/Internet Protocol (TCP/IP) and variants of SCSI.

A SAN allows different kinds of storage (mainframe disk, tape, and Redundant Array of Inexpensive Disk [RAID]) to be shared by different kinds of servers, such as Windows NT, UNIX, and OS/390. With this shared capacity, organizations can acquire, deploy, and use storage devices more cost-effectively. SANs let users with heterogeneous storage platforms utilize all of its storage resources. This means that within a SAN, users can backup or archive data from different servers to the same storage system; allow stored information to be accessed by all servers; create and store a mirror image of data as it is created; and share data between different environments.

By externalizing storage and taking storage traffic off the operations network, companies gain a high-performance storage network, shared yet dedicated networks for the SAN and LAN, and improved network management. These features reduce network downtime and productivity losses while extending current storage resources.

In effect, the SAN does in a network environment what traditionally has been done in a back-end I/O environment between a server and its own private storage subsystem. The result is high speed, high fault tolerance, and high reliability.

With a SAN, there is no need for a physically separate network because the SAN can function as a virtual subnet operating on a shared network infrastructure, provided that different priorities or classes of service are established. Fibre Channel and ATM allow for these different classes of service. Early implementations of SANs have been local or campus-based.

But as new WAN technologies such as ATM mature, and especially as class-of-service capabilities improve, the SAN can be extended over a much wider area. Despite the hype about the coming of unlimited bandwidth, WAN services remain costly today. However, as WAN technologies improve their quality of service, they will provide (even over public WANs) the robustness needed for each application, including networked I/O.

SAN Tools

In addition to reliability and performance, SANs promise easier and less costly network administration. Today, administrative functions are labor-intensive and IT organizations typically have to replicate management tools across multiple server environments. With a SAN, there is just one set of tools, and replication costs can be avoided. The traditional software functions of security management, access control, data management, and storage management will be mapped into the SAN architecture and performed differently than they have been in the past. For example, different security strategies have to be pursued when storage devices are more widely available.

Specialized I/O protocols such as Network Data Management Protocol (NDMP) are emerging, and the software functions will evolve much as LAN functionality has progressed in recent years.

Why Are SANs Important?

SANs will enable almost any application that moves data around the network to perform better. Just like conventional *subnets*, SANs add bandwidth for specific functions without placing a load on the primary network. In this fashion, SANs compliment LANs and WANs. SANs also enable higher performance solutions such as data warehousing. In fact, as Figure 1.2 shows, SANs are really pervasive and applicable to many networking environments [2].

SAN technology enables the network architecture of shared multihost storage, connecting all storage devices as well as interconnecting remote sites. This will soon be the standard configuration for centralized networks running mission-critical applications. Both disk and tape operations are centralized, attached via the SAN, and more resilient, as well as operating faster. As the IT community has learned in the database market, the key to application performance is usually the I/O network, not the disk drives themselves. SAN architecture holds the keys to the future.

The benefits of a SAN network architecture are huge and will cause many sites to adopt this methodology of attaching storage and transferring data. This list is indicative of the types of benefits seen in sites operating with SANs (see sidebar, "SAN Benefits").

SAN Benefits

Higher Application Availability

Storage is externalized, independent of the application, and accessible through alternate data paths such as found in clustered systems.

Higher Application Performance

Server and bus overhead degrades performance. Independent SAS arrays will outperform bus-attached arrays, as well as be compatible with performance clusters.

Easier Centralized Management

SAS configurations encourage centralization and the ensuing large management benefits.

Centralized and Consolidated Storage

Storage centralization and consolidation result in higher performance, lower cost of management, more scalability, flexibility, reliability, availability, and serviceability.

Practical Data Transfer, Vaulting, and Exchange with Remote Sites

Cost effective implementations provide high availability disaster protection (remote clusters and remote mirrored arrays)[2].

SAN Applications

Now look at SANs from an application viewpoint. At a high level, for example, Strategic Research Corporation has identified six application areas currently utilizing SAN architectures for data transfer as shown in Figure 1.3 [2]. This is not to mean there won't be more in the future. The purpose of Figure 1.3 is to explain how pervasive the technology is already [2].

As previously discussed, in the changing network architecture, externalized storage is a generic application, fitting a myriad of network-hosted applications with many benefits. Next, is clustering. Clustering is usually thought of as a server process

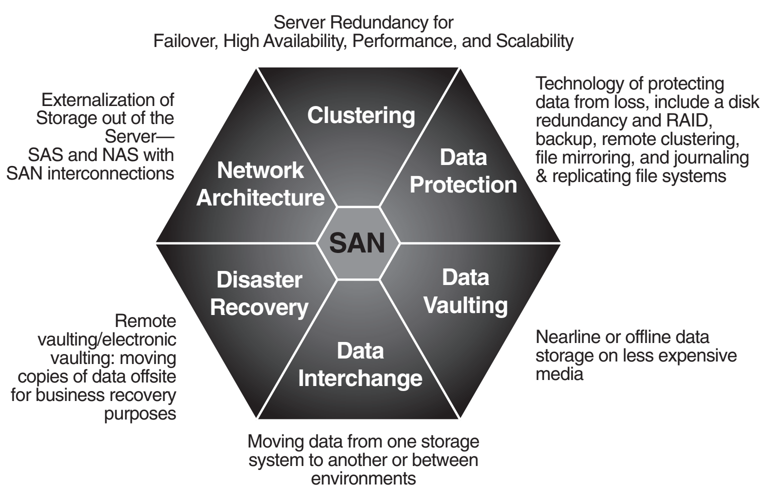

Figure 1.3
Applications utilizing SANs.

providing failover to a redundant server or scalable processing through using multiple servers in parallel. In a cluster, the SAN provides the data pipe, allowing storage to be shared. For example, Microsoft's ClusterServer, which is an availability cluster, shares a single array between two servers attached via a SCSI SAN. Next, data protection architectures operate through creating redundancy of storage on a dynamic basis. SANs provide the best interconnects, allowing storage mirroring, remote clustered storage, and other high availability (HA) data protection solutions because of the performance and independence as a secondary data path. SANs do not impact the primary network or the servers and they provide redundancy. Data vaulting is the process of transferring data, usually for the purpose of archive or logging, to a remote site. SANs make a very efficient transmission medium. Interchange and disaster recovery operations are very similar and use SANs the same way, whether local or remote, just for different purposes. SANs provide a very efficient pipe for moving data offsite or between sites. Disaster protection systems can be built on remote vaulting (backup) processes or with high availability remote array mirroring or clustering.

TYPES OF SAN OPERATING SYSTEMS SOFTWARE AND HARDWARE COMPONENTS

Recently, for example, EMC Corporation announced that their SAN software that's designed to manage storage devices made by some of its rivals, is a mixed-vendor support that is being done in reaction to pressure from systems administrators [1]. Users have been pushing for the ability to manage EMC's Symmetrix and Clariion disk arrays along with other storage products and connectivity devices. To start with, disk storage systems made by Compaq Computer, Hewlett-Packard, and Hitachi; and tape devices from Storage Technology have been qualified to work with the new software.

Network switches and other connectivity devices made by companies such as San Jose-based Brocade Communications Systems, and Broomfield, Colorado-based McData Corporation, can also be controlled. An addition to EMC's Enterprise Storage Network (ESN) product line, the new ESN Manager tool provides a single point of control for administrators to use in managing multiple *zones* of interconnected storage devices.

The development of ESN Manager should scare the hell out of other vendors. EMC is already clearly the storage king. All of a sudden, if EMC actually lives up to what they say and becomes an open systems management provider, they are really a lethal weapon.

EMC hopes to prompt users that don't have its devices now to migrate from single-vendor storage setups to mixed SANs that include Symmetrix and Clariion arrays. People claim EMC is proprietary, but they really are not.

EMC's announcement follows one made recently in which Sun Microsystems said it was teaming up with former rival Brocade. Sun said it would start selling Brocade's Silkworm switches, which act as data traffic directors in a SAN. Brocade in turn announced that it would begin using Sun's Jiro SAN management software with its devices.

EMC's increased spending on software research and development is a sign that the company is *genuinely interested* in doing more than selling disk arrays. EMC executives want to be more accessible to users who haven't fully committed to Symmetrix and Clariion.

Management tasks supported by ESN Manager include setting limits on which end users can access different devices on a SAN, and configuring logical pathways between various servers and storage subsystems. The software's base price is $24,000 per Symmetrix box, making it a relatively high-end offering in the storage management market.

Sharing Data in a SAN

EMC, for example, has also rolled out software which it claims will unite the competing worlds of NAS and SANs. EMC's HighRoad integrates new software processes into its Celerra File Server and other servers to improve file-sharing capabilities.

The software is aimed at web hosting, image processing and simulation, and modeling applications. It uses separate mechanisms for control actions and data delivery. HighRoad brings together SANs and NAS by routing files and other data directly from the SAN to the user without the intervention of a NAS server.

Customers need all their information to work together. It's about creating one unified infrastructure that builds on the individual and combined strengths of their data and storage networks.

Recently, EMC bought NAS software developer CrosStor, which was working on combining the two storage technologies. EMC also unveiled its latest Clariion IP4700 product, which is aimed at the low-end market dominated by rival Network Appliance. EMC claimed its product will be more reliable than Network Appliance's NAS clusters and cost half the price.

The storage industry has been very busy lately. Coinciding with EMC's announcement, IBM recently pledged to offer a universal storage system that works with all software and hardware systems. Sun Microsystems also announced plans to boost its storage unit by acquiring data storage management software maker High Ground Systems for $500 million.

Tools to Unlock SAN Promise

SAN software tools could bring IT managers closer to the data sharing they thought they were getting when they first bought SANs. People thought when they got a SAN, they'd be getting the ability to share data. And it's true, but they need management tools for that sharing to happen.

SAN resource management starts with knowing what components you have—a process that should be automated. BMC Software in Houston, Texas; Computer Associates International in Islandia, New York; and IBM's Tivoli Systems in Austin, Texas, all claim that their new tools offer varying degrees of *autodiscovery* capabilities.

Autodiscovery in theory does the same thing that Windows 98 does on your machine at home. Plug in some new equipment, and your system automatically sees it and understands how to manage it.

But, most people who have SANs are still working to put together and monitor all the pieces of the infrastructure. The promise of SANs runs far ahead of what users should expect today. It's rare to see anyone start from a SAN and go straight to application-level management. You have to crawl before you walk, and there's still some crawling to be done to get to interoperability and management at the component level. Underlying the success of the SAN concept is the key assumption that standards will be developed.

In an International Data Corporation (IDC) survey in 2000, more than half of IT managers responding indicated that they were considering purchasing a SAN. But 80% indicated that open standards are critical for any SAN implementation they would consider. See Chapter 3, "Standards," for more information.

We're still a long way off from having universal standards in SANs. However, SANs are following a path similar to what happened in the LAN environment.

The Case for SAN Hardware

Storage hardware is about as exciting to most IT and business managers as watching pet rocks sunbathe. But it's rapidly becoming the single most important element of e-business innovation. Just look at your company's e-business infrastructure.

There's only one proprietary component: the enterprise customer data. Never put that at risk in terms of availability, security, ability to scale customer relationship management (CRM) systems, speed of access, backup and archiving, or server consolidation.

Almost every other platform component is now a commodity; a company can substitute one excellent vendor's products for another's—low-end and mid-range servers, PCs, and Internet hosting services, for example. Or a company has some wiggle room: It can call in systems integrators and C++/Java wizards, or build front-end links to legacy systems. This is by no means easy, but none of these areas is the *giant bottleneck* that storage is now.

Here's the problem for IT: For decades, storage has been handled as just an add-on to IT strategy and as JBOD—storage professionals' acronym for *just a bunch of disks*. Whoever handles JBOD purchases says, "Your data warehouse is exploding again? Buy two clusters and call back next month."

Even the NAS versus SAN debates about how to best manage networked storage are typically handled in nonbusiness terms, centering on such concerns as response times and operating costs. In IT, there's often a wide gap in thought and knowledge between network and storage professionals.

Try asking your best telecommunications experts about Fibre Channel or backup and archiving. Then talk to the storage people about IP-based SANs. In most instances, you'll see blank stares. Look at the network architecture plans. See if you can find the storage architecture plans. Good luck. Then look at your company's many CRM activities and see if there's any discussion of their implications for storage beyond JBOD and *aspirin*. Again, good luck.

IT needs to raise the strategic discussion of storage in the same way, and to the same degree, that telecommunications moved in the 1990s from cables and boxes to e-business architecture; and in the same way that databases have moved from software to CRM. Storage vendors and buyers need to build an entirely new dialogue.

In the JBOD world, vendors are box salespeople, and IT organizations are box buyers. Both are in a commodity transaction, not partners in enterprise storage strategy. The JBOD suppliers come in with feature lists, prices, and service promises. That's fine for semicommodities such as low-end servers, PCs, and Internet hosting. But, it's inappropriate when the discussion is about the architecture for the firm's customer data resources or its e-business strategy and platform architecture—and recognizing the importance of never putting either at risk.

As the storage issue rises above JBOD, IT must redefine the vendor dialogue, and vice versa. Will EMC's powerful sales force and aggressive selling be the basis for your company's dialogue? Will Hitachi Data Systems' increasing dominance in pure technology and product leadership translate into architecture leadership? Will Sun be able to turn its e-business server strengths into comparable networked storage strengths? Until a year ago, \$3 of sales in servers meant \$1 of storage sales for Sun.

Now, it's the reverse. Dell, Compaq, Network Storage Solutions and Hewlett-Packard (which is mostly Hitachi with a different logo) all have good boxes.

Which will be the platform partner? That would be the next IT e-business agenda. Which would you choose? Probably Hitachi, because if your firm's customer data is your proprietary business edge, you would want the best hardware. But, don't take anyone's word for it. IT professionals must have their own opinions, shaped in their companies' best interests.

SAN MARKET DEMAND AND PROJECTIONS BY REGION AND COUNTRY

Japanese businesses, buffeted by a decade of economic anemia and suffused with the conservatism that permeates IT investment strategies, are cautiously coming around to the use of SANs. But new evidence that the economy may be slipping back into recession and centralized corporate bureaucracies may dampen adoption of SANs on a large scale.

Research institutes and educational customers are early movers to SANs. Central Tokyo University is using a SAN because they have an immediate need for more storage. But currently, SAN users remain a minority in Japan, where many companies use centralized storage and backup methods such as tape drives and hard drives.

Nevertheless, according to Gartner Group's Dataquest unit, the storage market in Japan is the fastest growing market in the Asia-Pacific region, where it is projected to grow at nearly 17% per year and reach $4.2 billion by 2003. That growth is being driven by the rapid expansion of communications infrastructure throughout Japan. Japan currently has nearly 61 million wireless subscribers; and, according to Dataquest, the country is Asia's biggest Internet market with total subscribers expected to hit 64.5 million by 2004.

One touted benefit of SANs is that they give users the ability to add more storage capacity without burdening the corporate data center. But some Japanese companies view SANs as the outsourcing of key business functions and remain reluctant to make that move.

Traditionally, Japanese companies like to make centralized decisions, and so moving small parts of a business to a SAN is not always feasible. This kind of change for a company is very complex.

There are some real cultural differences in the way business decisions about IT get made, and you have to respect them. And, while another factor in the slow adoption of SANs is cost, Japanese companies are being confronted by a unique twist on the need for more storage.

First there is the double-byte issue—referring to the need for added storage when using *kanji* characters. Each letter of their (Roman) alphabet takes a single byte of memory, but with Japanese (and Chinese or Korean) characters, you need twice as much storage capacity.

The second driving force for SANs are Internet service providers. With the explosion of Internet usage, service providers have to offer customers around-the-clock storage and remote backup. That wasn't something companies had to think about before. Tokyo-based ASAHI Net uses a Network Appliance. SAN provided by ITFOR to serve more than 200,000 subscribers and 190 domestic access points. But, current business and technology conditions continue to slow the growth of SANs in the region.

Business spending on capital equipment, one of the bright spots in 2000, as Japanese companies scrambled to buy technology, is sagging. The government projects a 2.1% decline in capital spending for the fiscal year, following a 9.4% jump in 2001.

EVOLUTION OF THE SAN MARKET ....

Thanks to the Internet and the rapid global expansion of computing, humans and their machines will create and store more information in the next three years than in the 300,000 years of history dating to the earliest cave paintings and beyond [1]. That was what researchers at the School of Information Management and Systems at the University of California at Berkeley forecasted late in 2000, much to the delight of EMC Corporation, the data storage giant that sponsored their work.

EMC was quick to pitch the study to Wall Street, adding it to analysts' projections that spending on data storage products is drawing even with spending on computers themselves and that it will account for 70% of information technology budgets by 2005. EMC also included its own projection that an individual (EMC likes to call him "Tommy" in its advertisements) could easily have a terabyte (the equivalent of 250 million pages of text) of stored personal records, photos, and other data by 2005.

The obvious outcome of such trends, of course, would be mind-boggling growth for data storage products and, more to EMC's point, full-scale storage systems intelligent enough to support the Internet's need for constant access to data. Data needs to live someplace. There's almost no value if it's just put away.

EMC's need to keep Wall Street awed is directly linked to its track record. It became the biggest gainer on the New York Stock Exchange in the 1990s by grabbing leadership of the market from IBM, producing both astonishing profits and sizzling growth. Then, as other technology giants stumbled in 2000, its shares gained another 10% to finish the year at $66.50. From such a pinnacle, convincing the Street the best is yet to come will be no mean feat. But, there are plenty of other companies beating the same drum, from giants like IBM, Compaq Computer, and Sun Microsystems, to fast-growing newcomers like Network Appliance, Brocade Communications, and Veritas Software.

Storage is becoming the heart and soul of all business. What you know about your customers, suppliers, and partners will differentiate you at the end of the day.

Trouble is, even if the vendors and analysts are right about the growth, investors may well have become unrealistically optimistic about how easily it will translate into profits. The publicly traded industry leaders tumbled in 2000, but are still trading at nosebleed levels that leave little room for earnings disappointments.

The impact on data storage of a general slowdown in technology spending is Wall Street's current fear, but storage companies also have to contend with tougher competition. New technology is driving down prices, just as in traditional computer markets. In addition, networking giants like Cisco Systems and discount computer specialists like Dell Computer are moving in, while rising interest from venture capital firms is spawning a steady stream of start-ups scrambling to define niches. They are fighting over a rapidly shifting landscape that International Data Corporation estimates was worth at least $70 billion in 2000, depending on which technologies and services are included.

Storage is getting to be as complex as servers and networks. One sector of the storage market focuses on data used in computations that are stored in caches on microprocessors, on memory chips, or inches away on disks inside a computer. At the other end of the technology spectrum is tape-based storage, which provides a low-cost if somewhat less convenient alternative to disk storage.

But the heart of the action these days revolves around disk-based storage systems outside the computer. The newest hardware building blocks are specialized file servers. Some, not much bigger than a VCR, allow users to add storage capacity directly to Internet networks without buying full-scale server computers. The workhorses for big enterprises, though, are refrigerator-size storage arrays of disks that support one or more mainframes or networks of smaller computers.

Steady advances in the disks and the software that manages them are producing astonishing performance gains. Remember Moore's Law, the longstanding rule that shrinking circuitry allows chip companies like Intel to double the processing power of processors every 18 months? Well, that amazing progression (from the room-size computing monsters of the 1950s to far more powerful fingernail-size chips) pales in comparison with advances in data storage. EMC, for example, says that the volumes of data it will stuff into shoebox-size devices by 2005 would have required covering an area the size of Argentina if 1950s technology were still in use.

For all that, the hottest storage battleground is not storage hardware but software, switches, and other components that meld the storage devices into intelligent networks and keep them online. Brocade Communications' market leadership in Fibre Channel, a specialized protocol for designing such storage networks, drove its shares from an initial public offering price in May 1999 of $2.38, adjusted for splits, to $133.72 in October 2000, though it has since retreated. Veritas's strength in software to manage incompatible storage products from numerous different vendors helped its shares climb from a split-adjusted initial price of 53 cents a share in 1993, to a secondary offering in August 1999 at $22.14 and a peak of $174 in March 2000.

Storage services are also booming as big data users hire consultants, rent outside capacity, or simply turn over the entire problem to technology management experts like IBM Global Services or new specialists like StorageNetworks, a two-year-old start-up based in Waltham, Massachusetts. International Data estimates that the service sector had revenues of more than $24 billion in 2000 and indicates its sales should top $40 billion in 2003.

If there is anyplace where the sometimes conflicting visions of storage's future intersect, it has to be the headquarters of EMC in the Boston suburb of Hopkinton. EMC, like IBM in the past, strives to design equipment that performs best with EMC software, so that customers become locked into it as a vendor. And, like IBM's mainframe business in the 1960s, EMC counts on its reputation for reliability and service support to make it the safe, if premium-price, choice for information managers. But, company officials say, any resemblance ends there and that no one will catch EMC off guard as EMC itself caught IBM in the storage business.

EMC's strategy assumes that information pipelines (bandwidth in the industry's jargon) will become so huge and fast that it will no longer be necessary to store data locally to ensure quick access. Such bandwidth, in EMC's estimate, will allow as much as 90% of data to be centralized in the kind of big businesses that have been EMC's prime customers. From medical files, to movies, to financial records, data consumers would download what they need when they need it, but would not necessarily store it on their own computers.

The best architecture for such data reservoirs is still up for grabs, however. Some data will reside in dedicated, maximum-security systems linked to particular computers. Some will be in cheap file servers (NAS) attached to the Internet. A lot of it is likely to end up in networks of storage devices (SANs) that would be linked to the Internet, computers, and tape storage systems through specialized servers.

How things develop depends on evolving network equipment and software as much as on the storage devices themselves. In areas where the landscape isn't as clear as EMC likes, they're placing multiple bets.

Big bets, too, judging from EMC's vow to invest $2.5 billion over the next two years in research and development, more than 75% of it in software. But what if projections like the Berkeley study prove to be wildly inflated? What if people become smarter about saving only what they really need? The industry's answer is another question: why would they bother? With storage prices headed from about 40 cents a megabyte today to less than a cent in 2005; and the industry moving toward making access to storage as easy as the universal dial tone on the telephone, it's going to take too much energy to throw things away. Besides, the Berkeley figures may well be too conservative, since they exclude any estimates for duplicate storage of information once it is created, one of the fastest-growing segments of the business.

THE VALUE OF INFORMATION ....

Finally, people have always placed a high value on information and knowledge. From the first cuneiform characters pressed into clay tablets to today's petabytes of data held on magnetic media, information has been protected and valued. Because it was valuable and hard to acquire or store, information was kept in the hands of a few experts for thousands of years. It was only with the advent of printing in the fifteenth century that an information explosion began that led to thriving new nations and burgeoning societies, such as those of North America and Europe.

As information and knowledge became more available from 1500 onwards, it became essential to learn to read. At first, only the rich and privileged had the time and the means to learn to read and access new information. However, as the need for skilled labor to build and maintain societies evolved, governments realized that an educated workforce was essential to national prosperity. An association of schools, libraries, vocational training establishments, and colleges appeared almost overnight in the newly industrialized countries, and created a knowledge and information system that catapulted its developers into the twentieth century, laying the foundation for today's astonishingly successful, connected world.

The lesson learned from these information-based developments is that sharing knowledge and data makes a society and its individuals and organizations better able to communicate and work together for mutual benefit. Just as money is much more useful when it is invested and in circulation, so is information a much more useful tool and broader resource when it is shared. The opening months of the twenty-first century show global business poised to benefit from shared electronic information in the same way that society was just beginning to see the real advantages of universal education around 1900.

The Business Information Landscape Today

The profusion of storage technology advances available now or soon to come to market is guaranteed to confuse IT professionals and leave business executives bewildered. While the objective is simple (to provide the best possible information systems), the method and technologies to be used are anything but simple.

For example, Tivoli Systems has refined the various available components into what it calls the Information Grid [3]. This combination of components consists of: SAN topology; new technologies, such as fiber channel hubs; switches and interconnects; new disk and tape technologies; and resource and data sharing techniques. The Information Grid promises extraordinary opportunities if organizations can deal with the inherent challenges.

Today's business information environment still consists largely of islands of information within an organization that have limited contact with each other and very little in the way of effective links and conduits across which information and data can flow to the people who need them. These islands may be in the form of large existing corporate systems that are based on mainframe architecture. They may be simple LANs that link the PCs or Microsoft Windows NT systems of a branch office or remote facility, or perhaps UNIX technology-based departmental systems in engineering, research, or accounting functions. Whatever form information takes, it has always been difficult, if not impossible, to share—and it is even more difficult to achieve the powerful benefits of harnessing information into a single, seamless environment.

Preparing for Storage Area Networking

By 2002, most experts expect the first complete implementations of SANs to be in place and running successfully (see Chapter 2, "Types of SAN Technology"). A large number of IT vendors are actively pursuing this goal. However, for the IT user community, deciding which company or companies to form successful relationships with will be critical.

SAN technology is part of the larger challenge of establishing a full storage networking management strategy. Careful consideration should be given to choosing a vendor or partner that is capable of implementing the full range of benefits of storage management. These benefits include:

- $24 \times 7 \times 365$ availability
- Scalability
- Data sharing across different architectures
- Storage access at all times and from all locations
- Better performance
- Cost reductions
- Improved security
- Significantly better data protection
- Verifiable and consistent data integrity
- Easy-to-use, consistent data management tools [3]

For example, Tivoli, in partnership with many of the leading storage management and SAN vendors, has developed and planned for the technology required to ensure successful, full implementation of SANs. Tivoli is developing new products that will be essential to exploit the Information Grid in all its capabilities. As highly IT-dependent companies develop increased competitive advantage through adept management of their information and data, they are turning to Information Integrity Initiatives (checking the reliability and quality of information) to provide the structure for their storage strategies. The key areas where installations must be prepared for SAN implementation and also for storage management fulfillment are:

- Application management
- Data management
- Resource management
- Network management
- Element management

Application Management

Application management is a vital component of the relationship between business systems and IT capabilities. It is also the obvious point at which strategic business goals can be furthered by the skillful implementation of enterprise applications. However, there are some considerations that affect application management in an enterprise environment. Large-scale business applications are usually very complex, involve custom and off-the-shelf software, and are linked to the three most common architectures—mainframe, open systems, and desktop environments. The picture is further complicated because all of these enterprise applications are essential to the financial health of a business.

Data Management

Data is the lifeblood of business. Just as governments and business discovered that money was an engine for economic growth only when it was available, so are businesses rapidly realizing that data assets are many times more effective and valuable when widely accessible. For example, in the Tivoli Information Grid, data management assures that data is available and accessible for applications; that data meets the specifications for application use; and that data is recoverable in case of a failure. Data management functions with all types of storage, whether remote, centralized, or removable. Data management functions at all levels, from large system servers to desktops.

Resource Management

Business executives who are intimately involved with IT are sometimes puzzled by decreases in the unit cost of processors, storage, communication, and the increased total cost of IT. Storage management, in general, and storage area networking, in particular, should provide some relief from this conundrum by managing pooled, fixed disk, and tape resources, as well as all removable media and the implementation of just-in-time storage management.

Network Management

Network management is probably the most critical part of the SAN and storage management challenge, in terms of the burgeoning e-commerce market. LAN and WAN resources already form part of every large enterprise and are familiar to all business executives. SANs are the logical extension of these proven networks and bring the same, or greater, value to the companies that deploy them. SANs will be as critical to business success in the future as LANs and WANs are today. SAN capabilities in the network management space include predictive capacity planning, connectivity mapping, performance, and error mapping, largely in the Fibre Channel area.

Element Management

Element management is the most detailed of the layers of storage management, and is usually overlooked by business planning. It involves the management and interaction of individual hardware elements within the SAN from different storage manufacturers, which enables SANs to integrate different storage architectures from different manufacturers.

The Challenge of Implementing SAN Technology

Although rapid strides have been made in SAN technology, and a great deal of implementation work is underway, knowledgeable IT watchers do not expect to see genuine SAN solutions up and running until the last quarter of 2002. To achieve a genuine SAN solution, the following are the basic conditions that a network must meet:

- Any-to-any interconnection of servers and storage systems
- Speed of recovery and magnitude of data loss after disaster
- Universal access and sharing of resources

- Centralized resource management
- Excellent information protection and disaster tolerance
- High levels of security and data integrity in system architectures
- Massive scalability to cope with the future explosive growth in information technology deployment [3]

SAN is not the ultimate cure for the challenges executives face due to the exponential growth of e-business and corresponding requirements for high storage capacity and data protection. However, SAN is an essential component in the future of business IT. Terms like multivendor tape resource sharing and LAN-free data transfer over established IP and Fibre Channel networks will soon become commonplace, as companies encounter the challenge of using SAN technology to provide business advantage.

Implementing SAN environments will consume many resources in the coming years. The selection of trading partners, such as IBM, Tivoli, or one of the other experienced technology creation and deployment companies, will be an essential prerequisite for success. As further enhancements (such as disk and data sharing) become common, the more important it will be to have a broad spectrum of experience and capability with technology integration.

The essential characteristics to look for in a SAN vendor or trading partner are not found in any one capability. It is immaterial whether that capability is in volume mapping, tape management software, or any other point. The key criteria for selecting a company to build a SAN or storage network management system is its ability to help make the best-of-breed technologies from all sources work together for the best total system. Most of the proven, smaller vendors have excellent products that can be considered at the element-management stage of creating a SAN.

The Storage Networking Management Vision

The future vision of storage area networking management is one that exploits both the Information Grid and customers to create storage solutions and networks that build business advantage (see Chapter 19, "SAN Solutions for Consideration"). By taking the best technologies and combining them with those from various integrators, it will be possible to assemble e-business, manufacturing, retail, banking, healthcare, or any other type of business IT solution that an enterprise will need. Islands of data will become transformed into a united resource that shares information and multiplies its worth to each organization. Storage area networks will vault their users through the twenty-first century by enabling the sharing of data and information, and multiplying their worth far beyond their value as discrete resources.

FROM HERE

This chapter covered types of SAN operating systems software and hardware, the driving forces behind SAN, SAN market demands and projections, the evolution of the SAN market, and the value of information. The next chapter discusses disaster recovery; I/O performance; high scalability and flexibility; technology platform, techniques, and alternatives; breaking tradition in video distribution; and SAP R/3 storage management.

END NOTES

[1] EMC Corporation, 35 Parkwood Drive, Hopkinton, Massachusetts, 01748, 2001.

[2] Michael Peterson, "Storage Area Networking: The Next Network," Strategic Research Corp., 350 So. Hope Ave. #A103, Santa Barbara, CA. 93105, 2001.

[3] "The Value of Information," Tivoli Systems Inc., 9442 Capital of Texas Highway North, Arboretum Plaza One, Austin, Texas 78759, USA, 2001.