IBM Storage Networking
June 2001

Demystifying Storage Networking

DAS, SAN, NAS, NAS Gateways, Fibre Channel, and iSCSI

by David Sacks
IBM Storage Consultant

Contents

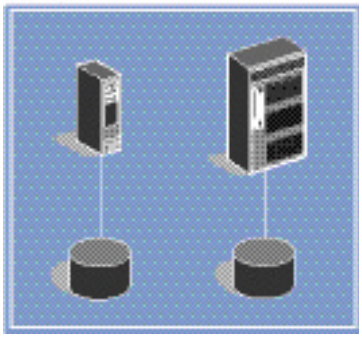
Options for connecting computers to storage have increased dramatically in a short time. Variations (and associated acronyms) for storage networking seem to be materializing out of thin air faster than they can be tracked. Storage networking offers significant capabilities and flexibilities not previously available, and understanding the technology basics is essential to making the best choices.

This paper provides an easy-to-understand comparison of the storage attachment alternatives you can select from to build the infrastructure to access your most important digital asset—your data. Information is presented beginning at a high level and slowly adding increasing detail. The focus is on connectivity options for midrange platforms such as IBM AS/400®, NetWare, Microsoft® Windows NT®, Microsoft Windows® 2000 and UNIX®.[1] Storage management and storage network management, while important topics, are not discussed in detail.

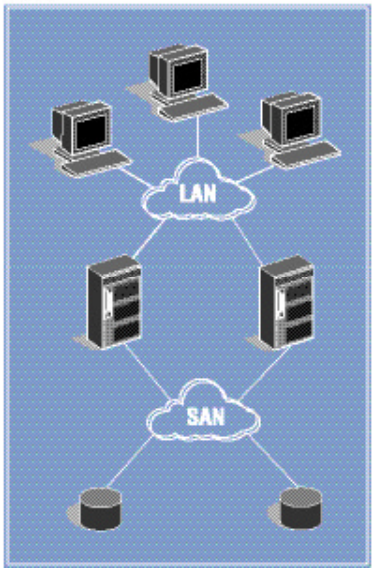
[1]"Midrange" is essentially shorthand for "non-mainframe, non-standalone PC."

In a Nutshell.

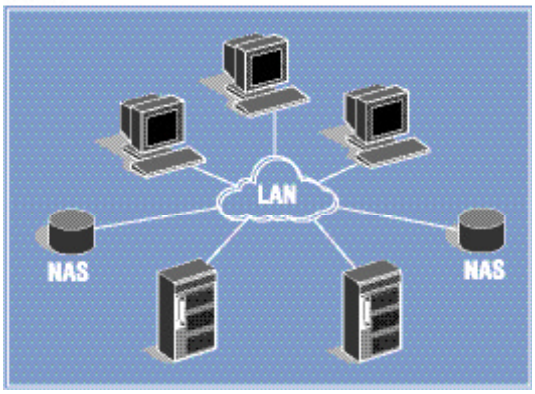
We'll start with a brief description of the major storage networking variations. The paper will then develop the concepts in a more structured manner.



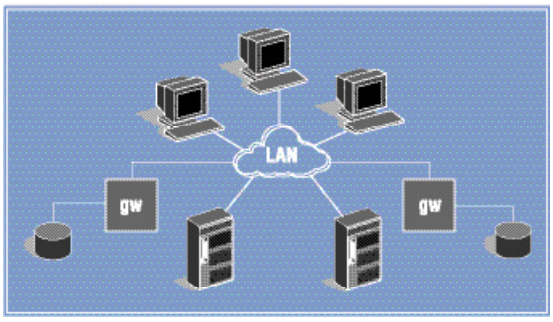
DAS: Direct Attached Storage. Storage (usually disk or tape) is directly attached by a cable to the computer processor. (The hard disk drive inside a PC or a tape drive attached to a single server are simple types of DAS.) I/O requests (also called protocols or commands) access devices directly.



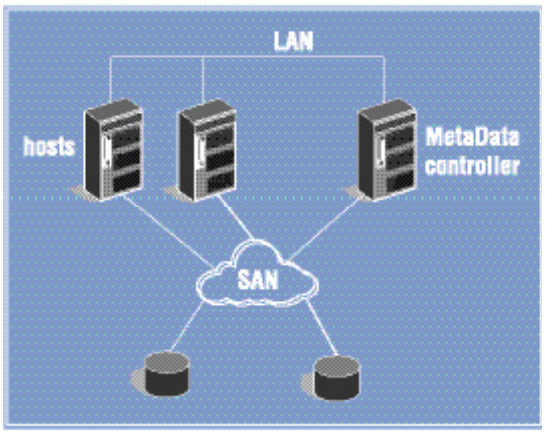
SAN: Storage Area Network. Storage resides on a dedicated network. Like DAS, I/O requests access devices directly. Today, most SANs use Fibre Channel media, providing an any-to-any connection for processors and storage on that network. Ethernet media using an I/O protocol called iSCSI is emerging in 2001.



NAS: Network Attached Storage. A NAS device ("appliance"), usually an integrated processor plus disk storage, is attached to a TCP/IP-based network (LAN or WAN), and accessed using specialized file access/file sharing protocols. File requests received by a NAS are translated by the internal processor to device requests.



NAS gateway: A NAS device without integrated storage (i.e., just the NAS processor). Instead, the NAS device connects externally to storage by direct attachment or by a SAN.



SANergy: SANergy is software from IBM and Tivoli that provides NAS-like file sharing, with data sent over the SAN rather than the LAN for improved performance. (IBM NAS gateways also include SANergy function.)

Why are there so many forms of storage networking? For one, new technologies emerge and evolve but don't replace the investment in previous technologies overnight. And no single storage networking approach solves all problems or optimizes all variables. There are tradeoffs in cost, ease-of-management, performance, distance and maturity, to name a few of these variables. For the foreseeable future, multiple storage network alternatives will coexist—often within the same organization.

The benefits of the major types of processor-to-storage connectivity can be briefly summarized as:

DAS is optimized for single, isolated processors and low initial cost.

SAN is optimized for performance and scalability. Some of the major potential benefits include support for high-speed Fibre Channel media which is optimized for storage traffic, managing multiple disk and tape devices as a shared pool with a single point of control, specialized backup facilities that can reduce server and LAN utilization and wide industry support.

NAS is optimized for ease-of-management and file sharing using lower-cost Ethernet-based networks. Installation is relatively quick, and storage capacity is automatically assigned to users on demand.

NAS gateways are optimized to provide NAS benefits with more flexibility in selecting the disk storage than offered by a conventional NAS device. Gateways can also protect and enhance the value of installed disk systems.

Tivoli SANergy is optimized for data sharing (like a NAS), but at SAN speeds. Tivoli SANergy is disk vendor-independent, and can be added to an existing SAN to enhance its value.

Introducing the Concepts.

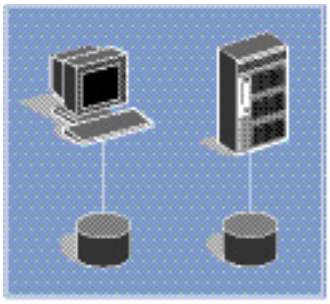
Let's step back and introduce the concepts that will lead to understanding the storage attachment alternatives. There are just three key concepts to be understood:

- **Connectivity:** how processors and storage are physically connected. Think of this as how the connections would be drawn in a picture.
- **Media:** the type of cabling and associated protocol that provides the connection.
- **I/O protocol:** how I/O requests are communicated over the media.

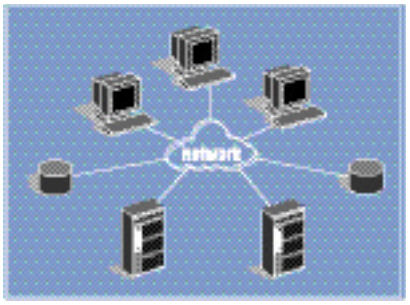
It is how these three items are combined in practice that differentiates the various ways processors (hosts) and storage can be connected together. Essentially, storage is attached to processors over a direct or network connection, and they communicate by the way of an I/O protocol that runs "on top of" the media protocol. Let's examine the three concepts one at a time.

Connectivity.

The pictures below illustrate the two basic ways to physically connect storage to processors.



Direct attach—a single storage device is connected to a single processor (host).



Network attach—one or more processors are connected to one or more storage devices.

The simplest form of direct attached storage (DAS) is a single disk drive or single tape drive connected to a single processor. Some disk systems allow the aggregate disk capacity to be "carved" into partitions (subsets) of capacity where each partition can be assigned to a different processor. Further, the subsystem may allow partitions to be manually reassigned from one processor to another.[2] This is essentially still a DAS approach to storage.

Direct attach can be thought of as a minimal network. For simplicity, and as is common in the industry, this paper will sometimes refer to storage networking alternatives without explicitly mentioning direct attach, but it should be considered as one such alternative.

Following industry convention, a cloud is used to indicate a network without showing the inner details of how cables, and devices such as hubs and switches, may be connected to form a particular implementation. Such implementations will vary from organization to organization and do not need to be understood in order to explain storage connectivity alternatives. The idea is that all objects connected to the same cloud can potentially communicate with each other. (Such any-to-any flexibility can be managed in practice to prevent undesired communications.)

[2] For example, the IBM Enterprise Storage Server disk system offers this flexibility.

Media.

The media is the physical wiring and cabling that connects storage and processors.

Media is always managed by a low-level protocol unique to that media regardless of the attached devices. A protocol is the rules for exchanging information between two objects. In computers, this specifies the format and sequence of electronic messages. In storage-to-processor connections, the following media and associated protocols are prominent. All are open, industry standards.

- **Ethernet:** Ethernet began as a media for building LANs in the 1980s. Typical bandwidths are 10Mbps, 100Mbps, and 1Gbps.[3] Ethernet is a media and its protocol. IP-based protocols such as TCP/IP generally run on top of Ethernet.
- **Fibre Channel:** Fibre Channel is a technology developed in the 1990s that has become increasingly popular as a storage-to-processor media (for both SANs and DAS). Bandwidth is generally 100MBps, with 200MBps expected in 2001.
- **Parallel SCSI (Small Computer Systems Interface):** (Pronounced "scuzzy"). Parallel SCSI is an evolving technology with origins in the 1980s. Typical bandwidths are 40MBps (also called UltraSCSI), 80MBps (also called Ultra2 SCSI), and 160MBps (also called Ultra160 SCSI). Parallel SCSI is limited to relatively short distances (25 meters or less, maximum) and so is appropriate for direct attach, especially when storage and processors are in the same cabinet, but is not well-suited for networking.
- **SSA (Serial Storage Architecture):** SSA is a media technology optimized for high-performance and used to connect disks together inside some disk systems. Bandwidth is 160MBps.

[3] MBps=megabytes/second, Mbps=megabits/second, and Gbps=gigabits/second. 1Gbps generally equals 100MBps since the (Ethernet and Fibre Channel) protocols involved use special 10-bit bytes.

I/O Protocols.

I/O processing uses specific protocols that run "on top of" the underlying media protocols. (In the case of Ethernet, I/O protocols generally run at some level on an IP protocol stack.) The following are the most common I/O protocols supported on midrange platforms.

- **SCSI (Small Computer Systems Interface):** The I/O protocol most prevalent in the midrange world. A SCSI I/O command might tell a disk device to return data from a specific location on a disk drive, or it might tell a tape library to mount a specific cartridge. SCSI is often called a "block level" protocol, or block-I/O, because SCSI commands specify particular block (sector) locations on a specific disk. Originally, SCSI I/O commands could only be sent over media called "parallel SCSI". Today, SCSI commands can be issued over different types of media such as Fibre Channel, SSA, and Ethernet, as well as over parallel SCSI.
- **NFS (Network File System):** A file-level (also called file-I/O) protocol for accessing and potentially sharing data. This protocol is device-independent in that an NFS command might just request reading the first 80 characters from a file, without knowing the location of the data on the device. NFS has its origins in the UNIX world.
- **CIFS (Common Internet File System, often pronounced "siffs"):** A file-level protocol for accessing and potentially sharing data. This protocol is device-independent in that a CIFS command, like NFS, might just request reading the first 80 characters from a file, without knowing the location of the data on the device. CIFS has its origins in the Microsoft Windows NT world.

With SCSI (block-I/O), disk volumes are visible to the servers attached to them. With NFS and CIFS (file-I/O), only files are visible to the attached processors, but the disk volumes on which those files reside are not visible to those processors.[4]

[4] While similar in principle, NFS and CIFS differ in many aspects such as user authorization and locking protocols. For the purposes of this guide those differences are unimportant. Other protocols that deal with files but not disk volumes include FTP (File Transfer Protocol) for transmitting entire files over a network, and HTTP (Hypertext Transfer Protocol) for transmitting Web pages over a network. These protocols are not further discussed in this paper, though they are supported by some NAS appliances.

The Storage Networking Acronyms.

Storage networking acronyms such as DAS and NAS can be viewed as various combinations of the three key concepts discussed above: connectivity, media and I/O protocol. Not every possible combination is implemented today, or may be implemented in the future.[5] Please refer to the figure on page 3.

- **DAS (Direct Attached Storage):** Storage is directly attached by a cable to the processor. The media could be any (i.e., Fibre Channel, SCSI, SSA, Ethernet). The I/O protocol is SCSI.
- **SAN (Storage Area Network):** Storage resides on a dedicated network, providing an any-to-any connection for processors and storage on that network. The most common media is Fibre Channel, but Ethernet-based SANs are emerging. (See iSCSI below). The I/O protocol is SCSI.
- **NAS (Network Attached Storage):** A NAS device is attached to a TCP/IP-based network (LAN or WAN), and accessed using CIFS and NFS—specialized I/O protocols for file access and file sharing. A NAS device is sometimes also called a file server, or "filer" or "NAS appliance". It receives an NFS or CIFS request over a network and has an internal processor which translates that request to the SCSI block-I/O commands to access the appropriate device only visible to the NAS product itself.
- **NAS gateway:** A NAS device with the internal processor but without integrated storage. Instead, the NAS device connects to storage by direct attachment or by a SAN. This term is most meaningful when there is a choice of the disk storage to attach to the gateway.
- **iSCSI:** Storage is attached to a TCP/IP-based network, and is accessed by block-I/O SCSI commands. iSCSI could be direct attached or network attached (i.e., DAS or SAN).
- **Tivoli SANergy:** This is a software product from IBM and Tivoli that provides NAS-like file sharing using NFS or CIFS I/O protocols, but with data sent over the SAN (using SCSI I/O protocols) rather than the LAN for improved performance. SANergy can run without a NAS appliance, and also is included with IBM NAS gateways to provide enhanced I/O performance.

[5] For example, it is possible to run TCP/IP over Fibre Channel and so use Fibre Channel as a LAN, and thus potentially use it for NFS and CIFS requests. However, this is rarely if ever implemented in practice.

Note that while the terms NAS and SAN seem similar, SAN refers to a dedicated storage network and NAS is a device on a LAN/WAN network (whether the network is shared or dedicated to storage). Occasionally, the industry uses the term "SAS" to refer to SAN Attached Storage. As you may realize, storage networking terminology is not intuitive, and isn't standardized; you may want to take care that you and others are talking about the same thing when using a given term.

A Tabular Comparison.

The various storage networking alternatives are summarized in the following table.

A Tabular Comparison

Processor-storage connection	Network	Media	I/O Protocol	Bandwidth	Capacity Sharing	Data Sharing
DAS	No	"Under the processor covers" wiring, parallel SCSI, Fibre Channel, or SSA	SCSI	40MBps up to 160MBps, depending on media	Manual or no	No
SAN	Yes	Fibre Channel is most common, with Ethernet emerging	SCSI	100MBps Fibre Channel, with 200MBps expected during 2001	Yes	Requires specialized software such as SANergy
NAS	Yes	Ethernet	NFS, CIFS	10Mbps to 1Gbps	Yes	Yes
NAS gateway	Yes	Ethernet	NFS, CIFS	10Mbps to 1Gbps	Yes	Yes
iSCSI	Yes	Ethernet	SCSI	10Mbps to 1Gbps	Yes	Requires specialized software such as SANergy
Tivoli SANergy	Yes	SAN media	NFS, CIFS, SCSI	SAN speeds	Yes	Yes

Legend.

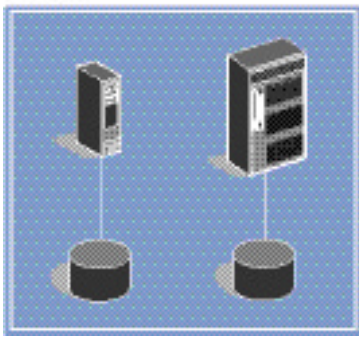
- **Processor-storage connection:** DAS or SAN or NAS, etc.
- **Network:** whether storage can be accessed by only one or by multiple processors.
- **Media:** the name of the media technology that connects the processors and storage. You can think of this as the cable and the basic low-level protocol to send data over the media.
- **I/O Protocol:** types of messages sent over the network media to access storage.
- **Bandwidth:** the bandwidths supported by the various media. Bandwidth is a technical specification of maximum potential throughput and does not indicate the performance a particular application will see. That performance will vary based on many factors beyond the scope of this discussion.
- **Capacity sharing:** the ability to pool disk space or tape drives for use by multiple processors. For disk systems, capacity can be divided into partitions assigned to specific processors. In a large disk system, it may be possible to manually reassign storage from one partition to another. For tape, a software-based management facility is used to ensure only one processor uses a given tape drive and cartridge at a given time.
- **Data sharing:** whether files can be shared concurrently among multiple hosts. This carries disk system capacity sharing to the next step—sharing of the data within the same partition by multiple processors at the same time. Benefits include reduced number of copies of data, access to current data, and reduced need to transfer copies of data between processors. In addition, by accessing data over a network using file-I/O protocols that are used for file sharing, processors and operating systems can be changed without having to reformat the data.[6]

[6] In general, every operating system, including every UNIX-based variant, stores data in a format that only that same operating system understands. File-I/O puts data on the network so that operating systems can access it using industry-standard protocols without any dependence on data format.

Exploring the Alternatives.

Let's explore each of the storage network variations one at a time.

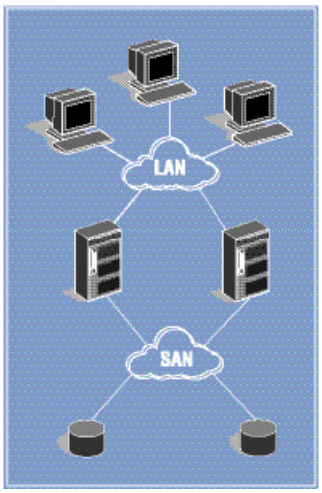
Direct Attached Storage (DAS)



Direct Attached Storage is storage that is generally restricted to access by a single host (processor); sometimes by two hosts in small cluster (failover or failback) configurations. Even an enterprise-class disk system such as an IBM Enterprise Storage Server™ (ESS) can be effectively configured as DAS by assigning portions (partitions) of the internal disk capacity to designated hosts. Each of these partitions is connected directly to the ESS by way of SCSI or point-to-point Fibre Channel paths.

For an individual, isolated processor, such as a laptop, a desktop PC, or a single server in a small business, disk storage usually resides inside the processor enclosure and is a simple form of DAS. When an organization has multiple processors, DAS may initially appear to be low cost from the point of view of each user or department. However, from the wider perspective of the entire organization, the Total Cost of Ownership of DAS may be higher than for networking approaches due to the difficulty of sharing unused capacity with other processors, and the lack of a central point of management for multiple disk systems.

Storage Area Networks (SAN).



A SAN is a dedicated network for storage devices and the processors that access those devices. SANs today are usually built using Fibre Channel technology, but the concept of a SAN is independent of the underlying type of network.

I/O requests to disk storage on a SAN are called “block I/Os” because, just as for direct-attached disk, the read and write I/O commands identify a specific device (disk drive or tape drive) and, in the case of disks, specific block (sector) locations on the disk.

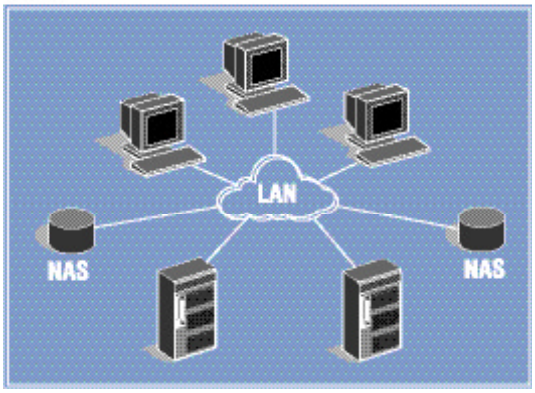
The major potential benefits of a SAN can be categorized as:

- **Access:** longer distance between processors and storage, higher availability, improved performance (because I/O traffic is offloaded from a LAN to a dedicated network, and because Fibre Channel is generally faster than most LAN media). Also, a larger number of processors can be connected to the same storage device compared to typical built-in device attachment facilities.
- **Consolidation:** replacement of multiple independent storage devices by fewer devices that support capacity sharing—this is also called disk and tape pooling. SANs provide the ultimate in scalability, because software can allow multiple SAN devices to appear as a single pool of storage accessible to all processors on the SAN. Storage on a SAN can be managed from a single point of control. Controls over which hosts can see which storage (called zoning and LUN masking) can be implemented.

- **Protection:** LAN-free backups occur over the SAN rather than the (slower) LAN, and server-free backups can let disk storage "write itself" directly to tape without processor overhead.
- **Data Sharing:** sharing data, as noted earlier, offers benefits such as reducing the number of copies of files, increasing accessibility to current data and reducing the need to transfer copies of data between servers over the network.

Because it uses a specialized network usually based on Fibre Channel, the initial cost to implement a SAN will generally be higher than for DAS or NAS. SANs require specialized hardware and software to manage the SAN and provide many of its potential benefits. Additionally, an organization must add new skills to manage this sophisticated technology. However, an analysis may justify the cost due to the long-term lower Total Cost of Ownership compared to an alternative connectivity approach.

Network Attached Storage (NAS).



A NAS is a device that resides on a network that may be shared with non-storage traffic. Today, the network is usually an Ethernet LAN, but could be any network that supports the IP-based protocols that NAS uses.

In contrast to "block I/O" used by DAS and SANs, NAS I/O requests are called "file I/Os". File I/O is a higher-level type of request that, in essence, specifies the file to be accessed, an offset into the file (as if the file was a set of contiguous bytes), and a number of bytes to read or write beginning at that offset. Unlike block I/O, there is no awareness of a disk volume or disk sectors in a file I/O request. Inside the NAS product ("appliance"), an operating system or operating system kernel tracks where files are located on disk, and issues a block I/O request to the disks to fulfill the file I/O read and write requests it receives.

A NAS appliance generally supports disk storage, and sometimes CD-ROM, in an integrated package; tape drives may often be attached for backup purposes. In contrast to SAN devices that can usually also be direct-attached (e.g., by point-to-point Fibre Channel) as well as network-attached by SAN hubs and switches, a NAS device is generally only a NAS device and attaches only to processors over a LAN or WAN. (NAS gateways, discussed later, offer some flexibility in combining NAS and SAN characteristics.)

Which is better, NAS or SAN? Neither and both. There are tradeoffs, and the best approach depends on the particular environment. Some organizations may implement a mix of NAS, SAN and DAS solutions. Consider the following.

Ease-of-installation.
NAS is generally easier to install and manage than a SAN. A NAS appliance can usually be installed on an existing LAN/WAN network. NAS manufacturers often cite "up and running" times of 30 minutes or less. (Customization procedures may take additional time.) Hosts can potentially start to access NAS storage quickly, without needing disk volume definitions or special device drivers. In contrast, SANs take more planning, including design of a Fibre Channel network and selection/installation of SAN management software.

Backup.
Most NAS appliances in the marketplace include a "snapshot" backup facility, to make backup copies of data onto tape while minimizing application downtime. For SANs, such facilities are available on selected disk systems or in selected storage management packages.

Resource pooling.
NAS allows capacity within the appliance to be pooled. That is, the NAS device is configured as one or more file systems, each residing on a specified set of disk volumes. All users accessing the same file system are assigned space within it on demand. That is certainly more efficient than buying each user their own disk volumes (DAS), which often leads to some users having too much capacity and others too little. So NAS pooling can minimize the need to manually reassign capacity among users. However, NAS pooling resides within a NAS appliance, and there is little if any sharing of resources across multiple appliances. This raises costs and management complexity as the number of NAS nodes increases. In contrast, an advantage of a SAN is that all devices on a SAN can be pooled—multiple disk and tape systems. So, at some point as total capacity grows, a SAN may be easier to manage and more cost effective.

File sharing.
NAS provides file sharing, but with products like SANergy discussed later, a SAN can do this as well. Many organizations install a NAS, not for file sharing, but for its ease of installation and management.

Performance.
How do NAS and SAN performance compare? It may depend on the particular configuration, but SAN is generally considered to be faster. This is mainly due to:

- SAN's use of a dedicated network (though this is possible with NAS).
- SAN network speed (100MBps Fibre Channel vs. 10Mbitps or 100Mbitps Ethernet, though Gigabit Ethernet at 100MBps is becoming more common).
- host overhead (Fibre Channel protocol handling is done in the host bus adapter, while TCP/IP protocol handling is done in host software and can add considerable overhead. There is work in the industry to offload TCP/IP protocol handling to host bus adapters, which will eventually help with the processor overhead problem.)

For relatively low amounts of activity, NAS and SAN may both perform acceptably well. Today, however, NAS will generally not scale as well as SAN in performance. It is not clear where the "break even" point is, but NAS devices often can handle several thousand I/Os per second with good average response time (e.g., under 10 milliseconds average for small random I/Os).

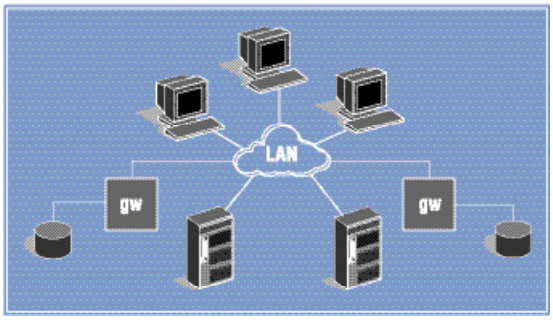
To summarize the comparison between NAS and SAN, while a NAS appliance is generally less scalable and less grandiose than a SAN, it can satisfy storage requirements in numerous environments ranging from small businesses to workgroups or departments in large organizations. NAS alone is, and will remain, a good fit in many environments. NAS and SAN hybrids (by way of NAS gateways, discussed below) will be a good fit in the largest environments, combining the best of both worlds.

NAS will generally cost more than DAS (because of its built-in file sharing intelligence), but has the following potential advantages: distance (because it is attached over a network), large number of users being able to access the same storage device, capacity pooling within the NAS appliance (sharing capacity among all hosts using the NAS), and file sharing (as opposed to data transfer or multiple copies on distributed hosts).

NAS appliances support standard file access protocols such as NFS, CIFS, and sometimes others, that run over an IP network. These protocols were developed before dedicated NAS “appliances” existed, and are often implemented in software that runs on most client and server processors. So, in fact, anyone could build their own NAS device by taking a server of any size and installing NFS programming on it, for example. NFS is actually supported directly by most operating systems, or is available from software vendors. The builder or integrator can use any disk products they want, even a single, internal disk for a small NAS built using a low-cost desktop PC.

Building your own NAS means flexibility. But buying an integrated NAS means less time, assurance that the “package” works, vendor support for the package, and usually specialized software tuned for the NAS environment and thus providing much higher performance than possible in a general purpose server and OS environment.

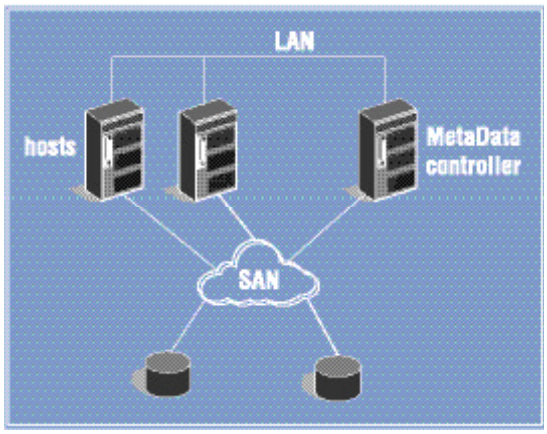
NAS Gateways.



A NAS gateway provides the function of a conventional NAS appliance but without integrated disk storage. The disk storage is attached externally to the gateway, possibly sold separately, and may also be a standalone offering for direct or SAN attachment. The gateway accepts a file I/O request (e.g., using the NFS or CIFS protocols) and translates that to a SCSI block-I/O request to access the external attached disk storage. The gateway approach to file sharing offers the benefits of a conventional NAS appliance, with additional potential advantages:

- increased choice of disk types.
- increased capability (such as large read:write cache or remote copy functions).
- increased disk capacity scalability (compared to the capacity limits of an integrated NAS appliance).
- ability to preserve and enhance the value of selected installed disk systems by adding file sharing.
- ability to offer file sharing and block-I/O on the same disk system. Disk capacity in the SAN could be shared (reassigned) among gateway and non-gateway use. So a gateway can be viewed as a NAS/SAN hybrid, increasing flexibility and potentially lowering costs (vs. capacity that might go underutilized if it were permanently dedicated to a NAS appliance or to a SAN).

SANergy.



In brief, SANergy is software from IBM and Tivoli that provides NAS-like file sharing, with data sent over the SAN rather than the LAN for improved performance.

Some in the industry are calling SANergy and similar facilities SAFS - SAN Attached File Systems.

SANergy has attributes of NAS and SAN, with additional flexibility. SANergy supports the NFS and CIFS protocols, but allows the installation to use virtually any disk storage they want (Fibre Channel, iSCSI, parallel SCSI, and SSA storage will all work.)

Here is a typical SANergy scenario. A set of processors run SANergy client software. The initial CIFS or NFS request for a file is intercepted by the SANergy client and sent over a LAN to a processor running SANergy Meta Data Controller (MDC) software which handles standard CIFS and NFS protocol functions such as authorization. The SANergy client dynamically transmits the actual I/O (data) traffic over the LAN or over the SAN, whichever is optimal.

Functionally, SANergy supports the protocols of a conventional NAS appliance but with significantly higher performance while not requiring the dedicated NAS processor front-end to the disk storage. Instead, SANergy sits as software in the client hosts (plus the MDC). See *www.tivoli.com/sanergy*

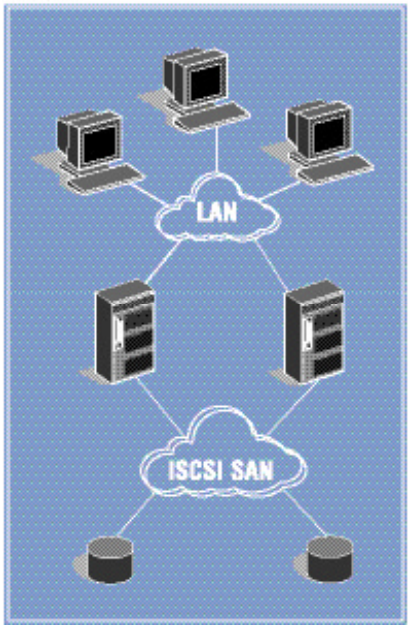
for more information.

IBM NAS gateways support running SANergy internally. This allows applications to access data using protocols supported by the gateway—CIFS, NFS, FTP, HTTP and NetWare File System—yet process I/Os at SAN speeds. An additional benefit is the ability to use multiple NAS gateways each with SANergy to access the same files, providing very high performance by scaling beyond the limits of a single NAS appliance. This can lower costs compared to adding NAS appliances each with dedicated disk storage.

Consider the following scenario that illustrates how IBM NAS gateways and SANergy can work together:

A Web server receives HTTP requests for Web pages and sends them to an IBM NAS gateway which in turn connects to disks over a SAN. Performance is degraded due to a large volume of Web pages being returned to the server over the LAN. So, the installation adds an adapter connecting the server to the SAN, adds SANergy client software to the Web server, and enables the SANergy MDC in the gateway. Now, Web pages travel from the disk to the Web server directly at SAN speeds. If traffic increases so that high server utilization becomes the bottleneck, then a second server with a SANergy client could be added, and connected to the MDC and the SAN similar to the first server. Both servers access the same Web pages at high-speed by using SANergy.

iSCSI.



iSCSI is a proposed industry-standard that allows SCSI I/O commands to be sent over a network using the popular TCP/IP protocol. This is analogous to the way SCSI commands are already mapped to Fibre Channel, parallel SCSI, and SSA media. The proposal was made to the IETF (Internet Engineering Task Force) standards body jointly by Cisco Systems, Inc. and IBM, and is expected to be ratified in mid 2001. The iSCSI standard is also supported by SNIA (Storage Networking Industry Association).

iSCSI connectivity can be implemented in different ways. Assume that an iSCSI device driver is installed in a server to accept application I/O requests and send them over a LAN using the iSCSI protocol. The target storage device could be directly attached to the LAN. An example of this configuration is the IBM TotalStorage IP Storage 200i disk system. An alternative to a native iSCSI device would be to use a router (protocol converter) that connects to the LAN, but has a Fibre Channel port on the "other side" so that it also connects to a storage device that supports Fibre Channel attachment. This allows storage products without native iSCSI ports to be accessed via iSCSI, and allows servers to access that storage without needing a Fibre Channel host bus adapter card. An example of this approach is the Cisco 5420 Storage Router connected to a Fibre Channel port on an IBM Enterprise Storage Server (Shark) disk system.

Because the concepts and products surrounding DAS, SAN and NAS preceded iSCSI, it is natural to try to understand where iSCSI fits in the world by comparing it to those concepts.

- **Definition.** iSCSI is a mapping of the SCSI I/O protocol to the TCP/IP protocol (which in turn usually runs over Ethernet). SAN and DAS are connection alternatives, while a NAS is a device.
- **Connectivity.** iSCSI can be used for DAS or SAN connections to devices. iSCSI devices could be placed on an existing LAN (shared with other applications), or on a LAN dedicated to storage I/O, or even on a LAN connected to only one processor (DAS). The same applies to NAS as well.
- **Media.** iSCSI and NAS devices both attach to IP networks. This is attractive (vs. the newer Fibre Channel) because of the widespread use of these networks, meaning they are already in place in most organizations and are supported by existing skills. The well-known early-life interoperability problems of devices on Fibre Channel SANs would seemingly disappear on networks using the familiar TCP/IP protocol. TCP/IP-based networks can potentially support longer distances than can pure Fibre Channel SANs.
- **I/O protocol.** iSCSI uses the SCSI I/O protocol. Therefore, it is block I/O-oriented like a DAS or SAN, rather than file I/O-oriented like a NAS appliance.
- **File sharing.** NAS supports file sharing while iSCSI SANs and Fibre Channel SANs generally do not. However, the SANergy product can add file sharing capabilities to iSCSI SANs and Fibre Channel SANs.
- **Management.** iSCSI is managed like any direct-attach SCSI device. iSCSI-connected disk volumes are visible to attached processors. Backup of data is done through any method that supports SCSI-attached volumes. A NAS appliance, because it “hides” disk volumes from its clients and often includes specialized backup facilities, may be easier to install and manage. Compared to newer Fibre Channel SANs, iSCSI benefits from using networks with established network management tools and people skills. SANs currently have more storage-related management tools than iSCSI, such as support for tape sharing for backup; this advantage will likely diminish as iSCSI matures and

the market demands SAN-like management for iSCSI devices.

- **Performance.** A performance comparison is difficult to generalize, because there are so many variables. That said, Fibre Channel at 100MBps (1 Gbps) is generally more efficient for I/O traffic than TCP/IP over Ethernet at equivalent bandwidth. iSCSI may perform better than NAS (both on Ethernet) due to reduced protocol overhead, since it handles SCSI directly rather than translating between file-I/O protocols and SCSI. Another performance consideration is impact on processor utilization. Fibre Channel SANs support SCSI commands mapped directly to Fibre Channel media, and processor overhead for this mapping is low. In iSCSI, handling of the TCP/IP protocol requires processor cycles at both ends. Therefore, at this early time in the evolution of iSCSI, it is likely best suited for situations of relatively low I/O activity. This point generally applies to NAS as well. ("Low" in this case can still be thousands of I/Os per second, but will be less than the highest performance levels a SAN could support.)
- **Cost.** Cost comparisons are difficult to generalize and will likely depend on particular products. An iSCSI SAN likely has a lower cost than a Fibre Channel SAN. For example, iSCSI network hardware such as Ethernet host adapters are generally lower cost than Fibre Channel host adapters; if iSCSI (or NAS) is attached to an existing LAN, no new host adapter cards may be needed at all. An iSCSI SAN can be built more quickly and with fewer new skills than a Fibre Channel SAN. An iSCSI disk device, all else equal, may be lower cost than a NAS appliance since the iSCSI device does not need to support file systems, file sharing protocols, and other facilities often integrated into NAS products.

The fundamental technical difference between iSCSI and NAS is that iSCSI is block-I/O oriented while NAS is file-I/O oriented. The fundamental technical difference between iSCSI and Fibre Channel SANs is that iSCSI uses TCP/IP networks. Therefore, iSCSI devices fill a void by uniquely supporting block-I/O applications over TCP/IP (usually Ethernet) networks.

The small table below summarizes this discussion. The columns show media alternatives, while the rows show how Block I/O and File I/O are supported on the media.

	Fibre Channel	**IP-Based media (Ethernet)**
Block I/O	Fibre Channel DAS or SAN	iSCSI DAS or SAN
File I/O	Not directly supported. Indirectly, SANergy reroutes File I/O data over Fibre Channel for improved performance	CIFS and NFS through NAS or NAS gateway, or through SANergy which reroutes I/Os over Fibre Channel for improved performance

An example where iSCSI would be a good fit is an environment with a database system that uses block-I/O to “raw” volumes without an underlying file system, and using Ethernet as the preferred connection media. Another good fit is an application that uses operating system logical volume facilities to control placement of data on specific disk locations (e.g., using outer vs. inner cylinders), and Ethernet is the preferred connection media. Some disk utilities, such as those that relocate data on disk to minimize seek times, likely use SCSI commands directly. Any program that issues SCSI commands directly rather than file system commands will not work with NAS, but will work with iSCSI or Fibre Channel SANs.

Future Directions.

The storage networking industry is moving so fast that any predictions should be treated cautiously. Certainly, higher speed media, both 200MBps Fibre Channel and faster Ethernet are expected soon.

The ability for organizations to implement "open SANs" and mix-and-match heterogeneous vendor storage and network components is increasing as experience with storage networks grows and as standards for interoperability evolve and are complied with. Tivoli's Storage Network Manager, for example, is a vendor-neutral SAN management product that adheres to open industry standards.

The industry is developing specialized chips and device adapters that will offload the TCP/IP protocol handling from the host and disk system processor, making iSCSI (and probably NAS as well) increasingly practical in more I/O-intensive environments. While iSCSI will likely start small, it is expected to increase in capability and popularity over time, providing SAN benefits such as scalability and storage network-oriented management tools, but without the need for a specialized Fibre Channel network.

NAS, SAN, and iSCSI will be increasingly converging. For example, if a NAS appliance is on a LAN dedicated to just the NAS storage traffic, it is SAN-like in its dedication to storage. A NAS gateway appears NAS-like to clients, but may attach to disks or tape through a backend Fibre Channel SAN. With iSCSI, a SAN can be built using Ethernet media, which is the media NAS generally uses today. Organizations will have increasing ability to customize storage connectivity to their particular needs, but the choices also mean more expertise is needed to make the best decisions.

iSCSI may accelerate the convergence of NAS and SANs. TCP/IP is already the entrenched vehicle for file-level protocols (such as CIFS, NFS, FTP and HTTP). Adding block-I/O to Ethernet by way of iSCSI appears to be a major industry direction, while adding file-I/O to Fibre Channel does not appear to have the same momentum (though it is possible since TCP/IP can be mapped to Fibre Channel media). To be clear, this does not mean Fibre Channel SANs will disappear anytime soon or are even declining in acceptance. Quite the contrary. Fibre Channel SANs still provide the fastest and most scalable network, offer pooling and other management functions not yet available for iSCSI storage, and there is extensive industry and customer commitment to Fibre Channel.

Today, different operating systems have different file system formats. NFS and CIFS hide the format, but have few if any management capabilities above file sharing and no pooling of capacity across appliances. IBM has previewed its plans to deliver IBM Storage Tank, a product based on work done by IBM Research. Storage Tank is planned to provide a common file system across multiple, heterogeneous storage systems, offering more efficient utilization of capacity and support of policy-based file placement to simplify storage management.

Selecting the Best Alternative.

Which storage networking alternative is best for a given organization may be obvious based on organizational objectives, current storage infrastructure and what the alternatives provide. Or, it may be a totally open question. Storage technology has clearly become more varied and sophisticated, and accordingly decisions have become more complex than ever. Choice means flexibility and that's good, but which choice to make is not always clear.

Some Rules of Thumb

If you knew nothing else, the following basic guidelines may help you get started:

If DAS, NAS, SAN, or iSCSI are currently implemented in the organization, and growth using that same technology meets requirements (including cost), then it is probably easiest (e.g., least disruption) to stay with what exists.

If a group of individual users with PCs needs to share disk storage capacity and perhaps also share files in that storage, then NAS may be easiest to install and manage.

If application servers need to share disk storage, and are each accessing independent (block I/O) databases, SAN or iSCSI may be appropriate. If a SAN already exists, it probably makes sense to integrate with it. For a small number of servers where no SAN exists, iSCSI may be less expensive and less complex. The larger the number of servers accessing a pool of storage, and the higher the performance requirements, the more likely SAN is a better solution than iSCSI today.

The table below identifies a few simple scenarios and perspectives on what may be effective storage connectivity approaches.

Situation	Solution Considerations
An organization has only a very small number of servers and low I/O loads, but wants to replace installed, aging direct-attach disk storage.	Either NAS, updated direct-attached storage, or iSCSI are likely best. A Fibre Channel SAN may not be justifiable or necessary. Compared to direct-attach and iSCSI, NAS offers better sharing of capacity even if there is no file sharing, and simpler management, but it will likely cost more than DAS. If disk system functions like "snap backup" are of value, that may tip the scale in favor of NAS.
An organization has an existing SAN using a variety of disk systems and wants to do some file sharing.	SANergy preserves the SAN and adds file sharing. Or, a NAS gateway could be placed in front of the disk system. Or, the files to be shared could be moved to a NAS, offloading some SAN traffic that is of value.
A large organization has heavy I/O loads including heavy database activity against a relatively small amount of capacity.	A SAN will likely provide the best performance.
An organization has an existing LAN that has a lot of unused bandwidth.	A NAS offers ease-of-attachment and avoids adding a new network. iSCSI may be attractive if data sharing is not needed or block-I/O is required.
An organization has an existing IBM ESS, IBM Nways® Multiprotocol Switched Services Server (MSS) or an IBM 7133 Serial Disk System and wants to do some file sharing.	Either SANergy or a NAS gateway will work, preserving the existing disk investment.
An organization wants to reduce the high costs of buying dedicated tape drives for backup whenever they buy a new server.	Several software backup products in the industry, such as Tivoli Storage Manager (TSM), can share a pool of tape drives among all clients to be backed up. Or, a single NAS appliance may support direct-attached tape for backup of internal files.
An organization has multiple departments making independent storage decisions.	Either leave things be for political reasons, or evaluate if cross-department storage networking solutions—SANs or NAS or both—might make better global use of resources, lowering Total Cost of Ownership.

Situation	Solution Considerations
An organization has few personnel with storage skills.	NAS and direct-attach will be simpler to manage than SANs. NAS may offer more function and ease-of-management compared to some direct-attach solutions. Managing a NAS appliance may be easier than trying to manage SAN or DAS volume definitions on many different servers. Built-in backup support with automated scheduling can further simplify NAS management.
An organization needs a large amount storage for a temporary project but does not have access to a SAN.	A NAS Gateway allows multiple users to access an existing SAN for available storage, without requiring direct access to the SAN (e.g., without installing Fibre Channel adapters on each host). After the project completes, the storage can be released back to the SAN for use by other users. Snapshot backup functions are also available through the gateway. An alternative would be to add iSCSI or NAS to an existing LAN, and later redeploy its capacity to other projects.
An organization wants to improve its disaster tolerance, and ensure a realtime copy of data is maintained in a remote location.	One solution would be to use disk systems, such as IBM ESS or IBM MSS, that maintain realtime remote copies (mirrors) of local data at a remote site. This offloads the process from the host systems. Alternatively, host-based mirroring is common in many operating systems and would allow the host operating system to write a copy of data in realtime to a disk attached at a distance using Fibre Channel or iSCSI, whether DAS or SAN. Some software products, such as The IBM High Availability GEOgraphic cluster (HAGEO) for AIX® and various third-party offerings, provide remote mirroring over LAN/WAN networks.
An organization wants to use its LAN for disk storage but has applications that use SCSI block-I/O protocols.	iSCSI supports this, allowing I/Os to flow over a LAN without the need to install a SCSI or Fibre Channel Host Bus Adapter in the servers.

Summary.

This paper has explored the exciting area of storage networks. If it has clarified what can be a rather complex subject, then it has been a success.

For more information

Please contact your IBM marketing representative or an IBM Business Partner. For more information about IBM Storage Solutions, visit: storage/ibm.com.

IBM Storage Systems Group
5600 Cottle Road
San Jose, California 95193

Produced in the United States of America
06-01

Please send comments to djsacks@us.ibm.com.

IBM hardware products are manufactured from new parts, or new and used parts. In some cases, the hardware product may not be new and may have been previously installed. Regardless, IBM warranty terms apply.