Chapter 16

Configuration Options for SANs

As was shown in Chapters 14 and 15, Storage Area Networks can be assembled and connected in a variety of ways. Switch options and device configurations are dependent on I/O workloads and recovery requirements. Once assembled and connected, the network as a whole has to be integrated into the data center infrastructure. Chapter 16 moves SAN configurations out of the abstract and into the data center.

Regardless of connectivity or configuration, viable SAN solutions are moving from single switch configurations to multiswitch configurations with numerous device attachments designed to deal with increasing workloads and even the simplest recovery and redundancy requirements. And though there are many ways to configure a switch, the basic starting-point configurations remain core/edge, meshed, and cascading configurations. All three present their own unique challenges when integrated into an existing infrastructure.

This chapter delves further into an external SAN connectivity beyond the node devices, which become the drivers for application usage. Storage configurations, the workhorses of SANs, are often called upon to provide multiaccess for the server population within the data center. One of the biggest integration challenges is in trying to throw the configuration into the support of heterogeneous storage components. Additionally, external forces driving new and enhanced storage media waiting to be included in the SAN can play a part in connectivity.

Connecting into the Data Center

Connectivity refers to how the SAN as a whole integrates into the data center. In this discussion of connectivity, we will draw a theoretical line between SAN components and server processing. As such, the SAN as a whole includes the components that connect to the FC switch with a dotted line drawn at the HBA. This allows us to position the FC SAN in a set of server populations and existing configurations. Even if only used as a logical model, it helps us visualize the inevitable problems. Figure 16-1 shows the logical demarcation line we will use throughout the chapter.

Implementing a SAN configuration into a data center presents three challenges right off the bat. The first challenge is the need to configure a multiswitch SAN. As I/O workloads, driven by supporting applications, are applied, the need to provide multiple connections, discrete zones, and paths for performance drive the configuration into multiple FC switches. There's just no way around it. Over and above this is the need to provide some level of recovery and redundancy for data availability and protection scenarios. Existing policies and practices will drive this challenge; if not, it's good to start thinking about recovery, redundancy, and single point of failure—of which, a single switch strategy presents.

The second challenge is in providing support to the server population within the data center, which can be driven by the justification for the SAN in terms of server consolidations and supported applications. Note, though, that it also moves into the

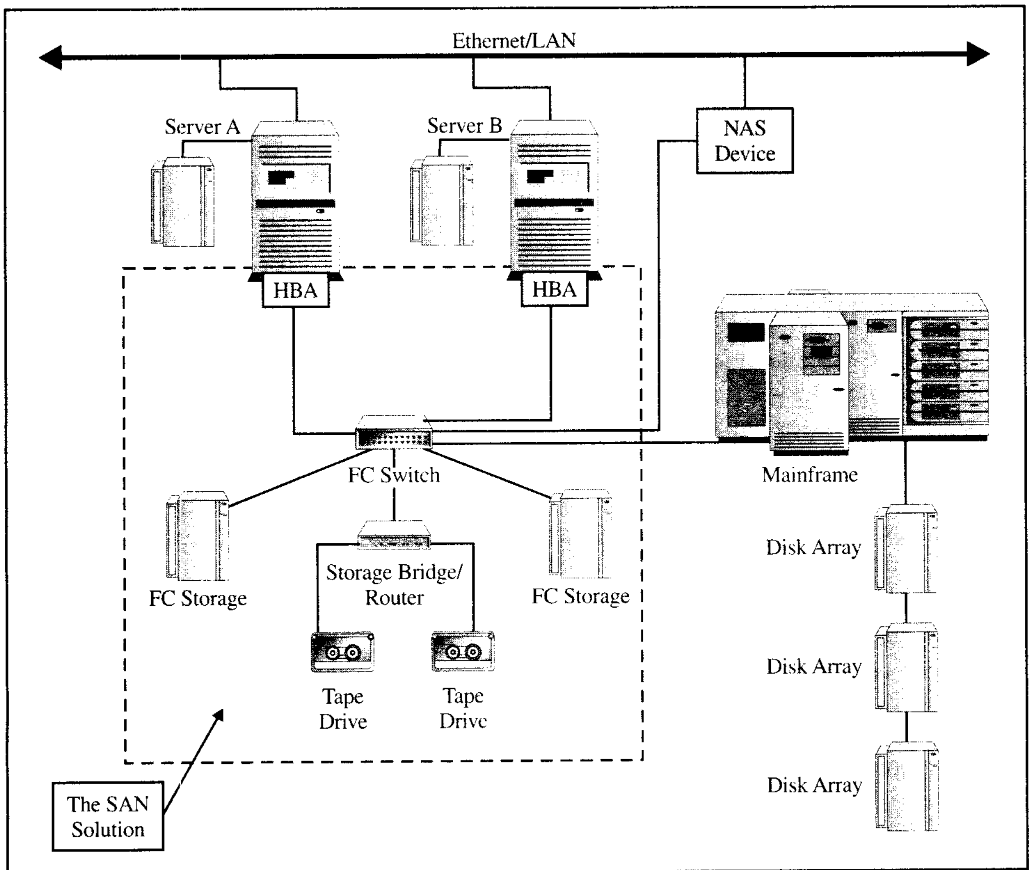

Figure 16-1. *Viewing the SAN as an entire solution*

realm of data ownership, production cycles and processing, and data sensitivity. More importantly, it pushes the need for a heterogeneous SAN supporting more than one operating system at the attached server level. In many cases, this becomes the big SAN imbroglio in which implementation becomes political and ceases being productive, or relatively anxiety-free—or, for that matter, fun in any way.

The third challenge is supporting external networks, which encompasses all of the issues mentioned previously. The ability to provide the data stored within the SAN to corporate networks is dependent on the external network points within the data center, as well as the capability of the attached server to distribute this data within the infrastructure it is connected to.

Implementing Multiswitch Configurations

The need to configure a multiswitch SAN solution will become evident when designing your first solution. Any solutions should encompass a rationale for the configuration orientation used (for example, cascading, core/edge, or mesh). Initial implementations often move from a simple cascading architecture to a multilayered cascading architecture without the benefit of, or options for, moving into a more complementary configuration. Keep in mind that SAN configurations are driven by supported I/O workloads. As such, the need to provide recovery and redundancy should play a significant role in the design. Here are a few ideas to chew on.

- **OLTP Workloads** Using a core/edge configuration enhances performance through storage access at the edge, while reducing instances of single point of failures for high availability transactions accessing multiple I/Os through the SAN configuration (shown in Figure 16-2).
- **Web Applications/Messaging Workloads** Using a mesh design provides alternate paths for high traffic, asynchronous I/O processing, while reducing instances for single point of failure. It relies on effective switching methods for high traffic management within the SAN configuration (shown in Figure 16-3).
- **Data Warehouse/Datacentric Workloads** Using a cascading design provides the necessary access and performance of datacentric transactional workloads. Here, availability requirements are reduced, but access to large data bases remains paramount when it comes to processing time and I/O intensive transactions (shown in Figure 16-4).

Coming up with the appropriate SAN design for I/O workloads also requires some thoughts on recovery. Though driven by the application's needs, a working design should encompass current backup/recovery processes that are in place within the data center. It's good to reconsider recovery issues relative to the storage capacities of the SAN and the recovery requirements of the data. In a normal backup situation, data is copied from the storage arrays to an attached server and shuttled onto the network so it can get to the backup servers attached to tape drives. In all likelihood, these are located close to the tape library. However, if the SAN starts out with capacities over 500GB of user data, this scenario might just overcome existing backup practices.

Consideration should be given to backup/recovery within the SAN infrastructure, which means the addition of external tape media—tape drives and tape libraries providing backup/recovery operations integrated into the SAN. Designs will have to be modified, regardless of configurations (core/edge, mesh, cascading), in order to accommodate the backup traffic and backup/recovery software—meaning the addition of a bridge/router component. This setup adds another benefit (which is largely dependent on recovery requirements and storage capacities): the evolving capabilities of FC fabric to provide a direct copy operation in support of backup software will eliminate a significant amount of I/O overhead. Often referred to as "server-free backup" or "server-less backup," these configurations will be covered again in Chapters 22 and 23.

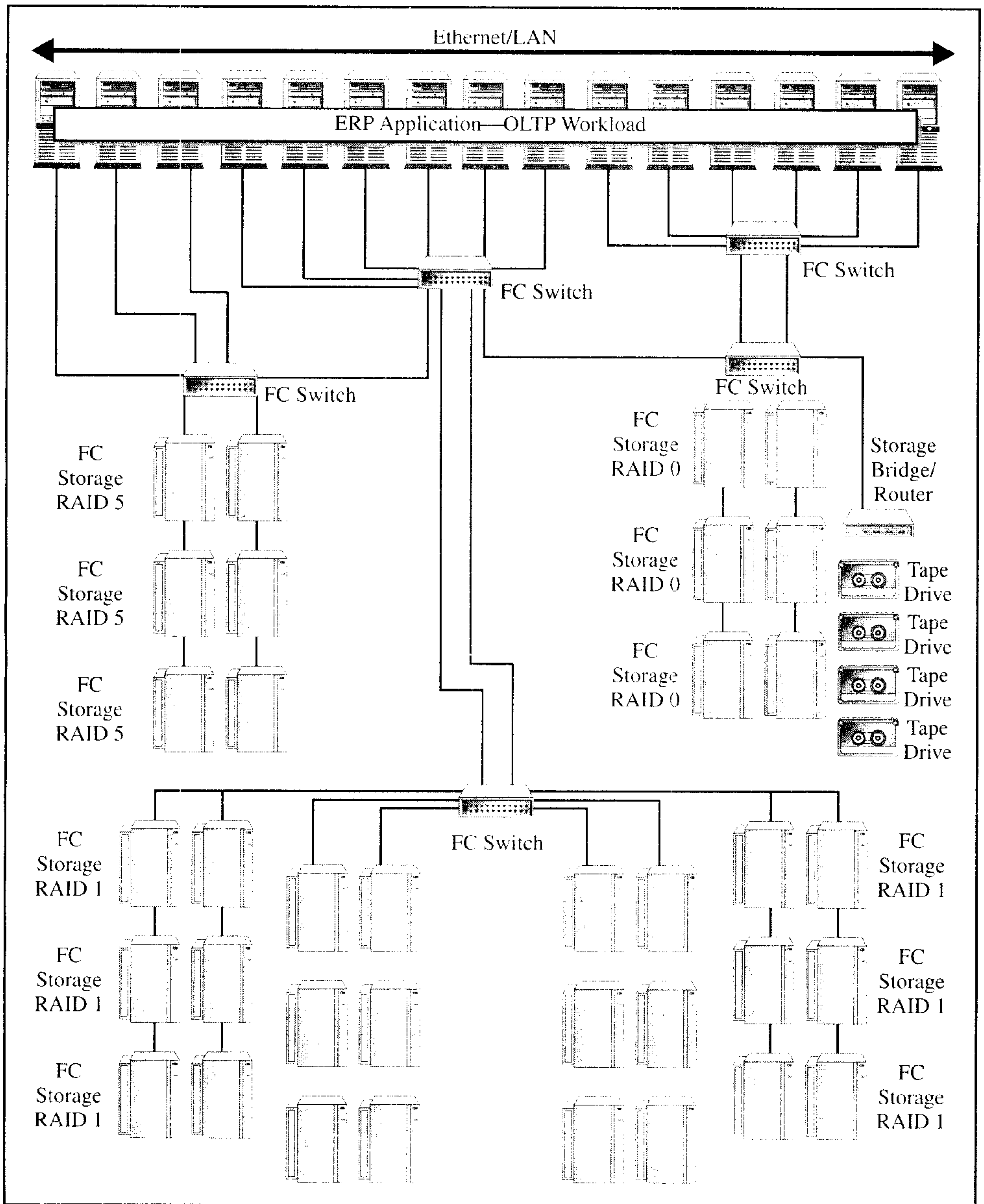

Figure 16-2. *OLTP workloads supported by a core/edge configuration*

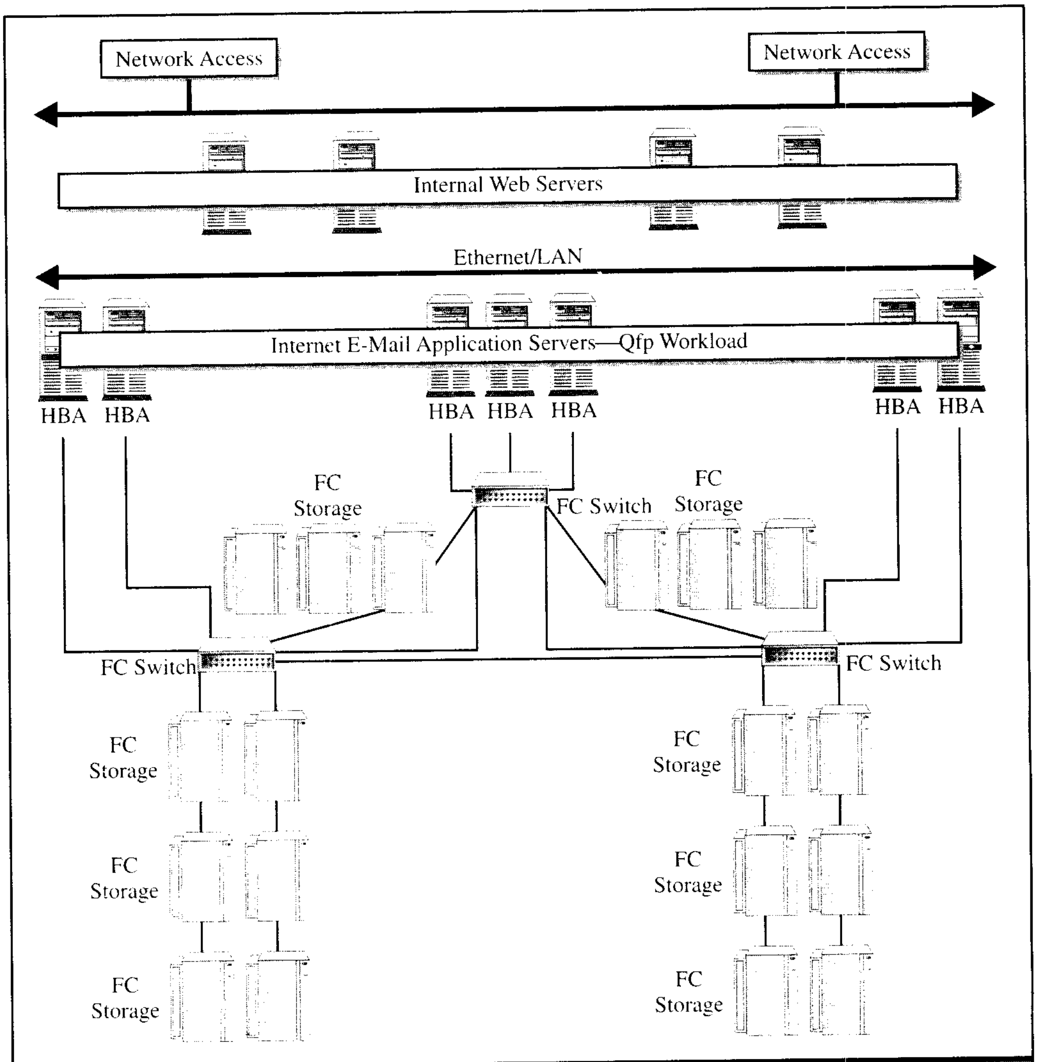

Figure 16-3. *Web and messaging applications supported by a mesh configuration*

Supporting the Server Population

The next challenge is determining which servers to support. Not only is this driven by what applications are moving over to the SAN, but also what strategy justified the SAN in the first place. If consolidation justified the SAN, then someone will be looking to retire or redeploy a number of servers through the consolidation efforts, which brings up an interesting set of activities.

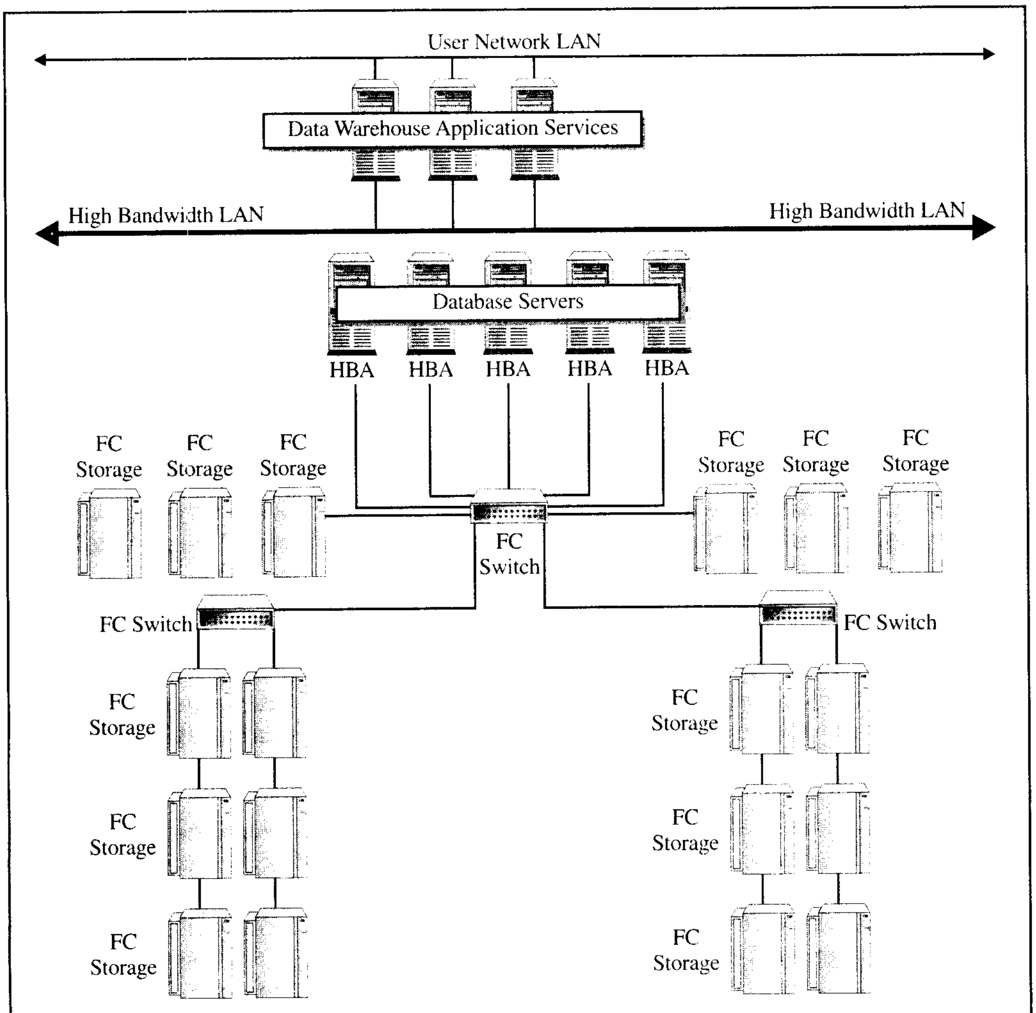

Figure 16-4. *A data warehouse application being supported by a cascading configuration*

First is the modification of existing capacity plans to calculate the requirements of new or redeployed servers, and to calculate the necessary I/O workloads as they contribute to the design of the SAN (some examples and guidelines for estimating I/O workloads are upcoming in Chapter 17, with specifics for SANs following in Chapter 18). This will very likely make you have to think about moving applications that are running on both UNIX- and Windows-based servers. And though this is possible, and many vendors provide a mountain of literature on the theory and ability to support this configuration, it significantly raises the cost of the SAN, not necessarily in terms of added hardware or software, but in complexities encountered implementing and supporting these configurations. Consider the following, carefully.

UNIX- and Windows-based servers do not access or handle storage in similar fashions. Even though there are multiple solutions based on POSIX standards, the capability to share devices within the SAN is not yet a reality, which means strict segregations have to be enforced with system-level zoning and LUN masking strategies that are both time-consuming and complex.

Applications supported by UNIX- or Windows-based servers have drastically different service levels. Design and configuration of the SAN will either add resources to the Windows area, which don't need it, or compromise the UNIX environment by not providing sufficient SAN resources. Plus, recovery requirements are going to be different, which means they will either share the tape media, likely to result in problems, or they will have to have their own tape media for backup.

We should stop here and make sure we're not overlooking one of the big-time workhorses within the data center, the IBM and IBM-compatible mainframes. The need to interconnect a mainframe to open computing has been with us since the first UNIX server deployed its data throughout the enterprise, and the Windows PC began to develop its own legacy data, while needing to be integrated with the corporate record on the mainframe. Unfortunately, mainframe solutions continue to elude vendors, IT systems organizations, and research agendas—not that disparate and site-specific solutions don't exist, they do, if only in order to support specific business applications. Integration of IBM mainframes into server environments remains elusive, even counting all the money IBM has thrown at making the old, but reliant, MVS operating system both compliant with POSIX as well as available to networking infrastructures like Ethernet and TCP/IP.

To its benefit, and its everlasting detriment, IBM mainframe-processing models predate many of the innovations we are now working with, including storage networking, data sharing, device pooling, and scalable configurations—both internally at the high-end SMP level as well as in the multiprocessing, multitasking, transaction performance that no other processing platform can beat. Sadly, and perhaps tellingly, the IBM mainframe-processing model was a proprietary system that resisted evolving into the client/server architecture too long and was thus eclipsed by its own inability to deploy and support new and distributed applications in a cost-effective manner.

There is, however, hope on the horizon, in the form of a serious revolution in the IBM mainframe world. The traditional operating environment has given way to a new operating system: the zOS. This next step in the evolution of MVS provides a multipartitioned operating environment in which you can run an IBM Linux system on user-defined partitions that leverage the power of the mainframe while ensuring that it remains cost-effective.

So what does this have to do with storage and, specifically, integrating a Storage Area Network into the data center? Quite a bit, if you want the truth. Because here's the first question you're going to hear: "Great! When can we connect to the mainframe?"

The ability to share storage resources within the complete populations of servers, keeping in mind that zOS facilitates a logical set of servers, is a big-time perk for data center managers. Now, here's the flipside. IBM mainframes are moving toward their own orientation to storage area networking and are connected to their data paths by a fiber optic connection, which the IBM folks refer to as *channels*. This evolution from legacy bus, tag cables, and enterprise connectivity (or ESCON cabling: a parallel bus implementation), not to mention the low-level channel link protocols to Fibre Connectivity (or FICON), supplies the first real direct connect to the IBM mainframe usable by other open storage networks.

This level of service in SAN requires a FICON connection at the FC switch port level. IBM has been the pioneer in connecting storage through a switched environment, having integrated director-level switches for some time. At the end of the day, providing switch-to-switch communication with IBM storage switches may be an alternative to the evolving mainframe connectivity conundrum. Entry through an IP address will also provide connectivity, but the bandwidth restrictions make this alternative unlikely.

Figure 16-5 illustrates the many alternatives to supporting a heterogeneous OS data center through a SAN infrastructure.

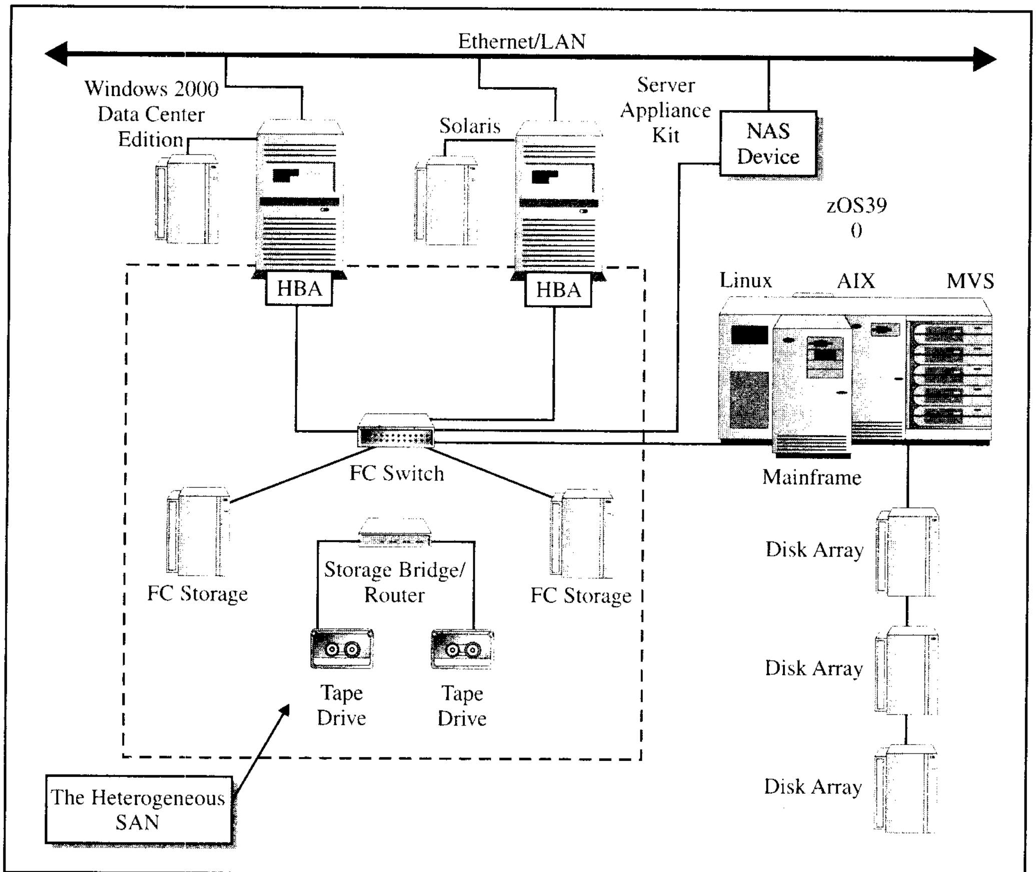

Figure 16-5. *A fully heterogeneous OS SAN-supported environment*

Supporting the External Networks

What if a SAN has to support an OLTP application for a highly available business process in three time zones? Not a problem. At least it's not *my* problem. I'm just a little ol' Storage Area Network.

Though maybe not.

The ability to access the SAN remotely, instead of going into the data center every time you need to access SAN operations and information, can be configured to meet administrators' needs through external network access. The SAN configuration discussed in Chapter 15 required that SAN operations be performed through a dedicated workstation that is directly attached to one of the SAN switches. Given this is where software configuration tools can be directly accessed for setup, maintenance, and initialization activities, most switch products are enabling users with IP access points that allow some of the SAN operational tools to be available remotely.

There are several innovations waiting in the wings that will drive advanced connectivity options for the SAN. Key among these is the need to further integrate IP-based storage solutions with FC networks. However this will require a more sophisticated usage of storage systems driven by user data that is increasingly distributed throughout a corporation. By analyzing current data usage scenarios, iterations of data must make several stops within an enterprise infrastructure in order to satisfy distributed application requirements.

The Evolving Network Connections

There are several innovations waiting in the wings that will drive connectivity options for the SAN. Among these is the need to further integrate IP-based storage solutions into FC networks, which will be largely driven by the integration of more sophisticated usage of storage as data is distributed throughout the enterprise. If we look at current data usage scenarios, data still has to make several stops within an enterprise infrastructure.

For example, many organizations still take it upon themselves to distribute data to remote locations, something largely accomplished by a series of data subset deployments that are both time- and location-dependent. Using a combination of data replication software services and remote storage devices, these configurations depend on general-purpose servers to communicate within the enterprise network so they can ship current data to remote locations. And though some of these can be cleaned up smartly with a NAS configuration, many remain datacentric and rely on sophisticated RDBMS functions at the corporate data center, as well as the distributed location. Figure 16-6 is a conceptual picture of how this works and how it lays the groundwork for future deployments and connectivity strategies for SANs.

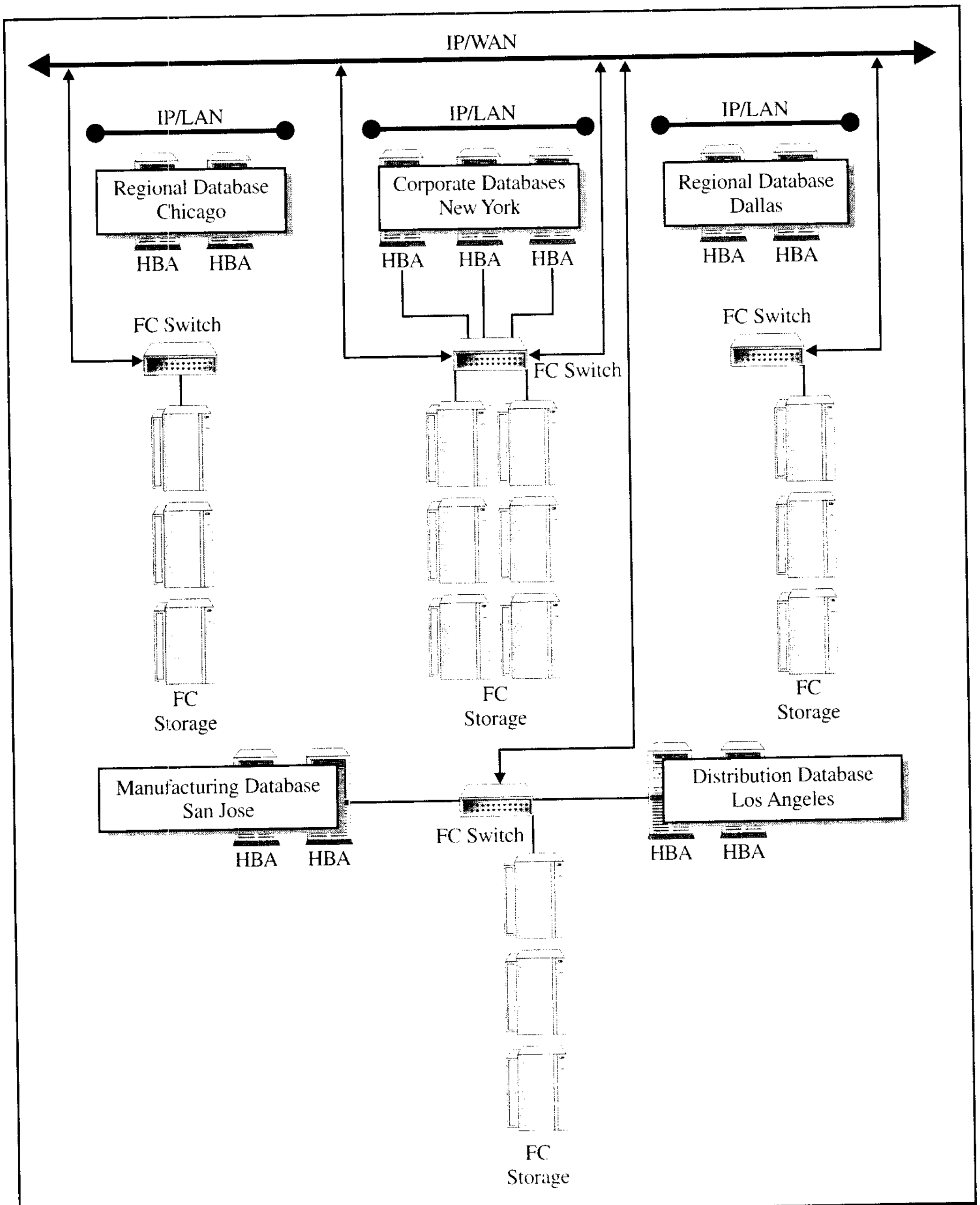

Figure 16-6. *A data replication strategy and potential SAN solution*

There are two basic scenarios that will push these networks into integrated, global SAN infrastructures. First is the integration through IP storage devices, which offer the capability to store data without the latency and overhead general-purpose server deployments and direct-attached storage components required. This is handled in two ways, either through an enhanced NAS device that supports direct communication with the SAN infrastructure, or with an iSCSI device that supports direct I/O communications with the SAN infrastructure. Either way, the devices have to be linked through an existing IP connection and be viewed as a node device within the SAN, which means additional port functionality at the SAN switch to handle the IP file or block I/O over the IP network.

The second scenario involves the installation of SAN configurations that are both local and remote and can communicate and be controlled remotely. Figure 16-6 also shows a conceptual view of how this is accomplished, facilitating timely data replication throughout a typical corporation. Remote configurations need further extensions of fabric software to encompass remote fabrics as extensions or as recognized external fabrics.

This SAN-to-SAN scenario assumes that switch port evolution will continue to support the extended functions of an E_Port communicating with a remote switch. As noted earlier, vendors must further enhance the switch as users demand full IP connectivity and FC processing between the storage node devices. Users will also require a management scheme to operate these configurations in a physically distributed fashion.

However, once this is established, the capability to provide a vast infrastructure of storage will be in place, providing yet another step in the ubiquity of data for applications. This will further enhance effective web services, thus allowing applications to be synchronized with data.

The feasibility and availability of these solutions is now becoming based on enhancements to existing network standards, including the addition of Fibre Channel throughout the enterprise, IP standards, and SCSI integration into IP specifications. All of which have added up to an initial wave of solutions that lead to the extension of SAN technology. Figure 16-7 shows the potential future IP storage integration and SAN-to-SAN communications supporting advanced data management strategies.

The Evolving Device Connections

Along with network connections, SAN connectivity options continue to evolve, encompassing more storage devices. In fact, integrating FC communications with the myriad tape solutions available has been underway for some time. Further enhancements to tape media will likely come in the form of optical and solid state disks, as well as advanced forms of caching mechanisms.

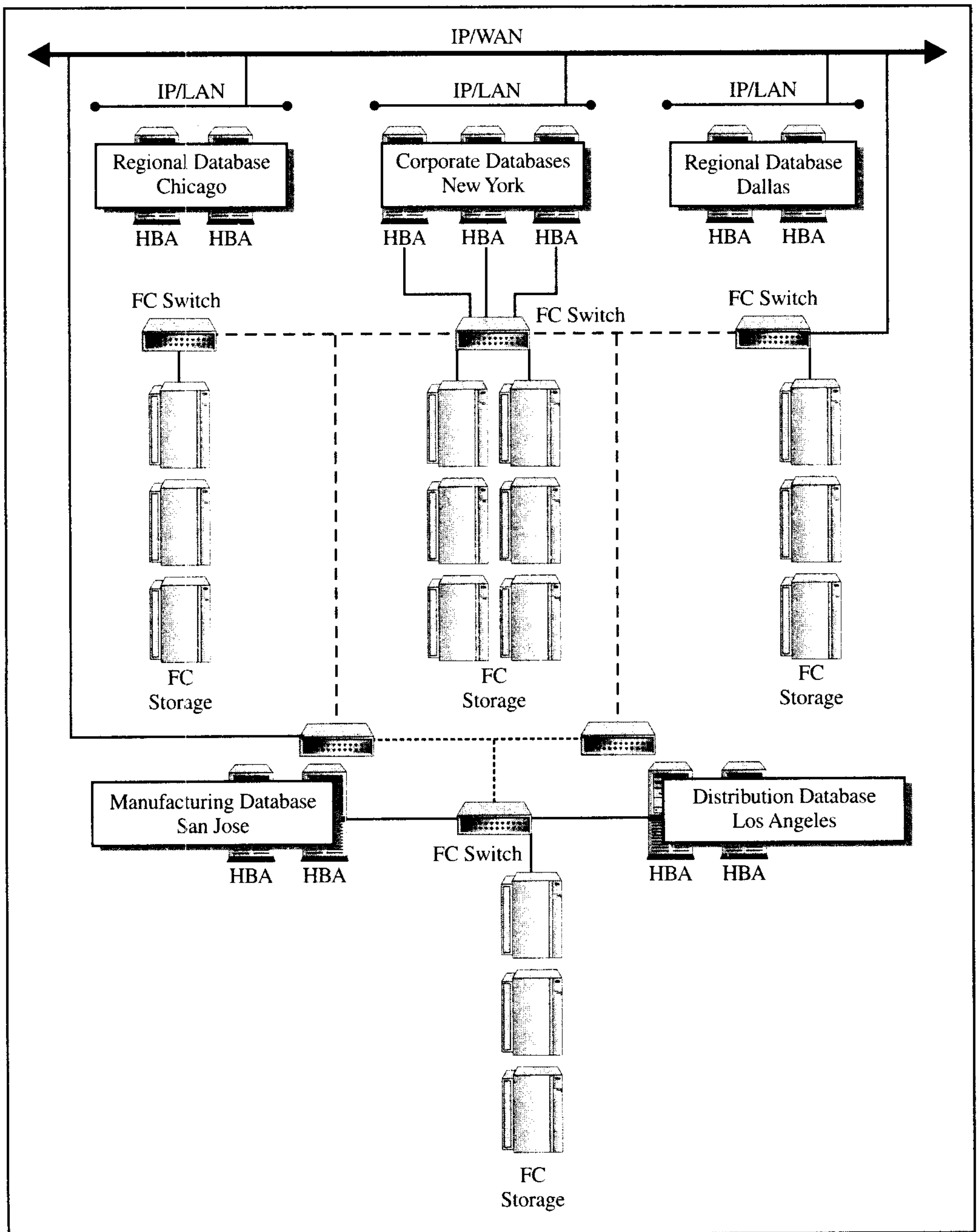

Figure 16-7. *A future SAN-to-SAN configuration*

Integrating Tape Solutions

Tape media is currently integrated using an external bridge/router solution (see Chapter 14). Bridge/router connections are increasingly integrating solutions whereby FC connectivity is built into the tape controller devices, allowing for a cleaner installation, not to mention a closer integration into the controller mechanism that takes advantage of individual tape manufacturer's differentiators and functions. This level of integration comes closer to allowing tape to function more efficiently by talking directly to the switch port. Developing tape media is an ongoing effort of the FC standards groups.

Tape media remains a particularly sensitive area within data centers, given the large investments that must be made in these realms. The capability to change and start additional library areas that support additional media is no small undertaking. One of the largest bottlenecks in support processing within data centers is backup, which gets performed on a daily, weekly, and monthly basis. This takes up an inordinate amount of time compared to the production processing that takes place in direct support of the company's business. (Additional details on backup/recovery processes are covered in Chapter 22.) Integrating tape library functions into the SAN is a significant task. Even though it may seem trivial at the beginning, the importance of the complete analysis of this integration greatly impacts the backup cycle—both positively and negatively.

The positive aspect is enhanced performance when moving backups to FC speeds. Even though you're throttled by the limitations of the tape device, it remains significantly faster than a slower and copy-redundant direct-attached model that can quickly congest an IP-based network infrastructure. There is also the potential value that a "server-less" backup scenario can be supported within your SAN configurations. Here, the elapsed time savings in running backup jobs can be as great as 50–75 percent, meaning that a typical backup cycle which runs seven hours would now take only three and a half. Because of this, significant overhead savings can be obtained over and above the amount of production processing time placed into the cycle.

The negative aspect is the increased cost to the facilities and recovery operations that result from integrated SAN tape configurations. The additional costs are needed to allocate tape drives for the FC SAN switch. More than likely, these will be dedicated given the non-sharing that FC SANs are currently encumbered with. In addition to this are the extra facilities necessary for extending the tape library functions—even though it may integrate into the current library system.

Another negative aspect is the consideration of the recovery operation that will ultimately have to take place. This requires the ability to locate proper backup tapes, log tapes, and control tapes. The tapes have to be dumped to the appropriate volume within the devices from the attached server, and then applied in the event of database recovery (which most may be). This isn't necessarily any different than normal recovery, other than that all of this has to be performed within the SAN environment, which, unless it's taken into consideration beforehand, can very quickly get out of hand.

Looking for an Optical Connection

Another popular media that will potentially integrate into the SAN is optical, or CD and DVD drives and libraries. Historically, these devices have seen limited usage as removable media libraries within enterprise network environments because of their performance as a write-once disposable media and physical read/write performance at the optical device level. Both of these conditions have improved to the degree that optical media and drives can be used for special applications of information distribution, reference, and, in some cases, effective backup media.

Optical devices can be integrated into a SAN using a bridge/router device to perform the same interconnection as Tape media. The use of SCSI optical drives with a SAN router can integrate the CD and DVD media sufficiently for production processing. The production of CD or DVD for distribution from digital sources, such as post processing of entertainment and educational packages, can benefit from the increased speed of SANs and allow it to process in a connection-oriented environment to facilitate the total bandwidth of the SAN.

Front-end processes of these applications, the preprocess production, where digital entertainment and educational products are edited and produced, have been an early adopter of SAN technology. This is largely due to bandwidth and storage capacities, as well as the ability to provide a Class 1 connection, which facilitates the streaming media production requirements for digital video files. These configurations can form either the production system for manipulating and creating digital video/audio effects, or the distribution system to move entertainment or educational files out to a network.

Actually, optical drives and autoloader packages have recently become available in NAS configurations. The NAS device is bundled with optical media drives and an autoloader system while maintaining the NAS file access front end and Plug and Play connection to an IP network. Given the future integration of NAS and SAN, the integration of NAS optical systems into SAN will become another alternative to SAN/ NAS integration and optical media usage with the SAN environment.

Connecting to the Future

Over and above the connectivity of SAN into existing data center infrastructures, SAN will soon need to integrate into the continuing evolution of external processing technologies, such as advanced clustering, InfiniBand, and blade-level computing. Each of these will require the SAN to address the compatibility issues of how the FC network interfaces with these evolving technologies.

Clustering has been an emerging solution for highly available configurations and is now employed to address availability levels for real-time transactional workloads. Although clustering solutions utilize the SAN in an effective manner, the ability to support both sides of a cluster is problematic, causing most configurations to support two distinct SAN configurations, which provides two distinct system images, with a SAN fabric supporting each.

The consideration regarding these configurations is the expense in doubling the processing environment to ensure that upon failover the supporting system can take over the workload and synchronize the data and continue processing. Secondly, the protocol used to communicate between the two systems, TCP/IP, adds quite a large overhead for the short communication paths between servers. Finally, the complexities of synchronizing the databases and data for continued processing upon recovery makes these solutions expensive and difficult to configure and maintain.

Look for clustering to take advantage of two systems sharing the same fabric, reducing the complexity of totally discrete systems to failover at the processing level with shared SAN access. This allows the data to be protected through a RAID-level mirror within the SAN, should it be needed at the storage level. Additional protection for hardware and data redundancy within the SAN configuration provides for failure recovery within the storage infrastructure, as shown in Figure 16-8.

If we take our clustering discussion and replace the system-to-system communications (now using TCP/IP) with a switched fabric network implementation using the InfiniBand protocol, we begin to see the benefits of InfiniBand. Let's take that one step further and implement a fabric-to-fabric interconnection, tying our SAN configurations together. Now we're really seeing the benefits of InfiniBand.

An InfiniBand configuration allows system-to-system communications to scale beyond the current 2-, 4-, or 8-way clusters that are expensive and unwieldy, but nonetheless available, and replace this with low-cost process node servers communicating through a switch fabric. This allows an infinitely larger number of servers to effectively communicate. In addition, the ability to tie in storage devices, such as the SAN, through an InfiniBand fabric to FC fabric interconnection allows some or all the servers to share the storage resources being managed by the SAN fabric.

InfiniBand technology allows for the creation of both a process infrastructure and a storage infrastructure, each with the capability to communicate though a common scalable interconnect in order to develop a dense computing and storage fabric.

If we take our Infiniband and FC SAN fabrics and key in on the process nodes, the application of a switched fabric internal to the process node provides yet another scalable resource within the process fabric. The use of internal switched fabrics, such as Rapid I/O and HyperTransport, form a foundation that allows internal components of the process node to scale new heights, bumping up the number of CPUs and addressable storage levels.

Given this configuration, we can begin to visualize the requirements for an internal storage function within the process fabric to support the internal functions of the process node. The communications between storage and process fabrics are enabled by the InfiniBand protocol.

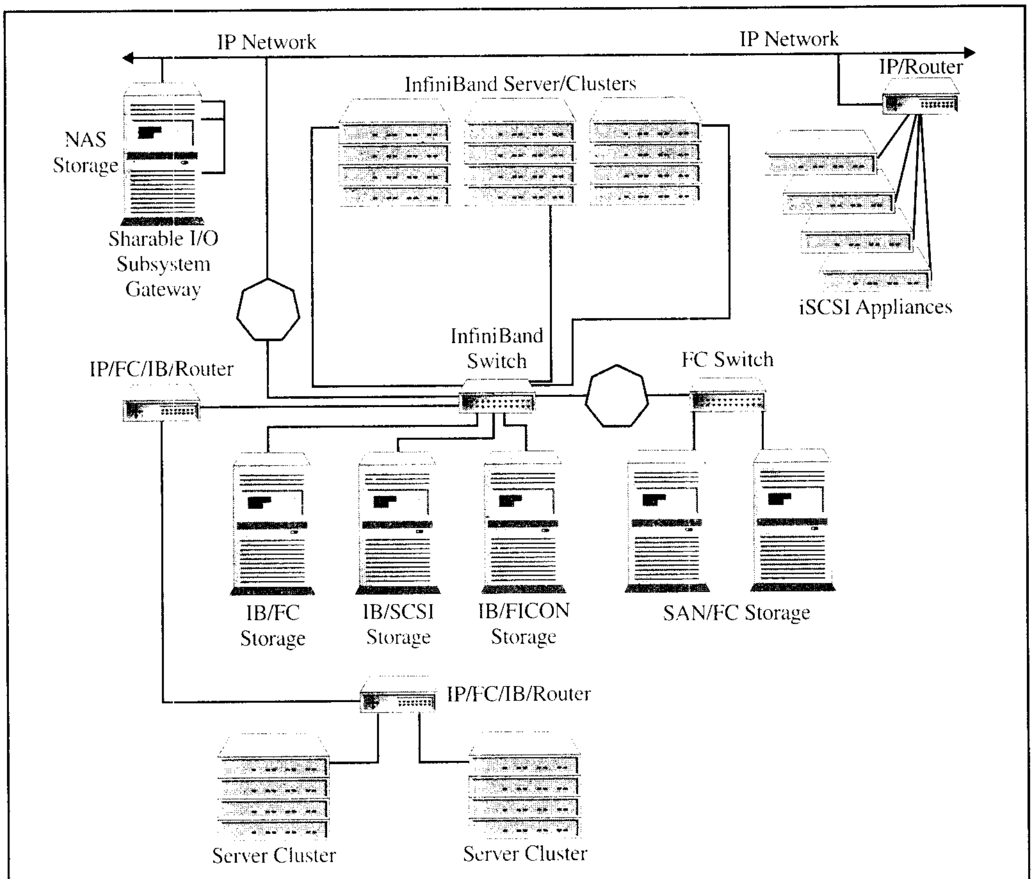

Figure 16-8. *Advanced clustering with shared storage SAN*

In conclusion, network technologies will continue to decouple the traditional computing configurations into process and storage networks. The data center is headed toward the next level of computing, driven by storage area network technologies that support all aspects of application sets and process environments: business, academic, scientific, and personal. The future data center is likely to become a logical set of storage and process fabrics, as illustrated in Figure 16-9, communicating through a public and private network infrastructure.

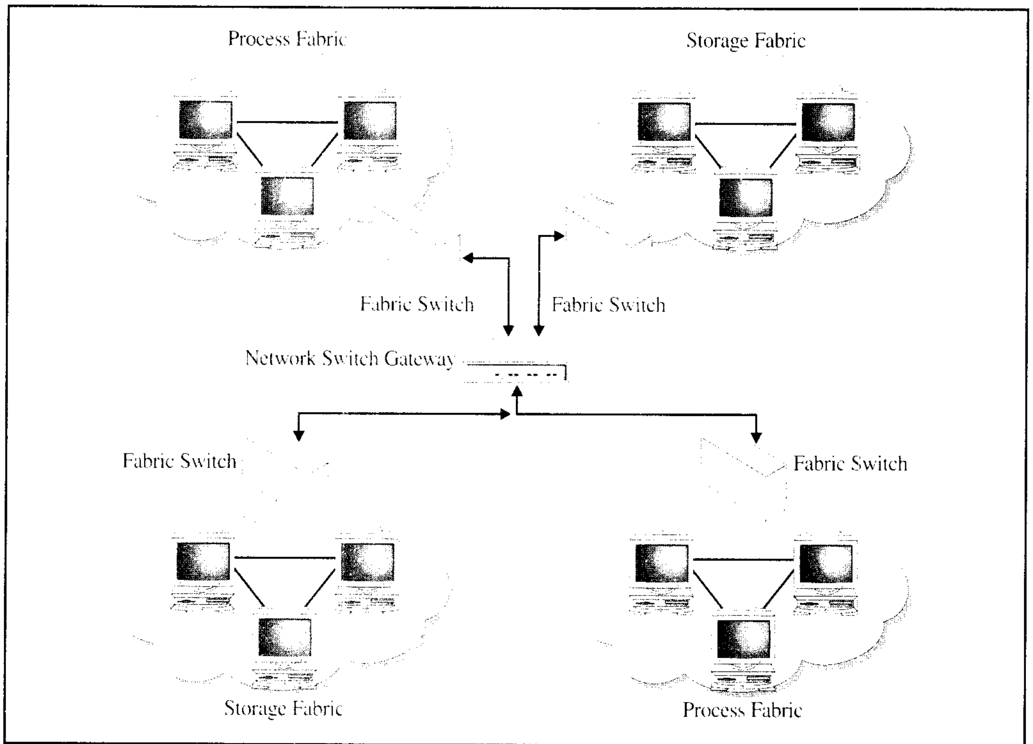

Figure 16-9. *The future: process and storage fabrics with a common interconnect*

SAN Configuration Guidelines

The capability of Storage Area Networks to integrate into the data center can become more complex than expected. However, given the right orientation, and a realistic view of current solutions, SAN offers a compatible solution to both UNIX and Windows environments. Not without its challenges, the ability to move beyond the homogeneous implementation can be successful, though more expensive and encumbering than the original SAN justification may have entailed. With the current set of products and solutions available, the data center can expect the following when connecting the SAN into the data center:

- *Multiswitch configurations will be a way of life.* Be prepared to work with multiple switch configurations. Again, there's just no way around this. Even if the simplicity of a single switch configuration may suit your requirements, the minimum attention to recovery and availability requirements will force a move to at least two switches for some level of continuous operations.

- *Matching I/O workloads to compatible configurations will require additional planning.* Part of the justification for the SAN will likely be server and storage consolidation. As you consider the combined workloads and their relative I/O, you will begin to see port count requirements, HBA data paths, and storage sizes themselves moving beyond a simple solution. However, for performance and availability, it is best to implement a SAN configuration that facilitates your workloads. Given the flexibility and numerous variables to work with, be prepared to spend considerable time upfront considering these design activities carefully.
- *SAN configurations will be subject to rapid change and enhancement.* One of the most important practices to consider is staying flexible, especially considering the evolutionary tracks SANs are on. Major enhancements will continue to be announced on a constant basis, so a reliable and consistent view and monitoring of the SAN industry will be an effective tool to understanding the evolution of SAN connectivity.

As Storage Area Network technologies and solutions continue to evolve, connectivity of components to key elements within the data center infrastructure can be terribly disruptive to existing operations. However, this same connectivity can also be the most productive in facilitating the scalability of a storage infrastructure as well as establishing a foundation for the evolving enterprise storage fabric.