NAS vs. SAN: Overview

Summary

Both network-attached storage (NAS) and storage area networks (SAN) have gained momentum in the last few years. But what are they, really? How do they differ? And will the models converge in the future?

Table of Contents

Technology Basics

Network-attached storage (NAS) and storage-area networks (SAN) are storage architectures that have gained momentum in the last three to five years and that aim to provide businesses with cheaper, more scalable, and manageable alternatives for their ever-growing data requirements. The NAS model can be considered a mature technology, with commodity products available for small workgroups to enterprisewide data centers. SANs represent a still-evolving technology that targets organizations with large-scale and complex data storage requirements. While solid SAN products are available from multiple vendors, the underlying technologies continue to undergo rapid change. The implementation of a SAN is complex, and great care must be taken in its design and implementation. As an emerging technology, it cannot be assumed that all products and components are interoperable, despite strong industry focus on standards. Especially on the high end of the market, NAS and SAN technologies are beginning to converge, leveraging the benefits of each approach. In the last year, the capacities of NAS and SAN products from each company have increased significantly, and the availability of 2 Gb/second fibre channel devices offers SAN environments a major performance boost.

In practically all business sectors, demand for storage continues on an explosive rise. This growth in storage requirements challenges the limits of traditional technology models. Server-attached storage, typically based on RAID technologies, may no longer scale to the immense capacities needed by many organizations. The current business environment requires storage systems with capacities of multiple terabytes and larger that can be easily expanded, with flexible options—all at the lowest possible cost. This report examines two recently developed technologies in the storage industry, network-attached storage (NAS) and storage-area networks (SANs) that address these needs.

The traditional approach to storage, called server-attached storage, relies on network servers running general-purpose operating systems, such as Unix (including Solaris, HP-UX, Linux, IBM AIX, SGI IRIX), Windows NT/2000/XP, or Novell NetWare. These servers provide access to file systems that can be shared among users of the network while also delivering a variety of other functions, such as distributed printing, directory services, mail delivery, Web hosting, database transaction processing, and many other network services. The classic network server uses some form of SCSI host adapter or RAID controller that in turn connects to groups of disk drives. The disk drives are under the exclusive control of the network server, which shares the data on the drives through protocols such as SMB/CIFS, NFS, or NCP. The operating systems for the classic network file servers are rich in features, have sophisticated user interfaces, and require technical expertise to install, configure, and maintain.

Server-attached storage continues to be appropriate for workgroup-level applications and enterprise environments with moderate storage requirements. The primary advantages of server-attached storage involve simplicity of implementation and modest cost. The disadvantage of this model is scalability. As storage capacity requirements increase to the multiterabyte level, and as data sets need to be distributed among multiple applications or servers, economy and simplicity both fall away. The shortcomings of the basic server-attached storage model led to the rise of the NAS and SAN alternatives.

Network Attached Storage

NAS devices embody many characteristics of the classic file server, but are stripped of all software features and hardware components except those related to file sharing. They include high-performance network interfaces, typically 100 MB/second Ethernet or Gigabit Ethernet (though we do expect to see NAS over Fibre Channel in the near future as well). The hardware architecture is engineered to deliver data from its associated storage devices as efficiently as possible.

Relieved of all other duties, a NAS device can be extremely simple to install and configure. Most can be fully functional in a matter of minutes. NAS devices typically do not come with monitor or keyboard. Users configure and manage these devices remotely through a Web browser or through a utility supplied by the vendor. NAS devices make their internal operating system transparent to the end user. Though they run operating systems—such as Linux, Windows, or proprietary software—they do not require any manual interaction with these environments.

Once connected to the network, NAS devices emulate standard network file servers. They typically appear as exported NFS file systems or as Windows shares through SMB or CIFS. HTTP and FTP access are also standard features. The file systems delivered to the network by a NAS device can either be accessed directly by client computers or can be mounted by application servers. Multiple servers or clients can access a NAS storage volume simultaneously. While NAS devices operate as independent servers, they can leverage the authentication and authorization capabilities of existing servers on the network, eliminating the need to re-create user accounts and access control lists.

A wide range of NAS devices is available. Low-end and midrange systems, with capacities ranging from 50GB to 500GB, target workgroup or departmental applications. High-capacity NAS equipment, with multiple terabytes of storage, appeal to enterprisewide data center environments.

The Pros and Cons of NAS

The key advantages of NAS lie in their lower cost per GB and reduced personnel costs, since they require little technical administration. NAS is a stable, mature model with a number of vendors engaged in vigorous competition to capture a rapidly expanding market. Capacities continue to rise as the cost per MB declines. While in today's market the lion's share of storage is packaged with servers, NAS is on the rise. There is a significant trend away from server-attached storage in favor of NAS.

One of the key challenges to data storage technologies in general, and to NAS in particular, involves data protection. As overall storage capacities grow into multiple terabytes, backing up data poses an enormous problem. Even with the fastest tape technologies, a comprehensive backup will take significant time and may result in network congestion.

Some NAS devices provide a degree of data protection by producing internal backups. These backup images use hashing or compression technologies to minimize the amount of space consumed. In a volatile data environment, data snapshots can be taken every hour. While this type of process does not protect from a system-level catastrophic failure, it does offer the ability to easily recover files lost through accidental deletion or software malfunction.

NAS devices provide storage to servers and client systems on the network through file-sharing protocols, such as NFS and CIFS/SMB. These protocols make NAS storage available to both Unix and Windows platforms. Few NAS systems support Novell's NDS architecture.

The NAS architecture allows files to be shared among multiple servers and applications. Volumes on a NAS appear on the network as NFS, CIFS/SMB, or NetWare shares, and multiple application servers or end-user clients can mount any given file system simultaneously. Provided the application implements standard record-locking and other data integrity features, directories and files can be viewed and updated by multiple remote systems. This file-sharing capability differentiates NAS from SANs, where such file sharing can be accomplished only indirectly.

Storage Area Networks

SANs aim to present solutions to many of the problems associated with large-scale data storage. SANs build on the server-attached model through the creation of a separate network of storage devices,

independent of the organization's LAN or communications network. Storage networks can include disk drives, RAID devices, tape libraries, and other storage equipment. Multiple servers—and even client systems—can participate in the storage network to gain access to these devices. The creation of a network of storage devices offers an organization many more options than the traditional approach of connecting storage devices directly to servers.

SANs move data efficiently without adding to the load of the communications network. While local-area networks (LANs) use protocols such as Ethernet and TCP/IP, SANs currently rely on Fibre Channel technologies (though SAN over IP is another upcoming technology). In its original and simpler form, Fibre Channel-Arbitrated Loop (FC-AL) connects a limited number of devices in a physical ring topology. More sophisticated SANs use hubs, switches, and routers to create a Fibre Channel fabric that can include a very large number of devices spanning extensive geographical locations.

The key principle of SAN involves offloading data transfers from the communications network. In a SAN environment, many data operations can be accomplished without having to traverse the LAN. Data communications protocols, such as Ethernet and IP, introduce significant overhead in the transmission, while Fibre Channel technologies dispatch large amounts of data with great efficiency.

The SAN is transparent to network end users, so client computers and Web users are unaware of the existence of the back-end SAN architecture. From the perspective of users on the LAN, there are no outward signs to indicate that the network servers rely on a SAN for access to storage devices.

The Pros and Cons of SAN

A major limitation of the traditional server-attached storage model lies in the restrictions in the numbers of storage devices that can be added to any given server and the level of difficulty in increasing storage capacities among groups of servers. The SCSI architecture does not scale well to very large-capacity storage systems. The number of devices per SCSI chain and slots available for host bus adapters per server are often too low to achieve the overall capacities needed by many organizations. A basic FC-AL system, however, supports up to 126 devices, while Fibre Channel fabrics can support an unlimited number of devices.

SANs allow servers to be physically separate from their storage devices. While SCSI cabling requires devices to be within a few feet of each other, Fibre Channel supports distances of many kilometers. This flexibility allows the centralization of storage equipment even when servers are distributed. Through a SAN, an organization can create large vaults of centralized storage that can be associated with servers that are housed centrally or are dispersed through the organization. SANs allow organizations with multiple data centers to distribute access to storage assets independently of physical location.

While the underlying technologies differ significantly, network managers will see many similarities between LAN and SAN architectures and equipment. Fibre Channel hubs, for example, allow a number of devices to be interconnected, but share the overall bandwidth of the unit. Fibre Channel switches deliver faster performance and offer each device dedicated bandwidth. In a large SAN, Fibre Channel switches connect to each other in redundant paths to form a physical mesh network that continues to operate even when individual links fail. These interconnected switches form a Fibre Channel fabric that allows each of the storage devices and servers to communicate with one another through redundant paths. Fibre channel bridges connect non-Fibre Channel devices, especially SCSI devices, to the SAN.

SANs can include tape libraries, which can be used to back up all the other storage devices within the fabric or loop. In a LAN-only environment, backup operations can saturate the network. SANs, however, enable backup operations to occur without moving any data on the network and can even occur without

the intervention of a server. Serverless backup, or the direct copying of data from a disk system to a tape library on a SAN, is one of the major benefits of this architecture.

The downside of SANs lies in their cost and complexity. IT departments will need to include technical staff with expertise to manage the SAN in addition to those that support existing LANs and wide-area networks (WANs). The fibre optic network and the communications equipment that compose the SAN require significant investment as well. Yet for many organizations, the overhead of implementing and managing a SAN is less onerous than the problems associated with expanding file servers with dedicated storage to keep up with the demands of an organization's large-scale storage requirements.

One of the major design considerations when implementing a SAN involves reliability and fault tolerance. A well-designed SAN includes many redundant components. Its internal disk systems will rely on RAID configurations that continue to operate even when one or more disk drives fail. Multiple controllers and power supplies provide additional protection. The use of multiple fibre channel switches creates a fabric of connectivity among devices that will survive switch or cable failures. Each level of redundancy increases overall cost.

Another limitation of the SAN model involves the inability for multiple servers to attach to the same device for direct file sharing. With NAS, file sharing is commonplace, since multiple servers—or even clients—can mount the exported file systems and access the same files. Application-level software takes care of file and record locking to prevent data corruption. New approaches are being developed to address this limitation in SAN architecture as part of the general trend toward convergence of SAN and NAS technologies.

Given the ever-diminishing prices relative to capacities in the disk drive market, it is easy to get a false impression of costs for storage managed through a SAN or NAS. In a SAN environment especially, but to a certain extent with NAS as well, the cost of storage involves much more than the raw disks. Switches, high-performance controllers, management software, internal processors, and other devices drive up costs as they increase performance, reliability, scalability, and manageability.

Technology Analysis

For many business environments, server-attached storage lacks the scalability and capacity for current needs and expected growth patterns. NAS approaches the problem with simpler, streamlined storage devices. Through its highly optimized storage design, and the elimination of superfluous hardware and software components, NAS offers lower costs per storage unit while decreasing systems administration costs. The NAS approach leverages the current trends of faster networks, larger disk drive capacities, and the scarcity and increasing costs of systems management staff.

SANs take the alternative approach of building complex and sophisticated storage environments to gain increased scalability, reliability, and manageability. The implementation of a large SAN requires significant initial effort and expense that may be recovered over time through higher reliability, more efficient management of storage assets, and decreased pressures on the LAN.

Though SAN and NAS are distinctive technologies, some synergies exist. High-end NAS equipment, for example, is beginning to incorporate SAN technologies. Several storage companies have begun to offer hardware and software products that blend the two models. In the long term, NAS and SAN will likely converge, or at the least be less differentiated. In fact, Gartner Dataquest forecasts that by 2005, 70.7 percent of the total multiuser RAID-based storage will be fabric attached storage (FAS), based on industry standard protocols. FAS is a new term coined by Gartner Dataquest to encompass both the emerging SAN and NAS architectures. (For more information on the FAS paradigm or Gartner Dataquest's storage

market share forecasts, see *2001 RAID-Based Disk Storage WorldWide Market Share and Forecast* [HWST-WW-MS-0106] dated 1 August 2001 by Roger W. Cox.)

The current generation of SANs rely on fibre channel connectivity. There is also interest in using SCSI over IP, taking advantage of the ever-rising throughput available over Ethernet. IP-based storage networks could be either separate, as with the current SAN approach, or could be merged with the communications network. While this approach compromises the basic tenant of SAN by re-introducing storage-related traffic to the network, it does bypass the need to maintain dual networks.

As the data throughput rates of networks and storage technologies increase, new bottlenecks emerge. With gigabit capacities emerging both in storage controllers and network cards, the current server architecture based on the PCI bus is becoming the weak link. The 64-bit, 66MHz PCI architecture delivers data at a rate of 532 MB/second. New generations of server bus technologies will emerge to accommodate these faster overall performance demands.

The emerging InfiniBand specification, driven by a wide range of hardware manufactures, will provide servers with drastically improved bandwidth to perform I/O operations, relying on a switched fabric for data operations instead the current shared bus architecture. Data throughput rates of up to 6 GB/second become possible. InfiniBand will both complement and compete with current SAN technologies. The InfiniBand Trade Association released Version 1.0 of the specification in October 2000. Early InfiniBand products emerged in 2001. A wider range of high-performance servers based on InfiniBand is expected in 2002 and beyond.

Business Use

Both SAN and NAS technologies are driven by the pervasive needs of organizations for increased data storage and for more flexible options to manage storage environments. On the high end, many businesses must deal with extremely large-scale storage demands. Multiterabyte data sets have become commonplace. The current business environment has created a general trend toward large, consolidated storage environments. ASPs that support business applications for multiple organizations, Storage Service Providers, the migration from departmental accounting schemes to Enterprise Resource Planning (ERP), data warehousing, knowledge management systems, and digitized multimedia content all represent the trend toward data storage consolidation.

The target audience for SANs includes organizations with very large data storage requirements with high performance and availability. In very large-capacity environments, overall costs for storage and the personnel needed to manage it can be significantly lower with SAN than with server-attached storage. SANs allow an organization infinite scalability of capacity and enable redundancy and failover options that would be extremely difficult to accomplish otherwise. In addition, backups can be accomplished without interruptions in service or degradation of network performance.

Standards

SAN and NAS rely on a number of existing standards related to data networking, including Gigabit Ethernet, TCP/IP, device interface protocols (such as Fibre Channel), and reliability standards (such as Extended Data Availability and Protection attributes [EDAP]). The emerging InfiniBand and Storage over IP (SoIP) specifications will likely have future impact on SAN and NAS developments as well.

Network Data Management Protocol (NDMP) addresses the issue of data backups. This initiative defines a standard protocol that storage management applications can use to back up data in a network of heterogeneous servers. NDMP defines a standard agent that can operate on any file server regardless of hardware platform or operating system. This allows manufactures of storage devices to more easily

deliver systems that are backup-ready for any environment and frees the developers of backup applications from having to develop different agents and procedures for each of the network operating systems and NAS storage engines. NDMP was proposed by Network Appliance, a NAS vendor, and Legato Systems, Inc., a developer of storage management and backup applications.

The complexity of SANs demands efficiency in management. SNMP (Simple Network Management Protocol), which is at the heart of management systems for communications networks, has recently found a home in the SAN as well. The availability of an SNMP MIB (Management Information Base) for fiber channel supports automated tools for monitoring a SAN for failures and performance problems.

Technology Leaders

The NAS market is dominated by two major leaders, with a number of additional companies beginning to offer products. Driven by demand for low cost per storage unit, the market is vigorously competitive.

The leading suppliers of NAS storage are Network Appliance and EMC Corporation. A number of other vendors offer products in this arena including Auspex Systems Inc.; Compaq Computer Corporation; Dell; Hewlett-Packard; IBM; Maxtor Corporation; Network Storage Solutions; Procom Technology, Inc.; and Snap Appliances.

Network Appliance (Sunnyvale, California)

Network Appliance develops and manufactures a line of NAS equipment, which it calls NetApp filers. Its F800 filer series, F740 filer, and F86 filer span a range of target environments ranging from the workgroup to the enterprise. The Network Appliance products rely on the Data ONTAP proprietary microkernel operating system.

EMC Corporation (Hopkinton, Massachusetts)

A dominant vendor in the general storage industry, EMC is also a leader in the NAS market. The EMC Celerra File Server targets enterprise-class storage environments. EMC also offers the CLARiiON IP4700 for the NAS market and a family of fibre channel connectivity products called Connectrix.

Other Vendors

Auspex Systems, Inc. (Santa Clara, California)

Auspex pioneered the NAS market, offering network storage products since 1987. Today, Auspex currently offers the NS3000 line of NAS products for all types of business, but describes its key markets as companies involved in software development, streaming media, ISPs, ASPs, oil and gas exploration, manufacturing, and semiconductors. Its products are based on its proprietary operating system, called NetOS, and support for NFS and CIFS file systems.

Compaq Computer Corporation (Houston, Texas)

Compaq, another major storage system supplier, offers products for the server-attached, NAS and SAN environments. Compaq's entries into the NAS market include the TaskSmart N2400 NAS Appliance, a rack-optimized solution with scalability up to 1TB that supports both CIFS and NFS file systems, and the NAS Executor E7000.

Dell Computer Corporation (Round Rock, Texas)

Dell's PowerVault line of NAS solutions includes four products. The PowerVault 715N is the entry-level deskside or rackmountable NAS appliance with support for up to 480GB. The PowerVault 750N (tower) and PowerVault 755N (rackmount) are the midrange offerings. The PowerVault 750N supports up to

438GB internally and the PowerVault 755N supports up to 219GB internally. Both models can scale to over 8TB using external storage. At the high-end of the line, Dell positions its co-branded Dell/EMC IP4700, which offers fibre channel to disk connectivity and scales up to 7.3TB.

Hewlett-Packard Company (Palo Alto, California)

Hewlett-Packard's NAS products include the SureStore NAS Virtual Array solutions, available in a variety of configurations with capacities scaling up to 15.4TB in a clustered solution, and the SureStore NAS XP for enterprise applications that scale up to 73TB. Hewlett-Packard also offers a range of products in its HP SAN family, including systems that support 2 Gb/second throughput and SANs that operate over IP.

IBM (Armonk, New York)

IBM stands as a major powerhouse in the storage market. The company's TotalStorage Solutions product line includes both NAS and SAN offerings. The NAS line includes the TotalStorage NAS 200, the TotalStorage NAS 300 and the TotalStorage NAS 300G. The SAN products include the Enterprise Storage Server (a.k.a. Shark) and the 7133 Serial Disk System. In the last year, IBM has also focused on Storage over IP technologies, leading to its iSCSI (SCSI over IP) products that deliver a SAN environment over IP networks.

Maxtor Corporation (Milpitas, California)

Maxtor's primary business is the manufacturing of hard disk drives, and with its April 2001 merger with Quantum HDD, it became the largest company in that arena. Maxtor's NAS offerings include the MaxAttach 4000 family (primarily targeting low-end and midrange environments with capacities up to 640GB), and the MaxAttach 6000 enterprise-class systems that scale up to 5.7TB. The MaxAttach products rely internally on Windows 2000 and support CIFS, NFS, NetWare, and Mac file systems.

Network Storage Solutions (Chantilly, Virginia)

Network Storage Solutions' (NSS) offerings include its Thunderbolt modular network storage products. Specific products include the Thunderbolt NASengine, the Thunderbolt uStore II, the Thunderbolt SPANstor GT and the ProStor. NSS also offers a storage router that serves as a gateway into a SAN to incorporate storage devices into the NAS. NSS products rely on their own SPANstorT NAS operating systems.

Procom Technology, Inc. (Irvine, California)

Prior to its shift to the NAS market, Procom specialized in CD-ROM servers and was a market leader in that field. The company's line of NetFORCE NAS filers spans both midrange and large-scale environments. On the lower end, the NetFORCE 1000 supports storage amounts up to 1.8TB, while the NetFORCE 3000 series scales up to 17TB.

Snap Appliances (San Jose, California)

This company was recently spun off as a wholly owned subsidiary of Quantum Corporation. Quantum acquired the Snap server technology through its purchase of Meridian Data Corporation in 1999. Meridian Data had been a leader in CD-ROM networking technologies prior to its entry into the NAS market. The company's Snap Server appliances are based on its optimized Snap OS and range in capacity from 40GB to 900GB.

The leading suppliers that offer complete SAN solutions coincide with the major players in the enterprise storage marketplace, including EMC Corporation (Hopkinton, Massachusetts), Hitachi Data Systems (Santa Clara, California), IBM (Armonk, New York), Sun Microsystems (Palo Alto, California), Hewlett-

Packard (Palo Alto, California), Dell Computer Corporation (Round Rock, Texas), and Compaq Computer Corporation (Houston, Texas). Each of these companies offers a variety of storage products and has expertise in designing SANs to meet complex business needs.

A variety of companies manufacture components for SAN infrastructure, such as Fibre Channel host bus adapters, Fibre Channel hubs, switches, and routers as well as fiber optic cabling plants. They include:

- Fibre Channel Host Bus Adapters—QLogic Corporation (Aliso Viejo, California), Emulex Corporation (Costa Mesa, California), Interphase Corporation (Dallas, Texas), ATTO Technology (Amherst, New York), JNI Corporation (San Diego, California), Hewlett-Packard, and Cambex Corporation (Waltham, Massachusetts).
- Fibre Channel Hubs and Switches—Brocade Communications Systems, Inc. (San Jose, California), Gadzoox Networks Inc. (San Jose, California), and High Velocity Systems (Danbury, Connecticut).

Technology Alternatives

Storage over IP (SoIP), also known as iSCSI (Internet SCSI), is a technology that has emerged in the last year that relies on Ethernet and IP network protocols to create a storage network rather than Fibre Channel technologies on which current SANs are based. SoIP leverages existing networks instead of requiring a separate SAN infrastructure. This approach was originally promoted by Nishan Systems, a San Jose, California-based company that offers SoIP switches, but is now also being targeted by several other companies, including IBM, Cisco, Falconstar, Dot Hill, Emulex, Adaptec, Agilent, and others. With SoIP, servers, tape libraries, and storage devices connect through special adapters and switches that deliver data without the inefficiencies associated with the full TCP/IP stack. Its promoters hope that SoIP will achieve greater interoperability and acceptance through the use of widely adopted protocols, like Gigabit Ethernet and IP, rather than Fibre Channel.

Insight

In the current marketplace, NAS and SAN provide a range of products to meet the data storage needs of businesses. NAS offers lower-cost storage with reduced requirements for technical systems administration. Organizations with large-scale data storage requirements and complex arsenals of storage devices can use SAN to create a more manageable and scalable environment. NAS products are low-risk commodities; SANs require careful planning and design due to their reliance on connectivity and storage products that must interoperate seamlessly. Trends to watch include products that blend aspects of both approaches and Storage over IP technologies that provide the same advantages of the SAN without the need for a separate physical infrastructure. Overall capacities will increase as disk manufactures continue to create higher-density drives, and performance will increase with the introduction of faster fibre channel and Ethernet communications devices.