The Complete Reference

Storage Networks

Part III

Network Attached Storage

Chapter 9

Putting Storage on the Network: A Detailed Discussion

Network Attached Storage (NAS) is the evolution of network servers. It was developed from efforts to configure, optimize, and operate a server directed only toward one thing: providing shared storage for users on a network. NAS is one of two pillars that make up storage networking infrastructures in today's data centers. It has evolved into a class of storage devices that address storage needs from departmental storage appliances to enterprise-class storage servers managing multiterabytes of data. This section will explain, discuss, and evaluate the NAS architecture, from the hardware and software required, to the necessary connectivity components. Developed using concepts from Part II, and expanding the NAS overview from Part I, this information will encompass a detailed view of the NAS internal structures.

The NAS Architecture

NAS is a specialized computer that provides file access and storage for a client/server network. Because it is a specialized solution, its major components are proprietary and optimized for one activity: shared file I/O within a network. NAS is considered a bundled solution consisting of prepackaged hardware and installed software regardless of vendor selection. The bundled solution encompasses a server part, a computer system with RAM, a CPU, buses, network and storage adapters, storage controllers, and disk storage. The software portion contains the operating system, file system, and network drivers. The last element is the network segment, which consists of the network interfaces and related connectivity components.

Plug and Play is an important characteristic of the NAS packaged solution. Given the bundling and configuration that comes with the product (which is dependent on the class of NAS solution you are implementing), it's out of the box and attached to the network in less than five minutes. Although this may be a fairly accurate portrayal for installation and identification to the network, the Plug and Play value must be considered in context with the size and scope of the NAS device. There are many implementation caveats to consider and ample thought should be given to existing server clusters and storage systems already in place.

However, for most small networks, providing a Plug and Play NAS solution is a productive solution. Though more sophisticated departmental solutions will take longer, the bundled characteristics of the NAS devices save considerable time in planning and implementation. Of course, enterprise-level solutions should have much more thought put into their planning and implementation. Obviously, with this class of storage, care should be taken not to fall prey to the "five-minute" Plug and Play temptation which could result in much larger infrastructure problems.

NAS can be a productive solution. Its simplicity is one of its greatest strengths. The entire perspective of the NAS phenomenon, however, should be considered prior to any implementation. The following points summarize the issues that are important to know and understand to facilitate a long-term productive configuration:

- NAS is a specialized server solution with proprietary hardware and software.
- NAS is optimized for file I/O; support of any other types of workload I/O will prove problematic.

- NAS is one of he easiest ways to accommodate increases to storage capacity for LAN-based shared files.
- NAS can be one of the most difficult to administer and manage when numbers of boxes exceed reasonable complexities.
- NAS uses existing network resources by attaching to Ethernet-based TCP/IP network topologies.
- Although NAS uses existing network resources, care should be taken to design expansions to networks so increases in NAS storage traffic do not impact end-user transactional traffic.
- NAS uses a Real Time Operating System (RTOS) that is optimized through proprietary enhancements to the kernel. Unfortunately, you can't get to it, see it, or change it. It's a closed box.

The NAS Hardware Architecture

NAS is a specialized server solution with proprietary hardware and software.

NAS products provide a *black box* style of computer externally configured as a server with extensible internal components configurable to the class of storage requirements you need. The black box computer style is available in both RISC and CISC or Intel processing chip sets. The computer system's hardware is proprietary to the manufacturer in terms of the internal component standardization, assembly, and the lack of interchangeable components.

The Server Part

Figure 9-1 is a generic layout of NAS hardware components. As you can see, they are just like any other server system. They have CPUs built on motherboards that are attached to bus systems. RAM and system cache, meanwhile, are the same with preconfigured capacities, given the file I/O specialization. What we don't see in the underlying system functionality are the levels of storage optimization that have evolved. NAS internal hardware configurations are optimized for I/O, and in particular file I/O. The results of this optimization turns the entire system into one large I/O manager for processing I/O requests on a network.

Given the black box orientation of the internal server components means you don't have to worry about further optimization and tuning for file I/O processing for the NAS solution. However, time spent on evaluating the hardware components of the vendor's solution is important, because once you install it, further changes will be only up to the manufacturer's extensibility of the model. Additionally, all file I/O workloads are not alike, as much as we would like to think so (see Part VI). Given the various and diverse products that exist, the evaluation of the speeds and capacities of the NAS hardware configuration are critical factors contributing to the effective storage and application performance.

Figure 9-1 shows the components that are optimized for file I/O, but does not speak to how they are optimized. This is done in the combination of the internal hardware and the proprietary operating system software as illustrated and discussed later in the software part of this chapter. However, it is important to note that extensive efforts have gone into the design and layout of the NAS internal hardware that provides the optimization for file I/O. It is unlikely we can find out how this is done unless we work for the particular NAS vendor. Needless to say, this work is proprietary to each vendor and is a significant part of the value proposition of their products.

The Storage Part

NAS is a storage solution, and as such, perhaps its most important aspect is the storage infrastructure it provides. NAS supports multiple storage types, capacities, and configurations, including select levels of RAID at the top end (for example, enterprise class NAS devices). However, support for RAID levels 1 (mirroring) and 5 (striping with parity) is beginning to become available already preconfigured in low-end NAS appliance devices. Support for incremental storage capacity upgrades depends on the vendor and flexibility of their models for upgrading.

Figure 9-1 is a generic representation of the storage part of the NAS configuration. Note that the proprietary nature of the hardware continues through to the storage connectivity and storage devices. Again, this part of the value proposition of the NAS bundled solution is optimized for I/O operations. There is an important design distinction to point out. Although largely assumed given the traditional I/O operations of the

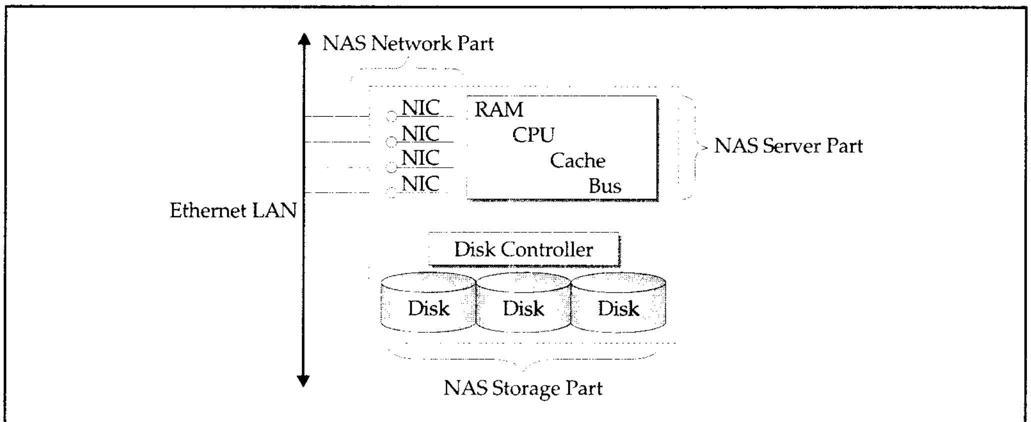

Figure 9-1. *The NAS hardware components*

NAS device as a server, it becomes lost in the focus on file I/O operations. The NAS device processes file I/O requests from clients within a network. As with any computer, these are processed into read/write requests through the OS and file system. Ultimately, this results in low-level disk operations performed at a block level, which means the NAS device still performs normal block level I/O operations within its internal processing configuration.

The sophistication of the storage system itself, and how it performs block I/O processing, depends on the type and class of NAS system. However, the storage infrastructure (the storage system, temporary locations, connectivity, and so on) is tightly coupled with the server part as well as the reconfiguration and capacity upgrades that depend on the flexibility of the vendor.

NAS devices are managed differently both internally and from an administrator perspective. The formatting of NAS disk devices is performed under the control of the proprietary operating system and, more importantly, its file system component. Consideration for cluster or block size may be out of the hands of the administrator, and thus limits the flexibility to the types of files the storage system will handle. Monitoring device activity is dependent on the tools provided by the vendor. Although additional third-party tools support limited levels of NAS activities, tools for the internal performance of system, storage, and network parts are lacking.

The Network Part

NAS supports Ethernet-based network connections, with all released standards supported (for example, 100-baseT through Gigabit Ethernet). Connectivity to the server bus is through standard NIC-type connections, however the bus connectivity from the adapter to the system bus is proprietary to the vendor. Most will find PCI connectivity standard even though processing chip sets vary from RISC-based systems to CISC-based Intel systems.

The NIC network components are available from single adapter configurations to extensible configurations that support multiple NIC connections such as those depicted in our generic NAS component diagram, shown in Figure 9-1. The adapters will support standard TCP/IP stacks with most connections being offered by NAS vendors through OEM products from various suppliers. Because of this situation, network performance can vary given the capacities and functionality of the NIC.

Departmental solutions are offered with only single NIC adapters, while mid-range NAS solutions offer multiple connections. However, even network connections for extensible mid-range solutions generally don't exceed a maximum of four NIC adapters. Enterprise NAS solutions will be extensible according to bus structures that are available from the vendor. This NIC connectivity and performance can be a limiting factor in deploying NAS in high growth areas. However, this is generally addressed by adding multiple NAS solutions. (More about these considerations is discussed in Parts V and VI.)

The NAS Software Architecture

NAS is optimized for file I/O; support of any other types of workload I/O is problematic.

NAS software consists of three major components: the micro-kernel OS part, a proprietary file system, and the network software drivers. These are depicted in Figure 9-2.

The Micro-Kernel OS Part

The operating system for NAS is a UNIX kernel derivative, as indicated in the server discussion. The system is developed with only OS components that enable effective file I/O operations. Components such as memory management, resource management, and process management are all optimized to service file I/O requests. This implementation as a black box architecture renders extreme limitations to initiating other processes, jobs, or activities through the OS. Besides the management tools that the NAS vendor or third-party independent software vendor provides, other processing activities within the server are not available.

Process management is programmed to handle the file I/O processes as the highest priority. Interrupt processing queues and handles other requests secondary to the I/O manager. Various scheduling methods and algorithms are used to optimize these tasks, such as "first-in, first-out" and random selection by queue. In this manner, all file I/O processes are completed prior to starting any new overhead tasks that the system may

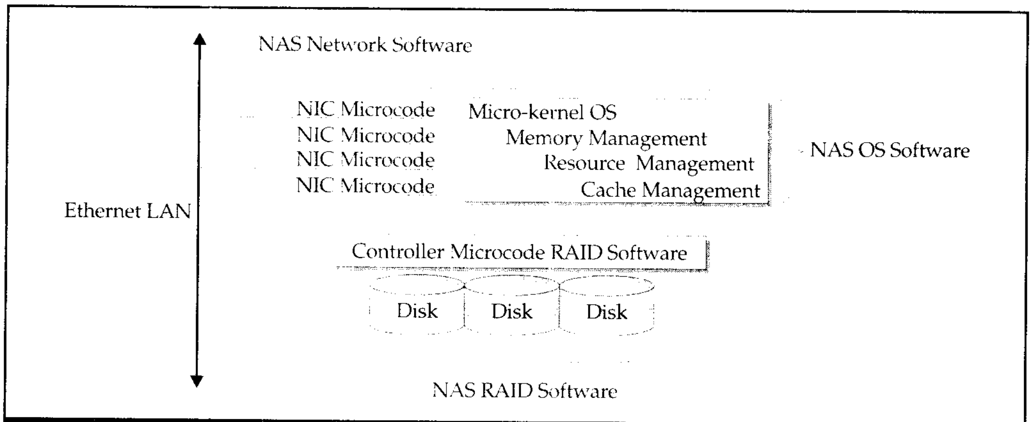

Figure 9-2. *The NAS software components*

require. Balanced among these is the need to manage network processes, as well as handling the overhead of the TCP stack. Given its reentrant functions, several network requests should be processing at any given time. This becomes a balancing act for the OS to handle sufficient network requests with I/O manager activities in order to complete block I/O operation to the storage.

Memory management will be optimized to handle as many file requests as possible—that means sufficient buffer space and cache management to facilitate the process management schedule. Capacity will play an important role in the capability of the OS to handle interactive requests. Most will be set for read-only operations to cache as many requests on the up-side (network input), and read-ahead buffer on the down-side (storage reads). Interchange will take place as the network process is completed in encapsulating the read request and passing control to the IP process and transmission. As we can see, a significant amount of time is used with TCP and IP process management.

Given the focused job the OS has to do, resource management is fairly static. I/O operations take over all the resources necessary to complete its tasks. Therefore, resource availability takes on a greater level of importance. In addition, the level of controller processing not only plays a significant role in completing the operations but is also fundamental in making resources available to the greatest number of processes. The operations of disk controllers and the further enhancements to controller processing, such as RAID levels and cache management, play significant roles in completing the I/O tasks.

NAS File Systems

The standard file systems included with NAS devices will support one of two types: the Common Internet File System (CIFs), the Network File System (NFS), or both. In many cases (as shown in Figure 9-3), the standard network file system interfaces with the NAS proprietary file system. Most NAS devices need to do this to manage their own storage. According to the vendors, the implementation of other types of connectivity protocols is supported inherently. As an example, the heterogeneous use of NAS devices to support multiclient access is evolving beyond CIFs through protocols such as SAMBA. Other proprietary network protocols such as Direct Access File System (DAFS) are beginning to emerge as well (see Part II).

Note *Samba is a freeware program that provides end users access to commonly shared resources, and is used primarily to share files and devices between Windows and UNIX systems.*

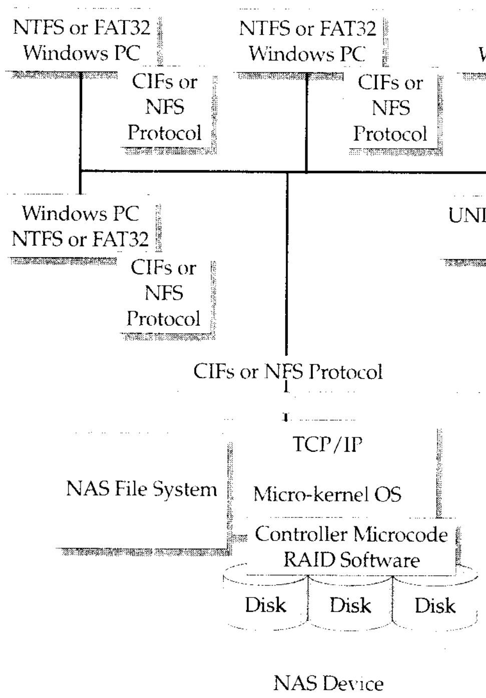

Figure 9-3. *NAS file system configurations*

Network Connectivity

The third of the three major NAS components supports standard TCP/IP communications within a diversity of transport media. The most popular topology remains Ethernet. Support here is evolving with increasing bandwidth in light of Gigabit Ethernet standards, sophisticated participation within enterprise networks with quality of server (Qos) and routing participation, and advancements in coexistence with other storage models, such as FC SANs.

A limiting factor of file I/O processing with NAS is the TCP/IP network environment and the processing of TCP layers. As we discussed, the communication between client applications and server I/O is processed within the encapsulation layers. Within the

NAS architecture, these layers (which are normally processed within the server) are now processed within the NAS box. To address these challenges, the TCP/IP overhead should be given great thought when implementing a NAS solution. However, TCP/IP performance issues are difficult to address, given the black box nature of the NAS software and the OEM supplier relationship of the NIC components.

An advancement that may help this condition is the development of TCP off-load engines (TOE). These are special NIC cards that off-load much of the TCP encapsulation processing onto the NIC card itself. Although this solution is just emerging, it reflects evolving technology directions that optimize latency within IP networks.

Note *TCP off-loads derive their value by processing TCP layers within the NIC card. Anytime software instructions are downloaded into lower-level machine instructions, it becomes highly optimized, operating at machine speeds. However, this is done at the sacrifice of software flexibility, given it's much harder to change micro-code or ASIC instructions than software that's operating under a high-level language within the OS.*

NAS as a Storage System

Putting all the parts together allows a highly optimized storage device to be placed on the network with a minimum of network disruption, not to mention significantly reduced server OS latency and costs, where configuration flexibility and management are kept to a minimum. The NAS solutions are available for a diversity of configurations and workloads. They range from departmental solutions where NAS devices can be deployed quickly and in departmental environment settings to mid-range and enterprise-class products that are generally deployed in data center settings.

The Departmental NAS Architecture

NAS is most often used to support the high growth of file servers that populate Windows client-based networks. An example is shown in Figure 9-4. It's an almost perfect solution for consolidating unoptimized file servers (for instance, Windows NT or Windows 2000 servers that operate only as network file servers for clients on the network).

The Internet NAS Architecture

An optimized solution is a more scalable and cost-effective answer as multiple servers can be consolidated into a single NAS device. NAS is also heavily used and is a good solution for interoperable access to files within web environments, as illustrated in

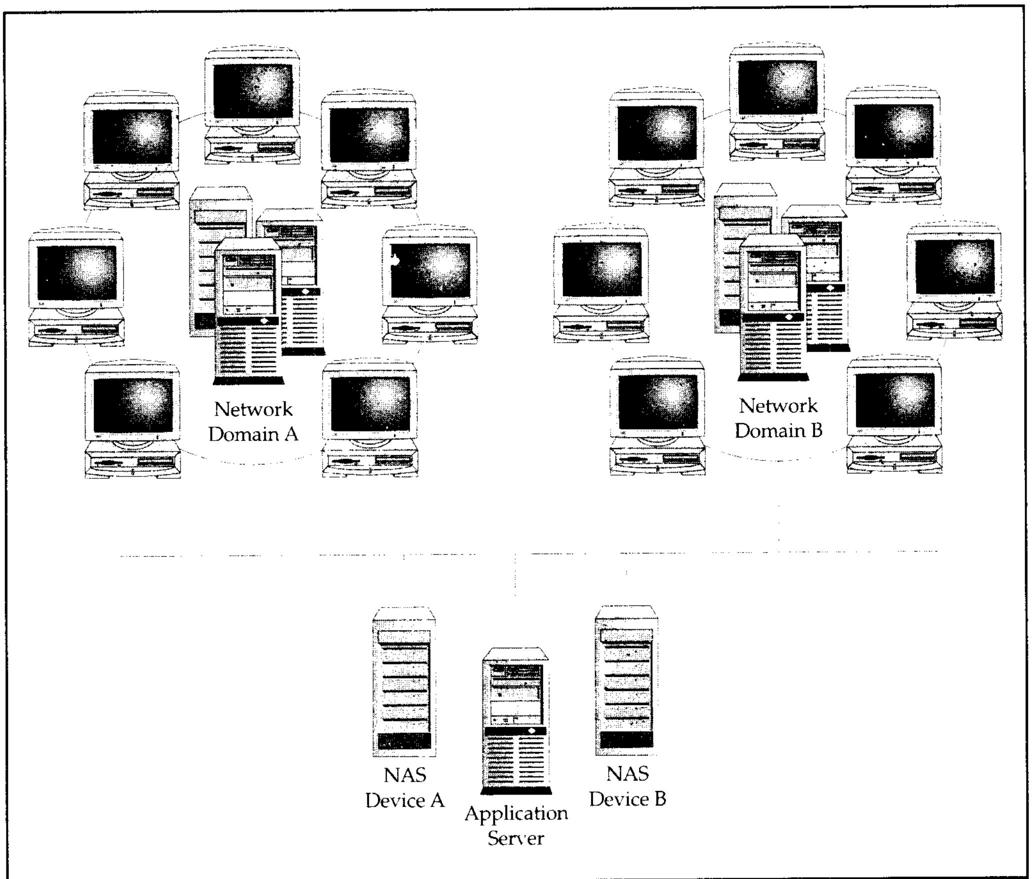

Figure 9-4. *The departmental NAS configuration*

Figure 9-5. NAS is the perfect storage solution for Internet connections to web-based files accessed through HTTP.

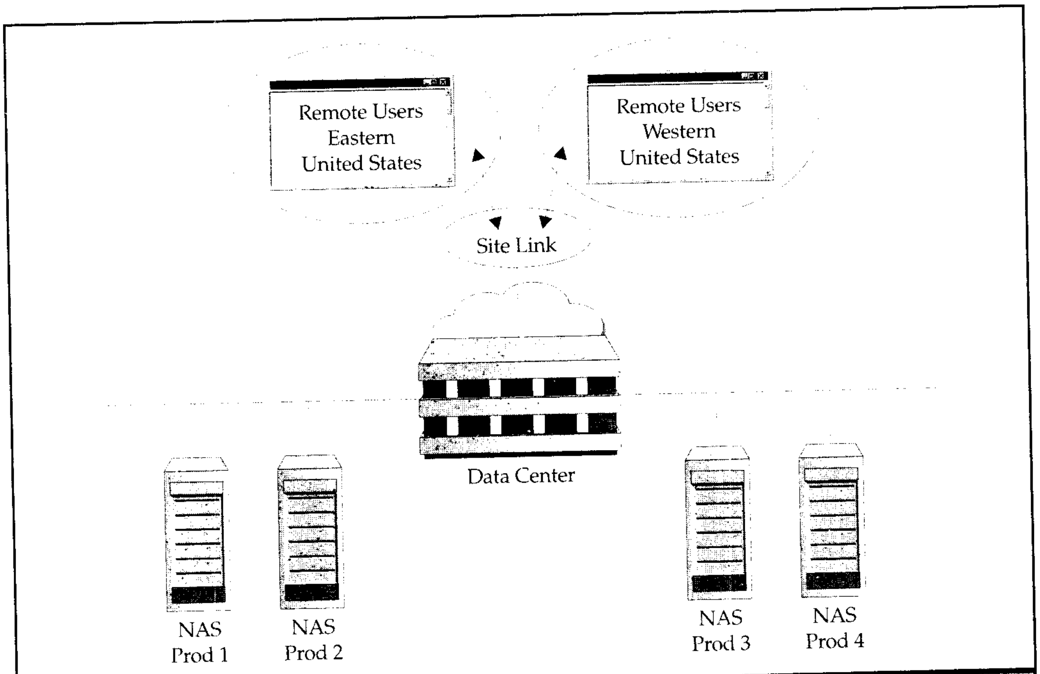

Figure 9-5. *The Internet NAS architecture*

The Enterprise NAS Architecture

At the high end are solutions that support large capacities of data, usually having reduced user transaction rates yet high I/O content. The storage and usage of unstructured data characterizes these workloads, although other structured data exhibiting these same characteristics can be NAS candidates. Figure 9-6 shows a specialized application being supported though high-end NAS devices.

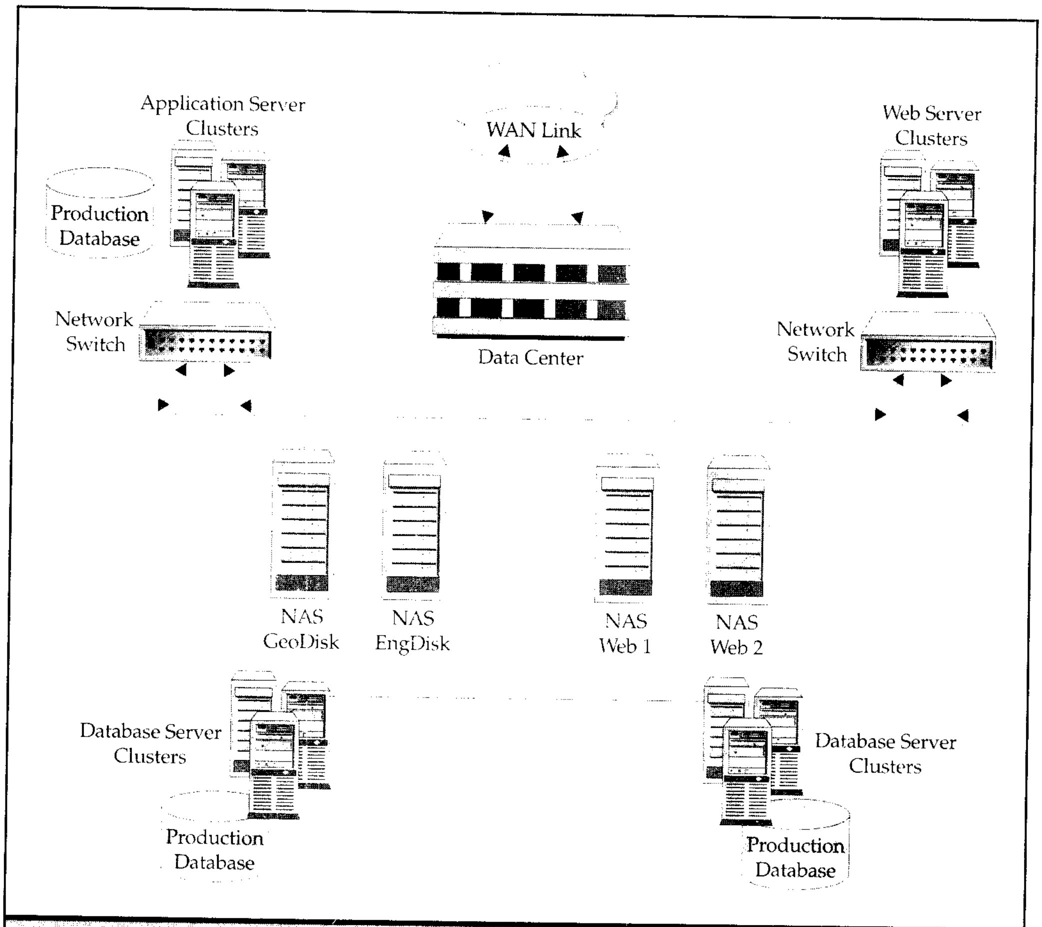

Figure 9-6. *Enterprise NAS configurations*