The Complete Reference

Storage Networks

Chapter 20

Considerations When Integrating SAN and NAS

Chapter 20 addresses the usage of SAN and NAS as an integrated solution. Discussions will cover current data-center realities and experiences but also look to the future uses of integrated SAN and NAS technologies. Because many of the challenges to integration are hidden in managing multiple I/O workloads, just the introduction of this subject can be a complex exercise. This chapter was developed with a broad scope in order to provide a meaningful orientation for those uninitiated to storage strategies as well as prove valuable to those experienced with enterprise storage and I/O workloads.

Integrating storage networking models is, in many ways, like developing a client/server-networking environment; it maintains some fundamental differences given the extreme architectures of SAN and NAS. The fundamental differences are block access in SAN operations versus file access in NAS—block I/O being an extension of the server OS, and file I/O as a functional extension to networking technologies. Beyond these visible differences, the challenges of microkernel compatibility, communications standards, and caching coherency issues will require server and storage vendors to cooperate with synchronization, technology integration, and yet-to-be-developed standards. The current evolution of business applications through the continuing enhancement of web-enabled solutions will incorporate all the attributes of OLTP, batch, and data-intensive I/O. These application attributes will be key drivers in the SAN/NAS product evolution. Taken in context with current data-center infrastructures, the evolution of comprehensive SAN/NAS integrated products will definitely be a future solution.

As combined SAN/NAS solutions support the drive to an application democracy existing in a universal network, the integration of major components of the I/O infrastructure becomes paramount with the development of integrated and faster I/O protocols. This chapter discusses some of the key future technologies that will begin to affect IT storage networking solutions, along the way viewing future I/O infrastructure technologies from two perspectives. First will be a view external to the server with innovations such as iSCSI and InfiniBand. This will be followed by a view internal to the processor components with advancements such as HyperTransport and Rapid I/O initiatives.

The Differences

This book, to a great degree, is about the differences between SAN and NAS architectures, technologies, and solutions. Looking at SAN and NAS products as an integrated solution within the data center, we must view both the differences and similarities from a different perspective—that is, how do we put such solutions together to provide value in the data center. Moreover, what are the applications and, more specifically, I/O workloads that would benefit from both technologies?

The most visible difference is the method in which SAN and NAS process an application's request for data. This represents the block I/O versus the file I/O orientation of each solution.

SANs provide a storage interconnection that enables an attached server to process an I/O as if it were communicating with a direct attached storage device. This allows the attached server's OS to execute a block I/O even though the operation is transported over the FC protocol within the storage network. Consequently, SANs enjoy the benefits and efficiencies of block I/O while enabling the flexibility and added bandwidth performance of being on a network.

From the perspective of the NAS device, the application's request for data is handled through a file request, which is in turn managed by the operating system, where the application is executed and the NAS operating system takes the file request. However, as the request is executed from the NAS operating system, it performs a block I/O to access the data on the NAS device. Block I/O is the native function in how the OS executes an application I/O, whether the application is local or remote (refer to Part II for more detailed information).

Figure 20-1 contrasts these differences to show SAN I/O as an *intra*system transaction and the NAS I/O as an *inter*system transaction. The SAN block I/O requires direct communication with the initiating operating system. The NAS file I/O does not require direct communications with the initiating operating system, but must communicate with another operating system where it obtains only the file attributes. Having accepted that information from the initiating OS, the NAS OS executes a block I/O on their behalf. When viewed from an application standpoint, the application program does not care where the data is, as long as there is some level of addressability to locate and read or write the data. The internals of that operation, which are a block I/O, are handled by the NAS OS where the data actually resides.

The other visible difference is the network where each solution resides. NAS uses existing Ethernet topologies and the TCP/IP-layered network software model, rendering it an IP-based storage transport. SAN, on the other hand, uses the Fibre Channel–layered network software model across a variety of media transports. Nevertheless, it remains a closed system that requires some method of translation to participate in other networks.

The differences of data encapsulation, speed, and reliability between transport protocols of TCP/IP and FC illustrate the challenges of integrating IP-based networks with FC-based networks. TCP/IP based upon data communications architectures is not well suited to the transmission of a native I/O, as indicated in Figure 20-2. The size of IP packets limits its value in the types of I/O workloads it can handle. Enhanced only through the increased media speeds, the limitations of TCP/IP for storage equate to many of the limitations of NAS products. However, the ability to leverage the wide area network capability of TCP/IP networks can provide access to remote storage locations.

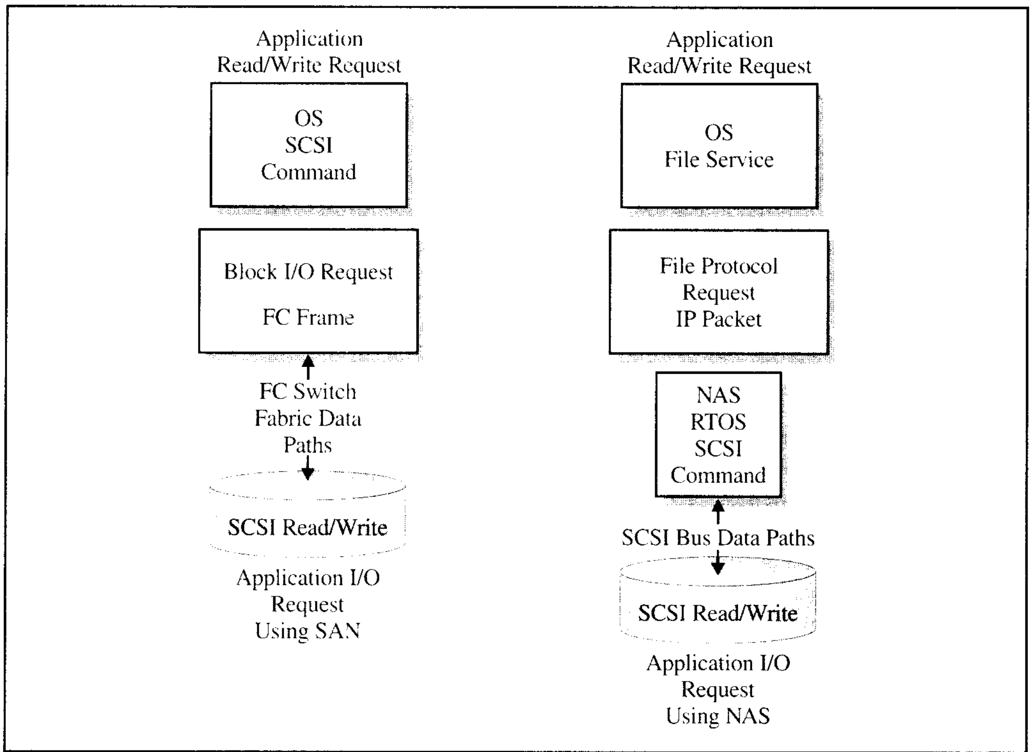
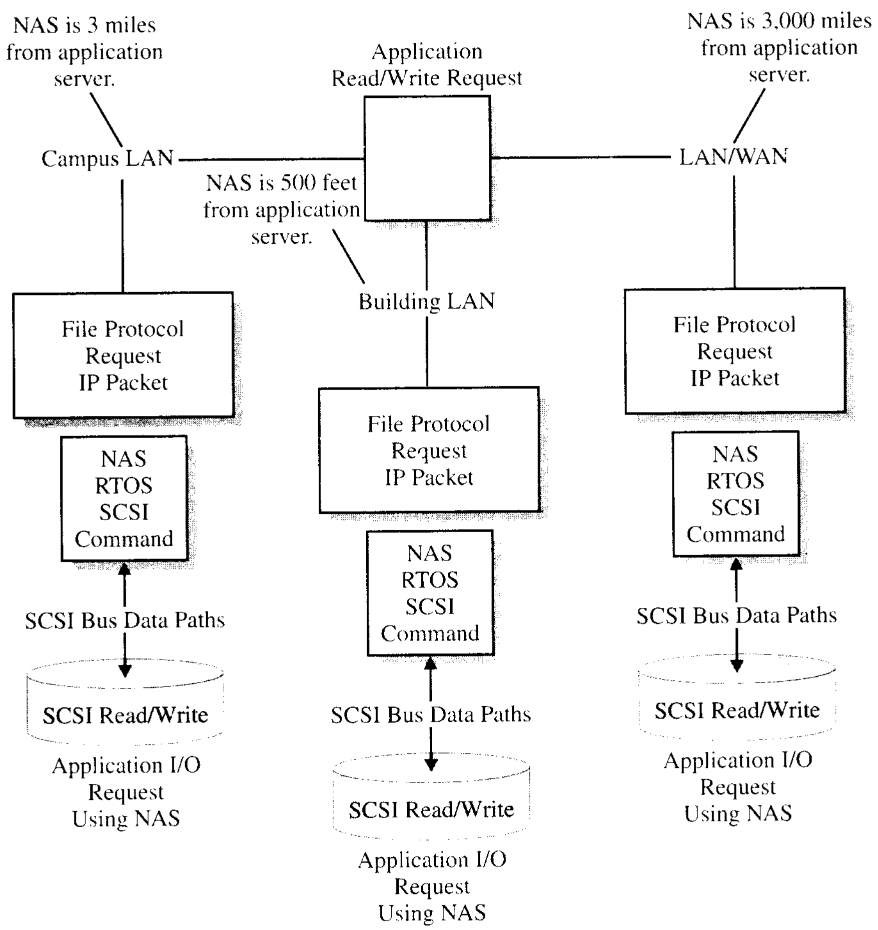


Figure 20-1. *SAN block I/O versus NAS file I/O*

A SAN's use of FC provides the basis for a channel-oriented architecture that facilitates its use as a storage transport (shown in Figure 20-3). The FC classes of service provide the levels of reliability to perform in environments where data integrity is important (for example, most commercial data centers). The frame architecture of approximately 2KB and its transfer rate of 1 Gbps moving to 2 Gbps and beyond extends FC's ability to handle a larger set of I/O workloads. The FC protocol functions as a transport for block I/O within the limitations of SAN networks. The issue with FC SANs is its limitation in providing wide area network usage when accessing remote storage locations.

The Similarities

The need to provide access to more storage resources characterizes the similarities of both NAS and SAN solutions. Both solutions utilize a network to facilitate this requirement. The NAS and SAN products provide multidevice access that interconnects both servers



NAS I/O uses the same overhead for local or remote I/O operations, resulting in TCP/IP and network latency for local I/O operations. However, this also allows applications to perform remote I/O from WAN distances.

Figure 20-2. *NAS I/O local limitations and remote strengths*

and storage. With respect to their local operations, they both perform block I/O, as described previously.

The use of a network ties these storage models together and provides a foundation for integration activities. Despite the evolution of data communications networks, the integration of disparate network topologies and protocols remains with us today. Accomplished through the use of bridging mechanisms at the lower levels of the

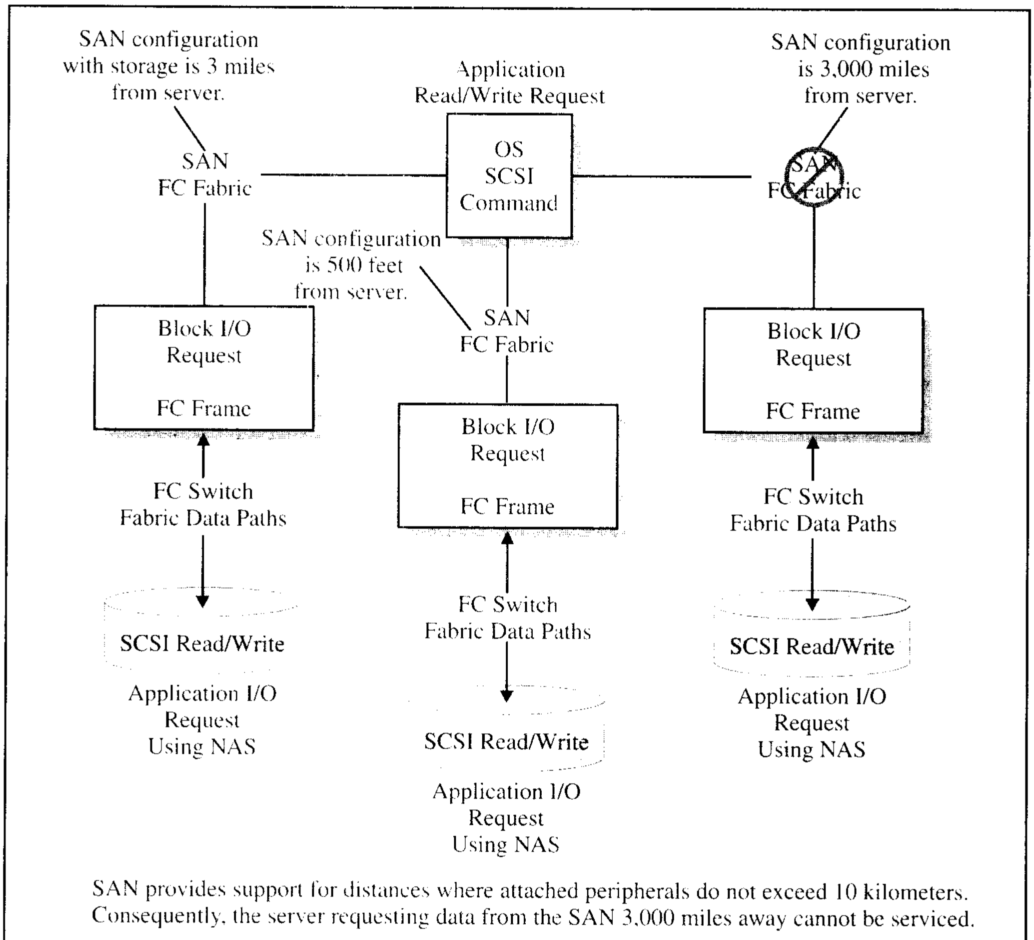


Figure 20-3. *SAN I/O local strengths and remote limitations*

network transport layers, data networks were able to interoperate, allowing the scalability of multiuser networks. The evolution of these elementary bridging technologies to router products facilitated the delivery of client/server data packets as well as sophisticated traffic management and communications reliability.

Although we have yet to see this sophistication within the storage networking world, the use of bridges and routers to accommodate disparate storage protocols in the SAN has been an accepted and widely used mechanism (see Chapter 15). Given their nature of communicating at the file level, NAS can leverage existing bridge, router, and switching technologies as an inherent characteristic necessary to utilize standard Ethernet TCP/IP networks.

In terms of value, they provide similar justification in what they offer—that is, storage consolidation, server consolidation, centralized management, larger storage capacities, multiple application access, and storage scalability. These value propositions center around the effective usage of each storage model, respectively.

This requires the ability to identify and analyze the characteristics of the I/O workload. As discussed in Chapter 17, specific I/O workloads are better suited to either NAS or SAN technologies. Which leads us to the underlying question: Why does anyone need to combine a NAS and SAN configuration? Part of the answer lies in the simple rationale as to why disparate data communications networks need to integrate. The reason is that those users with Macs need to access the same applications and data as PC users. Further, UNIX workstation users operating within an NFS environment need to access shared information on the PC networks. Users of mainframe applications require a separation of complexity from legacy applications that are increasingly integrated with front-end client/server systems. Ultimately, the Internet prompted uniform access to information both locally and worldwide.

The Need to Integrate

If we accept the premise that storage networking will evolve to some degree like data communications networks, then we have to accept the state of today's technologies as being at the forefront of this evolution. Consequently, elementary and sometimes proprietary and localized solutions give way to integration, allowing applications transparent access to storage resources regardless of the model.

If we view this from a pure user requirements perspective, the evolution of web services and the integration of local systems inside data centers to accommodate this will drive the need to integrate storage networks to a uniform access model. Not that any of the complexities will go away; actually, they will be more complex—just as they were with data communications networks. However, the necessity to consider lower-level access fundamentals, such as block I/O versus file I/O, will dissipate as increasing layers of storage access fold over the technical uniqueness of SAN and NAS architectures.

Web services will demand an integration of I/O workload types by their very make-up. Service that combines transactional synchronous responses will be coupled with secondary and tertiary transactions that respond with a series of asynchronous message-oriented requirements. Each of these will require a set of I/O workloads commensurate with its processing characteristics. However, each will be integrated in its application processing responsibilities. Thus, we may find the future of I/O workloads being driven by web services. Common personal and business applications will be executed by an integrated I/O system with the flexibility to route the appropriate I/O requests to the correct set of data using the correct access mechanism—for example, block, cache, or file.

These are the driving factors to integrate SAN and NAS. The flexibility to handle multiple workload requirements demands that storage networking have multiple access mechanisms. The industry is starting to experience this as NAS solutions begin to integrate both file access and block-level access. This is currently being accomplished through the integration of NAS server configurations with FC disk arrays. This sets the foundation for bilateral access to the data via file access through existing IP network requests and block I/O access through the FC disk array. There are many issues surrounding these early solutions, such as the need to provide extensions to OS and file system functions, application interfaces, and additional levels of standards, that are necessary to get everyone on the same page. However, it's a start.

Future Storage Connectivity

As the SAN/NAS core technologies evolve, they will be affected by additional external changes happening with the microprocessor and networking industries. These changes can be viewed from a perspective external to the server, many of which will be network driven. Other changes will occur from an internal perspective where technology advancements in chip-to-chip communications takes place within the microprocessor system.

iSCSI, LAN/WAN Storage

One limitation of the SAN I/O architecture revolves around distance. Performing a block I/O remotely requires the resources of the NAS solution and the inherent overhead of file I/O processing. If there was a method of extending the server I/O operation remotely, this would allow native I/O operations to take place without the necessity of an alternative network and required special devices, as in FC SANs, or the special server devices associated with NAS. This appears to be a very efficient and cost-effective solution for remotely accessing data in storage arrays.

This also describes an upcoming standard and product solution called iSCSI. The iSCSI solution (*iSCSI* denoting *Internet SCSI*) provides a standard for transmitting SCSI commands through an IP-based network. This would allow servers to send a block I/O request through an existing TCP/IP network and execute the SCSI storage read/write operations on a remote storage array.

The iSCSI configurations require special Network Interface Cards (NICs) that provide the iSCSI command set at the storage array end (as shown in Figure 20-4). This facilitates data and SCSI commands to be encapsulated into an IP packet and transmitted through an existing IP network, thereby bypassing file-level processing protocols and additional server-to-server communications inherent to NAS solutions. On the surface, this solution provides an effective mechanism for disaster recovery operations, data replication, and distribution.

However, this develops an increasing number of issues in implementing and managing an iSCSI configuration. First is the security of sending an unsecured block

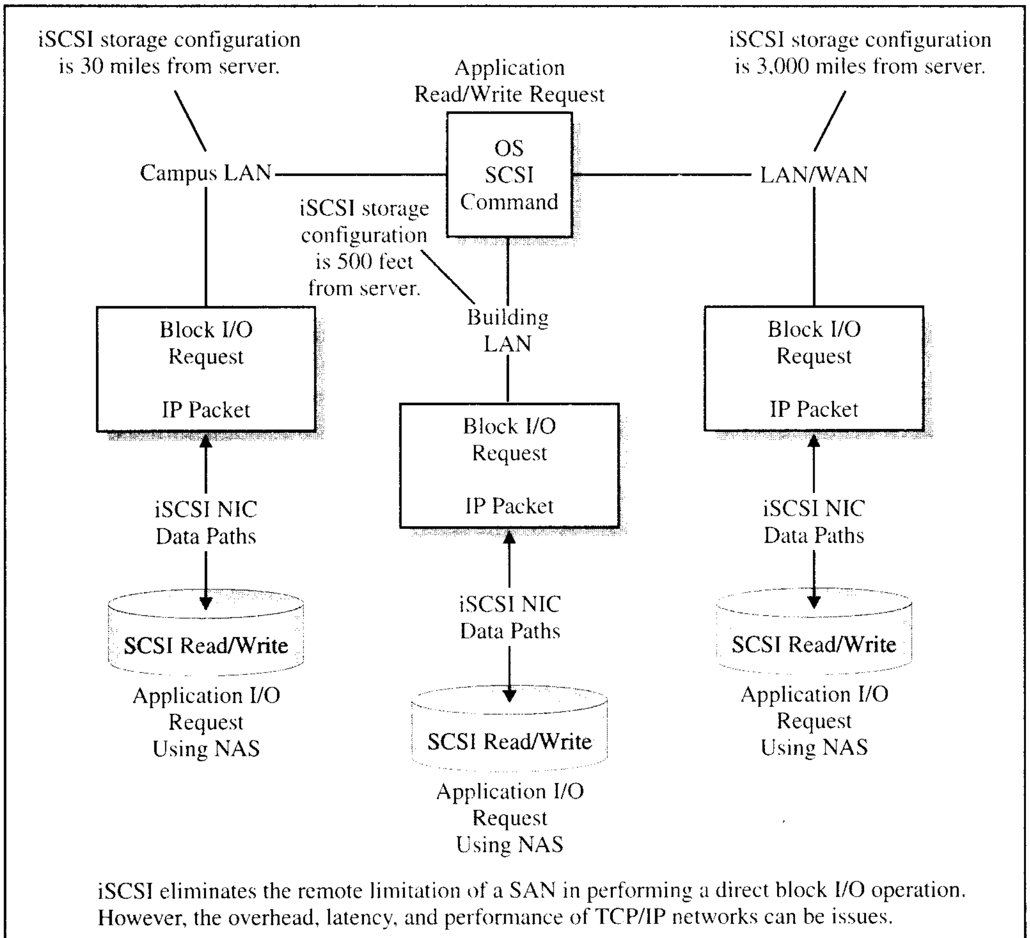


Figure 20-4. *iSCSI configurations showing remote block I/O operation*

I/O across a public network. An associated aspect is the reliability of the IP packet transmission process. Overall, the ability to access data within unsecured environments that can withstand packet transmission errors may be the I/O workload that suits this future technology.

Extending I/O functionality beyond the server has been a goal of system vendors for many years in their struggle to replace the aging bus mechanisms that form the foundation of our current computing platforms. Multiple initiatives suggesting that I/O should be disengaged from the computing elements have coalesced into the industry initiative of InfiniBand. InfiniBand provides a switched fabric environment very similar to the FC SAN switched-fabric environment. However, the InfiniBand standard encompasses all I/O from the processor into the external connectivity of networks and storage devices. InfiniBand on the surface can be viewed as a replacement for the aging PCI bus technologies (see Chapter 7). However, due to its scalable I/O infrastructure, it provides a shift in computing fundamentals by providing a switched-fabric architecture that allows devices to communicate with processor nodes with increased bandwidth, reduced latency, and throughput.

InfiniBand, the Universal Bus

InfiniBand is an I/O standard developed by seven of the computer industry's key players: IBM, Compaq, HP, Sun Microsystems, Microsoft, Intel, and Dell. InfiniBand replaces traditional PCI bus technology with a switched fabric (think network), which allows peripherals (such as storage and client/data networks) to communicate within a network (think switched fabric) of servers. It also allows InfiniBand-enabled servers to utilize the same fabrics (network) to communicate amongst themselves.

InfiniBand is poised to cause a fundamental change to the data center, as we know it. As its scalable I/O infrastructure is integrated, the total cost of ownership (TOC) models and metrics will change, as will how applications are deployed and managed. IT departments, especially IT management, must understand this infrastructure if they hope to leverage intelligent adoption strategies. InfiniBand vendors, on the other hand, must move toward systems solutions in order to overcome the initial wave of anxiety and reduce the complexities of integration and adoption.

As illustrated in Figure 20-5, the I/O fabric is connected from the server through a Host Channel Adapter (HCA), while connections to peripherals move through a Target Channel Adapter (TCA). These components communicate through a switch that routes to all the nodes that make up the fabric. In addition, the I/O fabric enhances communication within the fabric by using remote direct memory access (RDMA) methods to facilitate I/O applications operations. InfiniBand links are serial, segmented into 16 logical lanes of traffic. Given this architecture, the bandwidth for each link eclipses all other data transports now available.

Why is InfiniBand important? The traditional deployment of servers has come through vertical implementation. In most cases, expanding an existing application requires the installation of a new server. Each time this happens, total overhead and the latency needed to support the application increases. Left unchecked, the operating system services and network overhead required to support the additional servers can consume as much or more processing power than the application. Scalability of the application ultimately becomes non-linear as the number of servers increases.

From an I/O standpoint, the vertical deployment of servers poses additional problems. As application data grows, so must the required I/O bandwidth and paths needed to get to the data. If the server cannot provide either of these, the application suffers from having to wait on data to be processed. This same phenomenon happens when users are added to the application, or there is an increase in the users accessing

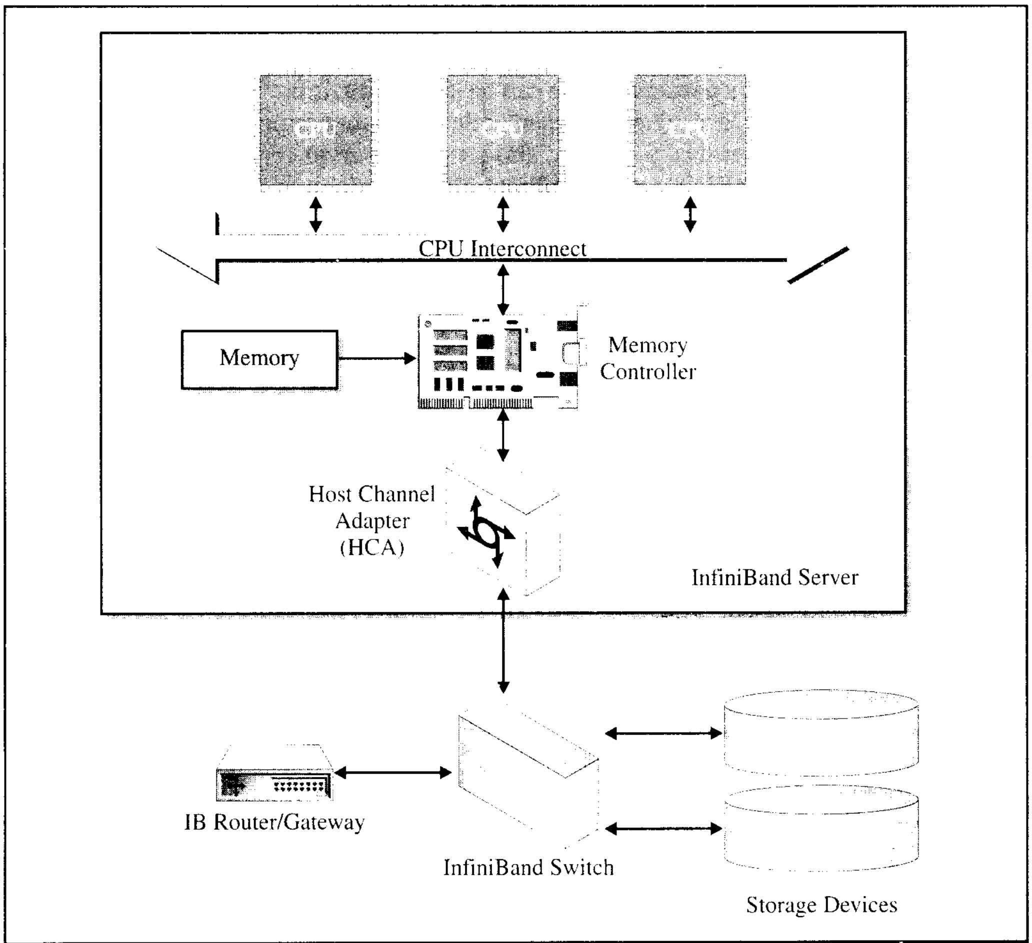


Figure 20-5. *InfiniBand's I/O fabric*

the application, as often happens with Internet applications. The required bandwidth and paths must be in place, otherwise the application suffers by having to wait on large queues of user requests.

These conditions are further exacerbated by the inefficient usage of resources—in both hardware and software—needed to provide an adequate configuration for supporting the workload. All of this places additional burdens on data-center personnel, who must handle the complexity of installation, maintenance, system upgrades, and workload management.

eclipsed by new serial connections that form a switched environment. HyperTransport is a vendor-sponsored initiative that proposes a standard set of specifications when building an internal switched bus to increase I/O flexibility and performance. Staging and moving data in and out of CPU processing units is just as important as the performance of external I/O.

The HyperTransport standard defines an I/O protocol that facilitates a scalable interconnect between CPU, memory, and I/O devices. As a switching protocol, the architecture is defined within a layered network approach based to some degree on the OSI network model. Consequently, there are five distinct layers of processing that direct how internal components communicate.

Figure 20-7 depicts an example of the deployment of HyperTransport architecture. In this type of configuration, an I/O switch handles multiple data streams and interconnections between CPUs and memory. The configuration is extended to support external bus and network connections such as InfiniBand, PCI, and Gbe Ethernet. This enables processing and memory components to have scalability and bandwidth flexibility similar to switched network environments like InfiniBand and FC SANs.

Storage Integration

Is it necessary to have a SAN and NAS model? In the near term, both storage networking models will be with us. As we continue to support the present application mix of OLTP- and batch-type workloads, it's necessary to provide an optimized access method. This will continue to drive the segregation of data between file and block level access and require increased duplication and overhead. The results will be a continuation of storage demand and growth, albeit an inflated use of resources.

The future of SAN/NAS integration will be driven by the increased need to utilize existing data for multiple purposes or applications. The slowing of duplication requirements is a critical requirement for any integrated front-end storage access method. As some applications require file access from remote locations, other local applications will require read/write access on a fundamental block I/O level. The ability to provide these access levels through a combined storage network utilizing a file or block access will be the next storage system.

What is missing today are the components to make that happen. Figure 20-8 shows a future configuration where an integrated file system serves as both file- and block-level access to a common storage network. Future components that will facilitate this solution will be a shared file system with storage network coherency, increased state-aware disk devices, and real-time fault resiliency in recovery fabrics. The storage network components will communicate with storage access appliances that serve

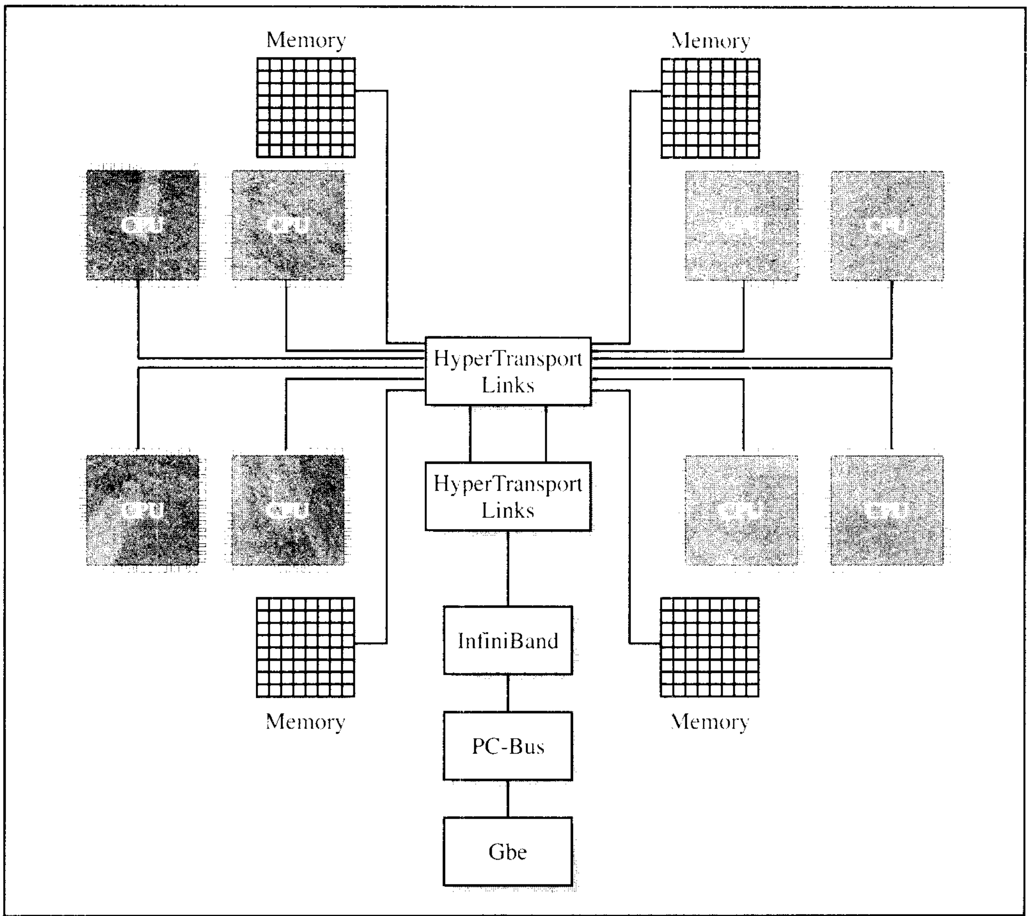


Figure 20-7. *HyperTransport: a future high-speed CPU interconnect*

applications with either block or file access. Applications will log in and register with the storage network regarding their preferences for recovery, access, and security.

In looking toward the future, use caution when developing storage networking configurations. The useful life of storage networking technologies has only been a year and a half in the short end and three years in the long term. The point must be made that we remain at the forefront of storage innovations that will be increasingly vulnerable to both changes in microprocessor technologies and the evolving universe of web services.

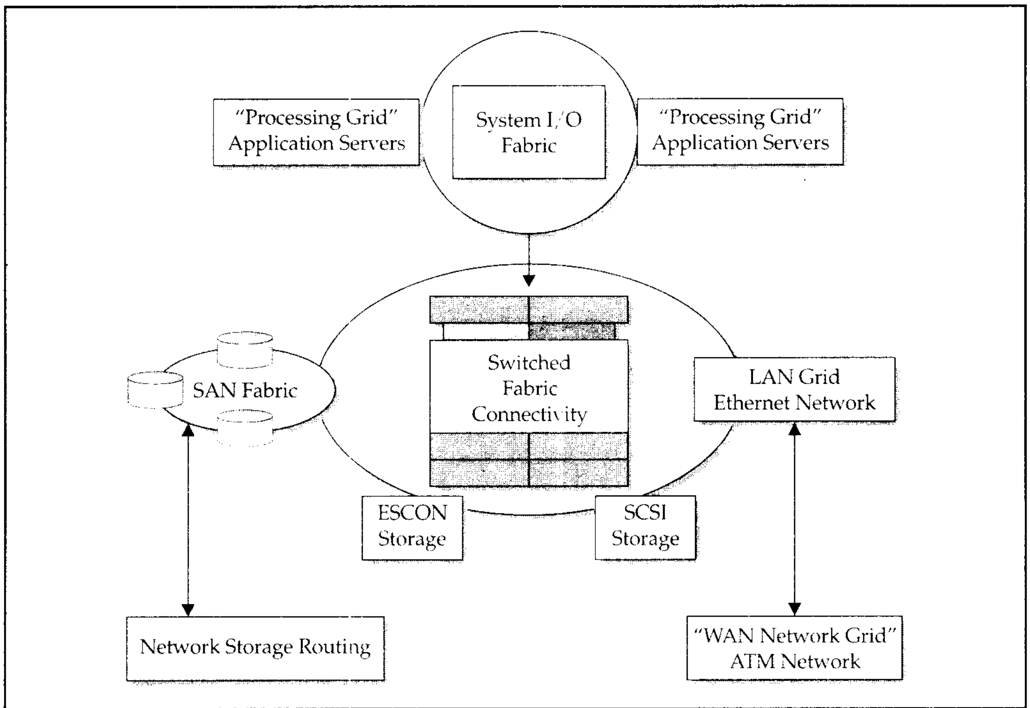


Figure 20-8. *A storage network of the future*