1 Introduction

DESIGNING STORAGE AREA NETWORKS requires an understanding of the upper-layer applications that depend on storage resources as well as the storage architectures available to satisfy them. Although the specific application requirements may vary from customer to customer, a framework has been developed to help storage architects determine which storage solutions may be most appropriate. The **SNIA Shared Storage Model** (SSM) provides a guide to understanding the relationships between host applications, storage networks, and storage facilities.

1.1 Using the SNIA Shared Storage Model

Sharing storage resources among multiple servers or workstations requires a peer-to-peer network that joins targets to initiators. The composition of that network and the type of storage data traversing it vary from one architecture to another. Generally, shared storage architectures divide into **storage area networks** (SANs) and **network-attached storage** (NAS). For SANs, the network infrastructure may be **Fibre Channel** or **Gigabit Ethernet**, and the type of storage data being transported is block **Small Computer Systems Interface** (SCSI) data. For NAS, the network infrastructure is typically Ethernet (**Fast Ethernet** or Gigabit Ethernet), and the type of storage data carried across the network is file-based. At the most abstract level, then, the common denominator between SAN and NAS is that both enable sharing of storage resources by multiple initiators, whether block-based or file-based.

Understanding the role of direct-attached, SAN, and NAS solutions and fitting them into a coherent IT storage strategy can be a challenge for both technologists and the managers who sign off on major storage acquisitions. The SNIA Shared Storage Model offers a useful framework for understanding the relationships between upper-layer applications and their supporting storage infrastructures. The ability to map current storage deployments to proposed solutions helps to clarify the issues that storage architects are attempting to address and also creates a framework for projecting future requirements and solutions.

As shown in Figure 1–1, the Shared Storage Model establishes the general relationship between user applications that run on servers and hosts and the underlying **storage domain**. Applications may support user activity such as processing online transactions, mining databases, or serving Web content. Storage-specific applications such as management, backup, cluster serving, or disk-to-disk data replication are grouped within the **services subsystem** of the storage domain. The model thus distinguishes between end-user or business applications at the upper layer, and secondary applications used to monitor and support the lower-level storage domain.

The storage domain subdivides into three main categories: the file/record subsystem, the block aggregation layer, and the block subsystem. The **file/record subsystem** is the interface between upper-layer applications and the storage resources. Database applications such as SQL Server and Oracle use a record format as processing units, whereas most other applications expect to process files.

Whether useful information appears to the upper-layer application as records or files, both formats are ultimately stored on disk or tape as contiguous

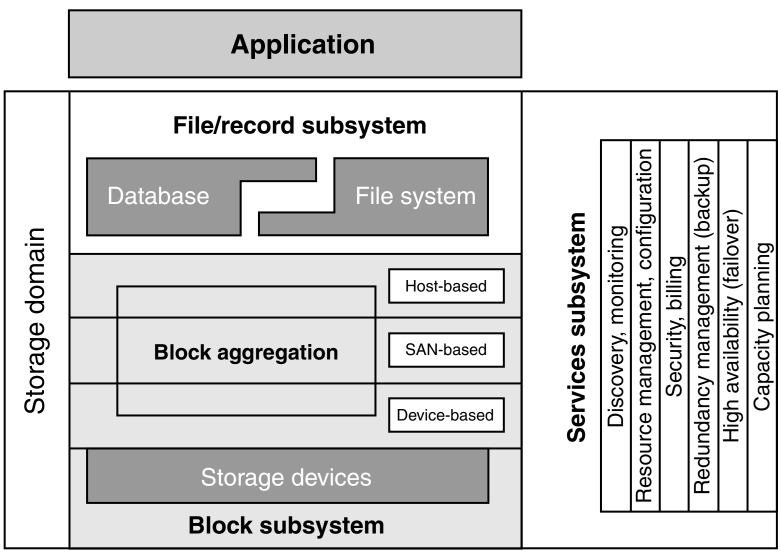


Figure 1–1 The SNIA Shared Storage Model overview

bytes of data known as **blocks**. The size of a data block can vary from system to system, as can the method of mapping records or files to the blocks of bytes that compose them. In all cases, the storage domain requires some means of associating blocks of data with the appropriate record or file descriptors. This function is depicted as the **block aggregation** layer, which may be based on the host system, within the storage network, or at the storage device. Having been identified with specific records or files, the blocks themselves are written to or read from physical storage, shown in the Shared Storage Model as the **block subsystem**.

Also part of the storage domain, but positioned as an auxiliary subsystem, the **services subsystem** contains a number of storage-specific functions, including management, security, backup, availability, and capacity planning. These services may appear as integrated functions in storage products or as stand-alone software applications used to monitor and administer storage resources. A particular block aggregation method may require a unique management service. NAS devices, for example, may perform backups somewhat differently than their SAN brethren, requiring a separate type of application in the services area.

Considerable engineering effort has been invested in creating products to reliably integrate the file/record, block aggregation, and block subsystems into viable shared storage solutions. Perfecting gigabit transport of block data to ensure data integrity over very high speed serial links, for example, took years of standards development and verification testing. Development of the services subsystem was not feasible until these basic infrastructure issues were resolved. Consequently, the creation of management, security, and other auxiliary services has trailed somewhat behind the deployment of shared storage networks. Today, the lack of enhanced management services inhibits further expansion of shared storage in the market, and the focus of the storage industry is shifting from infrastructure to auxiliary services that will make it easier to install and support shared storage solutions. In using the Shared Storage Model to position your own storage solutions, you should anticipate that new products will become available in the services subsystem to enable management, security, and other functions.

With the layered architecture of the Shared Storage Model as a guide, it is now possible to insert server and storage components to clearly differentiate between direct-attached, SAN, and NAS configurations. As shown in Figure 1–2, **direct-attached storage** (DAS) extends from the server to the storage target by parallel SCSI cabling. This is the most common storage configuration today, although shared storage is expected to gradually displace DAS

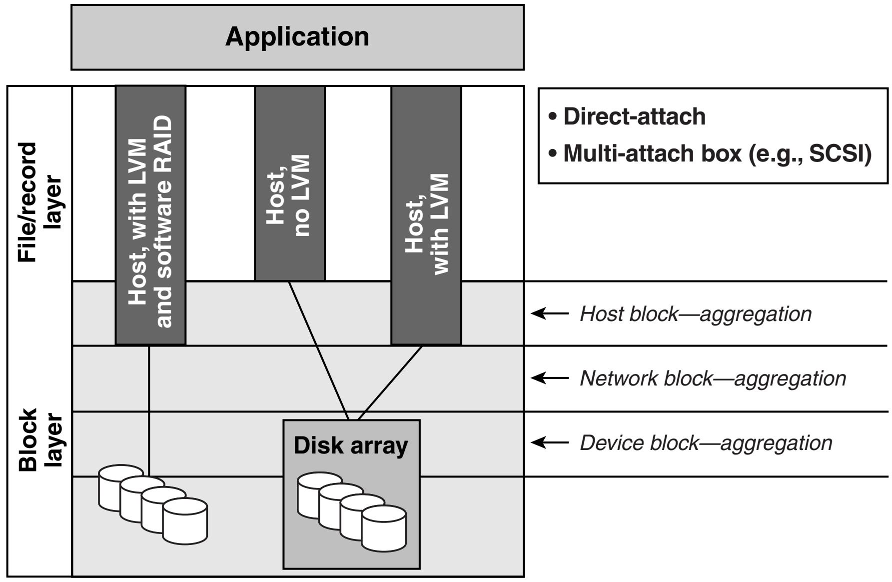


Figure 1–2 Direct-attached storage in the SNIA Shared Storage Model

over the next few years. In this example, the left side of Figure 1–2 shows a server with **logical volume management** (LVM) and software RAID (**redundant array of independent disks**) running on the host system. As the server receives information from the application that must be written to disk, the software RAID stripes blocks of data across multiple disks in the block layer. The host thus performs the block aggregation function. Whereas software RAID executes the mechanics to striping blocks of data to multiple disks, the logical volume manager presents a coherent image of the data to the upper-layer application in the form of volume (for example, M: drive), directories, and subdirectories.

On the right side of Figure 1–2, servers are shown that have a SCSI attachment to a disk array containing an integrated RAID controller. In this instance, the host is relieved of the task of striping data blocks via software RAID. Instead, the array itself performs this function. Consequently, it is shown overlapping the block and block aggregation layers.

SANs alter the relationship between servers and storage targets, as shown in Figure 1–3. Instead of being bound by direct parallel SCSI cabling, servers and storage are now joined through a peer-to-peer network. As in direct-attached storage, logical volume management, software RAID, and

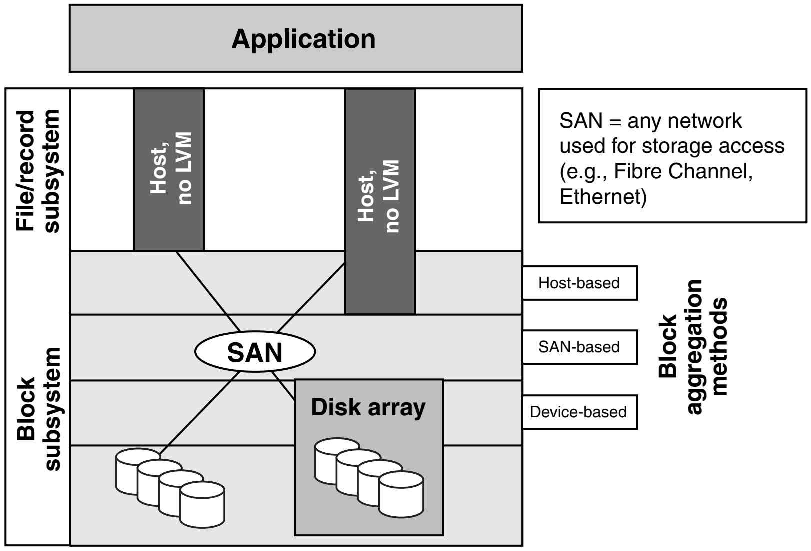


Figure 1–3 Changing the relationship between servers and storage via a SAN

hardware RAID still play their roles, but the connectivity between servers and storage devices now allows you to attach any server to any storage target. The exclusive ownership of a storage resource by a server, symbolized by the umbilical tether of SCSI cabling, is no longer mandatory. You can assign shared storage resources at will to designated servers, and you can alter the relationships between servers and storage dynamically to accommodate changing application requirements.

Replacing direct-attached connections with a more flexible network interconnection enables new storage solutions. For example, you can consolidate storage, share resources for both tape and disk, and cluster multiple servers for high availability. Depending on the SAN topology you use, you can also scale networked storage to higher populations of servers and storage devices. A large disk array, for example, can support tens or hundreds of servers in a single SAN, amortizing the cost of the array over more hosts and streamlining storage administration.

The SAN infrastructure can be Fibre Channel, Gigabit Ethernet, or, with the appropriate IP storage switches, both. To accurately represent actual customer deployments, this model could also depict multiple SANs servicing various upper-layer applications and providing attachment to common or distinct storage resources, as shown in Figure 1–4. For heterogeneous environments, it

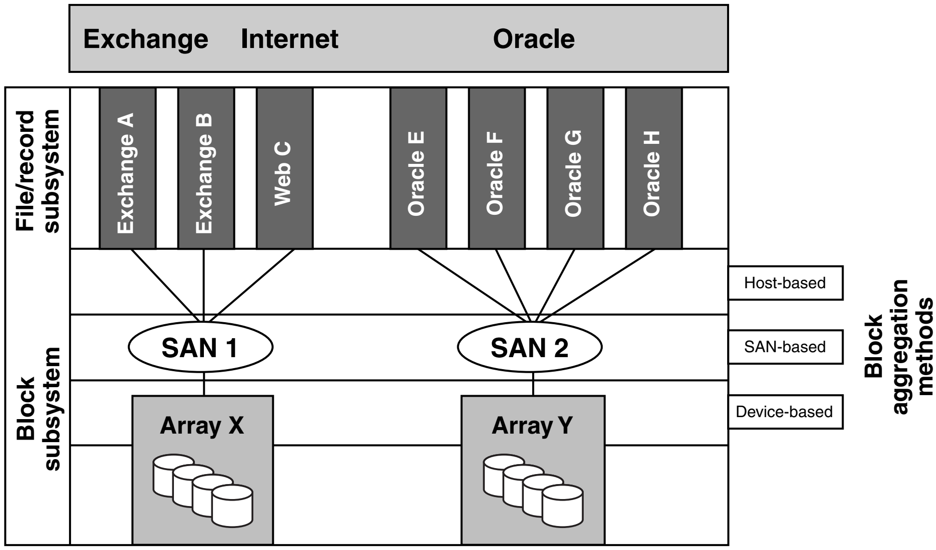
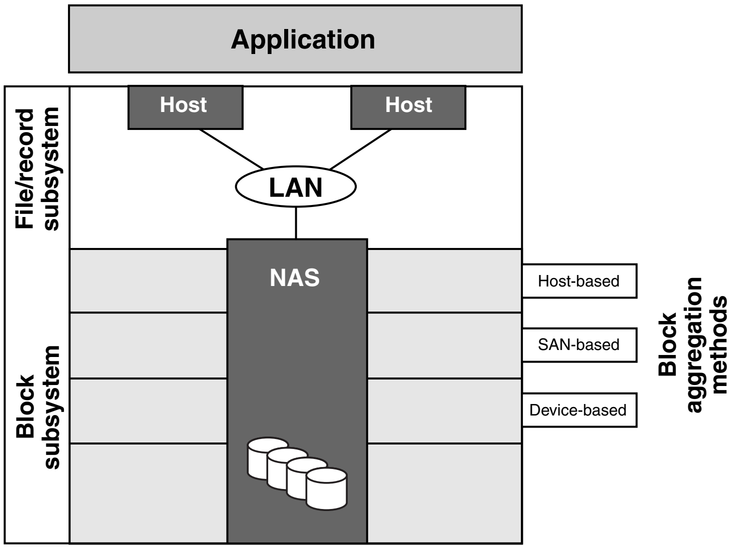


Figure 1–4 Multiple SANs within the Shared Storage Model

might be useful to call out which SANs use Fibre Channel fabrics exclusively and which involve a mix of Fibre Channel and IP components.

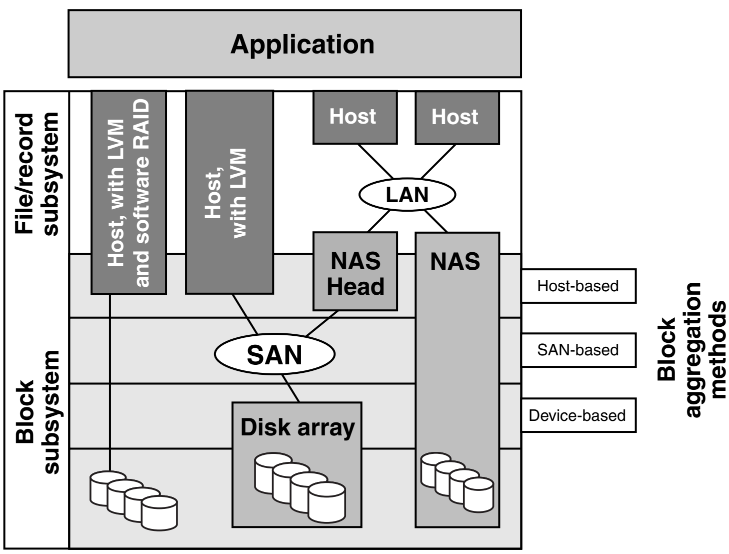
The Shared Storage Model positions NAS devices partly in the file/record subsystem layer, extending down to the block subsystem. NAS devices serve up files and so naturally include the block aggregation functions required to put data to disk. As shown in Figure 1–5, a NAS device is essentially a file server with its own storage resources. NAS devices typically use NFS (**Network File System**) or CIFS (**Common Internet File System**) protocols to transport files over the **local area network** (LAN) to clients. For NAS, the internal SCSI access transport between the NAS intelligence (head) and its physical disk drives is transparent to the user. The NAS disk drive banks may be integrated drive electronics (IDE) at the low end, SCSI attached, or Fibre Channel. Network Appliance, for example, uses Fibre Channel arbitrated loop disk drives for mass storage. This is in effect a SAN behind the NAS device, efficiently serving up blocks of data that the NAS head assembles into files for NFS or CIFS transport.

The Shared Storage Model in Figure 1–6 captures the general relationship among direct-attached, SAN, and NAS solutions. Because many enterprises have a mix of direct-attached, SAN, and NAS storage solutions, this



Copyright © 2000, Storage Networking Industry Association (work in progress)

Figure 1–5 The role of NAS in the Shared Storage Model



Copyright © 2000, Storage Networking Industry Association (work in progress)

Figure 1–6 Using the Shared Storage Model to show a mix of DAS, SAN, and NAS solutions is a useful framework for associating specific upper-layer applications with current and planned storage configurations. Applications currently running on direct-attached storage, for example, can be redrawn with SAN or NAS components, or multiple direct-attached storage arrays can be redrawn with consolidated SAN-based arrays and SAN-attached hosts.

The practical value of applying the SNIA Shared Storage Model is in defining the current user applications and supporting storage resources and mapping those to new infrastructures. This in turn provides a comprehensive overview of customer storage requirements and options and can be used as a blueprint for further development of the network.

1.2 Example: Carlson Companies

An excellent example of this practical application of the model is provided by the IT staff of Carlson Companies. Carlson Companies is one of the largest privately held companies in the United States, with more than 180,000 employees in more than 140 countries. Its IT division, Carlson Shared Services, acts as a service provider to its internal clients and consequently must support a spectrum of user applications and services.

In formulating a global IT strategy for networking and storage, the Carlson storage architects began with their current applications and the supporting server/storage infrastructure. Originally, Carlson's business applications were supported on a combination of direct-attached storage arrays and internal storage. As shown in Figure 1–7, database applications ran on servers attached to large storage arrays, and other applications ran on servers with internal storage.

Overlaid on the SNIA Shared Storage Model, Carlson Companies' original storage configuration clearly demarcates which applications access data via records, which via files, and which via device-level block aggregation or host-supported block aggregation. The first step in determining which applications might be better served by shared storage and which shared storage technologies are most appropriate for specific upper-layer applications is to diagram the upper-layer applications and their storage access methods.

By using the Shared Storage Model, the storage architects clarified the relationship between application requirements and storage resources and also revealed the administrative overhead incurred by dedicated storage units. The system required four large storage frames and six additional arrays to support Carlson's data center applications. Each storage unit required separate administration as well as maintenance contracts from the

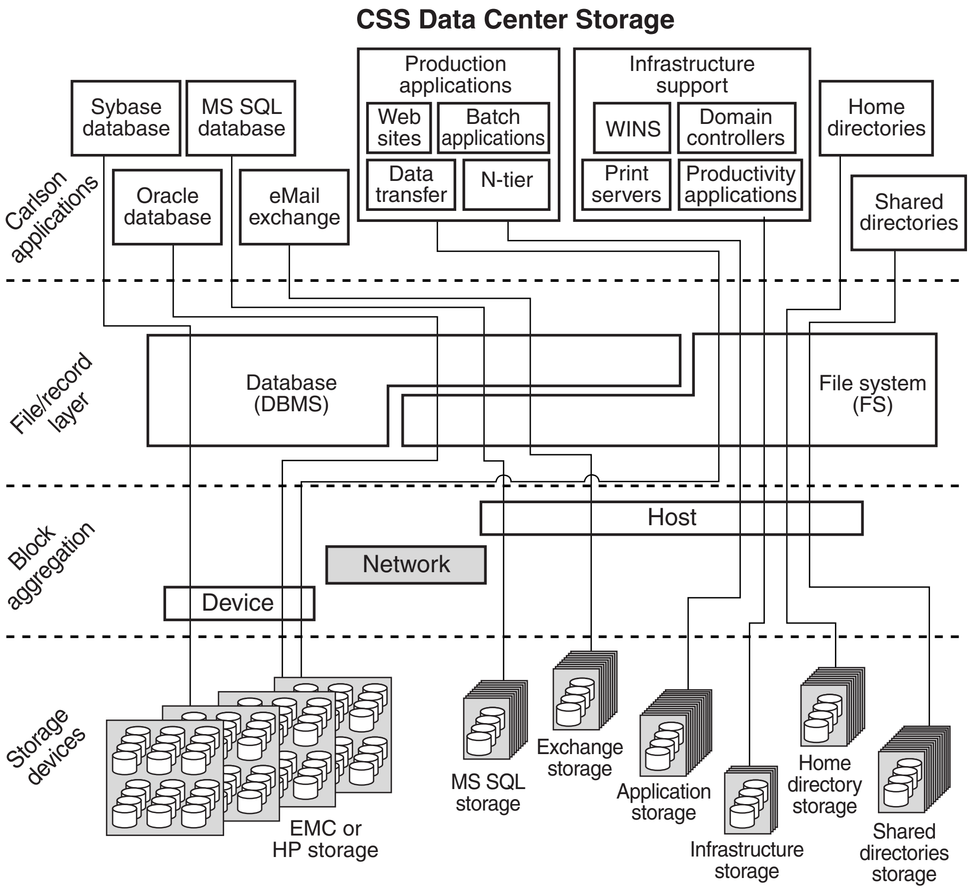


Figure 1–7 Carlson Companies' storage access prior to installing shared storage

supplying vendor. Because storage consolidation is one of the key benefits of SANs, the Shared Storage Model offered a framework for modifying the components in the block subsystem to reduce the number of storage devices while still serving Carlson's business applications.

As shown in Figure 1–8, analysis of Carlson's former storage deployment led to a proposed shared storage configuration that leveraged both SAN and NAS technologies to streamline storage administration and enable more efficient use of storage capacity. This design replaces the four large direct-attached storage frames with a single SAN-attached storage array supporting more than 10 terabytes of data. In this example, Oracle and other Carlson custom applications now share a single, highly available storage resource

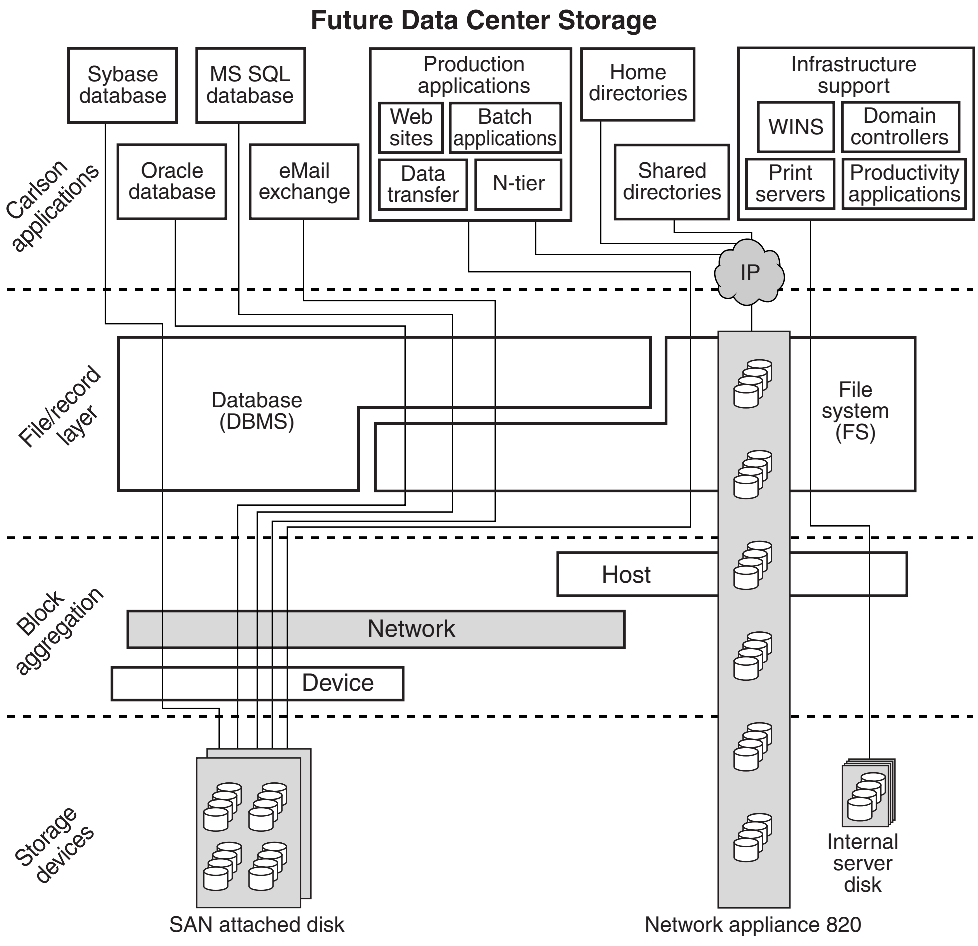


Figure 1–8 Carlson Companies' storage access after installing SAN and NAS storage

that is more easily administered and offers more economical maintenance support compared with the four direct-attached storage frames it replaced.

Shared directories and other internal applications have moved from direct-attached, dedicated storage to large NAS filers. In this instance, NAS was chosen for its cross-platform support and for its ability to serve remote NAS access for Carlson clients. Infrastructure applications that have no compelling reason to reside on either NAS or SAN remain on internal storage, as shown on the right side of the diagram.

Remote storage access is also a significant factor for organizations with geographically dispersed sites. In Carlson Companies' global IT strategy, safeguarding corporate data generated at remote locations is essential for business continuance. Relying on each remote office to back up its own data locally offers no guarantee that backups are successfully completed or that data can be restored in the event of failure. Consequently, Carlson decided to streamline backup operations by bringing remote backup operations into its central data center.

As shown in Figure 1–9, the Shared Storage Model is used to define four separate remote office backup scenarios. Remote offices with no local file servers have valuable corporate data on internal drives of laptops and desk workstations. Typically (and unfortunately) for most enterprise networks, this data is rarely backed up, resulting in loss of productivity if a laptop or desktop disk drive fails. For these sites, Carlson uses software that enables **block change backup**, meaning that only the data that has changed since the last backup is sent to the data center backup facility. This arrangement reduces the amount of data that must be sent across the wide area network

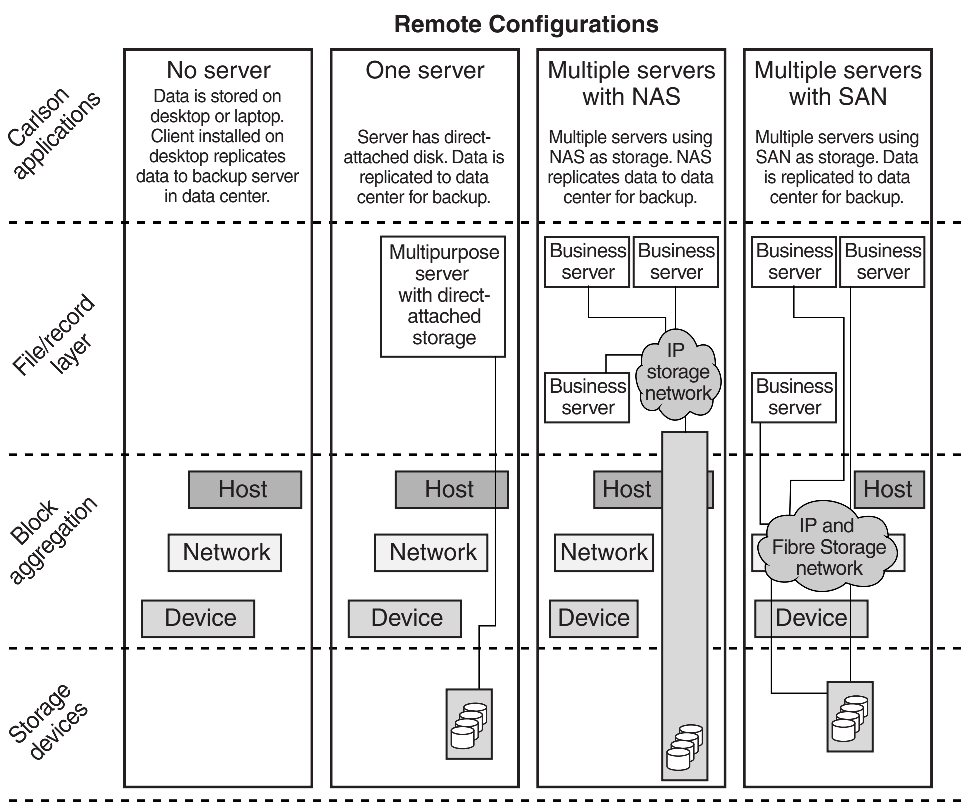


Figure 1–9 Carlson Companies' remote backup scenarios

and accomplishes the goal of securing dispersed corporate data assets. Similarly, remote sites with a local file server can be backed up using block changes, again reducing the amount of traffic generated over the WAN.

In the remaining two scenarios, a remote office may have a local NAS or SAN installed. For NAS, Carlson is leveraging vendor-supplied file backup utilities to replicate data from the remote site to the data center NAS equipment. At the data center, the NAS-based data can then be centrally archived to tape. For remote SAN installations, Carlson is using IP storage technology to transfer locally generated Fibre Channel SAN data to the corporate SAN. In Carlson's case, the data center SAN is built with an IP SAN core, thereby facilitating IP-based block backup between remote and central SANs.

For Carlson Companies, the SNIA Shared Storage Model offers a coherent framework for analyzing current and future storage requirements for both the data center and its dispersed remote locations. For storage architects, the Shared Storage Model is a tool for defining application and storage relationships while enabling drill-down to more detailed storage-specific issues. For CIOs and IT executives, the model provides a basis for understanding the composition of their storage deployments and seeing how proposed acquisitions contribute to broader strategic IT goals. Additional information on the SNIA Shared Storage Model is available in Appendix E as well as the SNIA Web site at www.snia.org.

1.3 Text Overview

The first edition of this work attempted to provide a comprehensive overview of SAN technology for both technical and nontechnical readers. Because SANs were relatively new arrivals on the storage scene, it was possible to accomplish this goal in a fairly short work by using a dense writing style that one reader has called "information compression." The current work necessarily includes additional discussion of new SAN technologies that have emerged over the past three years, but it attempts to preserve an information compression ratio that delivers more knowledge in fewer pages. Because few people in today's information technology space have the time to read a work cover to cover, it is hoped that this edition of *Designing SANs* will be a useful reference for readers who need ready access to specific information on various SAN topics.

The chapter progression begins with lower-level technical overviews of storage and networking primarily related to the underlying SAN plumbing that supports higher-level applications. This discussion includes both abstract standards and the products developed from them. SAN infrastructure components that were previously based on Fibre Channel alone now include IP-based solutions, so additional space in the current work is devoted to heterogeneous SAN design options. Following the infrastructure discussion, middle-layer SAN issues, such as storage-specific applications, management, and virtualization, are reviewed. With these additional concepts in place, it is then possible to discuss concrete implementations of SANs in customer environments. The text concludes with a general overview of outstanding issues yet to be resolved in the areas of standardization, interoperability, and management, along with speculation on the future direction of storage area networking as the adoption of SANs expands throughout the market.

Chapter 2 provides an overview of both storage and networking concepts and explains how the fusion of these technologies has created new means to solve data storage issues. As in any dialectical synthesis, the resulting product has new attributes quite distinct from its constituent parts. Storage has thus placed new demands on networking, and networking has fundamentally altered storage relationships. As a fusion technology, storage networking requires a basic understanding of networking and storage—a need addressed by this chapter—as well as new concepts addressed by later chapters.

Storage area networking has deep roots in Fibre Channel technology. Fibre Channel was the first serial transport to successfully solve gigabit transport, data encoding, optics, transport protocol, and topology issues required for high-performance shared storage applications. Chapter 3, Fibre Channel Internals, reviews the lower-level operation of the Fibre Channel physical and protocol layers, including framing, class of service, flow control, and naming and addressing conventions.

Chapter 4 provides the foundation for understanding Fibre Channel topologies that have been employed for SANs. Today, most Fibre Channel SANs are based on switched fabrics. Chapter 4 therefore focuses on fabrics and fabric switch features and devotes less space to point-to-point and arbitrated loop configurations. Arbitrated loop, however, is still commonly used for JBODs ("just a bunch of disks") and as back-end storage for some NAS products, and so this chapter also discusses loop principles.

Chapter 5 examines the Fibre Channel products that have been created to accommodate server and storage attachment to a SAN. This spectrum of products includes transceivers and cabling, host bus adapters (HBAs), Fibre Channel-attached storage and tape, loop hubs, and fabric switches. Because extension of native Fibre Channel over wide areas using tunneling or dark fiber is dependent on the same Fibre Channel behavior demonstrated in a local environment, native Fibre Channel extension and IP tunneling products are also grouped into this chapter.

Over the past two years, storage over IP has entered the SAN market to complement or sometimes compete with Fibre Channel solutions. Chapter 6 provides a brief overview of IP SAN technology. Native IP storage protocols include Internet Fibre Channel Protocol (iFCP) and Internet SCSI (iSCSI). As with Fibre Channel, IP storage must provide discovery and naming mechanisms, and these are reviewed in the sections on Service Locator Protocol (SLP) and Internet Storage Name Service (iSNS). Because IP SANs directly converge mainstream IP networking with storage, additional IP services, such as Quality of Service (QoS) and IP Security that can be applied to storage, are discussed. IP is also an enabling technology for metro and wide area SANs, and Chapter 6 concludes with a review of buffering and speed-of-light latency issues that a storage architect must consider when designing SANs over distance.

Chapter 7 examines various classes of IP SAN products that have been or will be introduced into the market. Much of an IP SAN infrastructure can be borrowed from mainstream IP networking. Gigabit Ethernet switches and IP routers, for example, are not specific to SANs but can be incorporated into SAN design without modification. Similarly, TCP offload engines (TOEs) have been developed to solve host processing issues historically associated with TCP but also have become enablers for high-performance IP SANs. IP storage gateways can perform Fibre Channel-to-iSCSI or Fibre Channel-to-iFCP conversion, or both. As in the discussion of Fibre Channel products, the focus here is on product functionality and not specific vendor implementation. Although it is tempting to review and compare individual vendor products, the market is moving too quickly for a book to capture capabilities of specific vendor products.

After establishing a foundation for understanding SAN plumbing options in the preceding chapters, Chapter 8 begins a higher-level discussion of SAN software applications. This chapter reviews storage-specific applications, including server clustering, backup, data replication, distributed file systems, and file sharing. Some of these solutions are software only; others can be embedded in storage hardware products.

Chapter 9 examines problem isolation in SANs. Problems can and do occur eventually and can occur at any layer in a SAN configuration. This chapter looks at some simple problem isolation techniques as well as hardware and software tools that facilitate identifying and isolating faults in a SAN.

Management of SANs is a broad topic because it includes management of the SAN transport as well as data placement and resource management. Chapter 10 discusses the major categories of SAN management and management protocols that can be used. This chapter also includes an overview of the Common Information Model (CIM) initiative. Although CIM has been under construction for some time, it has recently received strong industry support in the hopes that it will enable a coherent and unified solution for managing both data transport and data placement from a single platform.

Chapter 11 takes the reader into the Twilight Zone technology of storage virtualization. Like a quantum particle in a transient state, storage virtualization can both exist and not exist in the same moment. Elementary storage virtualization such as RAID certainly exists and has provided significant value for storage applications for some time. Advanced, application-aware storage virtualization has yet to appear and will require significant engineering effort to bring into being. The vendor and marketing hype aside, storage virtualization holds great promise for SANs and will enable customers to implement shared storage solutions throughout their networks.

Chapter 12 covers application studies of SAN solutions. From storage-intensive operations, such as full-motion video editing and prepress graphic creation, to more ubiquitous customer requirements for streamlined tape backup, SANs have successfully addressed a wide range of business needs. In discussing storage consolidation, server clustering, content distribution, disaster recovery, and other applications, the material in Chapter 12 attempts to offer examples that accommodate Fibre Channel, storage over IP, or both in a heterogeneous solution. Where possible, the SAN infrastructure is presented in a technology-neutral way so that the SAN architect can thereafter overlay the customer transport of choice.

Chapter 13 reviews some of the main areas of concern for SANs, including standardization, interoperability, and management. These are mutually dependent and overlapping issues that vendors and customers together must address. Standardization of a technology, for example, has little value for customers if it does not result in interoperability of vendor products. Similarly, even the most sophisticated management platform would be ineffectual if its cycles were consumed reporting errors due to standards violations or interoperability conflicts.

Chapter 14 is my opportunity to engage in wild speculation about the future of SAN technology. No one has a crystal ball in such a fast-paced technology, and the storage industry itself is being shaped by the push of emerging technologies and the pull of customer demands. Although InfiniBand may help transform the high-performance server market, storage virtualization and IP storage networking are now the main forces driving SANs to ubiquitous adoption.

This edition of *Designing SANs* also includes supplementary material in the appendixes to provide background and reference information on storage networking. In addition to resource and vendor listings, material is provided on the standardization process behind SAN development, including the work of NCITS/ANSI T10 and T11 groups and the Internet Engineering Task Force (IETF). The structure and activity of the SNIA are also included, as is the work of the SNIA Technical Council on the Shared Storage Model and the SNIA Education Committee's SNIA Dictionary. Some of my previously published articles on storage networking are grouped in Appendix G, SAN Essays.

1.4 Chapter Summary

Using the SNIA Shared Storage Model

- The SNIA Shared Storage Model provides a framework for understanding the relationship between applications and underlying storage resources.
- At the high level, the SSM separates applications from the supporting storage domain.
- The storage domain includes the file/record subsystem, block aggregation layer, block subsystem, and auxiliary services subsystem.
- The file/record subsystem defines whether data is passed to the upper-layer application as files or as records.
- Database applications typically use records.
- Files and records ultimately reside on storage disks as data blocks.
- A data block is a contiguous group of data bytes stored on disk.
- Data blocks are grouped into files or records at the block aggregation layer.
- Block aggregation can be performed on the host, within the SAN, or at the storage device.
- The block subsystem is composed of physical disk or tape devices that store data blocks.

- The services subsystem includes discovery, monitoring, management, security, backup, availability, capacity planning, and other storage-specific functions that enhance the operation of the storage domain.
- A SAN resides between the file/record subsystem and the block subsystem.
- SANs enable connection between multiple servers and storage resources in a peer-to-peer network.
- SANs serve data blocks to the host; NAS serves files to the host.
- The SAN infrastructure may be Fibre Channel, IP SAN, or some combination.
- A NAS device may span the file/record subsystem to the lower-level block subsystem.
- The Shared Storage Model can be used to overlay customer applications and storage and to facilitate planning for shared storage alternatives.
- The SSM is a convenient tool for both data center and remote site SAN design.