

# The Storage Area Network

**This chapter discusses:**

- **What is a SAN?**
- **Why are SANs needed?**
- **The signs of a SAN**

*We introduce and define the SAN, establish its value, and describe the underlying technologies.*

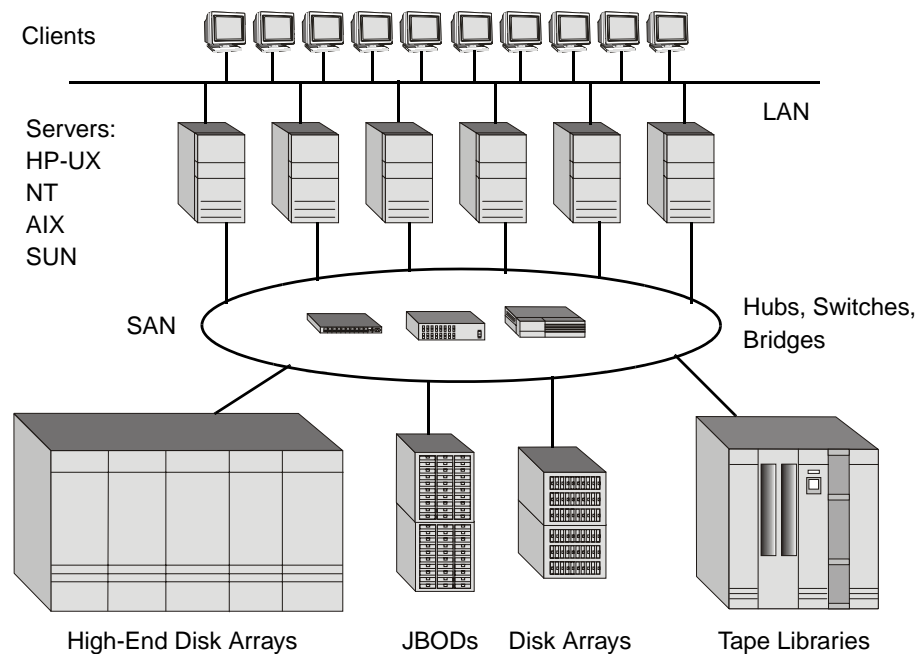
## 1.1 What Is a SAN?

Every book you will ever read about Storage Area Networks will start with this question. There is no one simple answer, but there are multiple simple answers (and they don't conflict with each other).

- A SAN is a mass storage solution, designed to provide enormous amounts of mass storage to an enterprise. It is fast, reliable, and highly scalable.
- According to Transoft Networks of Santa Barbara, California (recently acquired by Hewlett-Packard), the SAN is the next generation high-speed network architecture.
- According to Clariion, in a 1999 presentation by Ken Kutzer, Storage Area Networks Product Marketing Manager, a SAN is a "Network infrastructure that connects computers and devices, transports device commands, and runs value-added software."
- A SAN is identified by special connection architecture, known as Fibre Channel. Fibre Channel is (in the words of Dave Simpson, Editor-in-Chief of *INFOSTOR*) "the key enabling technology behind SANs."
- A SAN is identified by one or more (usually more) servers connected to an infinitely variable number and arrangement of storage devices, by means of Fibre Channel hubs, switches, and bridges.
- A SAN is identified by its components, usually high capacity, highly redundant (and therefore failure-resistant) storage devices.
- A SAN is a topology with three distinct features: 1) storage is not directly connected to network clients, 2) storage is not directly connected to the servers, and 3) the storage devices are interconnected.

Every manufacturer of computer hardware or software intended for Storage Area Networks has similar definitions.

If you read no further in this book, at least take a look at Figure 1-1. It shows a SAN and its common components.



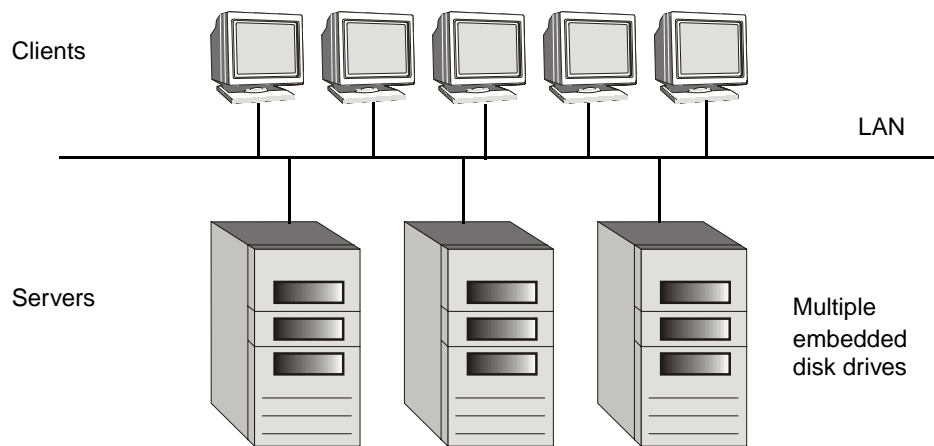
**Figure 1-1** A Storage Area Network

No server is connected to any one storage device, and all storage devices are potentially available to all servers. Connections between devices are made using hubs, switches, and bridges.

The SAN pictured is an ideal SAN, with a variety of storage devices, interconnection devices, and servers. The “loop” in the illustration is intended only to suggest the interconnection of the devices; it isn’t an actual connection scheme.

### 1.1.1 What a SAN Is Not

**A SAN is not embedded storage.** *Embedded storage* is not SAN storage. No matter how many disk drives there are, how large the capacity of the drives, or how many servers are attached to the LAN, it's still not a SAN. The disk drives are resident in the server.



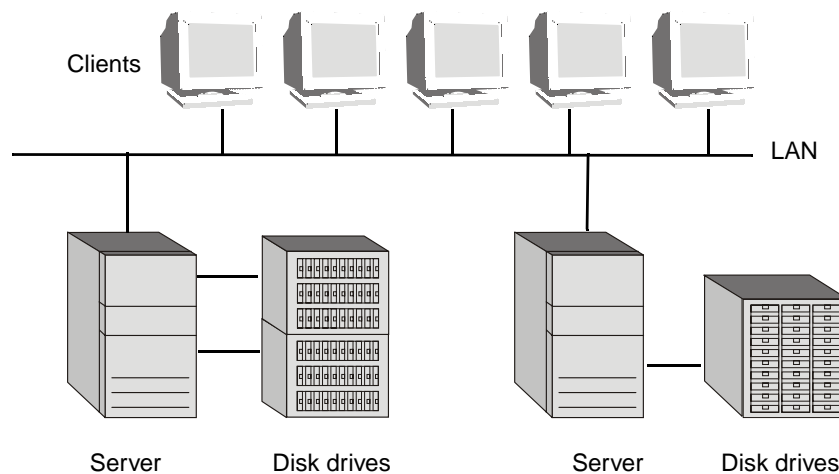
**Figure 1-2** Embedded Storage

In Figure 1-2, the amount of storage is limited by the server's capacity to accommodate it.

For smaller operations, there's nothing wrong with filling the server's drive bay with high-capacity drives, but there is a physical limit. Also, the arrangement amounts to putting all your eggs in one basket, and a server failure would make a great deal of data unavailable.

As you will see later, SANs are scalable. In theory, thousands of devices may be added to a SAN. In practice, SAN scalability is limited by performance issues and the current physical capabilities of hubs and switches.

**A SAN is not directly attached storage.** *Directly attached storage* (as shown in Figure 1-3) is a simple extension of embedded storage, with one or more JBODs (Just a Bunch of Disks) or disk arrays connected by SCSI directly to (and typically only to) one server.

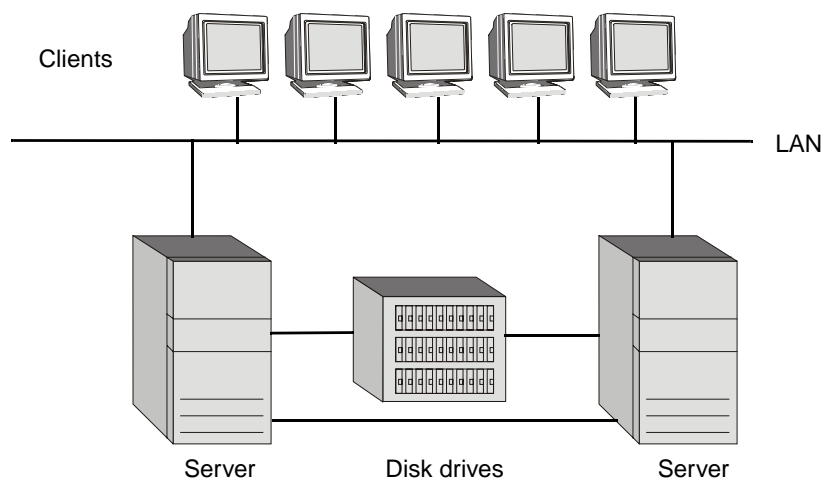


**Figure 1-3** Directly Attached Storage

No matter how many arrays there are, or the capacity of the drives, it's still not a SAN.

The scalability of directly attached SCSI-based storage is ultimately limited by the number of SCSI bus adapters and addresses available to the server.

Sharing SCSI-based storage between servers takes place over the LAN. This produces a performance penalty for the client workstations on the LAN. The clients need speedy server access, but this arrangement, they must share LAN bandwidth with servers requesting data from other servers' disk storage.



**Figure 1-4** Directly Attached Shared Storage

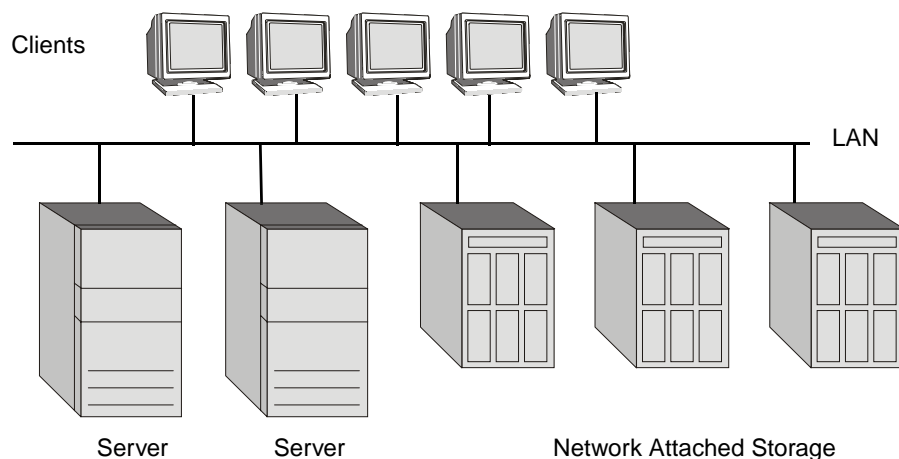
What about sharing storage directly between servers? Figure 1-4 shows a cluster of two servers sharing disk drives without moving disk I/O over the LAN. However, this configuration cannot readily be expanded to include additional servers. The scalability of directly attached storage, shared or nonshared, is limited.

**A SAN is not Network Attached Storage.** Network Attached Storage is highly useful in many business operations. It's amazingly easy to bring additional storage onto the LAN, and some manufacturers advertise it as a three step process—plug in the RJ-45 network cable, plug in the power, and turn the storage device on. That's not too far from the truth. In addition, the products often feature RAID technology for data redundancy and a tape drive for backup.

Well and good, but it's not a SAN. Storage is scalable, and data is highly redundant, but attaching storage devices to a LAN can degrade the performance of all processing. While NAS works fine in small and medium operations, in larger operations the performance math will eventually catch up to it.

Client access to data requires client-server and server-NAS interaction. A client request for a record means that: 1) the client requests the record over the LAN; 2) the server requests the record from NAS over the LAN; 3) the NAS device serves up the record over the LAN; and 4) the server delivers the record to the client over the LAN.

Because both client access and storage access interactions use the LAN, there is a quick buildup in traffic and “traffic penalties” on the LAN. The mathematics will vary depending on the kind of processing, but in general terms each new client adds a small traffic burden (its own I/Os) to the LAN, and each new NAS device adds a large traffic burden (everybody’s I/Os) to the LAN.

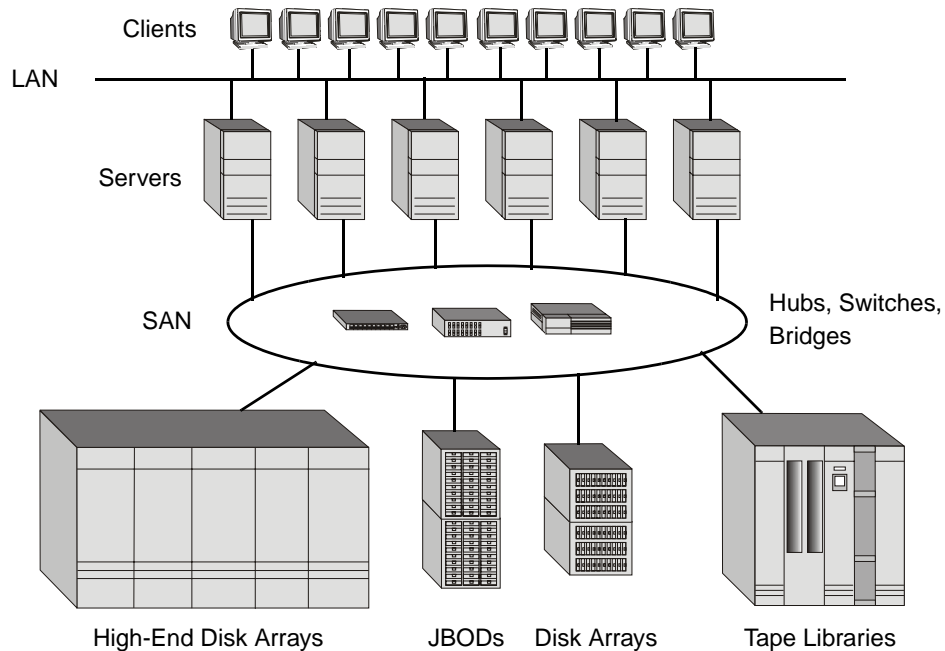


**Figure 1-5** Network Attached Storage (NAS)

Network Attached Storage (NAS) lives “in front of the server” (see Figure 1-5). In his book, *Designing Storage Area Networks*, Tom Clark accurately describes the SAN as being located “behind the server.”

That’s an important contrast. The SAN puts storage where I/Os don’t impact the clients. The significant contribution of NAS is the idea of networking storage devices. However, it takes a SAN to put the storage behind the server.

### 1.1.2 What a SAN Is



**Figure 1-6** Storage Behind the Server

The major distinguishing marks of a SAN are: 1) storage lives behind the servers, not on the LAN, and 2) multiple servers can share multiple storage devices.

In Figure 1-6, the LAN is the traditional connection topology, referred to by Xerox and other early LAN innovators as “bus” topology. However, the storage devices are not on the LAN bus. They have their own independent connection scheme—a Storage Area Network. It is shown as a loop to suggest that the storage devices are connected together, although the devices are not always connected in a loop. Multiple pools of connected devices are possible as well.



The interconnected group of mass storage includes high-end disk arrays, mid-range disk arrays, JBODs, tape libraries, and optical storage devices. The devices are accessible to all servers through hubs, switches and bridges.

There is client-server interaction over the LAN and server-storage interaction over the SAN. Neither group of devices needs to share bandwidth with the other.

Notice that more than one server participates in the SAN storage pool. The number of servers is limited only by the physical capabilities of the connecting devices. The same is true of the number of storage devices that can be attached.

### 1.1.3 SAN Servers

There has been a predictable and understandable pattern in the evolution of client-server configurations. With a single server, you'd begin by running applications and storing data for those applications on embedded disk drives. As storage requirements grew, you'd add attached storage. Eventually, there would be a need for another server, which would likely run different applications and access its own data stored on the second server.

That simple model didn't last long, and for these reasons:

1. The storage requirements of different applications grow at different rates. Despite the best planning efforts, it's hard to avoid server configurations where one server has disk space to burn and another is hurting for space.
2. Comprehensive databases contain a lot of data to be shared. The highly integrated, highly shareable database is one of the Holy Grails of Information Technology. However, because of size and the value of the data to multiple applications, a big database is better being placed on an external storage device than on a server.
3. Servers can fail, so it's not a smart idea to risk data becoming unavailable by placing it on only one server.

The point is that a SAN pools the data and offers relatively easy access to the data by multiple servers. This lessens the dependence on any one server. The possibility of a truly durable, failure-proof information system begins to emerge.

Further, it's very easy to add more storage, and it's immediately accessible by all servers. For that matter, it's very easy to add more servers.

In an enterprise containing servers with a common operating system type, such as HP-UX, connectivity from any server to any storage device attached to the SAN is relatively easy. This is a homogeneous server environment.

But what about mixed servers, from different manufacturers, with different operating systems? With the right equipment and a good design, a SAN can support heterogeneous servers. This can be a combination of HP-UX servers, Windows NT servers, and other open system (UNIX) servers. HP places its emphasis on HP-UX, Windows NT, AIX, Sun Solaris, and Linux.

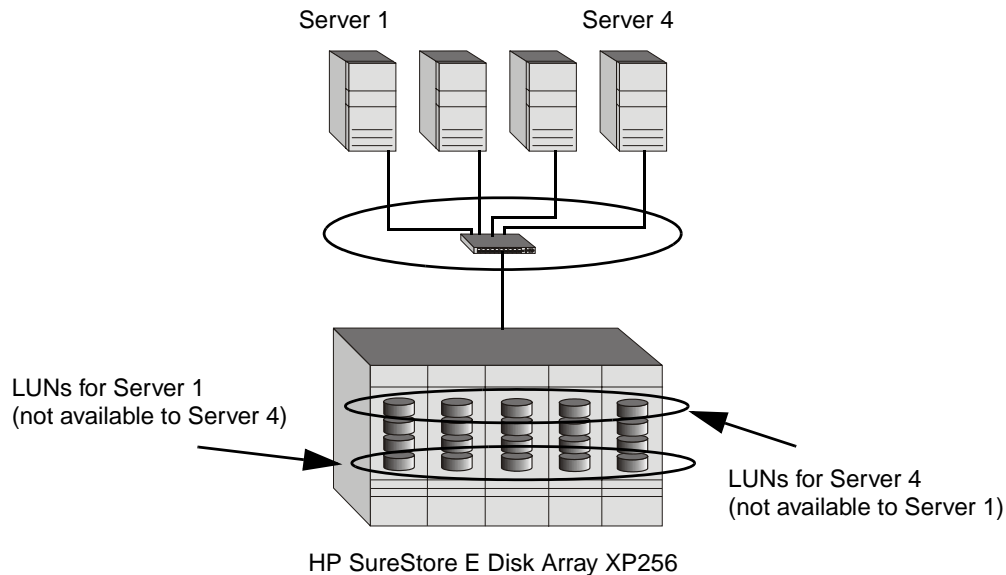
Other manufacturers have the same concerns about heterogeneous server environments. Also, Hewlett-Packard has announced its commitment to an open SAN architecture.

Mixing heterogeneous equipment is a "real-world" problem, because some IT departments have acquired quite a mix of servers. The goal of a heterogeneous-server SAN is to share data between servers running different operating systems.

Data sharing is possible to the extent that different operating systems can understand and use each others' file systems. This promise is not yet fulfilled, since Windows NT, UNIX, and IBM mainframe file systems are intolerant of each other. In time, however, data is likely to be shared with greater ease.

If data can't be shared directly, it can be converted. For example, Hitachi, IBM, and HP have software for "mainframe-to-open" and "open-to-mainframe" data conversions. This software typically operates on disk arrays that can emulate both IBM volumes and open system logical units (LUNs).

However, even if servers don't share (or can't convert) different data types stored on a SAN, there are still equipment cost and management benefits in sharing different disks on the same physical device.



**Figure 1-7** Sharing a Physical Device, But Not Sharing LUNs

With appropriate software (such as HP's Secure LUN feature used with the XP256 disk array) servers can "own" their share of disks (actually, logical units, or LUNs) on the same storage device (Figure 1-7). In fact, quite the opposite from sharing data, LUNs can be "zoned" to prevent interactions with unauthorized servers.

#### 1.1.4 SAN Storage Devices

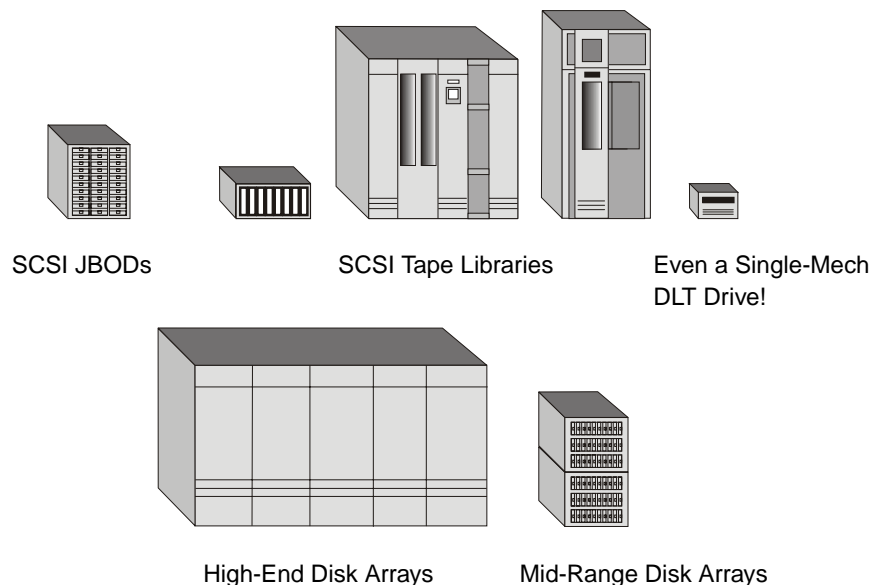
A distinguishing mark of the Storage Area Network is the wealth of storage devices that can be attached to the SAN.

The number of devices that can be connected is limited by the theoretical and practical limits of the hubs, switches, and bridges that interconnect servers with storage devices. As you will see, the theoretical limits are quite large, while the practical limits are a bit tamer.

Any disk device can contain numerous disk mechanisms (“mechs”). The HP FC30 has 30 disk drives, the HP FC60 has 60 drives, and the XP256 has up to 256 drives.

Any SAN-ready storage device can participate in the SAN storage pool with minimum difficulty, because no matter what its purpose (high-availability disk storage, plain old disk storage, near-line optical storage, tape storage), it will be identified by address to the servers that interact with it.

A SAN-ready device is a Fibre Channel device. SCSI devices (primarily legacy devices, but also the most current tape libraries) can participate in the SAN by means of Fibre Channel/SCSI bridges (Figure 1-8).



**Figure 1-8** SAN Storage Devices

**Disk Arrays.** The high-end, high-performance, highly managed disk array typifies the direction in which SAN storage devices are going. The Hewlett-Packard SureStore E Disk Array XP256 is a good example. It stores over 12 terabytes (TB) of data, using a maximum of 256 47 GB disks, and disk

capacity points seem to be rising every day. Also, as this book goes to press, HP has announced the introduction of the XP512 disk array.

Large IT operations are by no means limited to single large disk array. A bank or insurance company data center might run eight of these giant devices on a single SAN.

A high-end disk array is fast and highly redundant, with multiple fans, power supplies, and controllers. It also supports multiple paths to the SAN, eliminating a single point of failure.

The XP256 is “self-healing” to the extent that a number of components can fail, yet the device will keep on functioning. When something goes wrong, the disk array calls the Hewlett-Packard response center to report a need for service.

There are many other disk arrays, so you can find SAN-ready equipment that’s sized exactly right and has sufficient redundancy for your operation.

**SCSI JBODs.** SCSI JBODs can be connected to a SAN through Fibre Channel/SCSI bridges. Even a single-spindle mechanism (“mech”) can be a citizen on the SAN, although this is rarely seen. Non-HA (high availability) devices can be useful components on a SAN.

However, judging from business trends, we would expect the high-capacity, high-availability disk array to dominate disk storage over the next several years.

The business trends are e-business, rapid expansion, global consolidation, and round-the-clock operation. SAN-attached storage devices will answer the needs by serving up mountains of data and providing failure-proof delivery of information to the enterprise.

**Tape Libraries.** Tape libraries of virtually any scale can and should be part of a SAN. Your installation may be a single large tape library (using DLT or one of the new formats we’ll discuss later), or require multiple DLT tape libraries. In fact, in a recent issue of *INFOSTOR*, it was reported that one data center had a need for speedy multiple copies of tape backups. Multiple DLT tape libraries accomplished this, filling the need and producing a new acronym, RAIL. RAIL stands for “Random Array of Independent Libraries.”

A SAN can permit communication directly between storage devices, with minimum server interaction. This means that direct disk-to-disk and disk-to-tape backups are possible. This permits concepts such as “LAN-less” and “server-less” backup.

### 1.1.5 SAN Connection Medium

The connection technology for a SAN is Fibre Channel. As will be shown in greater detail in Chapter 3, Fibre Channel’s distinguishing marks are speed and distance.

**Speed.** Fibre Channel currently moves data at “gigabit speed”; that is 1 Gbps, or 1063 Mbps (some sources cite 1063.5 Mbps or 1064 Mbps). That’s approximately 100 megabytes per second. According to *INFOSTOR*, a number of vendors have demonstrated 2 Gbps Fibre Channel, and 4 Gbps Fibre Channel technology is planned for the future. This evolution clearly puts Ultra SCSI in second place as a speedy medium.

When you see a proliferation of thin orange (typically) fiber cables in a data center, starting to crowd out SCSI cables, it is certain that you’re in a Fibre Channel shop and chances are good that you are in a SAN shop.

What can be a little confusing is that Fibre Channel data transport is supported over both fiber cable and copper. However, copper is seen mainly in intra-cabinet connections between devices, and fiber cable is used far more widely.

Fibre Channel can be used for terminal-server interconnects, running at “quarter speed,” or 266 Mbps, but we’ve not seen any proliferation of Fibre Channel for this purpose. The main purpose of Fibre Channel is to connect mass storage devices and servers over a SAN.

**Distance.** Fibre Channel can connect devices over relatively long distances. For now, take it as a general rule that there may be a distance of up to 500 meters between a device and a hub, and up to 10 km between hubs.

Fibre Channel's ability to span distances makes SCSI, with its 25-meter distance limitation, a noncompetitor when it comes to moving data through a building or across a campus.

**ATM.** Asynchronous Transfer Mode (ATM) connections can appear in a SAN. This connectivity option will move data from site to site over distances of thousands of kilometers.

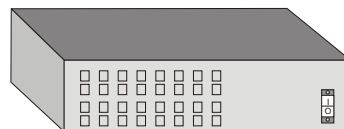
**SCSI.** SCSI connections appear in a SAN when a device does not have Fibre Channel capability. They can be connected to FC/SCSI bridges, and the bridges are connected to the SAN with fiber cable.

### 1.1.6 SAN Interconnection Devices

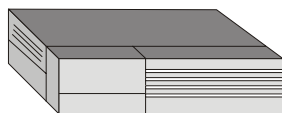
SAN interconnection devices are hubs, switches, bridges, and Fibre Channel host bus adapters (Figure 1-9). You'll learn more about the details of their operation in Chapters 3 and 4.



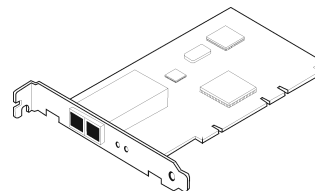
HP Hub S10 or Hub L10



Brocade 2800 or HP Switch F16



HP Bridge FC 4/2



A5158A FC Host Bus Adapter

**Figure 1-9** SAN Interconnect Devices

**Hubs.** Fibre Channel Arbitrated Loop (FC-AL) hubs are widely used to form the SAN, and there are cascading options to increase distances between devices and the number of ports available for connecting devices.

Although in theory up to 126 devices can participate in an FC-AL loop, as a practical matter the number of devices deployed on a hub will be limited. One limitation is the number of ports (10 on a typical hub, 18 when two hubs are cascaded). Another limitation is a decline in performance when too many devices contend for bandwidth in a loop (which is a shared, polling environment).

**Switches.** Fabric switches are gaining prominence, and are now replacing hubs in some implementations.

In theory, a switch allows over 16 million simultaneous connections (based on address) to the SAN. However, a typical switch will have fewer than 16 million ports. The Brocade 2800 fabric switch and the HP Switch F16 each have 16 ports, and there are cascading options.

At this time, even though a switch typically costs about four times as much as a hub, the performance of switches makes them the better interconnection device.

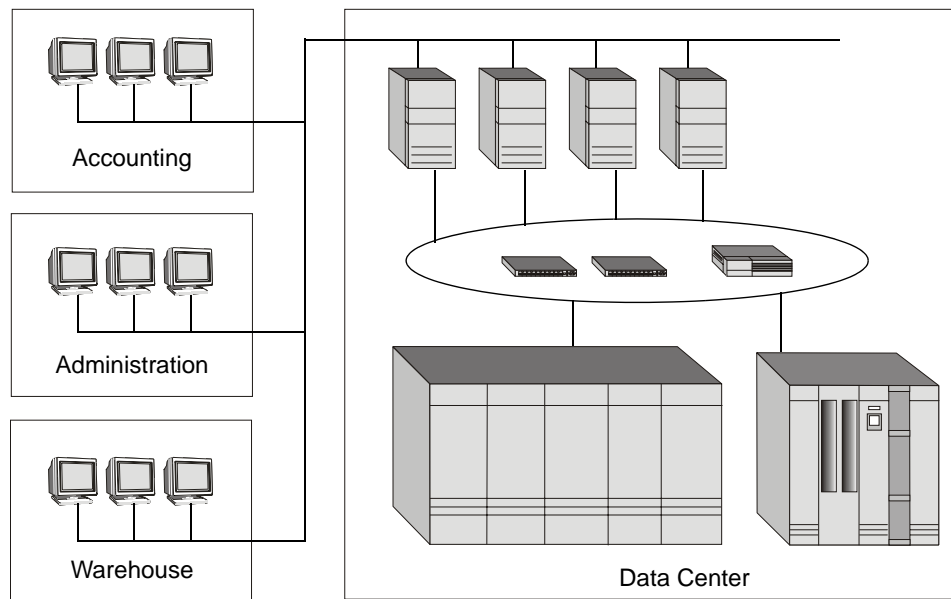
**Bridges.** The Fibre Channel bridge (for example the HP SureStore E Bridge FC 4/2) is essential for bringing SCSI equipment onto the SAN. The bridge is used to connect legacy or specialized SCSI equipment to the SAN. In addition, since virtually no tape libraries offered at this time are SAN-ready, the use of bridges on the SAN is essential for tape backup.

**Host Bus Adapters.** Fibre Channel host bus adapters live in the servers, and provide the connection to the SAN's hubs and switches. They come with HSC or PCI interfaces, single or double ports, and replaceable Gigabit Link Modules (GLMs) or Gigabit Interface Controllers (GBICs). Fibre Channel HBAs are usually based on Agilent Technologies' Tachyon or TachLite chip.



### 1.1.7 SAN Distances

SAN components may be located close to or relatively far away from servers. Typically, a SAN would have most or all of its storage devices in one room or in separate rooms on one floor of a building. The LAN connects workstations in different departments to the servers.

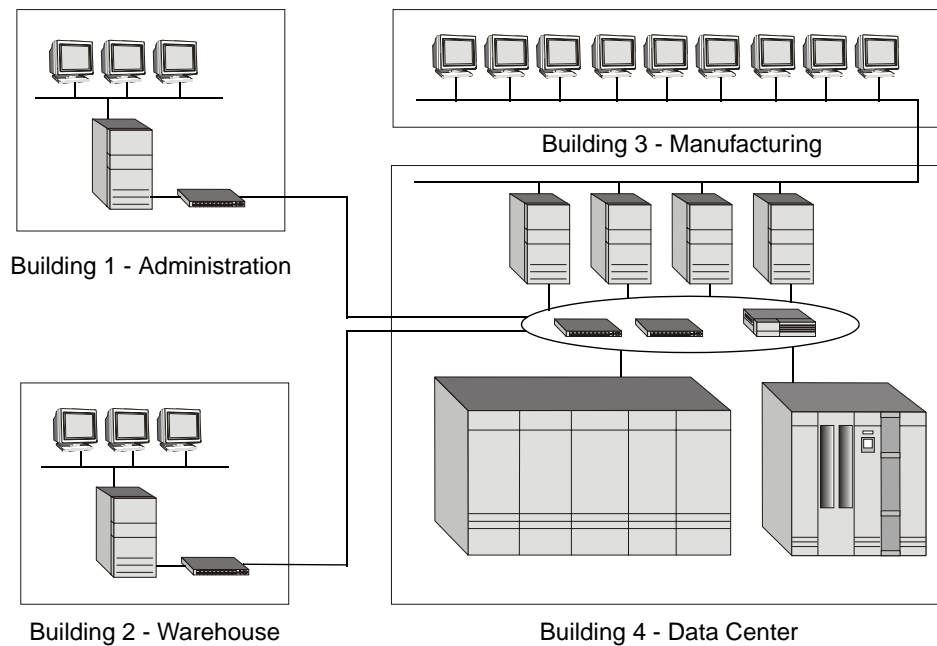


**Figure 1-10** An In-building SAN

The arrangement shown in Figure 1-10 is the same as most non-SAN data centers, and exhibits the same advantages: security, ease of management, and straightforward migration from SCSI to Fibre Channel.

Tape storage may be in a different room than disk storage or servers. The data center may be on a different floor of the building from other departments. Using even the relatively short distances available with short wave hubs (500 meters), there are many data center configuration options.

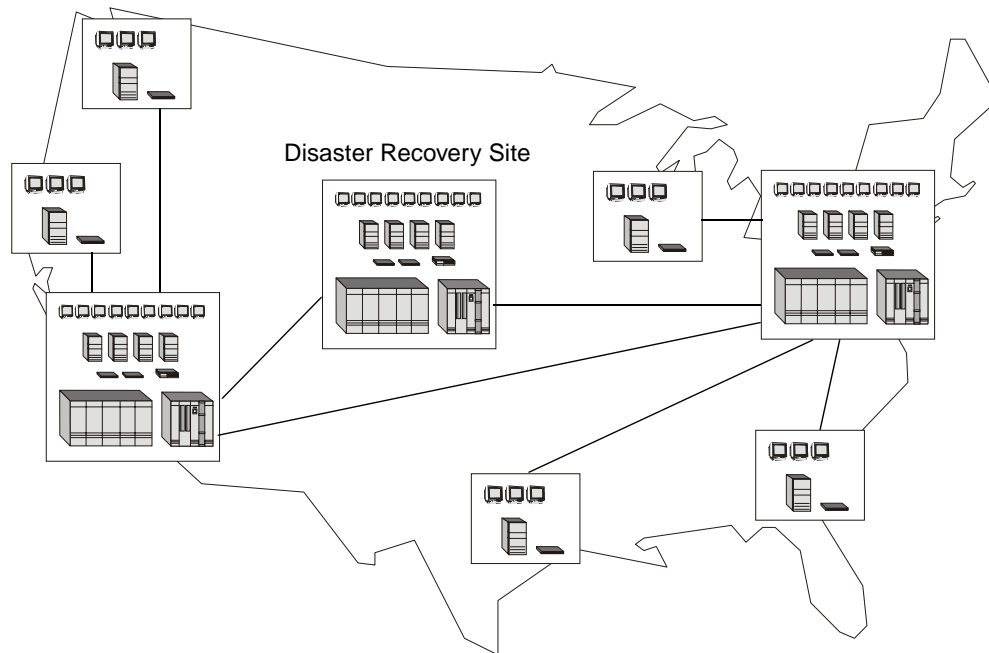
In some enterprises, “local” means “on the same campus”; in others, it means across town. Perhaps the servers are located in different buildings, but they are connected to same SAN (Figure 1-11).



**Figure 1-11** An On-campus or Crosstown SAN

Fibre Channel longwave hubs readily provide for connection distances of up to 10 km, so distance should not pose a problem. Many large corporations have campuses resembling small towns, and have worked through the engineering challenges of providing telephone and LAN connections between buildings. For crosstown connections, there are a number of leased-line options available.

The cross-country SAN (Figure 1-12) is accomplished by using additional Wide Area Network (WAN) hardware. Leased lines range in speed from up to 1.544 Mbps for T-1 to up to 622 Mbps for OC-12.



**Figure 1-12** A Cross-country SAN

The cross-country SAN is not only possible, but sometimes essential. For example, a company in earthquake-prone Los Angeles may find it prudent to mirror its data in Arizona or Nevada.

In addition, there are cost saving advantages in centralizing data in regional or national data centers. It takes fewer administrators to manage large amounts of data, maintain equipment, expand capabilities, and protect data.

Companies with offices worldwide can follow the cross-country model. They will need to design SANs to store data that must be both centralized and distributed.

## 1.2 Why Are SANs Needed?

Virtually every sector of commerce has an immense requirement for capacity, speed, and reliability in storing data. Government and education, although restrained by budget considerations, have the same requirement. And every enterprise is concerned with managing storage costs.

**Capacity.** In traditional business applications, simple growth of customer and manufacturing data drives the need for capacity. Business consolidations and worldwide commerce also drive the need. More data means more disk drives, greater capacity drives, and more densely populated enclosures to hold them.

New approaches to running traditional businesses drive a demand for storage. For example, data warehousing and data mining are methodologies that are producing gains for the businesses that use them. These are data-hungry constructs. In fact, at least one authority suggests that the warehouse and the data mine best serve the business as separate entities. That could amount to an instant doubling of storage requirements, and data warehouses typically aren't small.

In newer businesses, applications seem to need intense capacity and speed from their inception. Judging by the financial news, almost any e-commerce activity will ramp up from conception to reality in just months (possibly weeks), and some have become instant successes. Online retail sales, online business-to-business sales, and online auctions can attract millions of customers practically overnight. E-commerce is dramatic in its hunger for capacity, speed, and reliability in storage.

Video production, video-on-demand, and other imaging applications have storage requirements that make buying a book online look like child's play. It's fair to say that much wanted or much needed applications will not come into existence without the SAN.

Conventional storage methods can no longer satisfy the capacity requirements of mass storage. The scale of computing and the complexity of applications drive the demand for Storage Area Networks.

**Reliability.** In all business sectors, the cost of downtime is larger than ever. In a presentation by Computer Network Technology at HP World '99, the costs of downtime were cited (shown in Table 1-1). The source is the Fibre Channel Industry Association.

**Table 1-1** The Cost of Downtime

Business	Hourly Downtime Cost
Brokerage	\$6,450,000
Credit card sales authorization	\$2,600,000
Pay per view	\$150,000
Home shopping	\$113,000
Catalog sales	\$90,000
Airline reservations	\$90,000
Teleticket sales	\$69,000
Package shipping	\$28,000
ATMs	\$14,500

A SAN is Information Technology's best answer to avoiding downtime catastrophes, whether the "catastrophe" amounts to an hour less of selling books, or a week of nonoperation due to a major hurricane. A SAN lends itself to disaster recovery scenarios better than previous storage strategies.

**Costs.** In all business sectors, reduction of IT operating costs is all too often a dominant factor in driving IT choices. Business is always interested in a lower price per gigabyte for mass storage, and the hardware manufacturers are accommodating them. Business is also interested in lowering the labor costs for maintenance, management, and expansion, which can be intensive. Fortunately for IT management, a Storage Area Network can be maintained, managed, and expanded with relative ease by fewer people than required by non-SAN sites. There will still be \$90,000-a-year system administrators, but there will be fewer of them at a single site.

## 1.3 The Signs of a SAN

In summary, a SAN is a combination of hardware devices, interconnect strategy, and concept. Even as the capabilities of the hardware and speed of the interconnect change, the concept will remain stable.

A collection of servers, mass storage, interconnect devices, and interconnection media constitutes a Storage Area Network, if it exhibits most of the following characteristics:

- Storage is behind the server
- Storage devices are connected to each other
- Multiple servers are connected to the storage pool
- Heterogeneous servers may be connected to the storage pool
- Fibre Channel connectivity (fiber optic cable and FC host bus adapters) is used
- Fibre Channel hubs and switches are present
- Multiple paths to devices exist