2 The Value and Position of iSCSI

To the Reader

This chapter will take you through the different environments in which network-connected storage may be appropriate. Because market planners and engineers may have different views on the potential market for iSCSI products, we will discuss what the market looks like, in hopes of bridging the divergent views.

A taxonomy of the various environments is listed, and each area's potential relative to iSCSI is explained. This entire chapter is therefore recommended reading for marketers and hardware and software engineers.

Small installations are called **SoHo** (Small office, Home office) environments. In such environments customers will have one or more desktop systems, which they are tired of constantly opening up to install storage devices. They want easy interconnects that permit them to operate with external storage as fast as they can operate with internal storage.

The Home Office

Generally the home office will have computers that are connected on small locally attached Ethernet 100Mb/s links. These systems typically have more processing power than they have storage access capability. These installations will find value in placing their storage in a central location, dynamically adding it to their personal computer systems with a simple plug-and-play configuration, thereby obtaining additional storage without having to open up their systems.

The home office environment can use standard, low-cost, off-the-shelf networking components. The switches are readily available, and the desktop and laptop processors usually have more processing power than is needed for

the types of work they do. These home office systems usually come with Ethernet connections and do not need any special adapter cards. The only things they need are iSCSI software drivers and the low-cost IP storage devices that iSCSI will permit.

For the home office, vendors are bringing to market a simple disk controller that can attach from one to four (or more) low-cost ATA desktop-class drives. These controllers and drives will be purchased at local computer superstores for very low prices. Moreover, the customer will be able to buy one or two drives initially with the basic controller and then add drives whenever they wish.

Customers can now purchase 10/100/1000 Ethernet cards that can operate over the inexpensive Cat. 5 Ethernet cable already installed for their existing 10/100Mb/s network.

Prices for 10/100/1000 Ethernet NICs and switches are dropping rapidly. In early 2002, 10/100/1000 Ethernet NICs cost $60. The then current 10/100 Ethernet NICs cost only $30, down from $60 just a year previously. It seems reasonable to assume that the 10/100/1000 Ethernet NIC will soon be the default in most desktop systems, which means that the home office will have gigabit capability and processors fast enough to at least utilize 300 Mb/s. This will give desktop home office systems as much storage access as they can use.

Usually home offices obtain all their software either from the OS that came with the unit or from a local computer store. There is almost no software in this environment that knows how to work with shared files, and it is very rare to see a file server in this environment.

Home offices use peer-to-peer file-sharing functions that come with the OS to permit one user to operate on a file created by another (serial sharing). However, the storage is considered to belong to the system to which it is attached. As a rule, when one home system runs out of space, owners do not use the space on another system but instead upgrade the storage on each system independently as needed.

iSCSI will permit home office users to set up their own external storage pool connected via a LAN. The owner will then assign new logical or physical hard disk drives (**HDD**s) to each host system without needing to open or replace them. (See Figure 2–1.)

We have been talking as if the home office had more than one host system. This is because owners of home offices tend to keep their old systems, which have old data and applications that still work. Also, the home/family

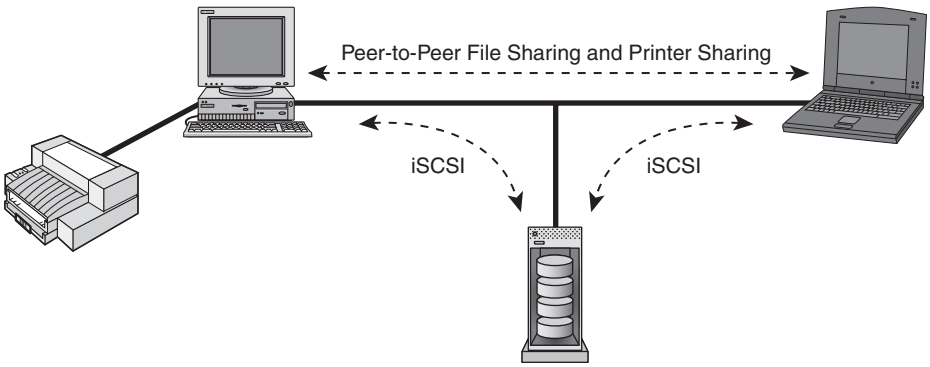

Figure 2–1 The home office and iSCSI.

use dynamic—the multiple computer family—is often at work. Many times there are two, three, or more computers in the same family—one for each adult and one for the children—but only one person is responsible for maintaining them all. In these environments a shared pool of storage is valuable for ease of both access and administration.

Even though multiple-system households are common, it is also true that home offices may have only a single desktop or laptop system. But even in single-system environments, the owner will want to have storage attached externally to avoid the problems of opening the system and adding new HDDs. Unlike other environments, these systems are often updated to add processing power not only for business but also for game playing. Whatever the reason, they will find iSCSI very useful in avoiding having to transfer all their key files from old systems to new systems, since the storage is all external. (And yes, iSCSI has boot capability.)

It has been pointed out that even in a home office, doctors, lawyers, and others would love to have their data placed on small, inexpensive **RAIDs** (redundant arrays of independent disks) because of their reliability. And iSCSI provides even this "upscale" home office the most appropriate, flexible, and inexpensive interconnect for that need.

The Home Office and Serial ATA Drives

At least one individual has made the claim that iSCSI's real competition is the new serial ATA (**S-ATA**)—a cabling protocol that travels from the controller chips on the motherboard directly to the HDD. This is not a true consideration today, since the S-ATA is currently 1 meter in length and has no sharing

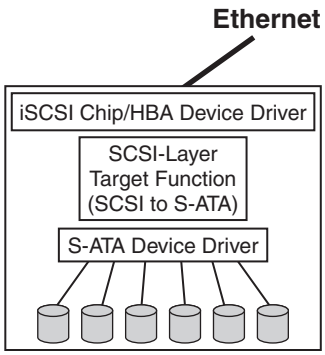

Figure 2–2 Attached serial ATA storage controllers.

capability. Therefore, it is unlikely that it will be used for a cabling interface that hangs out of a home PC for a general storage interconnect. There is a proposal for increasing the length, but this is for rack mount versions and it is not clear if that will ever affect the desktop or home market. Further, the S-ATA specification has not yet defined a technique for permitting the same storage device to be attached to more than one system at household distances. Within the iSCSI target, however, one may find iSCSI coming into a small box and then interconnecting to the ATA disks with S-ATA cables. (See Figure 2–2.)

The Small Office

Examples of the small office environment are dentists, accountants, construction, real estate, and the like—that is, any office that has a relatively small number of employees working in close proximity and a continuing need to grow its storage. Individual desktop systems in small offices have the same disruptive effect when adding storage as they have in home offices, but they are often connected to a file server. However, these file servers face the same problems that individual office or home systems face when adding additional storage. Their owners do not want to disassemble them to add storage and may instead replace the system altogether.

The real problem starts when the user has to migrate the data from an old file server to a new unit, a difficult and very disruptive process. There are technical solutions, of course. One of them is to buy another file server or a network attached storage (NAS) appliance. Thanks to iSCSI there is another, generally less costly approach—pooled storage. Pooled storage may

consist of simple iSCSI **JBOD**s (just a bunch of disks) or RAID controllers, all connected to the same network via iSCSI. With iSCSI, users can also begin small and add storage as needed, placing it wherever they have room and a network connection. Regardless of how many units they add, all of them are logically pooled and yet any of the individual systems in the office can have its own private storage portion.

With iSCSI all the major storage placement decisions are performed by the various host systems as if the storage were directly connected to them. Because of this fact, iSCSI is fairly simple compared to NAS and this results in low processing requirements in the iSCSI storage controllers. It is therefore expected that iSCSI JBODs and RAID controllers will be significantly less costly, and support more systems, than the same storage in a NAS appliance. Using iSCSI units in this way is *not* always the right answer; however, it does give the customer another option that may meet their needs for flexibility and price. In other words, for the same processor power an iSCSI appliance can support more clients and more storage than is possible with a NAS appliance.

A performance analysis carried out by IBM compared a NAS (NFS) server to an iSCSI target (with the same basic power and equipment). This comparison showed that the iSCSI target used much less processing power than did the NFS server. (Refer to the discussion of measurements in Chapter 3.)

Small offices are similar to home offices, except that they have more links and switches. In most installations, a switch can be used to attach either a new NAS appliance or an iSCSI storage controller. (See Figure 2–3.)

Like home office systems, small office systems are bound together with 100Mb/s links. As a rule, it will be possible for them to upgrade to 1000Mb/s links whenever their 100Mb/s links become congested. They can usually ease into iSCSI storage controllers with 100Mb/s links, but over time these links will be upgraded to 1000Mb/s. They will operate on the same (Cat. 5) Ethernet cable and provide all the bandwidth needed to support any normal demand for iSCSI access, NAS access, normal interactive traffic, and Web traffic.

In a small office that has NAS, iSCSI storage can be added to the existing network with no more network congestion than would exist if its current file server were updated or if a new NAS server unit were added.

A question that needs to be asked is when it is appropriate to add a NAS unit and when it is appropriate to attach an iSCSI storage controller. The answer depends on the customer application. Some small offices, such as

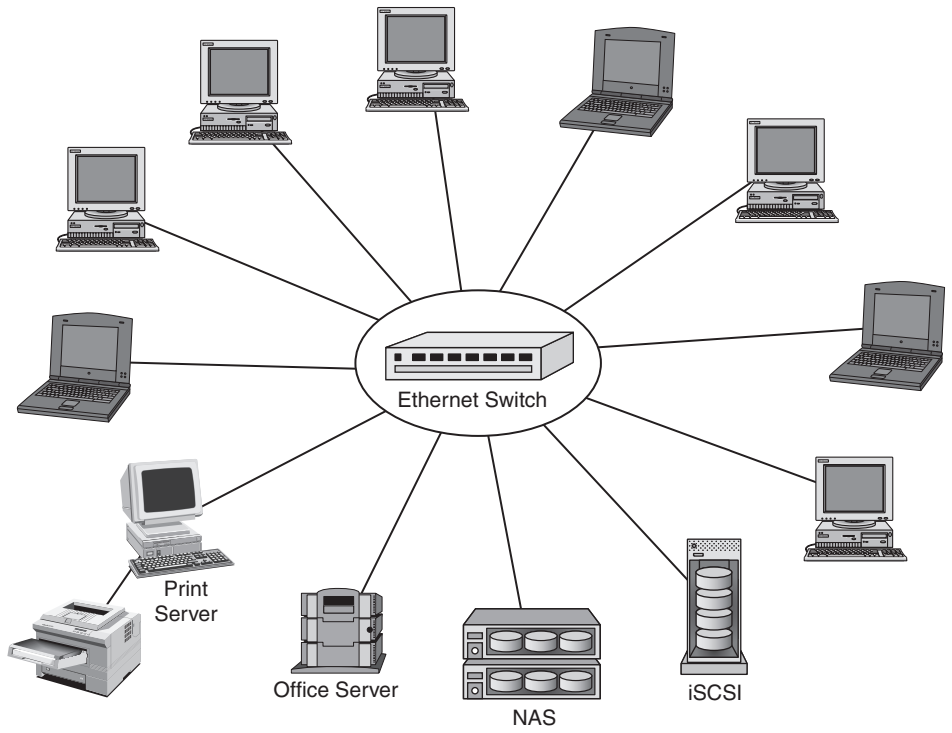

Figure 2–3 Small-office interconnect.

dentists and accountants, use specialty software, acquired from a salesperson who supplies them with other tools of their trade. It might be an appointment management system or a billing system, or sometimes a trade/business-specific system. Some are database applications; others are just file based.

The first thing to determine is whether the software is a database application or a file system application. If a database application, the normal choice for additional external storage is iSCSI. This is true whether or not the database uses a local file system or a "raw" partition (without file system formatting). In either case the most efficient process is to let the database or file system allocate its own data blocks and then use iSCSI to access them. Oracle and IBM DB2, for example, can operate efficiently with a raw partition, but they also have been set up to interface efficiently with many local file systems. The use of iSCSI exploits that fact. As indicated above, the iSCSI storage controller can handle more loads, at higher performance, than a NAS system with similar hardware can.

The clear trend in the industry is to tie applications to databases. When this is done, for the most part the question of sharing files becomes academic since almost all major databases* use "Shared Nothing." In this model the storage is not shared with other systems. This is true even in larger "clustered" installations, where the database query is "parsed" and divided up among the different database systems in the cluster that "own" the data. In this way the more database servers an installation has, the less it needs to share files and the more useful iSCSI becomes.

If the application uses a file system instead of a database, we need to ask if the application is capable of sharing data with other users in either serial or parallel manner. If so, then we need to ask if the user actually does any file sharing. The answers to these questions will generally indicate that 90% or more of the data at the site is not shared. In fact, the answer with the highest probability in a small office is that *no* file sharing occurs.

Normally the only data sharing in a business environment is via a database. The main exception to this is in an engineering environment. For example, physical design engineers—such as those working at a company like Boeing—may share design files that need to be integrated from time to time so that full simulations can be done. The wing designer might need to share his design file with the engine mount designer, and both might need to share files with the engine designer.

We can find a similar situation with chip engineers, who need to share macro files and their own design files with other designers and the simulation system. Software engineers are still another example. They share their development libraries, their resultant components, and even their documentation library.

The engineering environment is the primary environment where files are shared. Much of the rest of the world shares by sending things around in e-mails. In general, they usually do not have a file sharing need.

When incidental serial file sharing occurs, the small business office, like the home office, is more apt to use peer-to-peer sharing and not put in a NAS server. However, this probability varies with the size of the office; the larger it is, the higher the possibility that it will devote a system to file serving or

*Oracle offers a special version of its database that uses shared disks (LUs) in clustered systems. This is called Oracle Real Application Clusters, but it is usually found on high-end enterprise database servers and is not normally used with NAS servers.

install a NAS appliance. Even when the file sharing is not very important, it often finds use for the file server/NAS as a backup device.

Once file-sharing requirements are understood, we can determine if a NAS unit or an iSCSI storage controller best meets the need. If we have a large file-sharing requirement, NAS is usually the best fit. If not, the best fit is usually iSCSI. Then, in the rare case when file sharing is needed, the peer-to-peer type is appropriate.

If the installation has a lot of sharing and nonsharing requirements, we determine the minimum number of NAS systems needed for file sharing and the number of iSCSI systems needed for nonfile sharing, buying the minimum of both. If all workload can fit into a single inexpensive NAS unit, often this is the best approach. In general, it is best for an installation to fit the entire file-sharing workload into the fewest number of NAS units that can support it. Then it can either try to fit some amount of the nonfile sharing storage access into the unfilled NAS capacity or leave it for subsequent NAS growth. The rest of the storage access load, which will probably be a large majority of the total storage requirement, should exploit the efficiencies of iSCSI storage controllers. This approach is usually much more cost effective than pushing all workload into NAS units.

One other solution is the single unit that supports both NAS and iSCSI functions, which I call a "dual dialect" storage server. With this type of system the installation does not need much precision in its NAS/iSCSI planning. As an example, let's suppose the installation only needs 10% of its clients, or applications, to have file sharing. In this case, one dual dialect storage server can devote 70% of its processing capability to that 10% and use the remaining 30% for the 90% that have no file-sharing requirements. (See Figure 2–4.)

Of all environments, a small office environment is the least capable of anticipating and designing the storage layout that can optimally support its workload. I therefore expect dual dialect systems to be very successful in small office environments that include both engineers and nonengineers. Environments without heavy engineering design work will generally find iSCSI storage controllers to be the most appropriate equipment for providing IP-based storage.

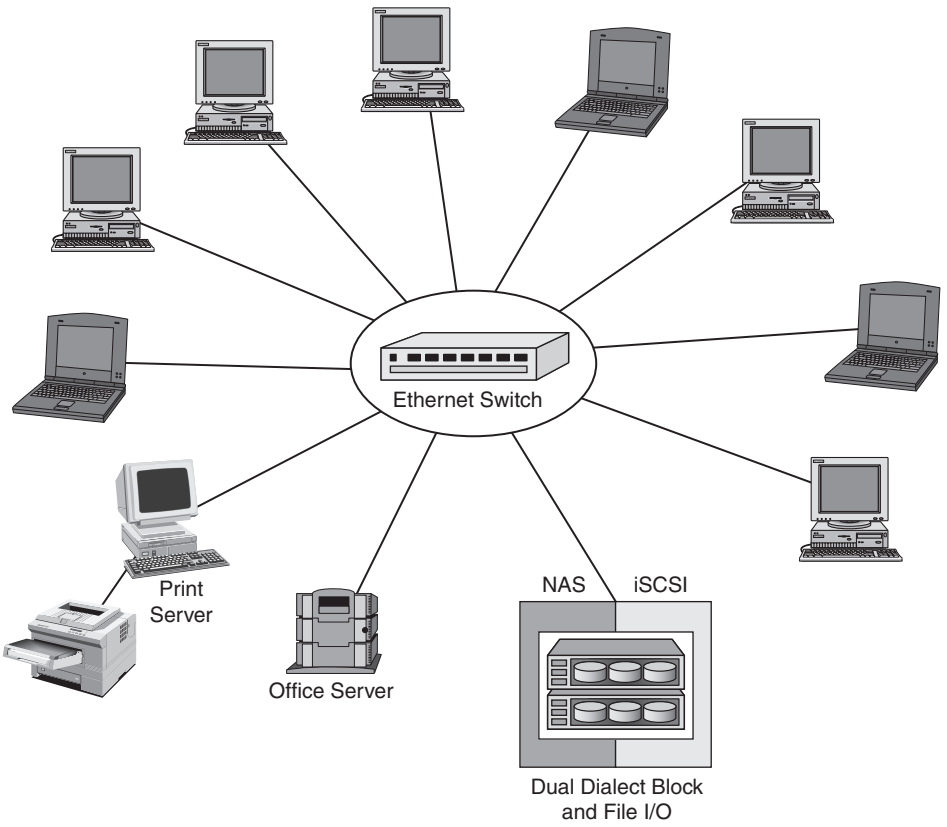

Figure 2–4 Storage combo—NAS and iSCSI.

The Midrange

Moving up to the midrange company environment, we find multiple server systems. These are unlike desktop or laptop systems, which usually have more processing power than they can use. Servers are heavily loaded performance-critical systems that consume all the CPU cycles they have and often want more. They need access to storage with the smallest amount of lost CPU cycles possible. In an FC, or a direct-attach environment, these systems expend approximately 5% processing overhead to read and write data from and to a storage device. If iSCSI is to be competitive in the server environment, it needs a similar overhead profile. This requires that the processor overhead associated with TCP/IP processing be offloaded onto a chip or host bus adapter (HBA).

Offloading requires a TCP/IP offload engine (TOE), which can be completely incorporated in an HBA. All key TOE functions can be integrated on a single chip or placed on the HBA via the "**pile-on**" approach. The pile-on approach places a normal processor and many discrete components on the HBA (along with appropriate memory and other support chips) and includes normal TCP/IP software stacks. The integrated chip and the pile-on technique both permit the host processor to obtain iSCSI storage access without suffering the overhead associated with host-based TCP/IP processing.

We will be seeing a number of HBAs of all types from a collection of vendors. These will include not only a TOE but also in many cases full iSCSI offload. We will also see a pile-on HBA that can support 70% to 100% of the line speed while operating at close to 100% CPU utilization on the HBA. The customer of the server processor will care not how hard the HBA is working, but only that it can keep up with line speed and offload the iSCSI overhead (including TCP/IP) from the host processor.

The pile-on HBA approach will have higher latency than an HBA that has TCP/IP and iSCSI processing integrated onto a single chip. Even if the pile-on HBA can operate at line speed (1 Gb/s), the latency caused by this type of adapter is unlikely to permit its ongoing success in the market. That is because HBAs with full iSCSI and TOE chips will permit not only operation at line speed but also very low latency. We should consider the pile-on approach to be a time-to-market product that vendors will replace over time with faster and cheaper HBAs using iSCSI and TOE integrated chips.

The goal of iSCSI HBAs is to keep latency as close to that of Fibre Channel as possible (and it should be close when using integrated chips) while keeping costs significantly under those of Fibre Channel.

Some people have argued the price issue, saying that Fibre Channel can easily lower its prices to match iSCSI's because an FC chip will have less silicon than an iSCSI TOE chip. This is of course an important consideration, but sales volume is the key, and iSCSI has the potential for high volume with a technology that operates in an Ethernet environment. This includes operating at gigabit speeds with normal Cat. 5 Ethernet cable attachments so that the customer doesn't have to install and manage a new cable type.

As stated previously, I do not believe that FC vendors will give up their high margins in the high-end market in order to fight iSCSI in the low-end and midrange markets. This will only occur when iSCSI is considered a threat in the high end, but by then iSCSI will have large volumes in the rest of the market and will be able to push the price envelope against Fibre Channel. Also

remember that TCP/IP (and Ethernet) connections will always be needed on these systems anyway. Therefore, since FC is always a "total cost adder," whereas iSCSI will have much of its cost supported by other host requirements for IP interconnect, price advantage will clearly go to iSCSI.

There has been talk that FC vendors will attempt to move their 1Gb offerings into the midrange while keeping their 2Gb offerings at the high end. However, the total cost of ownership (TCO) to the midrange customer will still be higher than iSCSI because of the shortage of FC-trained personnel, the use of new special cables, and, as mentioned above, the fact that Fibre Channel is always a total cost adder.

The goal is for the midrange environment to be able to obtain iSCSI-block I/O pooled storage, with performance as good as that of Fibre Channel but at lower cost. However, the midrange customer will still face the dilemma of iSCSI versus NAS. The same consideration and planning should be done in this environment as in the small office environment. The only difference is in the capabilities and price of the competing offerings.

In addition to the normal NAS and iSCSI offerings in this environment, there will be dual dialect offerings also. The difference is that the iSCSI-offload HBAs and chips can be employed to reduce the iSCSI host overhead to a point where they are competitive with Fibre Channel and direct-attached storage. This is not currently possible with NAS.

The other consideration in midrange company environments is that they have desktops and laptops that feed the server systems, which will also, from time to time, need additional storage. Their users will want to get the additional storage and have it managed along with the server storage. This is similar to the needs of SoHo environments: Instead of spending time upgrading internal disk storage, they want to get their additional storage via the network they are already plugged into.

With the new copper 1000Mb/s Ethernet adapters, users can have both a high-speed interactive network and a high-speed storage network, all without changing the Cat. 5 Ethernet cable already installed throughout their company. iSCSI storage controllers can supply the needs of both servers and client desktops and laptops.

Still, the argument is often made that a NAS solution can address the needs of desktops and laptops. This is true, but at a higher cost. As pointed out earlier, in the small office environment many applications are being written to use databases. They generally use a "shared nothing" approach and therefore provide an information-sharing environment in which NAS is not required. Again, if files need to be shared, NAS is appropriate; otherwise,

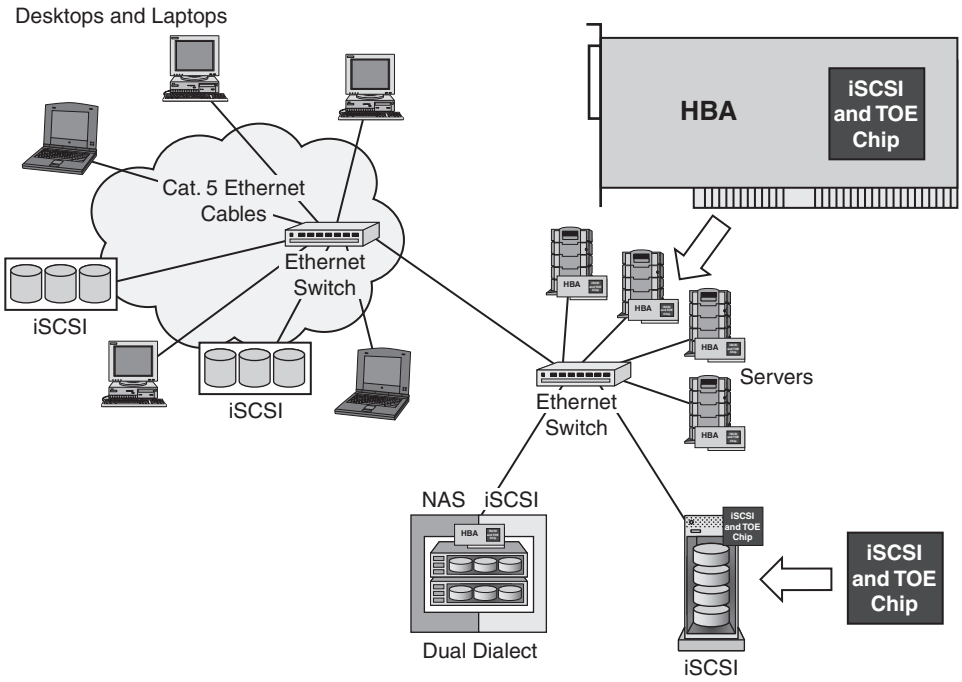

Figure 2–5 The midrange environment.

a block I/O interface best meets the requirements. iSCSI is the most cost-effective approach for non-shared pooled storage.

Many of these midrange companies will be building iSANs. These are logically the same as FC SANs but are made up of the less expensive iSCSI equipment—less expensive because the entire Ethernet and IP equipment market is relatively low priced (at least low priced when compared to Fibre Channel). Even iSCSI HBAs are cheaper than current FC components. Intel, for example, has declared that its HBA will be available at a street price of under $500. It is further expected that iSCSI HBAs and chips will have even lower prices as sales volumes go up.

One significant difference between the midrange and small office computing environments is that the I/O requirements of the various servers can be as demanding as that found in many high-end servers. Therefore, in the midrange one tends to see more use of iSCSI HBAs and chips in various servers and storage controllers, and a smaller reliance on software versions of iSCSI. (See Figure 2–5.)

The High End

High-end environments will have the same processor offload and performance requirements that midrange environments have. However, they will probably be more sensitive to latency, so it is expected that the pile-on type of HBA will not be very popular. Because of the never-ending throughput demand from high-end servers, it is in this environment that HBAs with multiple 1Gb Ethernet connections and 10Gb implementations will eventually find their most fertile ground.

Another important distinction is that the high-end environment will probably have some amount of FC equipment already installed. This means that the cohabitation of iSCSI and Fibre Channel will be required.

Because of the usefulness and flexibility of iSCSI-connected storage, and because high-end servers are probably already connected with Fibre Channel, it is expected that high-end environments will first deploy iSCSI in their campus, satellite, and "at-distance" installations. (See Figure 2–6.) The following sections will break out each of these subconfigurations and then address the central facility.

The Campus

The campus is the area adjacent to the central computing site—within a few kilometers—where private LANs interconnect the buildings containing local department servers as well as the desktops and laptops that are also spread throughout. The different department areas are analogous to the midrange and small office environments. Their general difference is that, with the use of iSCSI, they can exploit the general campus IP backbone to access the data, which may be located at the central computing location.

Often these department areas have policy or political differences with the organization that runs the central computing complex, and so they want their own independent server collections. Generally they want the flexibility that a storage area network (SAN) can provide (such as device pooling and failover capability), but they do not want to get into the business of managing an FC network.

In spite of their independence, these departments want access to the tape libraries at the central location. They want access to these robust backup devices, which they consider essential but which they do not want to service, manage, or maintain. (See Figure 2–7.) The departments also want to access disk storage at the central location as long as they do not have to abide by

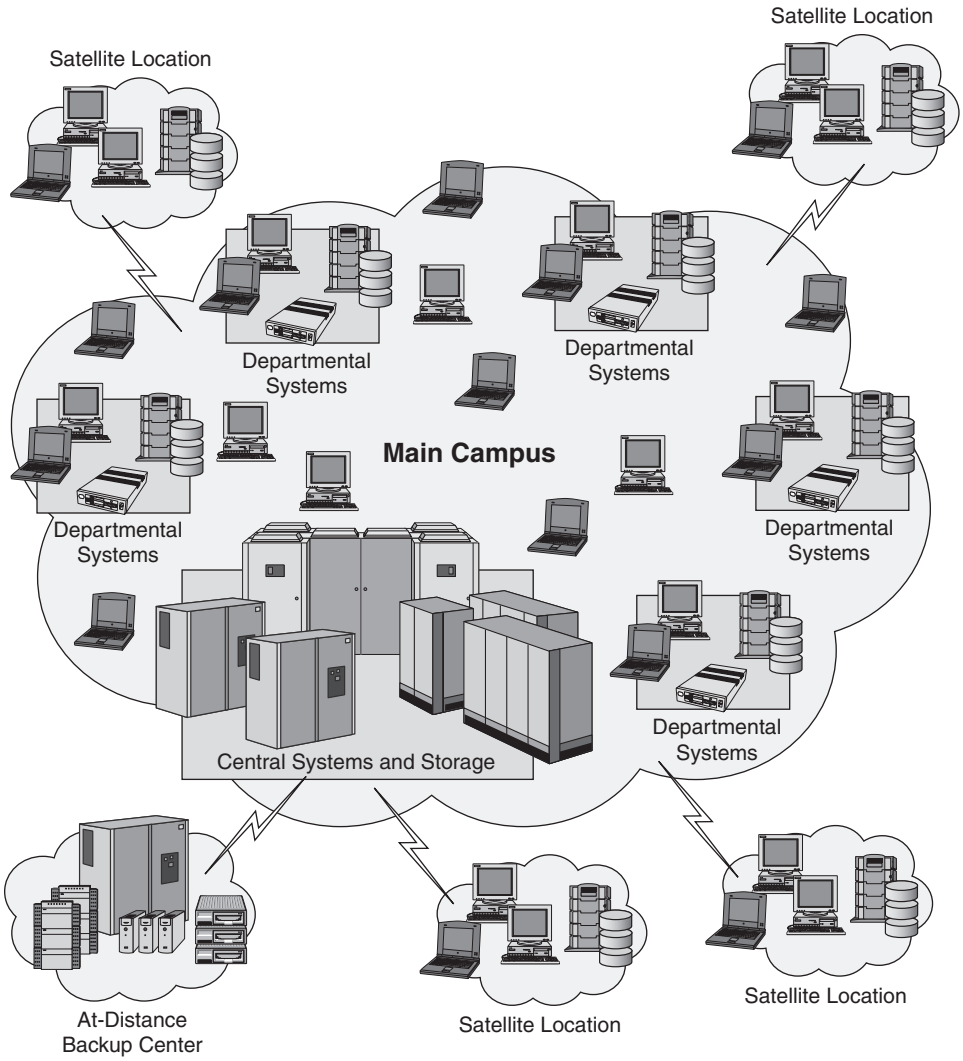

Figure 2–6 The high-end environment.

what they perceive as excessive centralized regulation and control requirements.

Today, even if the FC cables could be pulled to the various campus locations, since Fibre Channel has no security in its protocols, the access control demands of the central computing location may be more than the departments want to put up with. iSCSI, on the other hand, has security built into the basic protocol (both at the TCP/IP layer and at the iSCSI layer), which permits fewer invasive manual processes from anyone, including the disk

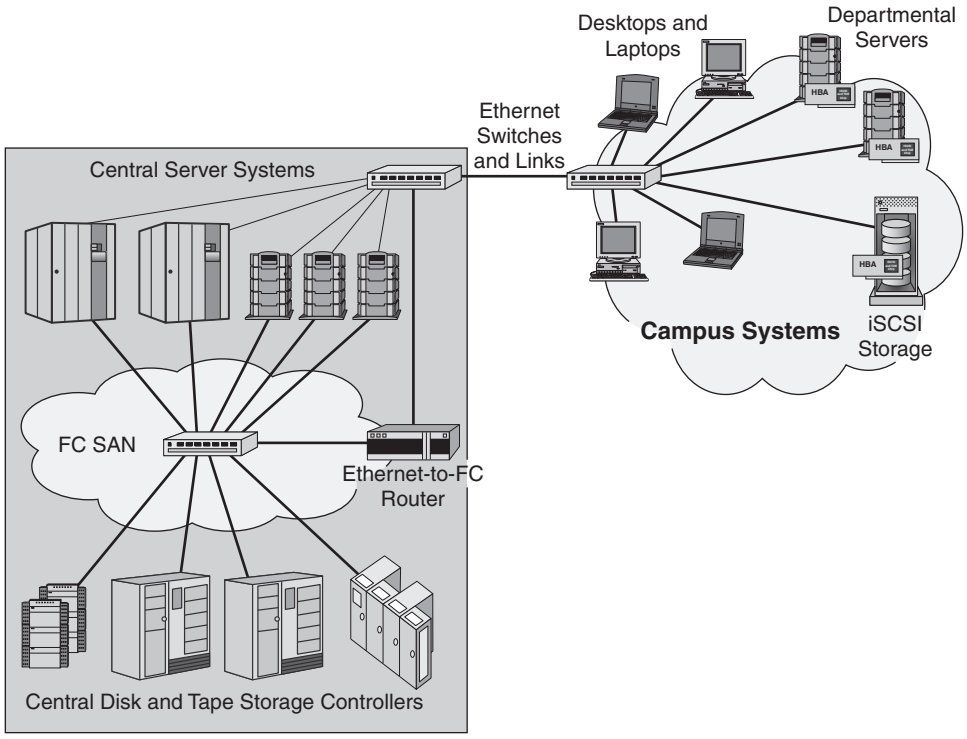

Figure 2–7 Campus and central system/storage.

storage administrator at the central location. iSCSI also permits the department servers to be booted as often as necessary, while still getting at central storage, something that probably would not be done if located within the main computing center.

Because of their needs and desires, campus departments are very likely to view iSCSI as key to their strategic computing direction. However, the campus environment is made up of more than just the department servers. It also has individual desktops and laptops distributed throughout that look like home office systems. A major difference, however, is that their users are not encouraged to modify them. Instead, every time they need additional storage, they have to justify it to either the department or the central computing location. Then the central or department "guru" who handles the system, must schedule time to come out and do an upgrade. Since these guru types handle many different users, they take approaches that can be unpleasant for the end user, often causing the loss of data or carefully constructed "desktop screens." Gurus are in a no-win scenario. They do not like taking

end users' systems apart, especially since users can be abusive about procedures, scheduling, and so forth.

Some installations have been known to just upgrade the entire system whenever new storage or processor power is needed. In the past this was often a reasonable approach since the need for processing power was keeping pace with the creation of storage. Now, as a rule, this is not the case. The 1-to-2 gigahertz (**Gh**) processors seem to have reached a plateau where the productivity of the office worker does not benefit from the additional speed of the laptop or desktop. However, one can still generate a lot of storage requirements with these processors, and it is beginning to occur to many companies that replacing systems just to upgrade the storage is a waste of time and money. Further, it greatly disturbs employees when they lose their data, their settings, or their visual desktop. Even when things go right in the backup and restore stages of bringing data from one system to another, the process is lengthy and tedious. Companies that believe time and productivity are money dislike these disruptions.

Both the end user and the guru will love iSCSI. To get additional iSCSI storage the end user just has to be authorized to use a new logical volume, and the issue is done. Often this can be accomplished over the phone.

Over time, desktops will be "rolled over" for newer versions, which will come equipped with the 10/100/1000BaseT (**gigabit copper**) IP adapter cards. 1000BaseT-capable adapter cards permit desktop performance of up to 1 Gb/s, which will greatly improve the performance of iSCSI storage. Note that most installations use Cat. 5 copper cables for 10/100Mb/s Ethernet connections, and these same Cat. 5 cables are adequate for gigabit speeds. Therefore, installations do not have to rewire in order to get gigabit iSCSI storage access for their ubiquitous desktop systems.

Since iSCSI also supports remote boot, one can expect many desktop systems only to support storage connected via iSCSI in the future. The desktops can then be upgraded as needed independently of the data.

The Satellite

A remotely located office, known as a satellite, will have an environment similar to that of a campus. It often functions like a department or small office. Satellites have their own desktop systems and sometimes their own servers. They generally suffer from the lack of adequate "remote support," which often means slow response to their needs.

As in the small office, satellite users do not usually touch the system but instead get a guru to come to the remote location to fix things. With the use

of iSCSI many satellite installations can have their storage-related needs handled via the phone. As they need more storage, they can call in the storage administrator, who enables more logical volumes for their use. This is possible since with iSCSI they are connected to the central location via a virtual private network (**VPN**).

A VPN is provided by a combination of carrier and user equipment. A carrier or ISP delivers some type of "IP tone" to the remote location, and the remote office uses encrypting firewalls and the like to secure access to a central computing facility, even across the Internet or other public infrastructures.

When the various satellite offices are located in a metropolitan area, a VPN becomes very attractive, since there will not be a large problem with "speed of light" latency issues. These network types are called metropolitan area networks (**MANs**). However, the greater the distance, the more local (iSCSI) storage will be deployed at the satellite location and the less central storage will be used for normal operations. These more remote locations will like the feature of local pooled storage that they get with iSCSI, without having to learn Fibre Channel.

When metropolitan area satellite offices need more iSCSI-based storage, they just ask the storage administrator at the central installation to logically attach more virtual volumes to the user's iSCSI access list. All this is possible without significant effort at the satellite location, assuming, of course, that adequate bandwidth exists between the central location and the satellite office.

In the past, satellite office connections required private or leased phone lines, but it is now becoming prevalent in many areas for carriers to offer "IP tone" at a much lower cost than leased lines. Thus, the customer is now more likely than before to have high-speed connections between the satellite office and the central office.

Satellite locations may also have local servers and storage requirements, and will want the flexibility offered by a SAN. They will find iSCSI a more cost-effective solution than Fibre Channel, especially since the network management can still be handled at the central location.

The satellite installation, like the campus environment, will also want to be able to use centralized tape units for backup without having them located at the satellite location. This also is an ideal exploitation of the capabilities of iSCSI. (See Figure 2–8.)

The At-Distance Site

"At-distance" reflects the needs of many centralized computing locations that use storage equipment placed at remote locations.

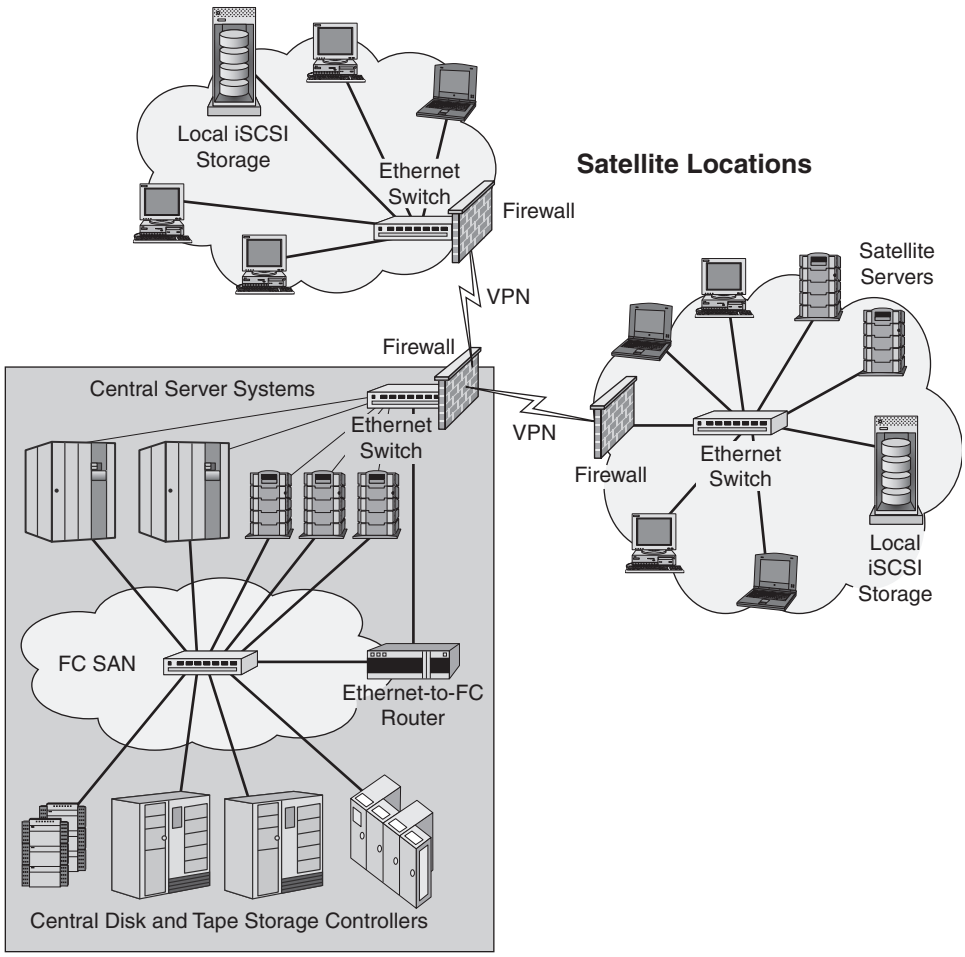

Figure 2–8 Satellite and central system/storage.

The remote equipment can range from small RAID arrays to large disk RAID arrays, tapes, and tape libraries. Even the central location prefers to have the tape library "at distance" from the central site. Often this is a tape vaulting area (such as "Iron Mountain"). iSCSI permits "natural" access to such remote units, without undue gateways, routing, or conversions from one technology to another. (See Figure 2–9.)

Part of this "natural" access is the ability of either the servers, the storage controller, or some third-party equipment to create dynamic mirrors at remote locations, which can be "spun off" at any time and then backed up to tape. This permits remote backup without impacting online applications. The

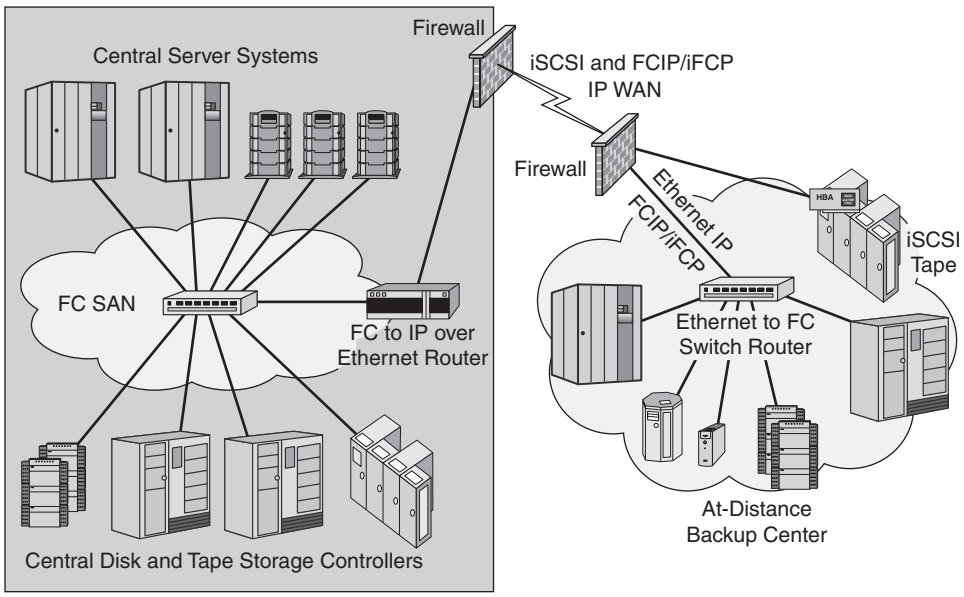

Figure 2–9 The at-distance environment.

remote mirror can be located at the tape-vaulting site, or iSCSI can be used to send it to another remote location. This type of process, though possible at a local site, seems to be very valuable when located at a secure remote site.

Today this remote storage access is done with proprietary protocols, devices, and often expensive leased lines. In the future it will be done with standard IP protocols, primarily iSCSI, often utilizing carrier-provided IP tone interconnects.

The Central Site

The central site will receive iSCSI storage requests from campus department servers, from desktops and laptops, and from satellite locations. Likewise it will issue storage requests to remote locations for backup and disaster recovery. (See Figure 2–10.)

The central environment is considered to be the high-end processing and storage environment. It has the highest speed requirements not only on the processor itself but also on the I/O network it uses. Further, it has an overarching need for high reliability, availability, and serviceability (**RAS**).

In the central site iSCSI must be able to perform as well as Fibre Channel and match the same RAS requirements. These difficult but attainable requirements dictate that hosts use top-of-the-line iSCSI HBAs and that these

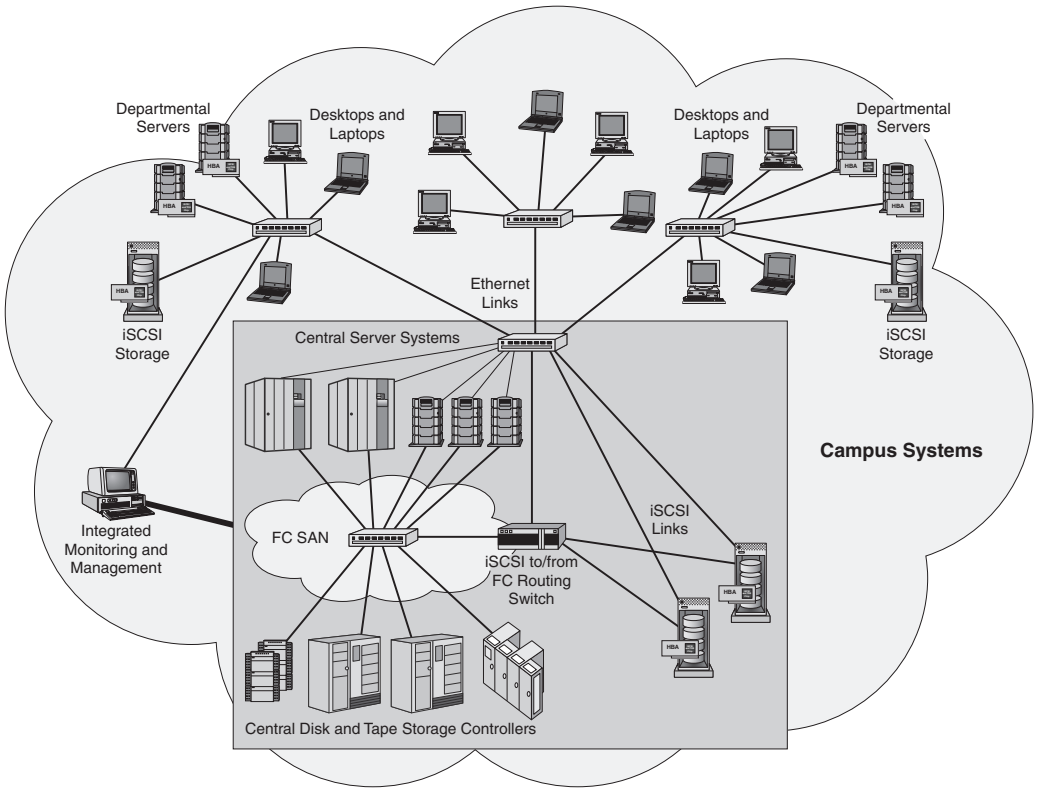

Figure 2–10 The campus environment.

HBAs be configured to operate in tandem such that failover is possible. Also, they need to be usable in a parallel manner such that any degree of throughput can be accomplished. The iSCSI protocol has factored in these requirements and supports a parallel technique known as Multiple Connections per Session (**MC/S**). This permits multiple host iSCSI HBAs to work as a team, not only for availability but also for maximum bandwidth. The same set of capabilities within the iSCSI protocol also permits iSCSI target devices to perform similar bandwidth and availability functions.

The high-end environment will have both the FC and the iSCSI storage controllers needed to service it. And since Fibre Channel is already there and can't be ignored, the installation must be able to interconnect the two storage networking technologies and expect them to "work and play" well together. The installation will have the problem of how to begin and how to integrate the two networks. Customers will want to invest in iSCSI storage controllers and yet continue to capitalize on the FC SAN investments they already have.

Various vendors offer "bridge boxes" that convert iSCSI host connections to FC storage connections. Some boxes convert FC host connections to iSCSI storage connections. Both of these functions are accomplished via routers, gateways, and switches (**switch routers**).

The thing that will actually make all this interconnection capability work is the management software. Probably there will be storage network management software that can operate with all FC networks and similar software that can control the iSCSI network. Clearly, though, there is a need for storage management software that can manage a network made up of both FC and iSCSI.

Even though multi-network software is sophisticated, some vendors are bringing it to market now. Luckily, the iSCSI protocol has defined a set of discovery processes that can be shipped with each iSCSI device which will permit a full iSCSI discovery process. This process, when used in conjunction with the FC discovery processes, will permit the interplay of iSCSI and FC SANs.

Since iSCSI and Fibre Channel share the SCSI command set, most existing LUN discovery and management software will continue to operate as it does in SCSI and Fibre Channel today. Therefore, there should not be significant changes to the SAN LUN management software.

The key problem is that the combined "SCSI device" discovery processes need to be carried out when there are both FC and iSCSI connections. It needs to be done both to and from the hosts and to and from the storage controllers. When an FC network manager performs its discovery process and detects an FC device, which just happens to be available to an iSCSI host (via a gateway device of some kind), it is important that the iSCSI network manager also know about the device. Therefore, an FC/iSCSI network manager needs to combine the results of the FC discovery process with its iSCSI discovery process so that all appropriate devices can be offered to the host systems as valid targets.

In addition to knowledge about each other's hosts and storage controllers, there has to be a melding of names and addresses such that iSCSI hosts can actually contact the FC target storage controllers and vice versa. Luckily, a companion protocol/process called **iSNS** (Internet Storage Name Service) deals with this problem by mapping the names and address between the FC and iSCSI views. In this way, with the appropriate surrounding management software, both networks can be seen as a seamless interconnected SAN.

To sum up, the high-end environment contains all aspects of the low-end (SoHo) and midrange environments, plus additional requirements for high

availability and large bandwidth, along with campus and WAN (intranet) connections. It also requires seamless interconnect between FC and iSCSI networks.

FC and iSCSI

Many have asked, "Will the integration of iSCSI and Fibre Channel continue indefinitely?" Since that question has an open-ended timeframe, the answer has to be no. A more important question, however, is, "What is the timeframe for when the customer will more likely buy iSCSI products than Fibre Channel products?" Said another way, "When will the sales of iSCSI equipment surpass the sales of Fibre Channel equipment?" My guess is 2006–2007. Other analysts have said 2005 and even 2004, but this seems to me to be wishful thinking.

I believe that the volumes will tip in favor of iSCSI (regardless of the timeframe) because iSCSI can perform all the functions that Fibre Channel can. In addition, iSCSI will operate on the wide area network (WAN), the Internet, and campus LANs up to the individual desktop and laptop level. Thus, when an installation is considering what to purchase next, it is probably going to choose the most flexible technology, especially since that seems to be the technology with the lowest projected price. This applies to both the initial costs to purchase and, as addressed earlier, the ongoing cost of management, training, and the like.

I did not say that Fibre Channel was going to go away in 2006–2007, just that iSCSI would begin to outsell it. That is because iSCSI will make significant gains by playing to a larger market, not just by displacing current Fibre Channel. I think Fibre Channel will continue to evolve for a while and continue to support its customer base. In other words, if I could project myself to 2010, I would still see sales of FC equipment. This means that not only will there be **iSANs** (SANs made up of only iSCSI network entities), but there will also be long-term business in the area of integrating iSCSI and FC SANs via various gateways and switching routers.

Chapter Summary

This chapter focused on the various environments where iSCSI would be of value. In it we discussed **SoHo** (Small office, Home office) environments and where iSCSI fits in them.

Home offices:

- Purchase software as part of the OS.
- Purchase software from a local computer store.
- Seldom share files.
- Share files via e-mail or peer to peer.
- Find iSCSI solutions useful for easily adding storage without opening systems.

Small offices:

- Sometimes install specialty software, which often uses a database.
- Sometimes have NAS servers.
- Can use iSCSI to ease the impact of storage growth.
- Find software iSCSI initiators and software-based iSCSI targets acceptable.
- Are attracted to dual dialect servers.

Midrange environments:

- Use dedicated server systems to process their business needs.
- Prefer HBAs instead of software iSCSI device drivers in their server systems.
- Require their target storage controllers to have hardware iSCSI HBAs or chips to ensure maximum capacity and performance.
- Have desktop and laptop systems that use software iSCSI device drivers.
- Use dual dialect boxes.

The size of the midrange market will be so large that the price of iSCSI adapters and chips will bring significant cost reductions to the iSAN market. iSCSI will tend to dominate the markets

- At the midrange.
- In the low end.
- Over time in the high-end environment.

The high-end company has

- A central computing location.
- Well-trained personnel.
- Backup tape drives.
- Tape libraries.
- Many storage controllers.
- A campus made up of:
  - Many small fiefdoms of key managers who need to *own* their own systems.
  - Subenvironments that in many respects map to SoHo and midrange environments, including desktops and laptops.
- Satellite locations.
- At-distance installations.

Satellite offices find that

- VPNs and iSCSI are valuable technology partners.
- They can access central storage as if it were located at the satellite location.
- They will be able to operate in a manner similar to the operations of small offices and midrange companies.
- They can have their own local iSCSI SANs with local storage.
- They can have all their disk storage requirements coordinated with a central location.
- They can use the central tape library for backup.

The bigger the satellite office, the more it will function independently, but even a small office will have its logical SAN extend across VPNs to the central location. This will be especially true in metropolitan area networks (**MAN**s). MANs will permit central iSCSI storage to be used as if it were local to the satellite location.

The high-end environment will be composed of "at-distance" environments:

- These environments want access to devices such that tape can be located in some remote area (called tape vaulting).
- Backup to tape at a remote location will be one of iSCSI's "killer apps."

  The central location holds all the big servers and big storage controllers.
- This is the primary location for FC devices.
- The key requirement is to ensure that iSCSI and Fibre Channel "work and play" well together.
- A centralized technique is needed to
  - Perform discovery.
  - Manage the total storage environment.
  - Permit both Fibre Channel and iSCSI to work as a single transport network.
- Fibre Channel is not going away, or at least not quickly.
- iSCSI does not have to take on all FC deployments to be successful.
- iSCSI will be considered a very successful transport if it can meet all the needs of SoHo and midrange environments.
- Routers, gateways, and switches will permit access by campus systems to the FC storage in the central computing center.
- Any amount of native iSCSI installation in the main computing center should be considered to be net added value to the iSCSI business plan.