The
Complete
Reference

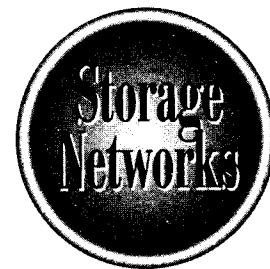Storage
Networks

Chapter 10

NAS Hardware Devices

NAS devices form an effective Network I/O manager for remote applications. From the perspective of the NAS hardware, functioning as an I/O manager, careful consideration should be given to the components and their corresponding characteristics that provide this complex functionality. This chapter will discuss NAS hardware components and their characteristics along with the issues that will confront these elements. Within these discussions will be definitions of key hardware components of the NAS device, configuration and capacities, and typical NAS I/O workloads as they relate to NAS hardware.

NAS hardware working in conjunction with related components function as an I/O manager. The NAS hardware components described in this chapter are the processor elements, RAM and related buffer management, and, of course, the storage system elements. We cannot overlook the communications and input-processing components largely made up of NIC interface elements and related bus connectivity options. These elements will be described and discussed in Chapter 12, which deals with NAS connectivity options, but it is important to recognize them in relation to the NAS hardware.

Closely related to NAS hardware are the effects of NAS processing models. Three models of NAS processing exist which drive NAS configurations and related component capacities:

■ Simple extensions of a client's file system, as a shared storage resource, make up simple file processing (Sfp).

■ Remote file acquisition with added reliability from a heterogeneous set of clients constitute quality file processing (Qfp).

■ Enterprise-level processing of files using NAS within a transactional type of system with heterogeneous access and multitier applications form a complex file processing (Cfp) configuration.

With the NAS hardware functioning as an I/O manager, the performance of its components will be characterized by the workloads it processes. NAS workloads are made up of file-oriented processing activities performed on behalf of various applications. These workloads consist of I/O operations with disparate I/O content. Client network files transferring office application files offer a dramatically different I/O content contrasted against the unstructured data such as image, video, and audio. On the other hand, web-based and other specialized application file structures exhibit everything from variable length files to compressed and encryption-encapsulated, structures.

NAS Hardware as the I/O Manager

As we discussed in the previous chapter, the NAS devices perform a singular directive with a duality of function, network file I/O processing and file storage. A more concise term would be that it functions as an I/O manager. So, what is an I/O manager?

The I/O manager is a major component of computers and is made up from a complex set of hardware and software elements. Their combined operations enable the necessary functions to access, store, and manage data. A discussion of many of these components can be found in Part II. A comprehensive view of the internals of I/O management is beyond the scope of this book, however an overview of a theoretical I/O manager will help you understand the hardware operations of NAS.

Figure 10-1 shows the typical components that make up a theoretical I/O manager.

The I/O subsystem is made up of hardware, the processor instruction sets, low-level micro-code, and software subroutines that communicate and are called by user applications. The I/O manager is, in effect, the supervisor of all I/O within the system. Some of the important aspects of the I/O manager are the following:

- **Hardware Interrupt Handler**   Routines that handle priority interrupts for hardware devices and traps associated with error and correction operations.

- **System and User Time**   A system clock that provides timing for all system operations. I/O operations called by the user and executed in kernel mode take place in system time and are performed by the clock handler routines.

- **I/O Device Addressing**   A system whereby hardware devices are addressed by a major and minor numbering scheme. Major device numbers are indexed to a device table. The minor device number identifies the device unit number. For disks, the minor number refers to the device number and the partition number. For tapes, the major number refers to the controller slot number, and the minor number refers to the drive number.

Figure 10-1.   *The I/O manager components*
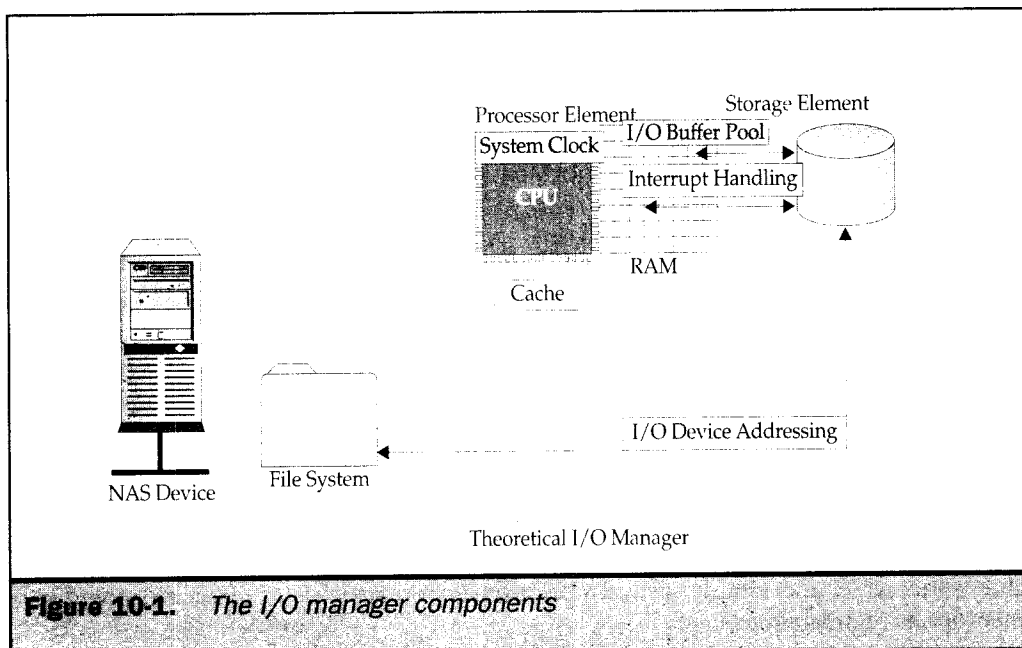
■ **I/O Buffer Pool** Retaining data blocks in memory and writing them back only if they are modified optimizes disk access. Certain data blocks are used very frequently and can be optimized through buffering in RAM.

■ **Device I/O Drivers** Block I/O devices such as disk and tape units require instructions on how they must work and operate within the processing environment. Called device drivers, they consist of two parts. The first part provides kernel service to system calls, while the second part, the device interrupt handler, calls the device priority routines, manages and dispatches device errors, and resets devices and the status of a device.

■ **Read/Write Routines** System calls that derive the appropriate device driver routine, which actually performs the read or write operation to the device.

■ **I/O Priority Interface** System routine that handles requests for the transfer of data blocks. This routine manages the location of the block, making sure it is handled properly if in the buffer, or initiates an operation to locate disk address and block location within a device.

■ **Raw I/O on Block Devices** A system call and method used for very fast access to block devices, effectively bypassing the I/O buffer pool. These types of low-level operations of the I/O manager are very important in order to better understand the architecture and usage of database implementations using I/O manager functions.

NAS Hardware I/O Manager Components

If we view the NAS hardware functioning as an I/O manager, the critical components will be the processor element, the memory/cache elements, and the storage element. Certainly, the connectivity to the outside world is important and as such will be handled through its own chapter (Chapter 12 in this book). However, the internal hardware aspects of the NAS devices determine how fast it can process I/O transactions, how much data it can store, the types of workloads it can handle, and how it can recover from failing components.

The NAS Processor Element

The NAS processor component is critical to the power required to handle specific types of workloads. Although most vendors have moved to a CISC model (for example, using Intel chips to keep costs down), many of the leading vendors offer RISC-based processor components. The difference, although seemingly slight, can be substantial when evaluating the performance required for complex workloads.

RISC (Reduced Instruction Set Computing) versus CISC (Complex Instruction Set Computing) continues to be an on-going discussion within chip designers and system vendors. The differences are based upon different architectural designs in how the CPU component processes its instructions. The differences are depicted in Figure 10-2, which shows that RISC uses simple software primitives, or instructions, which are stored in memory as regular programs.
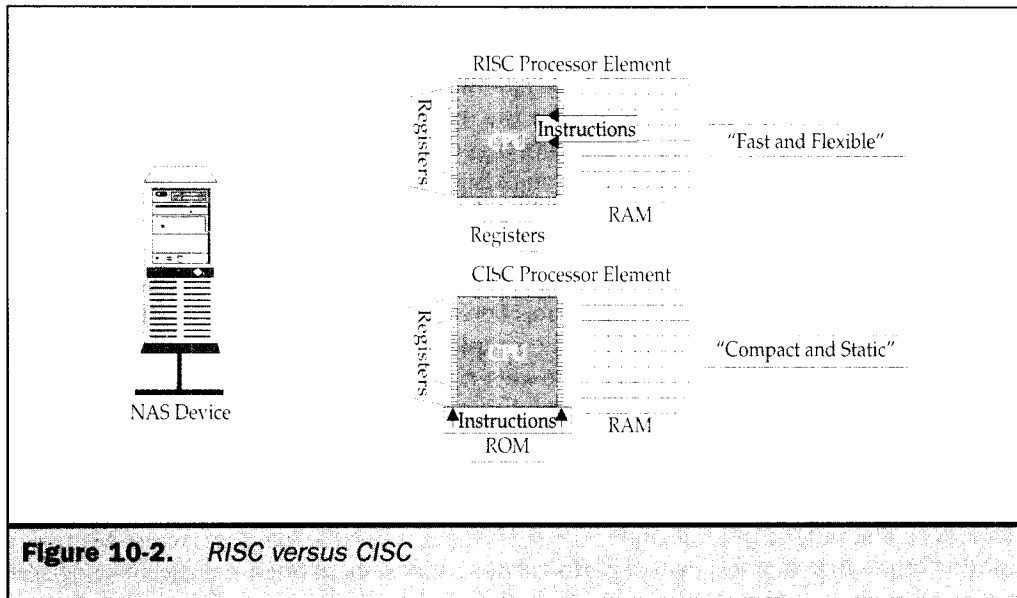
Figure 10-2. *RISC versus CISC*

CISC provides the instruction sets using micro-code in the processor ROM (read-only memory). The RISC processors have fewer and simpler instructions and consequently require less processing logic to interpret—that is, execute. This results in a high execution rate for RISC-based processors. Consequently, more instructions will produce a gross higher processing rate with everything being balanced at a higher throughput. Another important characteristic of RISC-based processor architecture is that they have a large number of registers compared to CISC processors. More instructions that are closer to the ALU unit means faster gross processing and processor throughput.

Given the architectural differences, the capability to provide high system performance, with repetitive instructions, against a focused workload puts the RISC-based system on top. However, this is not achieved without cost. RISC-based systems are more costly than CISC-based systems, specifically Intel commodity processors. Therefore, a balance must be achieved between the performance sophistication of the NAS device versus the processing complexities of the I/O workload.

This is important as decisions made on cost alone result in non-scalable CISC-based NAS devices, or devices that may have a short-term usefulness in supporting a complex I/O workload. On the other hand, making a purchase based on performance alone may lead to an overallocation of processor performance which could be handled just as effectively with CISC-based systems performing in multiples (for example, deploying many NAS devices).

This is also important because it defines the type of Micro-kernel operating system the vendor supplies. All RISC-based systems are UNIX-based micro-kernel systems, while the CISC-based systems are a combination of both UNIX and Windows micro-kernel systems.

When evaluating NAS devices, you should at least know the type of processor architecture that's being used. Looking at technical specification sheets beyond the capacities of the disks and storage RAID capabilities, it's important to understand the scalability of instruction set computing for the workload being supported. You will find that all CISC-based processing will use Intel-based processors with the most recognizable RISC-based processors being offered through Sun, HP, and IBM.

The NAS Memory Element

RAM is an important factor in processing I/O requests in conjunction with available cache and cache management. Caching has become a standard enhancement to I/O processing. Cache management is a highly complex technology and subject to workload and specific vendor implementations. However, most caching solutions will be external to the processor and utilize cache functionality within the storage array controllers. As NAS solutions are bundled with storage arrays, any caching will be integrated into the NAS device using both RAM and storage controller cache. Caching solutions within the NAS hardware will remain vendor specific and available through the configurations. Note that some vendors have included a NVRAM (Non-Volatile Random Access Memory) solution that will provide a level of performance enhancement but a significant safeguard in data integrity to writes that occur within the workload.

The functions of the I/O hardware manager enable buffer management, which is a key element in the effective handling of I/O requests. As requests are passed from the TCP stack to the application, in this case the software components in the I/O manager, they are queued for processing through buffer management. Given the slower performance of the disk elements compared to RAM, the RAM I/O buffers provide an important element to quick and orderly processing of requests that may come in faster than the disk I/O can handle. Therefore, the more buffer capacity you have, the more reque' 's can be processed in a timely fashion.

For data operations inside the processor element, the functions of system cache can also provide an additional level of efficiency. Our discussion regarding RISC versus CISC showed the efficiencies of having software instructions loaded as close to the CPU as possible. Having a combination of RISC-based instructions and user data blocks can provide additional levels of efficiency and speed. This can also prove effective in CISC architectures given the cache usage for user data blocks and the redundant nature of the I/O operations performed by the processor element within the NAS devices.

A popular solution to the esoteric I/O bottleneck problem is to add memory. Although this is an overly simple solution to a complex problem, in the case of NAS simple file processing (Sfp), it can generally be effective. The caveats (refer to Chapter 17) are many and lie within the read/write nature of the workload. Given that many NAS implementations are for read-only high-volume I/O transactions, the RAM element is an important part of the solution for increasing capacities for I/O buffer management.

The other effect RAM has on NAS devices is on the input side (for example, TCP/IP processing). Given that similar requests may develop faster than the processor element can execute them, the capability to queue these requests provides a more efficient environment in processing the TCP stack. In addition, the same will be true on the

outbound side as requests for data are queued and wait their turn to be transmitted to the network.

NAS devices are configured with static RAM/Cache options in the low-end appliance level devices Sizes of 256MB are common and are balanced to the small workloads these devices are configured for. Sizes and options increase dramatically in the mid-range and high-end models, with RAM options able to reach limits of one to two gigabyte levels. The configuration of system cache is highly dependent on the vendor and their proprietary implementation of the hardware processor elements.

The NAS Storage Element

Being that NAS is a storage device, one of its most important elements is the storage system element. As stated earlier, NAS comes with a diverse set of configurations for storage. These range from direct attached disk configurations with minimal RAID and fault resiliency support to disk subsystems complete with FC connectivity and a full complement of RAID options. Again, choice of NAS type is dependent on your needs and the workloads being supported.

At the appliance level there are two types of entry-level configurations. First is the minimum Plug and Play device that provides anywhere from 100 to 300GB of raw storage capacity. These devices provide a PC-like storage system connected to an Ethernet network using IDE connected devices with no RAID support. These configurations provide a single path to the disks, which range anywhere from 2 to 4 large capacity disks. As shown in Figure 10-3, the system provides support for a workgroup or departmental workload as Windows-based network file systems. However, these limited devices also do well with applications that support unstructured data (for example, image and video) within a departmental or engineering workstation group. Given their price/performance metrics, their acquisition and usage within an enterprise can be very productive.

The other ends of the appliance spectrum are the entry-level enterprise and mid-range NAS devices, as shown in Figure 10-4. These devices are also Plug and Play. However, capacities start to take on proportions that support a larger network of users
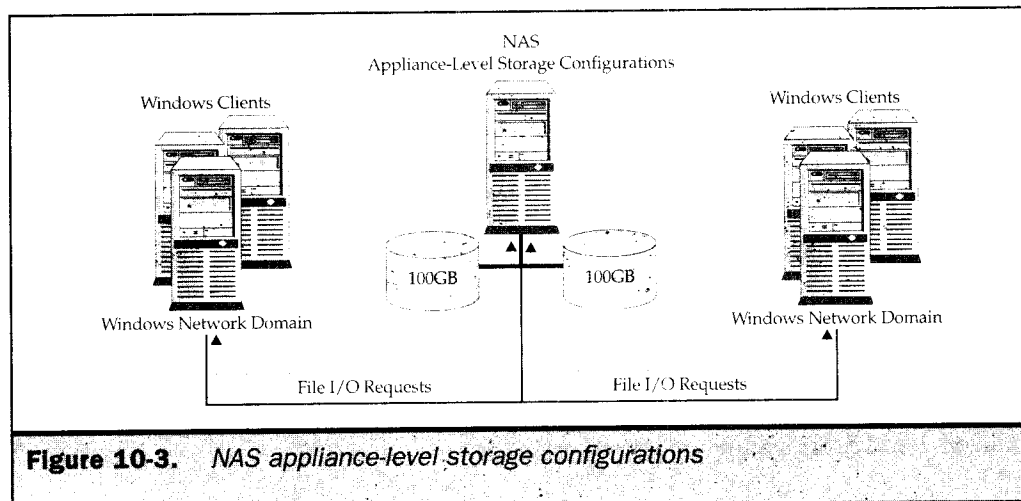

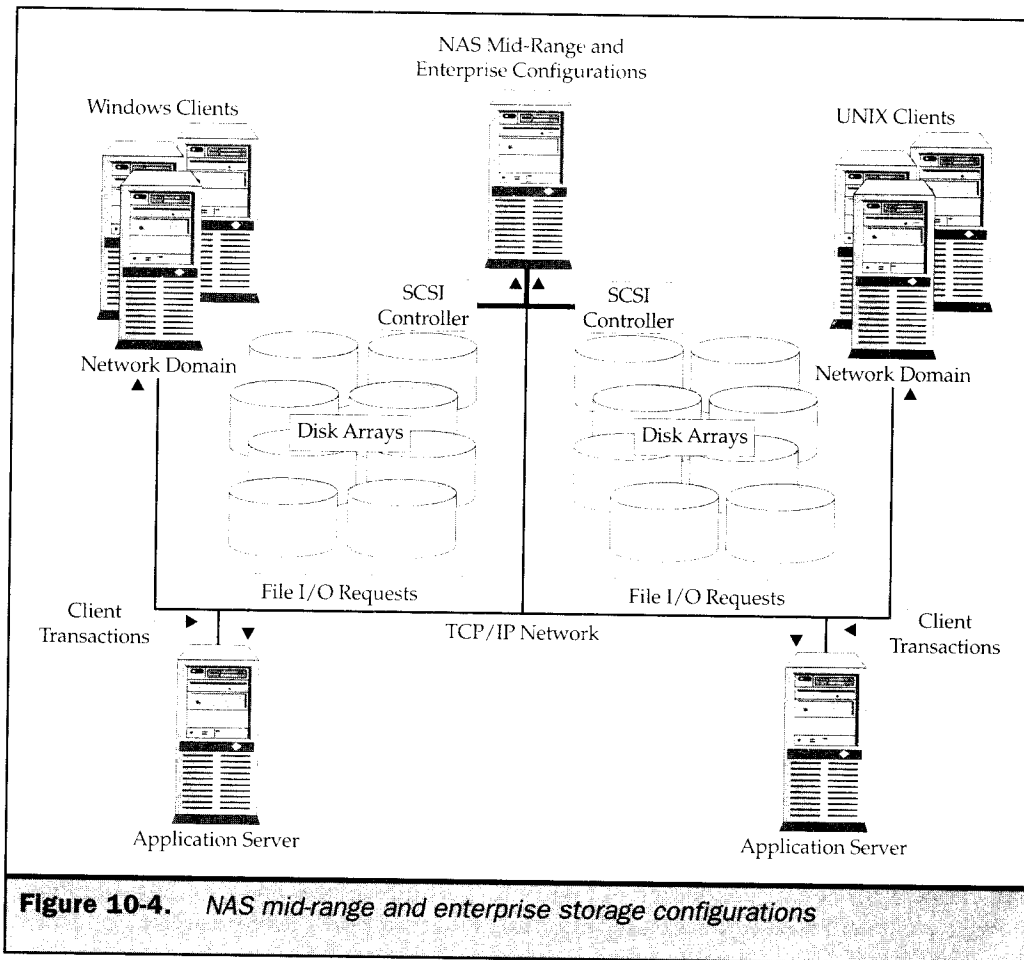**Figure 10-3.** *NAS appliance-level storage configurations*

Figure 10-4. *NAS mid-range and enterprise storage configurations*

and workload types. Storage systems capacities range from 300GB to 1 terabyte in size with RAID functionality at levels 0, 1, and 5. At these levels, the storage system is SCSI-based with options for the type of SCSI bus configuration required. With this, SCSI controller functionality becomes increasingly sophisticated. Although most use a single path into the SCSI controller, the SCSI bus functionality, because of its architecture, has multiple LUNs connecting the disk arrays.

The enterprise NAS devices extend the previous configurations with storage capacities, SCSI controller functionality, SCSI device options, and extensibility. Capacities extend into the terabyte range with multiple path connections into the disk arrays. The number and density of disks will depend on the array configuration. However, the top NAS configurations can support over 1,000 disk drives within a single device. At this level, RAID functionality becomes common place with pre-bundled configurations for levels 1 and 5, depending on your workloads.

Across the mid-range and enterprise offerings of NAS are three important hardware factors worth considering. These are fault resilience, FC integration, and tape connectivity. The NAS architecture renders these devices as functional units operating in a singular and unattended capacity. Therefore, the capability to provide a level of hardware fault tolerance particularly for the density and operation of the storage system has become a characteristic of the NAS solution. While Storage Area Networks provide another pillar to the storage networking set of solutions, the requirement for NAS to interoperate with SANs is increasing as related user data becomes interspersed among these devices. Finally, one of the key drawbacks to keeping critical user data on NAS was its inability to participate in an enterprise (or departmental) backup and recovery solution. This deficiency has slowly abated through the increased offerings of direct access to tape systems bundled with the NAS devices.

The ability to recover from a disk hardware failure has become sophisticated and standard among enterprise storage systems. These features have become integrated into the higher mid-range solutions and enterprise NAS devices. Look for full power supply and fan redundancy at the high end, and elements of both as these recovery options evolve from appliance-level solutions. This can be beneficial, as the environmental factors will differ greatly from closet installations to crowded data center raised-floor implementations. As one failure perpetuates another, this is an important place to consider fault-tolerant elements, because the inoperative power supply will bring the entire disk array down and potentially damage disk drives in the process. On the other hand, an inoperative or poorly operating fan will allow drives to overheat with potentially disastrous effects.

Another important aspect to the fault-tolerant elements is the capability to recover from a single inoperative drive within one of the disk arrays. NAS devices at the mid-range and enterprise levels will offer hot-swappable drive units. Used in conjunction with RAID functionality, either (software- or hardware-based) will allow the device to be repaired without the data being unavailable. Performance, however, is another consideration during a disk crash recovery period and is dependent on the workload and I/O content (additional discussion on this condition can be found in Part VI).

As with any recovery strategy, the importance of backing up the data on the storage system can't be underestimated. As stated previously, this is one area that has been lacking in the configurations and operation of NAS devices: how to back up the data within existing enterprise backup/recovery operations. Look for NAS devices to provide attachments to tape units. These will be offered as SCSI bus attachments where enterprise-level tape devices can be attached directly to the NAS device. Although a step in the right direction, there are two important considerations to make before launching into that area. First is the transparent implementation of yet another tape library segment. Copying data from the NAS device to an attached tape unit requires some level of supervision, ownership, and operation. In addition, the software required to provide the backup, and more importantly, the recovery operation has to run under the micro-kernel proprietary operating system of the NAS device. So, be prepared to support new tapes and new software to leverage this type of backup and recovery. (Additional discussion of this condition can be found in Part VI.)
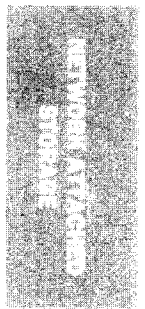
User data is now stored on disks directly attached to servers, using Storage Area Network and NAS devices. The applications using the storage infrastructure will largely ignore the specifics of the device infrastructure being used. That means that more interplay will happen with data that is stored between these different architectures. This condition is accelerating the need for NAS and SAN to communicate and provide some level of data access of each other's device. Look for high-end NAS devices to provide additional enhancements to participate within an FC-based storage area network. This will be offered as a host bus adapter from the NAS device with the appropriate software drivers running within the NAS OS. Using this type of configuration, the NAS device can service a file I/O and a block level I/O into the SAN. (Additional discussion on FC and Network connectivity can be found in Chapter 12.)

The sophistication and operational usage is just evolving for these types of integrated I/O activities. Data integrity, security, and protocol considerations (to name a few) are being developed as of this writing. In Chapter 17, we'll discuss some of the factors driving these requirements and various considerations to take note of when implementing applications within these configurations.

The NAS Processing Models

NAS uses three modes of processing. Understanding these modes helps determine and identify the growth of NAS workloads from the appliance level into the enterprise solutions. If we continue to view NAS as the I/O Manager for file-level processing, the modes in the following sections can be expressed.

Simple File Processing (Sfp)   This mode of processing simply extends the client's F/S and storage capacity into the NAS device. As depicted by Figure 10-5, this provides a simple transactional level of communication through a single type of protocol. Generally supporting Windows-based systems that require additional storage capacity for shared files or multiuser access to workgroup files, Sfp is handled by appliance-level type NAS devices with capacities that range up to 300GB at the high-end with minimal RAID support. Due to the limited transactional level (for example, access to shared files versus number of application transactional users), levels of RAM at or above (but not below) 512MB will generally support and balance the I/O operations for these configurations. Said configuration cross over to entry level and mid-range NAS devices as two things evolve: first, the number of users and consequently the size of user data; and second, the number of protocols supported. This can manifest itself in the heterogeneous support of client computers (for example, the addition of UNIX and Mac clients).

Quality File Processing (Qfp)   As user requirements move into supporting multiple protocols and increasing data capacities, the capability to provide quality processing increases. Quality processing, as shown in Figure 10-6, is defined as the capability to provide a level of data reliability to withstand component hardware failures, as well as support transactional level processing and heterogeneous sets of users. Qfp is supported
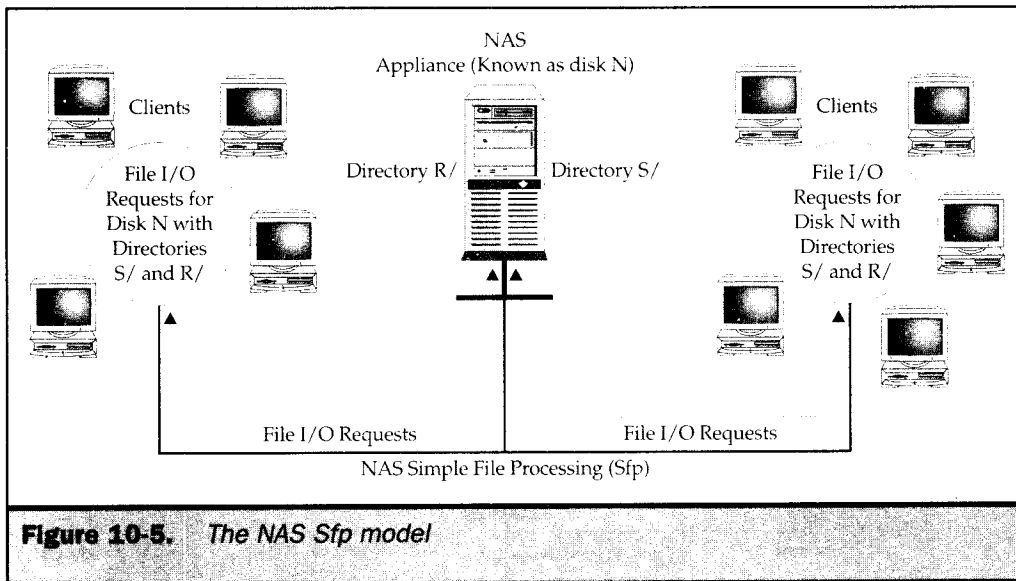
Figure 10-5. The NAS Sfp model

from entry level and mid-range level NAS devices with components such as fault-tolerant hardware components (power supply, fan redundancy, and hot-swappable drives), RAID functionality, and multipath disk array extensibility. Qfp is depicted in Figure 10-6 where a larger user base is supported along with transactional-like support through protocols such as HTTP. Storage capacities range from 300GB to 1TB, with RAM moving into the gigabyte range. These configurations cross into the area of enterprise level NAS solutions as the work begins to scale into multitier client/server applications with a heterogeneous user base and multiprotocol support.

Complex File Processing (Cfp) Figure 10-7 illustrates the complexity of requirements that NAS can grow into. The multitier processing of application servers and the capability to service high volume Internet traffic forms the basis for Cfp. Within these configurations, the NAS devices must support quality processing and handle the transparency of transactional redirection. User requests coming into an application server that provides web services are redirected to the NAS devices where the files are physically stored. A more complex processing configuration is the integration of storage area networks. In these configurations, user transactions must be redirected to NAS file systems and further into the level of database block I/O contained in the SAN storage network. Cfp will clearly stress the high end of NAS devices and push the limits of what vendors offer. Clearly, this requires the features of fault-tolerant hardware given its participation within a highly available transactional system. The terabyte levels of storage capacities and RAM gigabyte levels notwithstanding, the uniform and reliable performance becomes paramount in Cfp. Consequently, additional importance will be placed on recovery options (for example, tape connectivity and backup/recovery software, management tools, and overall enterprise data center integration).
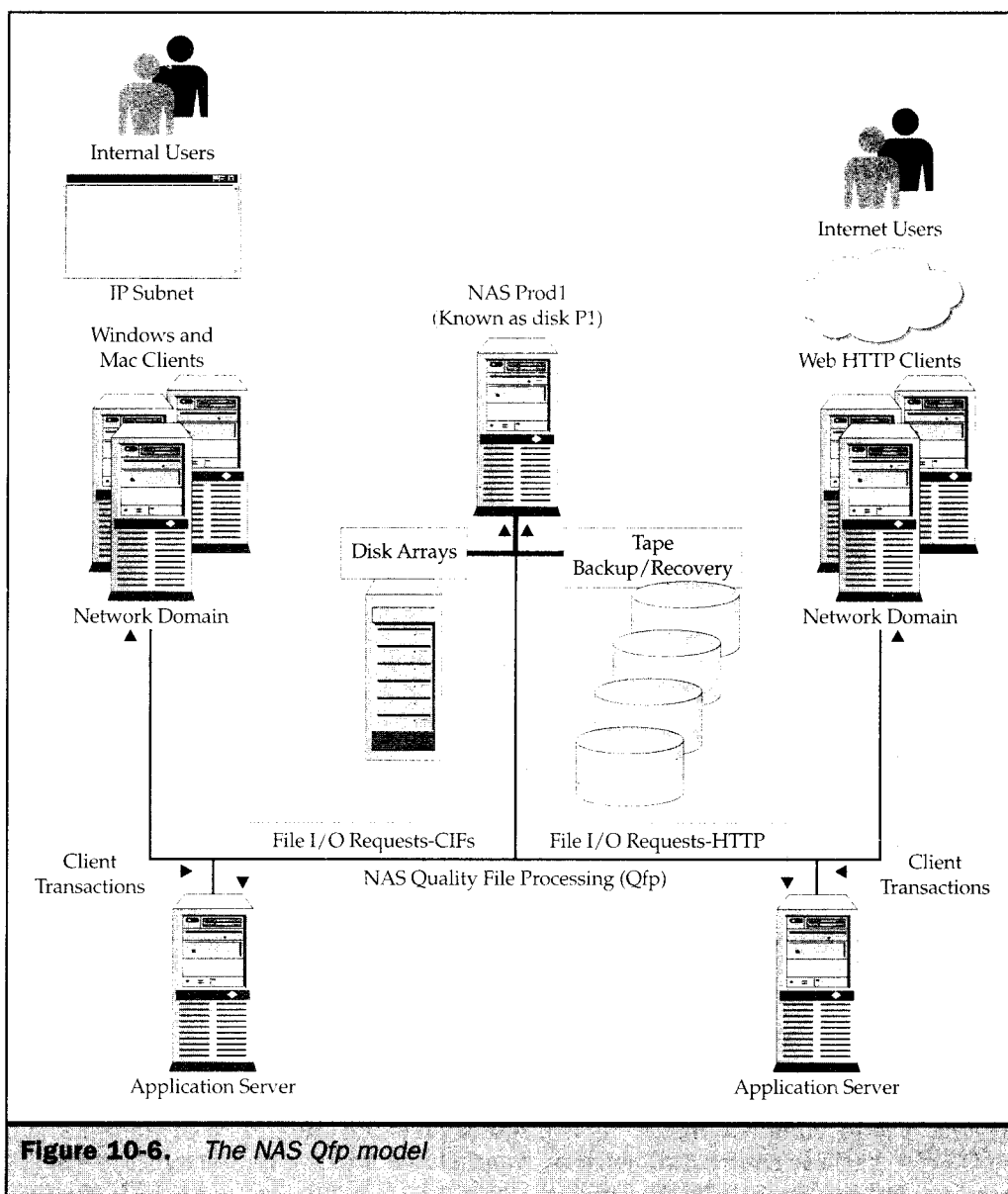
Figure 10-6. *The NAS Qfp model*

NAS Workloads and NAS Hardware

The NAS workload is characterized as file-oriented processing for various applications that end users require. Typical among these are Windows client files using personal applications such as office applications, shared forms, and templates. Other typical applications can manage unstructured data such as image, video, and audio files.
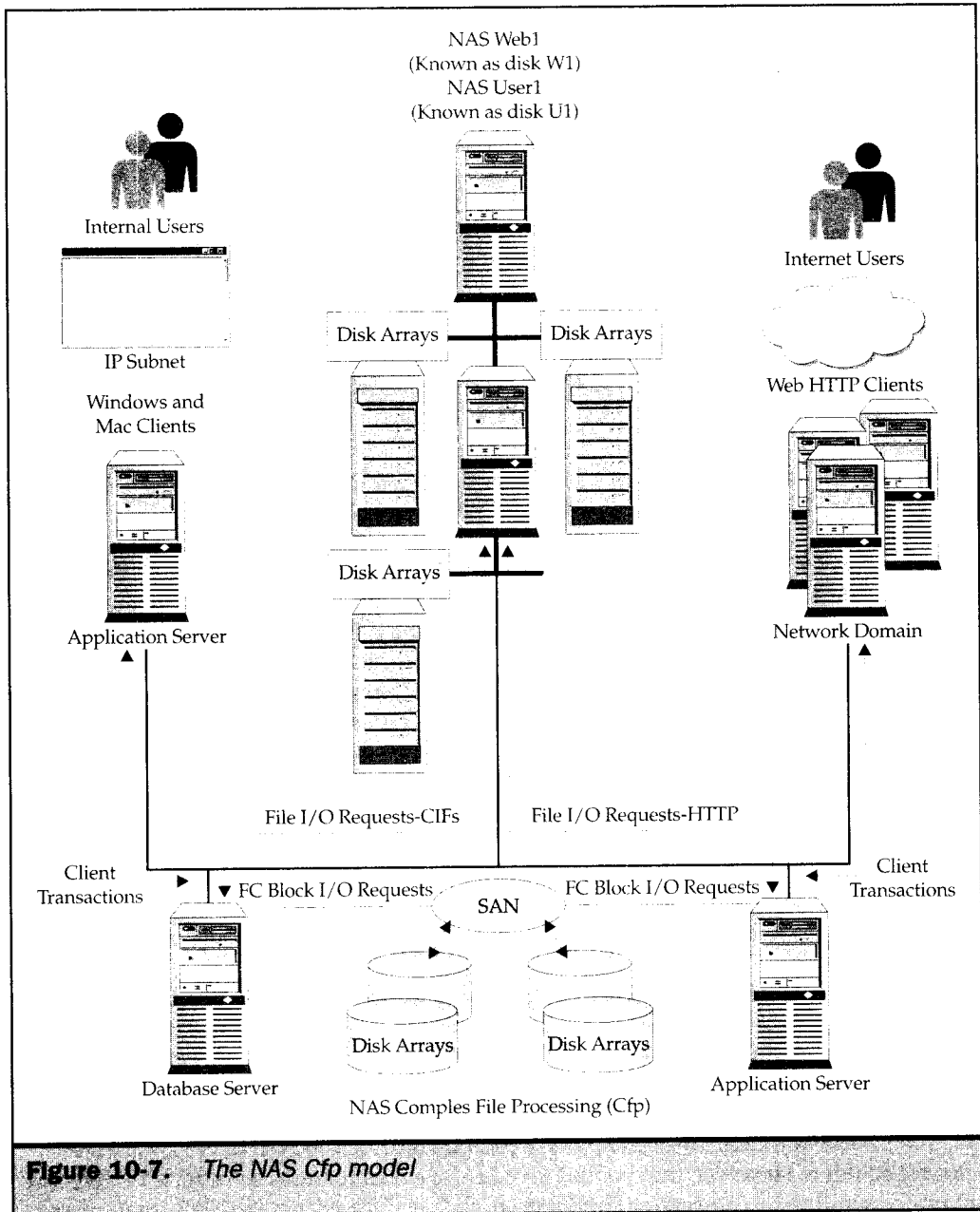
NAS Web1
(Known as disk W1)
NAS User1
(Known as disk U1)

Internal Users

IP Subnet

Windows and
Mac Clients

Disk Arrays — Disk Arrays

Web HTTP Clients

Internet Users

Disk Arrays

Application Server

Network Domain

File I/O Requests-CIFs   File I/O Requests-HTTP

Client
Transactions   ▼ FC Block I/O Requests   FC Block I/O Requests ▼   Client
Transactions

SAN

Disk Arrays   Disk Arrays

Database Server   Application Server

NAS Comples File Processing (Cfp)

Figure 10-7. *The NAS Cfp model*

Larger scale enterprise applications are centered on relational database tables while
enterprise Internet applications are grounded in new web-based file formats. The
application notwithstanding, the workload that NAS has to contend with centers on

the characteristics of the files they are storing, or to put it another way, the I/O content they have to satisfy.

Within the context of NAS, the I/O content is determined by the type, format, and characteristics of the file. For example, the content of a word document is very different than the content of a JPEG image. The reason is not only self-evident but also simple: the size and structure of the data. If we look at the access of these contrasting file types, we see the effect each has on the NAS hardware components.

If we take the next step and observe the I/O operations necessary to access these diverse file types from the NAS storage system, we find that the data content transferred within each I/O operation can differ greatly. Contrasted against the operating system, we find that the I/O operation is dependent on many things, such as the number of paths to the controller, the number of LUNs, and disk density. Consequently, the number of I/O operations, commonly referred to as IOPS, is a number without a corresponding level of efficiency.

Note   *IOPS is defined as the number of I/O operations performed per second by a computer system. It's used to measure system performance and throughput.*

It's possible, as we illustrated in Figure 10-5, that the transfer of a large text document can theoretically take more I/O operations, resulting in an inefficient transfer of data with each I/O. Conversely, the I/O operations for the JPEG could be done with less I/O operations, resulting in an efficient transfer of bytes of data with each I/O—very efficient! The difference is not only the density of the disk, the cluster size as set with the operating system, but also the capacities of the internal components of the NAS hardware elements themselves. The point is that the workload determines the capacities for individual, as well as groups of, NAS hardware specifications. Ensure that workloads and relative I/O content are given due consideration prior to setting the record for the fastest NAS install.

Given the black box orientation of the NAS device, look for efficiencies in the I/O manager that will provide levels of flexibility in these contrasting workloads. Operational evaluations of NAS devices should also cover the effects of a *black box* architecture on other data hierarchy locations, including system-level cache, RAM, and types of processor architecture.

Additionally, careful evaluation is necessary in maintaining the NAS devices for data recovery, disaster recovery, and data archiving. Operational discussions should also cover the challenges involved in participating in larger storage infrastructures, where NAS must interface with various storage devices such as tape, optical, and other server disk storage systems.

NAS hardware constitutes commodity components bundled into a Plug and Play storage solution. However, the evaluation of NAS devices requires some level of diligence in understanding the processing model used, the need for particular NAS components, such as processor types, device extensibilities, and capacities. More important is the type and extensibility of the storage system used with the bundled solutions. Given the rapid growth of end-user storage, configurations should ultimately be scalable to both the enterprise and workload.