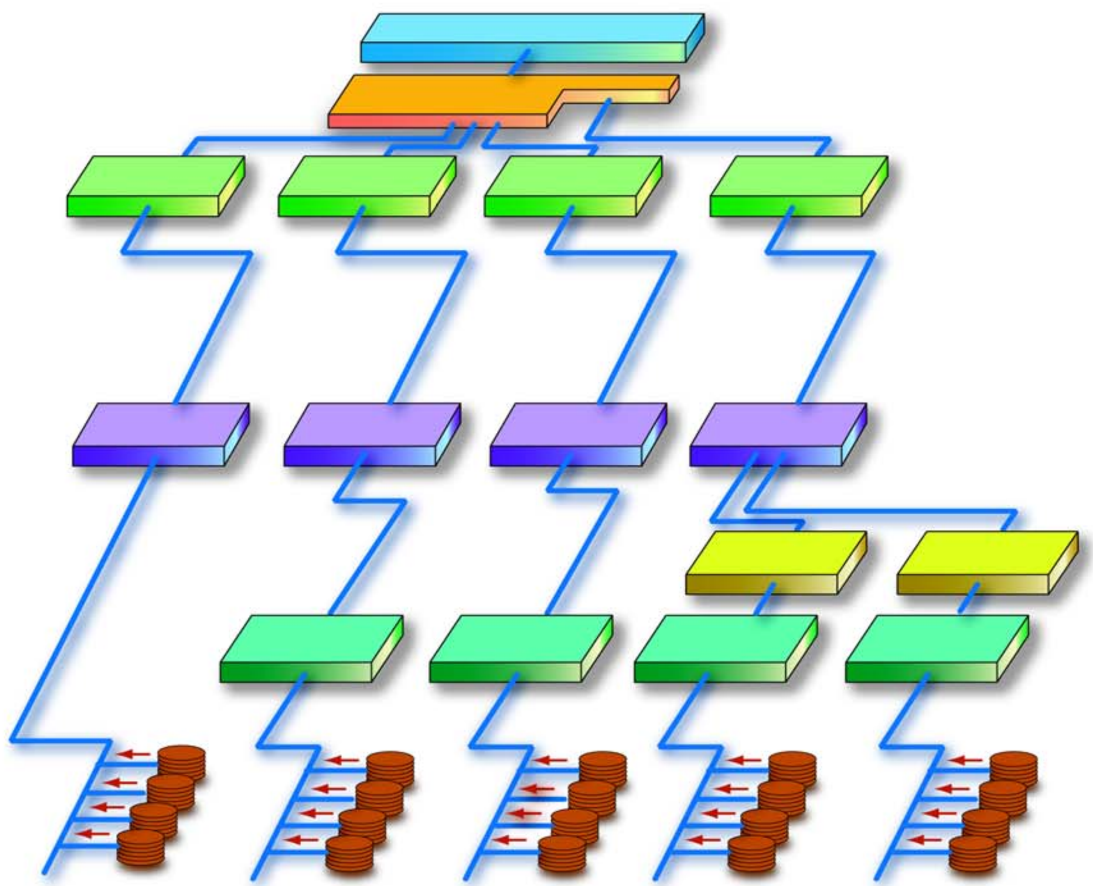
A Storage Architecture Guide - Second Edition

A business and technical discussion of the integration of NAS and SAN



AUSPEX

Auspex Technical Report

A Storage Architecture Guide - Second Edition

A business and technical discussion of the integration of NAS and SAN

The importance of choosing an effective storage strategy for eventual NAS and SAN integration

Understanding the technical differences between DAS, NAS and SAN

Deciding when to use DAS, NAS or SAN

Unifying NAS and SAN for Enterprise Storage

Selecting the right NAS product to prepare for NAS and SAN integration

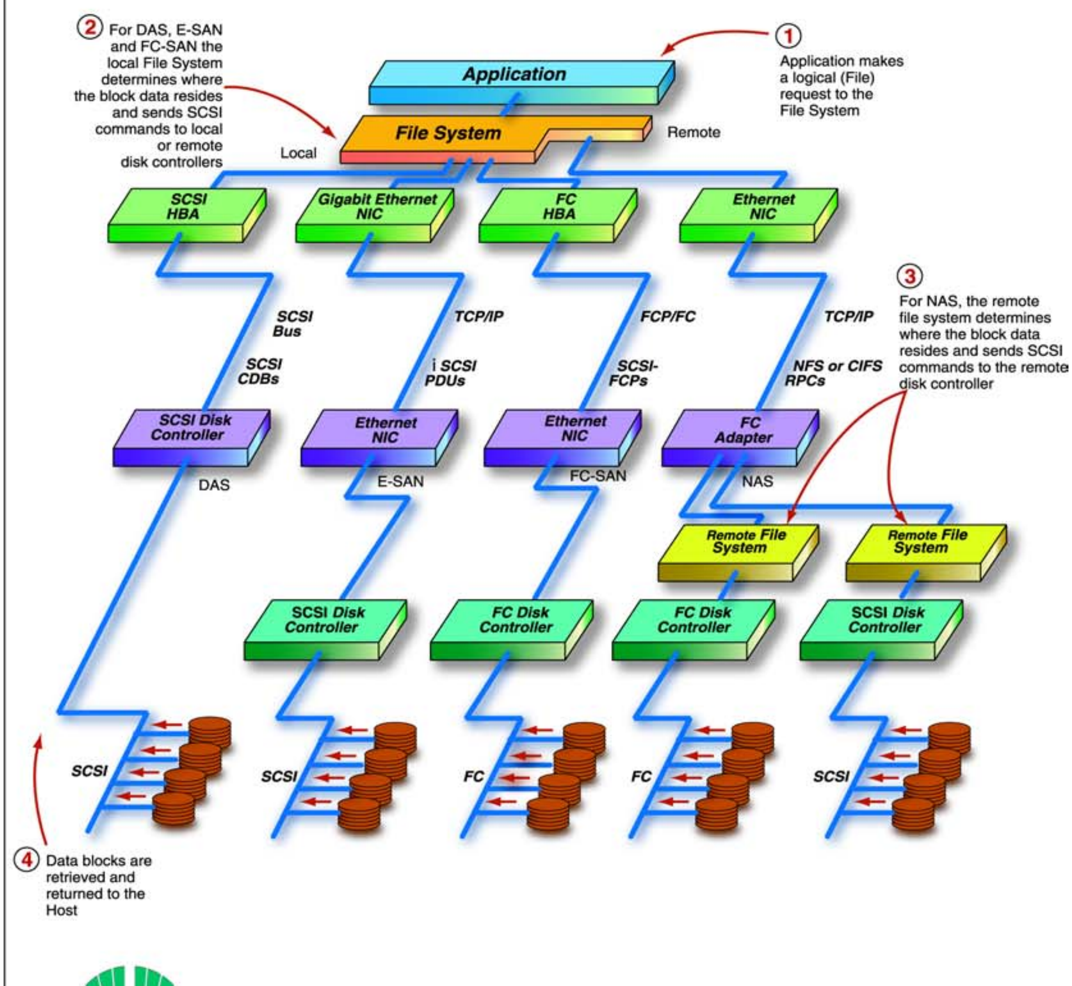
Summary and Conclusion

NAS offloads work from the Local Client or Server to the Remote Server

The inefficiencies of emulation software in data sharing

A taxonomy of network protocols

Glossary of Terms



AUSPEX

Auspex Technical Report

Copyright 2001, Auspex Systems, Inc.™ ®
Copyrights and trademarks apply.
Contents represents the opinion of Auspex Systems Inc. and includes accurate information as far as can be determined at the time of publication. Auspex is not responsible for inaccuracies resulting from changing technologies.

ClearCase is a registered trademark of Rational Software Corporation.

EMC and Symmetrix are registered trademarks and HighRoad, Celerra and SAN are trademarks of EMC Corporation.

IBM, Tivoli and OS/2 are registered trademarks and SANergy is a trademark of International Business Machines Corporation.

i-man is a registered trademark of UGS Corporation.

Microsoft, Windows, Windows media, Windows NT, and Windows 2000 are either trademarks or registered trademarks of Microsoft Corporation in the United States or other countries.

Oracle is a registered trademark of Oracle Corporation.

Quicktime and OS X are registered tradmarks of Apple Computer, Inc.

RealNetworks, RealPlayer, and RealAudio are the registered trademarks of RealNetworks, Inc.

Sun Solaris and Netra are registered trademarks of Sun Microsystems.

UNIX is a registered trademark in the United States and other countries, licensed exclusively through X/Open Company Limited.

This document is for informational purposes only, and does not set forth any warranty, express or implied, concerning any equipment or service offered or to be offered by Auspex Systems, Inc.

Table of Contents

Figures

Tables

The Importance of Choosing an Effective Storage Strategy for Eventual NAS and SAN Integration

1

Introduction

The cover and title page of this report illustrate the four basic methods of retrieving data from disk. Examining the details of these four methods can serve as the basis for an enterprise to decide whether their storage architecture is going to be well positioned for the eventual integration of NAS and SAN and to make decisions to achieve this goal. Both NAS and SAN share the common objectives of providing cost efficient and centralized management of all enterprise data with full sharing of any data by any computer to a single lock protected data image. Both NAS and SAN have three common goals: 1) data sharing by both UNIX and Windows client and server systems, 2) using network interfaces to accomplish this and 3) simplified and cost effective systems, network and storage management. As we examine each one of these common objectives it will become clear that optimal NAS and SAN integration is not there yet and that it is important for an enterprise to choose an architecture today that will provide the best strategy for their eventual convergence.

NAS and SAN have three common goals.

You cannot buy a fully integrated NAS and SAN product today! This is in spite of many confusing vendor claims to the contrary. A lot has been written about this subject in recent years and the literature has gone through two distinct phases. In the late 1990's, the first phase discussions focused on whether NAS *or* SAN is better for a particular application. Starting in 2000, the literature for storage vendors' products entered a second phase that is marked by vendors claiming or implying that they offer an integrated NAS and SAN architecture. Vendors have been doing this in order to protect their market share and to conceal characteristics of their storage offering that are less than optimal. This report objectively addresses the current state of NAS and SAN integration and how an enterprise can best position for the ultimate benefits of a truly unified architecture.

Building an effective storage strategy on the best elements of both NAS and SAN

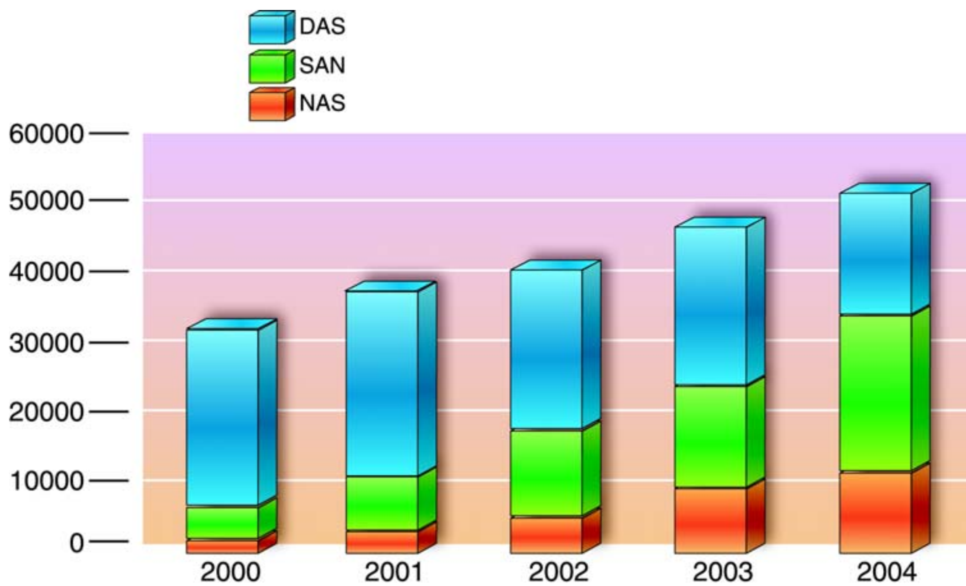
It is logical to build an effective storage strategy on the best elements of both NAS and SAN.

Common sense dictates that if two architectures are going to merge, then it is logical to build an effective storage strategy on the best elements of both. In **Chapter 6** we recommend effective strategies after first discussing what comprises the *best elements* of both NAS and SAN. Using these strategies, minimum re-configuration and incremental investment will be required when the actual NAS and SAN convergence does occur. As can be seen from this report, all NAS systems are not the same and involve different technical implementations of the most important NAS benefit of multilingual (UNIX® and Windows®) file sharing. Likewise, all SAN systems are not the same and can be constructed using either Fibre Channel (FC) or Gigabit Ethernet (Gig-E) network infrastructures. Since the major benefit of SAN is in the area of low cost and centralized management of all data, it is important for the enterprise to decide whether to implement the more expensive Fibre Channel SAN architectures available today (FC SANs). Or should the enterprise wait until more cost effective and high performance SANs based on the standard Gig-E technology (E-SANs) become available over the next few years. Similarly, it is important to select a NAS architecture that offers the best file-sharing scheme for a single large-scale data image, since universal data sharing is perhaps the most difficult technical challenge for a NAS vendor. Auspex believes that universal data sharing will form the cornerstone of an effectively integrated NAS and SAN architecture of the future and that it is wise to choose NAS products carefully to avoid future configuration expense when NAS and SAN do converge.

Storage expenditures are trending to NAS and SAN

Both NAS and SAN architectures are projected to grow at explosive rates over the next five years. International Data Corporation (IDC), a respected technology analyst firm, projects the three major storage architectures 1) Direct Attached Storage (DAS), 2) Storage Area Networks (SAN) and 3) Network Attached Storage (NAS) to be almost equal in market size by the year 2004 as shown in **Figure 1.**

Figure 1 – DAS, NAS and SAN will be almost equal in size by 2004.



Source: IDC, March 2001

NAS and SAN growth drivers

This growth is being driven by six common objectives of both NAS and SAN architectures: 1) The ability to handle explosive growth through the easy capacity and performance scaling of storage, 2) Deployment flexibility regardless of application, 3) Separation of the storage and server purchase decision, 4) High availability of information and tolerance of both manmade and natural disasters, 5) Simplified centralized management with the minimum skill level necessary to accomplish this on the part of network and storage administrators, 6) True UNIX® and Windows® file sharing from a single data image with secure locking to both environments. This is illustrated in **Figure 2.**

With major shifts such as these being projected for the future, it is increasingly important for IT professionals to develop comprehensive strategies designed to optimize network infrastructure with storage solutions that will enable scalability, reliability, performance, availability, affordability and manageability.

Figure 2 – Storage expenditures are trending to NAS and SAN and away from DAS.

Compounding the challenge of explosive storage growth, there are two major technology shifts IT professionals must consider when developing an enterprise storage strategy: first, the FC vs. GigE impact of networking technology on SAN architecture and management; second, the impact of universal data sharing on the design of NAS products.

The Auspex Storage Architecture Guide–Second Edition is designed to help CEOs, CIOs and network administrators understand these architectures and know the best use for each, while it provides suggestions for designing an effective storage strategy for workgroup, enterprise and ebusiness.

Direct Attached Storage (DAS) involves variations of SCSI and FC

Today, the majority of all computer storage devices such as disk drives, tape devices and RAID systems are directly attached to a client computer through various adapters with standardized software protocols such as SCSI, Fibre Channel and others. This type of storage is alternatively called *captive storage, server attached storage* or *Direct Attached Storage* (DAS) as illustrated in **Figure 3**.

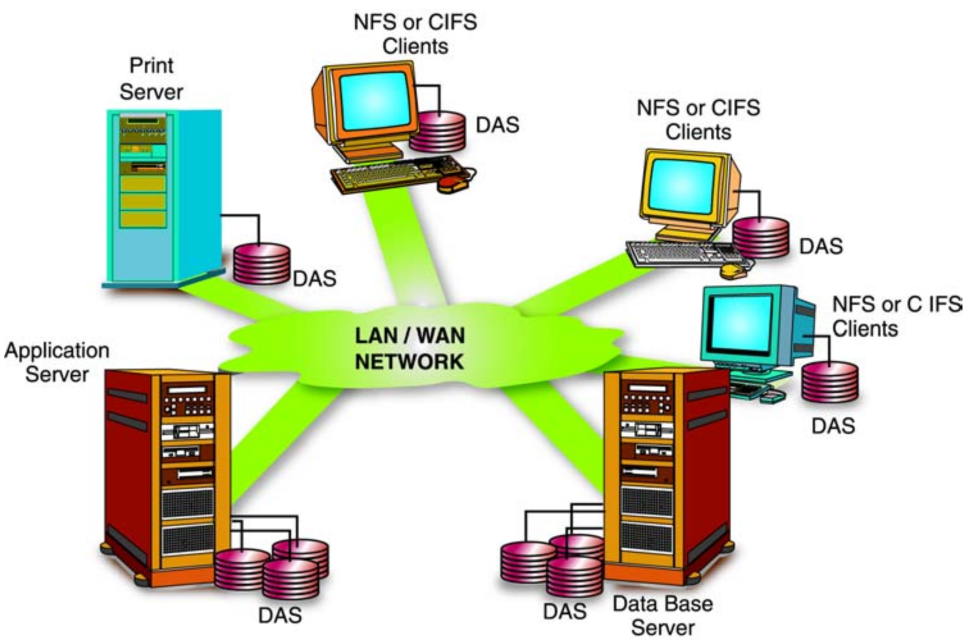


Figure 3 – Direct Attached Storage (DAS) topology.

Storage standards therefore are based on weak standards.

The committees that established these standards, however, allowed such wide flexibility in interoperability that there are many non-interoperable implementations of SCSI and Fibre Channel (FC) for the many available UNIX® and Windows NT® systems. For example, there are seven variations of SCSI, and most vendors implement FC differently. This is because storage was local to a specific server when these standards were defined and server vendors implemented variations that were not compatible. *Storage standards therefore are based on weak standards* and driven by component considerations. In other words, the problem with storage standards is that there seem to be so many of them.

As a result of these weak storage standards, third-party DAS vendors such as EMC and Compaq Corporation need to re-qualify their products with each revision of a server's operating system software. This can often lead to long lists of supported operating systems for SCSI or FC interconnects to different hosts. Each interconnect often requires special host software, special firmware and complicated installation procedures.

Network Attached Storage (NAS) is based on open networking standards

In contrast, *network standards are strong standards* that are driven by system considerations. There are two true network standards for accessing remote data that have been broadly implemented by virtually all UNIX® and Windows NT® system vendors. Developed and put into the public domain by Sun Microsystems®, Network File System (NFS) is the defacto standard for UNIX®. Developed by IBM® and Microsoft® and Common Internet File System (CIFS) is the standard for all flavors of the Windows operating system. As a result of these broadly accepted standards for network data access, storage devices that serve data directly over a network (called Network Attached Storage or NAS servers) are far easier to connect and manage than DAS devices. Some NAS servers, such as the Auspex NetServer 3000™, support *Universal Data Sharing (UDS)* between NFS and CIFS computers, which together account for the vast majority of all computers sold (see **Figure** 4).

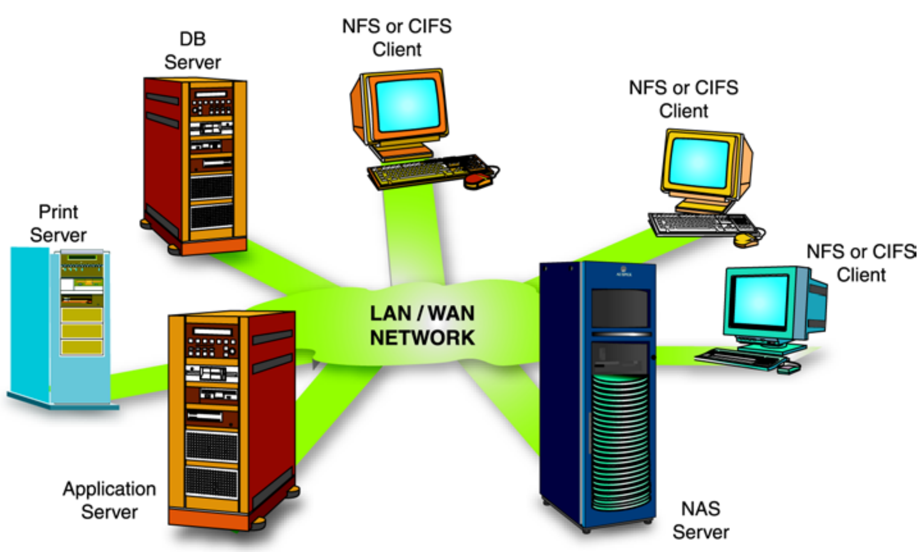


Figure 4 – Network Attached Storage (NAS) topology.

Like DAS, SANs must support many variations of SCSI and FC

As server vendors have implemented a variety of specialized hardware and software schemes to encourage the sale of DAS storage with their processors, SAN vendors have followed the same strategy. These architectures are alternatively called Storage Networks (SNs) or Storage Area Networks (SANs).

SAN actually adds network latency to the DAS storage model.

Instead of putting the storage directly on the network, the SAN concept puts a network in between the storage subsystems and the server as shown in (**Figure 5**). This means that SAN actually *adds* network latency to the DAS storage model. SAN standards are still in the formative stage and may not be established until well into this decade. EMC has announced a proprietary Enterprise Storage Network (ESN), and Compaq has announced a proprietary Enterprise Network Storage Architecture (ENSA). As with UNIX® and SCSI, SAN is likely to become a collection of proprietary architectures that are not based on strong standards. This may create major roadblocks to successful NAS and SAN integration and data sharing between heterogeneous platforms.

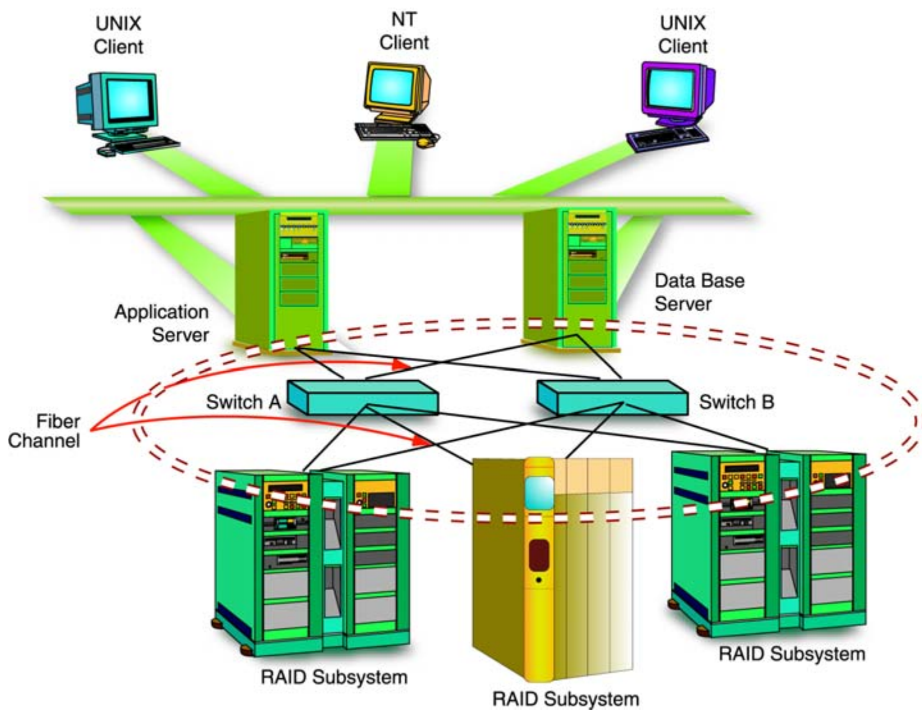


Figure 5 – Storage Area Network (SAN) topology.

Should you put Business Benefits or IT Infrastructure Benefits first when planning for NAS and SAN integration?

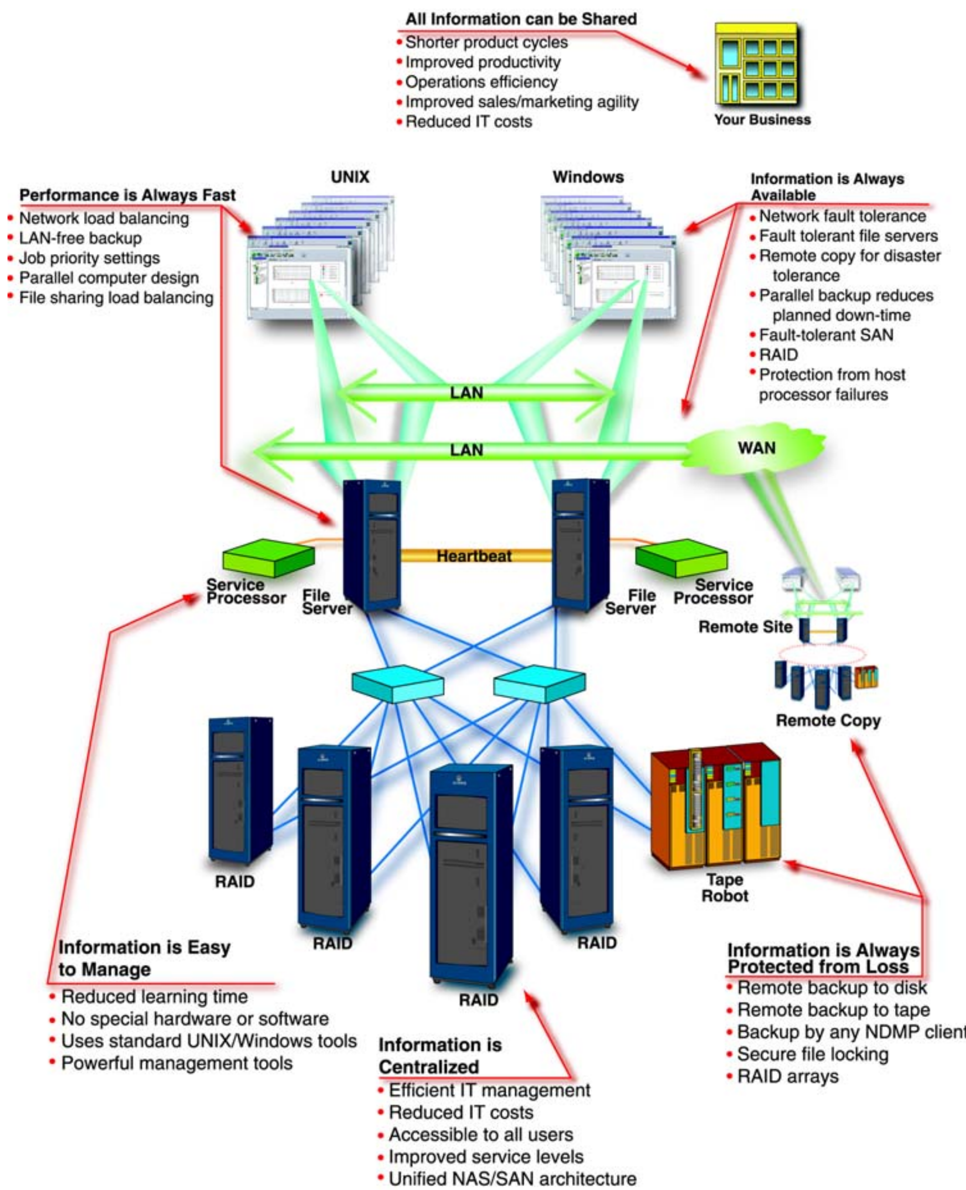
First, select the best data sharing solution and use this as a cornerstone in planning.

There is no disputing the importance of infrastructure and cost benefits of SANs to an IT department. In fact, low cost and centrally managed IT resources are a business benefit to the enterprise. However, the business benefits to non-IT departments are generally regarded as more important than IT infrastructure benefits for the enterprise to remain competitive in today's markets and achieve their business goals. When the six major benefit areas of storage architecture are considered, only the information sharing characteristics of NAS provides direct business benefits to non-IT departments. Therefore, it makes sense to first select the best data sharing solution and use this as a cornerstone in planning for eventual NAS and SAN integration. This is because the technology to accomplish Universal Data Sharing (UDS), as in the Auspex NS3000™ File Server, is the most difficult to achieve technically. The large-scale universal sharing of *any* information to *any* client has not been accomplished by *any* NAS vendor other than Auspex. The technical basis for understanding the specific reasons for the unique benefits of UDS are discussed in **Chapter 5** of this report in greater detail. In fact, Auspex solutions provide virtually every conceivable infrastructure benefit an enterprise could want, so there is no need to compromise infrastructure benefits and business benefits when selecting the best NAS solution to position for the future convergence of NAS and SAN. Auspex is committed to the integration of NAS and SAN and has recently taken the first major step in this direction with the announcement of the NS3000™ Product Family.

The Six Major Business and Infrastructure Benefits of the Auspex NS3000™

The Auspex NS3000™ architecture is considered the most advanced product design available for the specific task of serving network files with market leading performance, consistent data availability and robust security. This is because the NS3000™ architecture is the only parallel hardware and software product design available that distributes processing workloads to many processors working in parallel. This design improves system performance and provides for consistently high performance to users with 99.99+% data availability. Importantly, extra storage capacity can be added seamlessly without on-line system interruption or any decline in performance or availability of data. There are six major business and infrastructure benefits of the Auspex NS3000, which are shown in **Figure 6** and discussed in detail in the Auspex technical report titled the ***Auspex NS3000™ Series Product Guide*** that can be downloaded in PDF format from www.auspex.com.

Figure 6 – The business and infrastructure benefits of the Auspex NS3000 Series.



1. All Information can be shared
2. Performance is always fast
3. Information is easy to manage
4. Information is centralized and scaleable
5. Information is always available
6. Information is always protected from loss

All information can be shared

Excellence in information sharing is Auspex's commitment to its customers. The Auspex expertise in supporting file sharing applications and implementing file sharing infrastructure is unsurpassed in the market today. Unlike the NS3000™, alternative solutions partition storage or do not allow full read/write access to all files. Regardless of whether your employees and business partners have UNIX® or Windows® computers, they can't help but become more productive with an Auspex file server.

Performance is always fast

The commitment of Auspex is to look at the entire network, storage and computing performance environment to optimize the fast delivery of files to users regardless of the computers they are using. The parallel design of the Auspex NS3000™ file server allows multiple nodes to serve files in parallel. This provides for linear performance increases without having to add additional servers. In addition, the Auspex file server design prevents backup traffic from using network bandwidth except in the case of remote operations. This provides users with consistently fast file performance. Auspex File Servers also provide load balancing within each server for file sharing accesses. File locking and file system processing are distributed to different processors to further improve parallel performance, especially for mixed UNIX® and Windows® environments.

The Auspex NS3000 file server allows multiple processors to serve files in parallel.

Information is easy to manage

Because Auspex file servers allow the use of all native UNIX® and Windows® management tools, there is virtually no learning time on the part of Network and System Administrators. Auspex file servers also allow for easy integration of powerful management tools into existing enterprise system management frameworks through the support of industry standard protocols. No special hardware or software is required.

Information is centralized

It is well accepted in the industry that centralized files can be more efficiently and cost effectively managed. The Auspex NS3000™ file server design is ideal for the centralization of files because additional storage can be easily added without degrading performance. In addition, these centralized files can be managed with existing UNIX® and Windows® tools that are well known to IT administrators. This feature improves IT service levels for availability and performance since fewer administrators using familiar tools have a reduced risk of human error.

Information is always available

To achieve high availability computing, it important to avoid not only *unplanned* interruptions to file and network availability, but also *planned* interruptions such as scheduled downtime for servicing and backup. Whether planned or unplanned, downtime means lack of information availability for employees and business partners.

Whether planned or unplanned, downtime means lack of information availability for employees.

Many network file serving and file sharing applications require files to be "always available" day and night all the time. Whether engineers are designing a new product using network file sharing or a manufacturing team is working three shifts to meet demand, availability of the latest information can be critical to effective business operations in all departments. For many companies if files are not available, employees cease to be productive. The cost of downtime can be very high!

Information is always protected from loss

Whereas all of the availability features discussed in the prior section contribute to the protection of files against the possibility of loss or corruption, the ultimate protection from loss of files is multiple backup copies in a remote location.

Should you wait for Ethernet-based SANs (E-SANs) to emerge or deploy FC-SAN today?

Fibre Channel is a pseudo-networking technology.

The second critical decision in positioning for the eventual integration of NAS and SAN architectures requires an understanding of networking trends and the likely way in which NAS and SAN will converge from a technical network and I/O data flow point of view. The data flow technical differences between DAS, NAS and SAN are discussed in **Chapter 2** of this report and the likely basis for true integration of NAS and SAN are discussed in **Chapter 4.** As a preview to this material it is important to remember that Fibre Channel is a pseudo-networking technology that was conceived and implemented by the storage industry for use in SANs. As a result of this circumstance, Fibre Channel SANs (FC-SANs have security and congestion control deficiencies and a 10 Kilometer distance limitation (**Chapter 3**)). Since most SAN implementations are proprietary, the security problems are not easily fixed given lack of vendor cooperation. There is a "tunneling" networking transfer protocol known as Fibre Channel over IP (FC/IP) that proposes to send Fibre Channel encapsulated SCSI commands over IP networks but this approach is inherently inefficient. A more efficient networking approach would be to send SCSI command directly over IP networks as is proposed by the emerging iSCSI standard that forms the basis for Gigabit Ethernet-based SANs (E-SANs). Furthermore Ethernet networks are expected to surpass the bandwidth capabilities of Fibre Channel and will not require adoption of a separate network infrastructure for Fibre Channel. However, if an enterprise needs the benefits of SAN today, or has already implemented FC-SAN technology, the luxury of waiting for Ethernet-based SANs (storage over IP) may not be an option. This would be unfortunate since the cost advantages of managing and procuring only one type of network (Ethernet only instead of Fibre Channel AND Ethernet) would not be possible. In other words, whereas it may be ideal to wait for Ethernet-based SAN technology to emerge, practical considerations necessitate accommodation of Fibre Channel as it exists or is needed in the short term.

FC/IP is inherently inefficient.

It may be ideal to wait for Ethernet-based SAN technology.

Other information sources available from Auspex

Auspex is the originator of NAS.

Being the originator of NAS, Auspex is widely considered by customers and analysts alike to be the authority in both storage and networking. Since the topic of NAS is new to many customers, Auspex is committed to provide the best public information available on optimizing the flow of accurate information and support on both a pre- and post-sales basis. Auspex sales and system engineering teams will often recruit additional technical support from Auspex resident specialists, who are experts in each of the areas mentioned in this report. As with any IT architecture decision, probably the most important issue is the selection of a vendor/partner with the best "total" solution. This means not only choosing a vendor who remains at the forefront of technology with the most advanced parallel architecture, but also making sure the vendor can supply the most knowledgeable professional services, consulting services and support personnel.

Understanding the Technical Differences Between DAS, NAS and SAN

2

Physical vs. Logical I/O

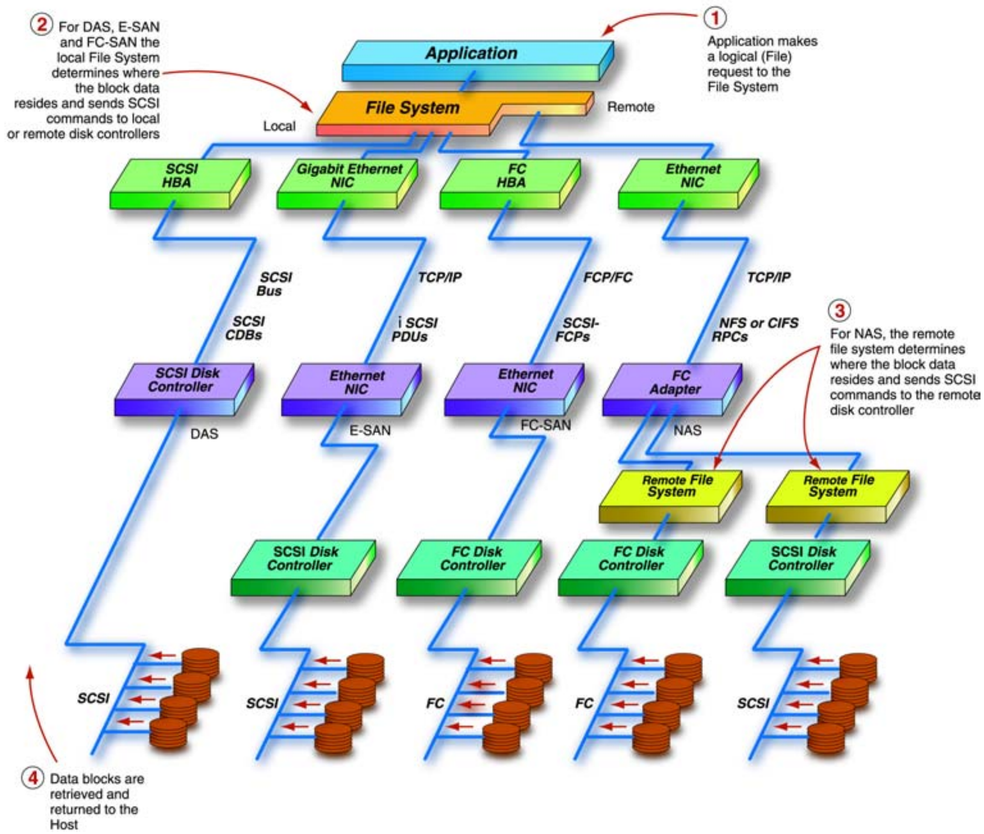
DAS and SAN deal with physical block I/O and NAS deals with file I/O.

It is important to understand the differences between physical and logical I/O requests in order to understand the differences between DAS, NAS and SAN. DAS and SAN deal with physical block I/O and NAS deals with file I/O. Disk drives are electro-mechanical devices that physically organize data into blocks, sectors and tracks on a spinning magnetic surface. Disk drives only understand requests for physical block data and have no concept of file data or logical I/O requests. The blocks are organized on a disk drive in terms of sectors and tracks. Multiple blocks, sectors or tracks can make up a file but the disk drive has no knowledge of this. The translation between the logical file data, that all application software programs require, and the physical organization of this data into blocks on a disk drive is the job of the file system. For DAS and SAN the file system is part of the UNIX®, Windows NT® or Windows 2000® operating systems. In the case of NAS the remote file system resides on the NAS Server. For DAS and SAN, the translation of files to data blocks is done at the file system level on the client or server system. In the case of NAS this translation of file requests to block data is done by the NAS file system thereby offloading work from the client or server system[1]. This is shown in **Figure** 7.

NAS offloads block retrieval processing from the local client or servers file system.

[1] Appendix A contains a detailed discussion of the block retrieval processing that is offloaded to the NAS server from the local client or server.

Figure 7 – I/O Data Flow differences of the major storage architectures.



As can be seen in this illustration, the application software program first makes a logical (File) request to the file system of the client or server. For DAS and SAN (whether it is Ethernet- or Fibre Channel-based), the local file system determines where the block data resides and sends SCSI commands to local or remote disk controllers.[2] These SCSI commands are packaged into a SCSI Control Data Block (CDB) and sent to a locally attached SCSI controller for data retrieval and return to the local file system. SAN architectures insert an Ethernet or Fibre Channel network in place of the SCSI bus and encapsulate the SCSI command using either Ethernet or Fibre Channel Host Bus Adapters (HBAs)[4] to package the command and manage network transmissions. So what is sent over the SCSI bus in the case of DAS is the same SCSI command that is sent over the network in the case of SAN whether it is Fibre Channel FCP/FC (FC-SAN) or Ethernet TCP/IP (E-SAN) based. For NAS, the remote NAS file system on the NAS server sends SCSI commands to the remote disk controller. In the case of NAS all data blocks that comprise a particular file request are sent back to the requesting client or server system over the network. So in the case of SAN architectures the traffic over the network is "block requests" both to and from the disk drives. In the case of NAS, the network traffic is a "file request" one way and a "stream of data blocks" (representing the requested file) on the return trip.

The network and file system location also defines the storage architecture

At a higher level Direct Attached Storage (DAS) can be viewed as a three-element model, while NAS and SAN can be viewed as four element models (see **Figure 8**). The

[2]Although different types of disk drives (SCSI or IDE) respond to different types of block level requests, most enterprise storage systems use disk drives based on SCSI. Therefore SCSI is used in this example.

fourth element is the addition of a network in the I/O path. Unlike DAS, both Network Attached Storage (NAS) and Storage Area Networks (SAN) have in common the fact that I/O requests pass over a network before arriving at the disk where the desired data resides. In addition, both NAS and SAN have in common that the data itself passes over a network a second time before being delivered back to the requesting application.

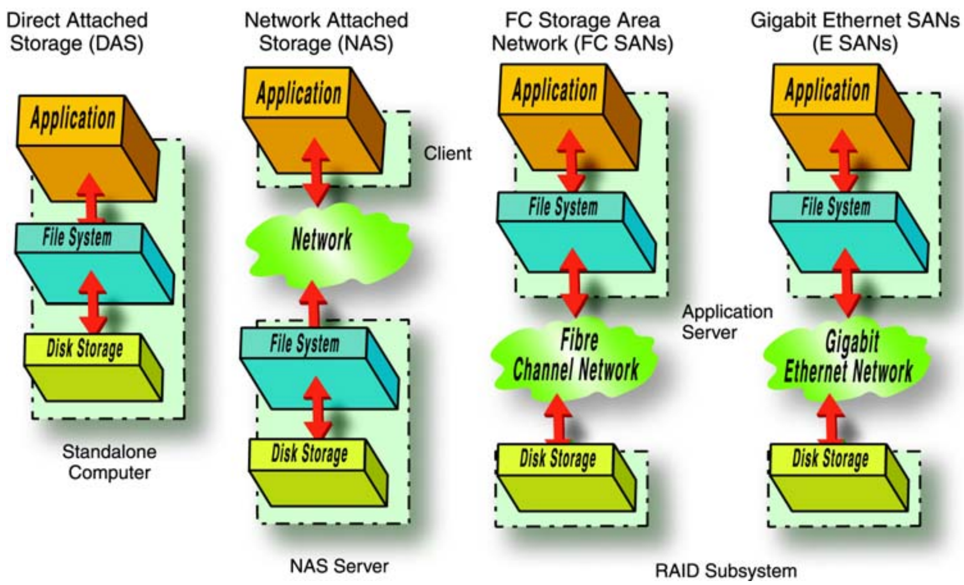


Figure 8 – Location of the network determines the storage model.

Both the request and the returned data pass through many layers of software as the I/O makes its way from application to disk and back again. NAS and SAN, however, differ with respect to where the network is placed relative to the file system. In the DAS and SAN models, the file system remains with the application, while with NAS it is remote.

The fundamental differences between DAS, NAS and SAN, therefore, can be explained in terms of 1) whether or not a network is involved in the I/O path, 2) the placement of the network relative to the file system and 3) whether logical physical data requests are made over the network.

The Direct Attached Storage (DAS) model

The Direct Attached Storage (DAS) model can be thought of as the way computer systems worked before networks. The DAS model contains three basic software layers: application software, file system software (which is part of the UNIX® or Windows NT® operating system software) and disk controller software. The elements are usually located close together physically and operate as a single entity. In the DAS model, the UNIX® or Windows NT® application software makes an I/O request to the file system that organizes files and directories on each individual disk partition into a single hierarchy. The file system also manages buffer cache in UNIX®.

When database applications are installed, the database software sometimes bypasses the UNIX® buffer cache and provides its own cache as with Oracle's System Global Area (SGA). The file system or database software determines the location of the I/O requested by the application and manages all caching activity. If the data is not in cache, the file system then makes a request to the disk controller software that retrieves the data from its disks or RAID array and returns the data to the file system to complete the I/O process.

The Network Attached Storage (NAS) model

The NAS model was made possible because NFS for UNIX® or CIFS for Windows® allows a file system to be located or mounted remotely and accessed over a network, instead of residing on the application server. In the NAS model the application software makes a network request for I/O to the remote file system mounted on a NAS server. The file system on the NAS server determines the location of the data requested by the client application and manages all caching activity. If the data is not in cache, the NAS file system then makes a request to the disk controller software, which retrieves the data from its disks or RAID array and returns the data to the NAS file system, which returns the data to the client across the network. Compared to DAS, NAS servers off-load all of the functions of organizing and accessing all directories and data on disk and managing cache. This frees the server's CPU to do additional work, thereby reducing potential CPU bottlenecks.

The Storage Area Network (SAN) model

In the SAN model, the file system continues to reside on the application server. As in the case of the DAS model, the server performs its normal file system functions of organizing and accessing all files and directories on each individual disk partition and managing all caching activity. Unlike NAS, there is no reduced workload for the client or server processor because the "logical to physical" translation process has not been offloaded. However, a SAN does offer the benefits of storage resource pooling and LAN-free backup.

Technology maturity: NAS standards versus SAN vision

When compared to DAS, both NAS and SAN strive to bring the benefits of networking technology to storage architecture planning and data management. However, SAN of either type (FC-SAN or E-SAN) is not as mature as NAS. FC-SAN has weaknesses in security, and congestion control is not automatic. E-SAN and i-SCSI offer promise but are not yet proven.

Deciding When to Use DAS, NAS or SAN

3

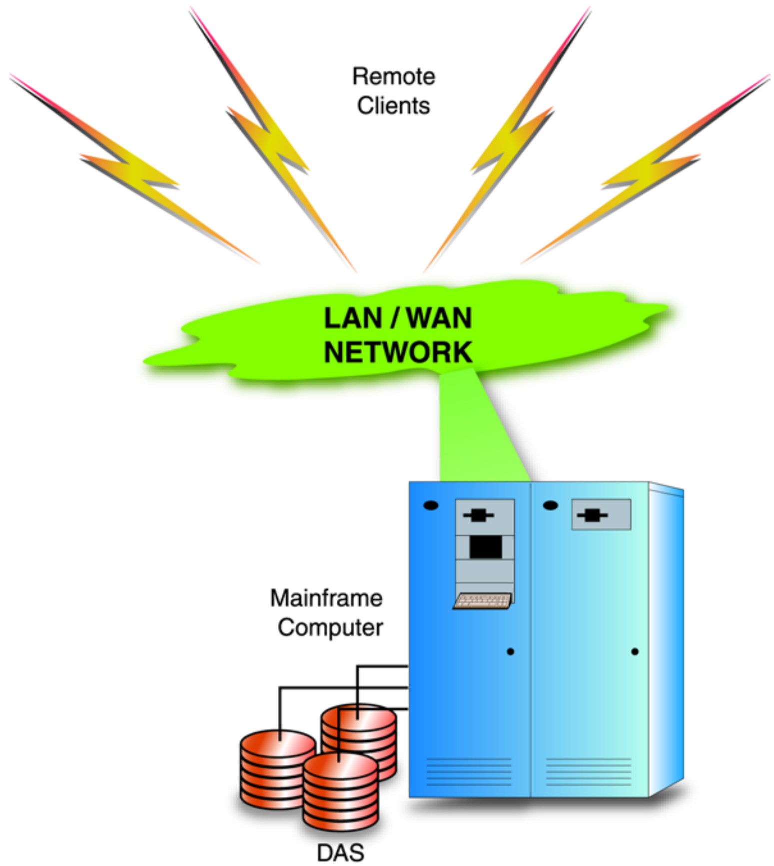
DAS, NAS and SAN have important differences

Most industry analysts agree that all three storage models serve different and complementary roles in the enterprise storage architecture. In addition, most analysts believe that NAS and SAN will coexist and eventually be integrated into a unified architecture. Understanding the differences between the three models can help identify which architecture is appropriate for different enterprise applications (See **Table 1**). Until NAS and SAN are integrated, the enterprise can realize maximum non-IT department business benefits from NAS and the business benefit of streamlined IT operations and management from SAN.

Storage Characteristic	DAS	NAS	SAN
Use of standard network file sharing protocols	N/A	Yes - NFS/CIFS	No
Single image data sharing across heterogeneous OS	No	Yes	Requires Gateways
Installation	Varies	Very easy	Very difficult
Centralization of Management	No	Yes	No
Ease of Management	Varies	Web-based	Usually difficult
Storage focus	Server centric	Network centric	Channel centric
Intelligence	Varies	Yes	Future file systems
Server independent backup	No	Using NDMP or direct back end connections	Requires special software
Disaster tolerance	Custom Solutions	Emerging Standards	Proprietary Solutions
Offloads work from application server	No	Yes	No
Removal of backup data flow from the LAN or SAN	No	Yes	Varying degrees of difficulty
Server free backup	Yes, but requires network bandwidth	Yes	Yes, but block data only

Table 1 – Key differences between DAS, NAS and SAN architectures

Figure 9 – A performance critical Customer Reservation System (CRS) or Financial Database



DAS applications are appropriate for very low-end and high-end applications

Low-end environments

DAS is most suitable for very low-end applications, like in home office environments and with portable computers. In this case it can be difficult and expensive to add networking to the storage architecture.

High-end environments

Some high-end environments use mainframe computers with large disk farms of DAS devices that are located close to the application server. This can provide acceptable performance for I/O intensive applications like customer reservation systems and large financial databases that contain many small records.

NAS applications business and infrastructure benefits of UNIX and NT file sharing applications

NAS is the only form of storage that optimally supports both NFS and CIFS network file system protocols for sharing storage between UNIX® and Windows NT® hosts. Since they offer standardized, reliable and integrated file locking, NAS servers are suitable for *many applications where business advantage can be gained from sharing data between UNIX® and Windows NT® clients.* It is important to note that file sharing applications provide not only business benefits to non-IT departments but also provide a major infrastructure benefit to the IT department – increased storage utilization. This is accomplished because only one copy of file information is required instead of multiple copies as occurs with data exchange schemes that are sometimes used to make a translated copy of information from one operating system to another.

Why Auspex File Sharing is "Best of Breed"

The NS3000 supports both NFS and CIFS network file sharing protocols to allow ***Universal Data Sharing (UDS)*** between UNIX® and Windows NT® hosts. This is important because ***Universal Data Sharing (UDS)*** can often result in a business advantage for the enterprise. For this reason, about 75% of Auspex users share data between UNIX® and Windows NT®. The NS3000 ***Universal Data Sharing (UDS)*** feature offers any user full "read/write" file privileges to any file on an NS3000™ Series system. This is a major advantage compared to competitive products where emulation software is installed or files need to be partitioned where certain users have "read only" access. In a software development environment or CAD environment, some engineers may use UNIX® workstations whereas others use NT® to access and update the same data.

There are four basic approaches for supporting mixed UNIX®/NT® environments.

- Separate UNIX® and NT® servers accessed by different clients.
- Client-based emulation.
- Server-based emulation.
- Bilingual Network Attached Storage (NAS) with universal file sharing such as the Auspex NS3000™.

These approaches are discussed in detail in Appendix B.

Excellence in information sharing is Auspex's commitment to its customers. The company's expertise in supporting information sharing applications and implementing information sharing infrastructure is unsurpassed. Regardless of whether users have UNIX® or Windows® computers, they can't help but become more productive with an Auspex file server. Heterogeneous information sharing between UNIX® and Windows® hosts from a single image of data is perhaps the most important feature to evaluate when selecting a file server for file sharing to enable business benefits. The Auspex Net OS for CIFS software solution for UNIX® and Windows® data sharing represents the *best of breed* available in the industry. This is due to Net OS for CIFS concurrent native file locking routines for both UNIX® and Windows® universal information sharing protocols **and** an Auspex unique performance load-balancing feature between UNIX® and Windows® users so that all users have equal priority of access.

Information sharing between employees and business partners has been shown to dramatically increase productivity across all functional departments of a business. If an enterprise's *Engineering* department designs products using computer software form UNIX® and Windows® platforms, then Nett OS for CIFS ***Universal Data Sharing (UDS)*** can be of critical importance to employee productivity. ***Universal Data Sharing (UDS)*** also helps a business increase the efficiency of *Manufacturing* operations in the areas of Supply Chain Management (SCM) and Business to Business Integration (B2B) systems. File sharing solutions can increase *Marketing* and *Sales* agility relative to competitors to generate more business through effective collaboration. The business benefit of the Auspex infrastructure solutions is in streamlining *IT* operations and reducing costs.

Universal Data Sharing can be critically important to employee productivity.

UNIX® and Windows® clients can share any file with full "read/write" privileges

Unparalleled performance.

Net OS for CIFS is an Auspex software solution that allows UNIX® and Windows® clients to share any file securely stored and managed on Auspex file servers with full "read/write" privileges. The functionality of this software is shown in **Figure 10**. Since only one physical copy of data is stored on the Auspex file server, costs are reduced and true high performance multi-lingual Universal Data Sharing (UDS) is enabled. Since no client emulation software is required, access is transparent to users and there is no performance penalty as in implementations with Samba® or other intermediate software solutions. Administration is facilitated through native UNIX® or Windows® tools so all existing IT skills are preserved. Through an industry unique "One-World View" of UNIX® and Windows® files, and the distribution of file sharing operations to multiple processors on the Auspex file server, unparalleled performance and data security are possible. This file sharing load balancing is also industry unique.

In addition, Net OS for CIFS requires only one physical copy of data, which increases data reliability and ease of administration dramatically. Since no client software is required, access is transparent to users. Through an optimized implementation, the Net OS for CIFS Universal Data Sharing (UDS) load-balancing design provides even performance, reliability and scalability to all users.

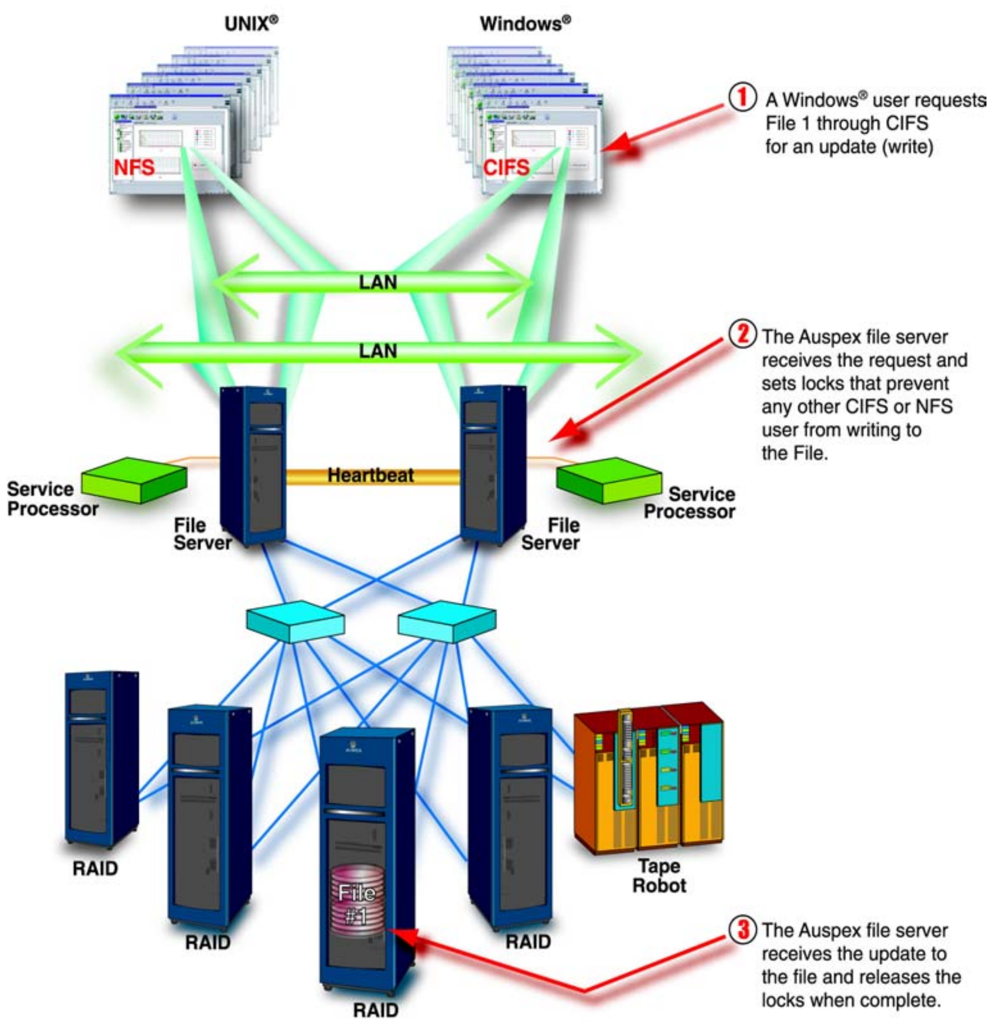


Figure 10 – Net OS for CIFS provides secure Universal Data Sharing (UDS) with load balancing.

In the early days of the Internet, Web site content was limited to static text and images.

With "One-World" permissions, changes on a file from either Windows® or UNIX® clients will result in the proper permission changes being made to the other file access protocol.

Streaming Media applications

In the early days of the Internet, website content was limited to static text and images. As high bandwidth DSL, cable and fibre optic connections became available, video and audio media content could be delivered effectively over the Internet. Movies, music, video-based training and live video could be provided to a global audience for both consumer and enterprise applications. This is no easy task since most types of multimedia content are still recorded in analog and must be converted to large digital files and then delivered to many remote users without interruption to the picture or audio. To accomplish this, a media content supplier must create and transfer huge amounts of data over the Internet in the most fault tolerant and efficient manner possible.

The Auspex streaming media NAS solution gives providers of multimedia Internet content the three things they want the most. 1) An unbroken multi-media user experience. 2) Virtually every conceivable option for high availability and high performance. 3) Development capability on UNIX®, Windows® or Mac OS X® platforms with UDS from a single image of data. From the primary file copy at the content provider's website to the client's multimedia player, the Auspex streaming media solution provides benefits not available from any other vendor. Auspex engineers will help provide for the proper deployment of all network, server, storage and telecommunications equipment so that complete fault tolerance is achieved for the entire computing environment. Streaming media is an ideal application for Network Attached Storage (NAS) and the Auspex NetServer™ family of hardware and software products.

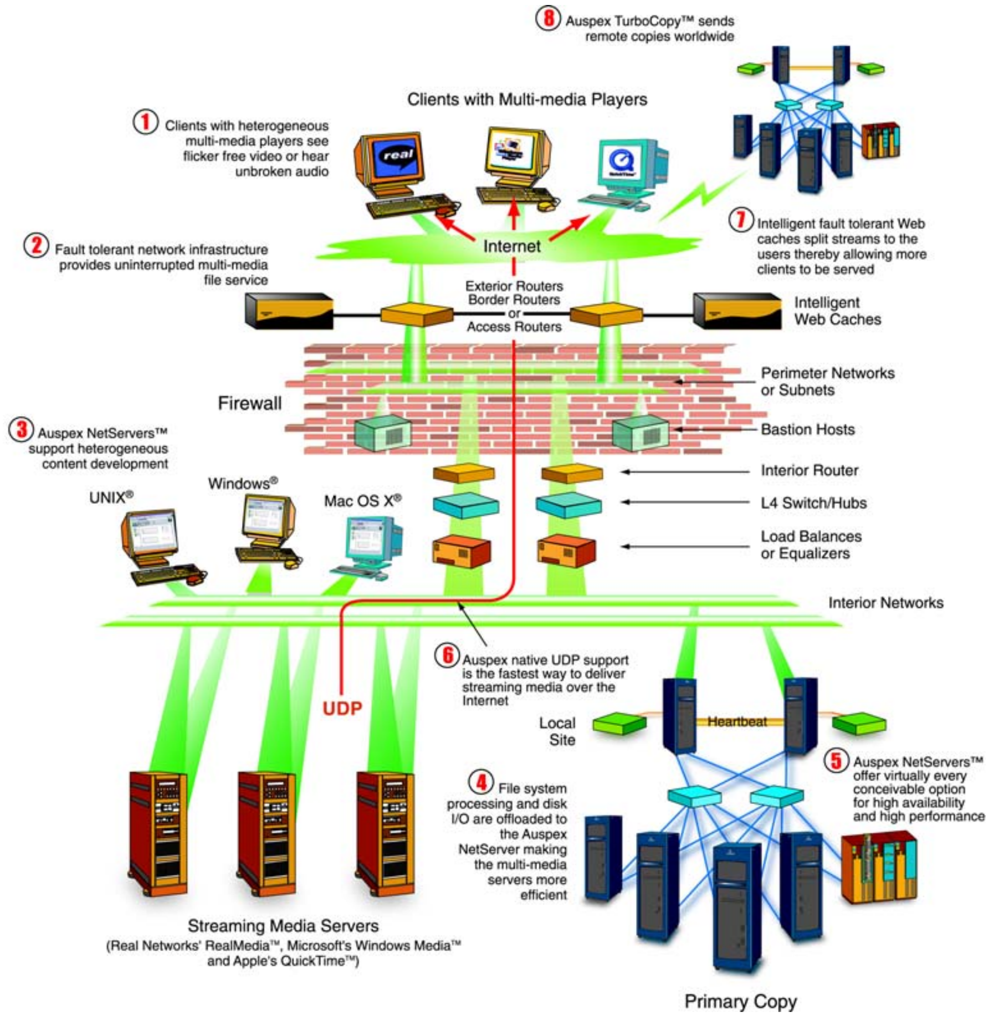
The end objective for a streaming media content provider is to make sure clients enjoy an unbroken video or audio experience while having a cost effective and flexible environment for content development. To provide uninterrupted file service, it is important to design a completely fault tolerant computing infrastructure. For the important job of content development, Auspex NetServers™ support heterogeneous content development on UNIX®, Windows® and Mac OS X® platforms. The Auspex "best of breed" NAS design offloads both file system processing and disk I/O from the media servers thereby making them more efficient with very high performance in terms of data throughput. Auspex NetServers™ also offer virtually every conceivable option for high availability from the disk to the network so that any possible failure mode can be tolerated. A very important and competitively unique feature of the Auspex streaming media solution is that both TCP and UDP protocols are supported natively in the file server hardware. UDP is highly recommended for streaming media because it does not break the media stream in the event of lost packets, as happens with TCP. Whether utilizing UDP or TCP, however, the NetServer™ delivers streaming media more efficiently through the local network and out over the Internet to enhance the client's viewing and/or listening experience. To improve performance by reducing the impact of firewall delays, intelligent fault tolerant Web caches split media streams to the users thereby allowing more clients to be served simultaneously. To further improve performance by reducing Internet hops, remote copies of the media files are often pre-positioned at the edge of the Internet or on local LANs for large multi-site enterprises. Auspex TurboCopy™ software is ideal for this task and sends remote copies worldwide to reduce the distance penalties of streaming files from only one location.

The Auspex streaming media NAS solution gives providers of multimedia Internet content the three things they want the most.

The end objective for a streaming media content provider is to make sure clients enjoy an unbroken video of audio experience.

UDP is highly recommended for streaming media.

Figure 11 – The Auspex solution gives streaming media providers the three things they want the most.



Consolidated file-serving applications

Another common deployment of NAS is in application environments where *storage can be consolidated* from numerous and distributed UNIX® and Windows NT® servers to a reliable NAS platform such as the Auspex NetServer 3000™ (see **Figure 12** and **Figure 13**). This helps reduce operating costs by centralizing data management functions. It also improves performance and availability by consolidating environments where user content would otherwise be spread over 100 or more application servers.

In these highly decentralized environments, server node failures could lock users out of critical data if the content is not adequately protected. Robert Gray of International Data Corporation (IDC) discusses these cost and management advantages in a white paper titled ***Network Attached Storage: A Compelling Story for Storage Consolidation***, available for download at http://www.auspex.com.

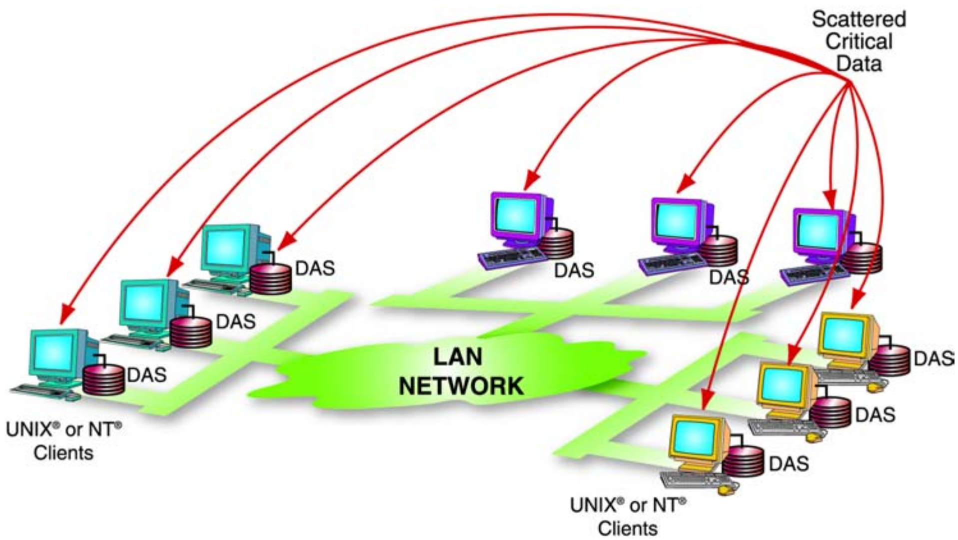


Figure 12 – Distributed storage, before consolidation.

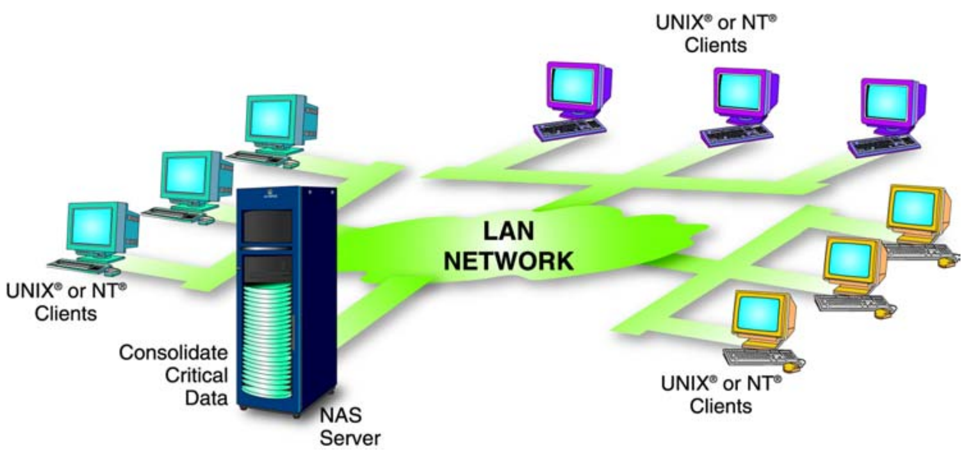


Figure 13 – Distributed storage, after consolidation.

Technical and scientific applications

NAS is appropriate for technical engineering applications like oil and gas exploration, computer aided software engineering (CASE) and mechanical engineering, where a large file or group of files are simultaneously accessed by multiple engineers (see **Figure 14**). In this environment, geoscientists would be able to simultaneously perform data analysis on large databases or graphics.

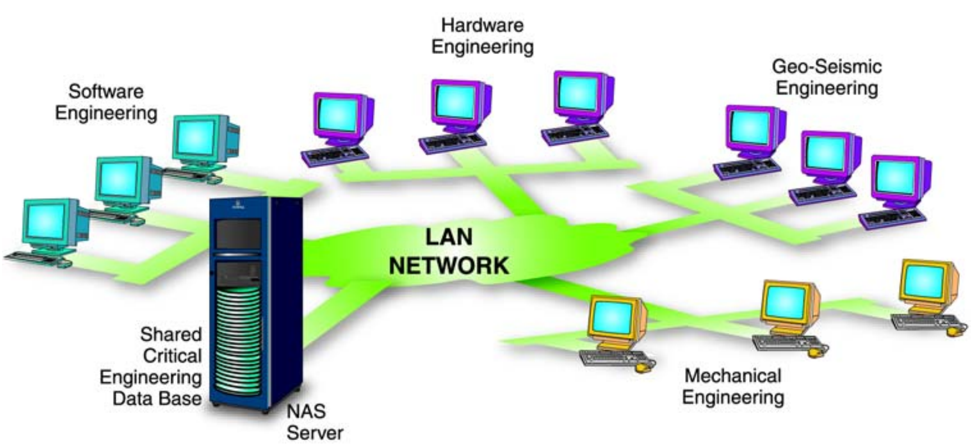


Figure 14 – An MCAD, ECAD, software development or geoseismic engineering application.

NAS is also ideal for *software development* environments in which engineers check code daily during development, then return it to the NAS server periodically throughout the day. *Mechanical CAD* (e.g., airplane, automobile or machine design) and *Electrical CAD* (e.g., circuit board design) are similar types of technical applications in which large technical objects are exposed to a development process that requires high availability, high performance, consistent and predictable file service and UNIX® and Windows® data sharing. An added benefit of off-loading file serving in these environments is increased performance. This results from removing file serving workload from compute-intensive application servers onto the NAS server. Off-loading file serving with a dedicated NAS server frees processing power for high-end, high-cost, servers to perform critical, CPU intensive calculations, such as 3-D rendering in automotive design.

Product Data Management (PDM) applications are ideally suited for NAS

Consisting of the integration of CAD, CAM and CAE, the PDM (Product Data Management) market is ideally suited for NAS according to many industry analysts. This is because of a NAS system's ability to fulfill the multi-department business needs for reliable storage with fast delivery of graphically intense data from a single data image. In addition, this large industry was heavily based on UNIX® and is now rapidly moving to NT®, particularly at the client level. PDM is a term that applies to both the discrete and process manufacturing industries and denotes the systems and methods that provide an electronically integrated structure for all types of information. PDM systems can be used to define, manufacture and support how products are stored, managed and controlled.

PDM applications usually comprise huge amounts of data.

PDM applications are ideally suited for NAS.

The PDM system manages the product development process as well as the data. PDM systems control product information, states, approval processes, authorizations and other activities that impact product data. PDM systems also provide data management and security, and ensure that users always get and share the most recent, approved information. PDM applications usually comprise *huge amounts of data*, and are therefore ideally suited for NAS servers such as the Auspex NetServer™ product family. PDM applications are typically implemented on top of a shared database such as Oracle. The first tasks in utilizing PDM tools are to develop a “product definition/structure” to describe what the product is and a “process plan” to describe how a product flows through the departments of a company. It is important to note that the flow is NOT a serial process, and as such, PDM is sometimes described as an enabler of Concurrent or Collaborative Engineering (CE).

At the beginning of a discrete manufacturing product life cycle, products are MCAD (Mechanical Computer Aided Design) systems. MCAD is a term used to denote the software and hardware used by mechanical design engineers to generate an electronic model/simulation of mechanical parts and assemblies. The parts can be combined into assemblies and checked for interference, proper tolerance, form and function. They can be passed to engineering analysts for structural, thermal and dynamic compliance and in parallel to manufacturing for tool building and production (Computer Aided Manufacturing or CAM). They can be viewed by planning and purchasing departments for MRP (Material Resource Planning) activities. Parts can also be viewed by other organizations such as the technical publications department. “Alerting” software is used to notify users when a document of assembly has changed. Parts and assemblies can often be shared with suppliers for work that is outsourced or viewed by customers for preliminary approval.

The CAD/CAM/CAE/PDM market consists of a highly diverse universe of OEMs, partners, suppliers and customers with vastly different computing platforms, users, budgets and technical needs. Many vendors are competing solely on new product functionality and price with low-end Windows® computers. To address these platforms, some vendors have developed or bought mid-range products. However, none of these mid-range offerings are compatible with the customers' high-end UNIX® systems. Although some vendors view the Windows® market and high-end CAD/CAM/CAE as two wholly distinct markets, Auspex partner UGS views them as one emerging marketplace. UGS offers a graduated series of CAD/CAM/CAE/PDM products that have been optimized for each level of user throughout the organization including i-man®. i-man® is an Internet-centric software product used by manufacturing enterprises to manage product content for collaborative commerce. i-man® automates the management of data and documents throughout a product's life cycle, assuring that access to appropriate information is given to authorized personnel and that everyone works with up-to-date files. The emphasis is on ease of use, concurrent collaboration and security through best practice industry techniques. Whatever a person's job function, he or she can easily retrieve needed information, readily understand its status in the development cycle and immediately update it for further use.

i-man® is an Internet-centric software product used by manufacturing enterprises to manage product content for collaborative commerce.

Within an i-man® environment, an i-man® file system daemon processes program variables and graphical images. Specifically, it provides read or write access permissions to data residing on mass storage devices. Since the access time it takes to locate a single byte of information on a mass-storage device is affected by the location of the file system daemon, locating the i-man® daemon on the Auspex NetServer™ is the most efficient possible way to reduce access time. This fact and the fact that Auspex is the ONLY certified NAS vendor for i-man® applications validates Auspex's "best of breed" multilingual file sharing capability. Since i-man® applications often involve 10's of terabytes of data, the Auspex large-scale multilingual file system with Universal Data Sharing (UDS) becomes a critical business advantage. **Chapter 5** discusses the importance of a file system's ability to provide *any* user with full read/write access from *any* network to *any* information, *any*time, *any*where from a single large scale data image.

Auspex is the ONLY certified NAS vendor for i-man® applications.

Software Configuration Management (SCM) Applications

Driven by the need to speed delivery of high quality software products, many organizations are moving from a "waterfall" development process to an iterative approach. This enables rapid, continuous software development cycles for multiple projects and includes quality testing throughout the cycle. It also demands robust workflow management and a network and data storage infrastructure that meets the increasing data flow requirements created by individuals and development teams working in parallel. For example, Software Configuration Management (SCM) vendor Rational Software Corporation's ClearCase[3] product line optimizes workflow by applying best practices to version control, workspace management, configuration management and change control. Auspex File Servers™ support and enhance ClearCase™ functionality by enabling the implementation of a robust, high availability, high performance network and centralized storage infrastructure that delivers optimal data flow with simplified system management. Together, Rational ClearCase™ and Auspex network and storage solutions provide a comprehensive integrated solution enabling fast, iterative development of high quality software products.

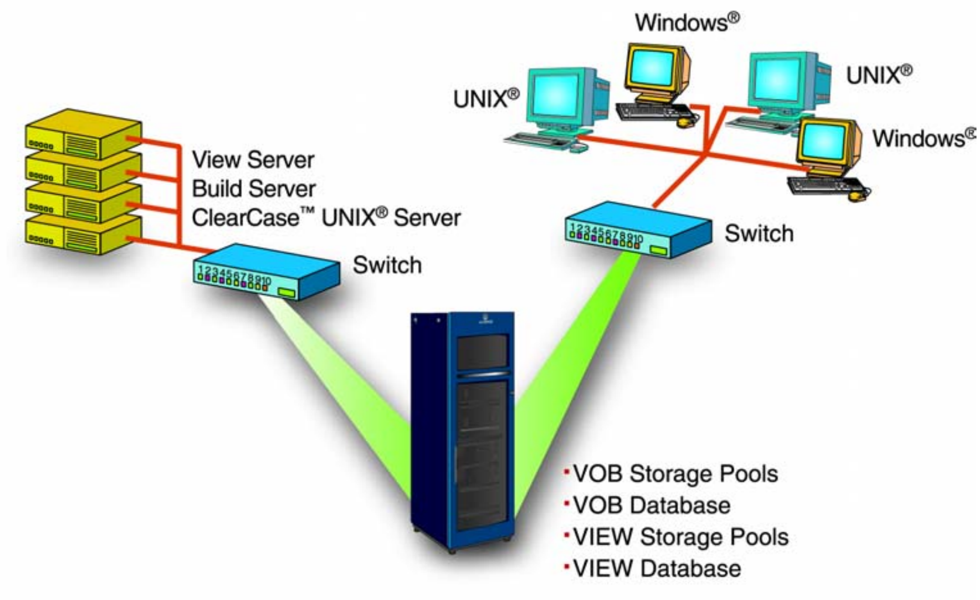
Enables rapid, continuous software development cycles.

As illustrated in **Figure 15**, the Rational ClearCase™ software and Auspex NetServers™ are configured in a way that all ClearCase™ data, including source code, build data and database tables are stored and accessed on Auspex File Servers™, while VOB/View serving and database processing are executed on separate UNIX® servers. This approach provides maximum data flow performance since dedicated machines are optimally applied for specific functions – data I/O and application/database computing. However, there is a great deal of configuration flexibility, which, for example, allows users to locate the Views local to the client workstations.

All ClearCase™ data, including source code, build data and database tables are stored and accessed on Auspex File Servers™.

[3] Auspex File Servers™ are certified to support local and multi-site deployments of Rational Software Corporation's ClearCase® Software Configuration Management (SCM) software for both UNIX®-only and UNIX®/Windows® heterogeneous environments.

Figure 15 – Auspex Net Servers™ are certified to be interoperable with Rational ClearCase™ Software products.



By storing all ClearCase™ data on the Auspex server, all developers, regardless of whether their workstation is UNIX® or Windows®-based, can quickly and seamlessly access and share all project files. The Auspex server supports a true "one-world view" for Windows® and UNIX® clients, which enables sharing of single files while maintaining access permissions and security for both environments. This capability also eliminates the need for client-side or server-side NFS emulation to support a mixed UNIX®/Windows® client environment. (See also Appendix B for a discussion of the inefficiencies of emulation approaches to file sharing.)

Internet and Intranet network attached storage applications

The computing architectures of ISPs, SSPs, ASPs and corporate websites have grown up as quickly as the Internet to support the booming Business to Business (B2B) and Business to Consumer (B2C) applications that comprise E-commerce. Network Attached Storage servers such as the Auspex NetServer™ family of products can optimize website architectures for the increasingly critical requirements of 24x7, flexibility, speed and manageability. Recent well-publicized security attacks and system outages due to server failures have underscored this need. As shown in **Figure 16**, Network Attached Storage (NAS) should be deployed for recommended application in the Internet site along with other website components in order to ensure the best results.

With the recent growth in the storage-on-demand market, new storage solution providers (SSPs) in particular can benefit from NAS. The speed and flexibility of the Auspex NS3000™ family of products provide unique value for SSPs who serve customers with different requirements for data protection and data access. Auspex FMP and standard-based networking options provide additional value for SSPs who deliver data over high speed IP networks to their "on-demand" customers.

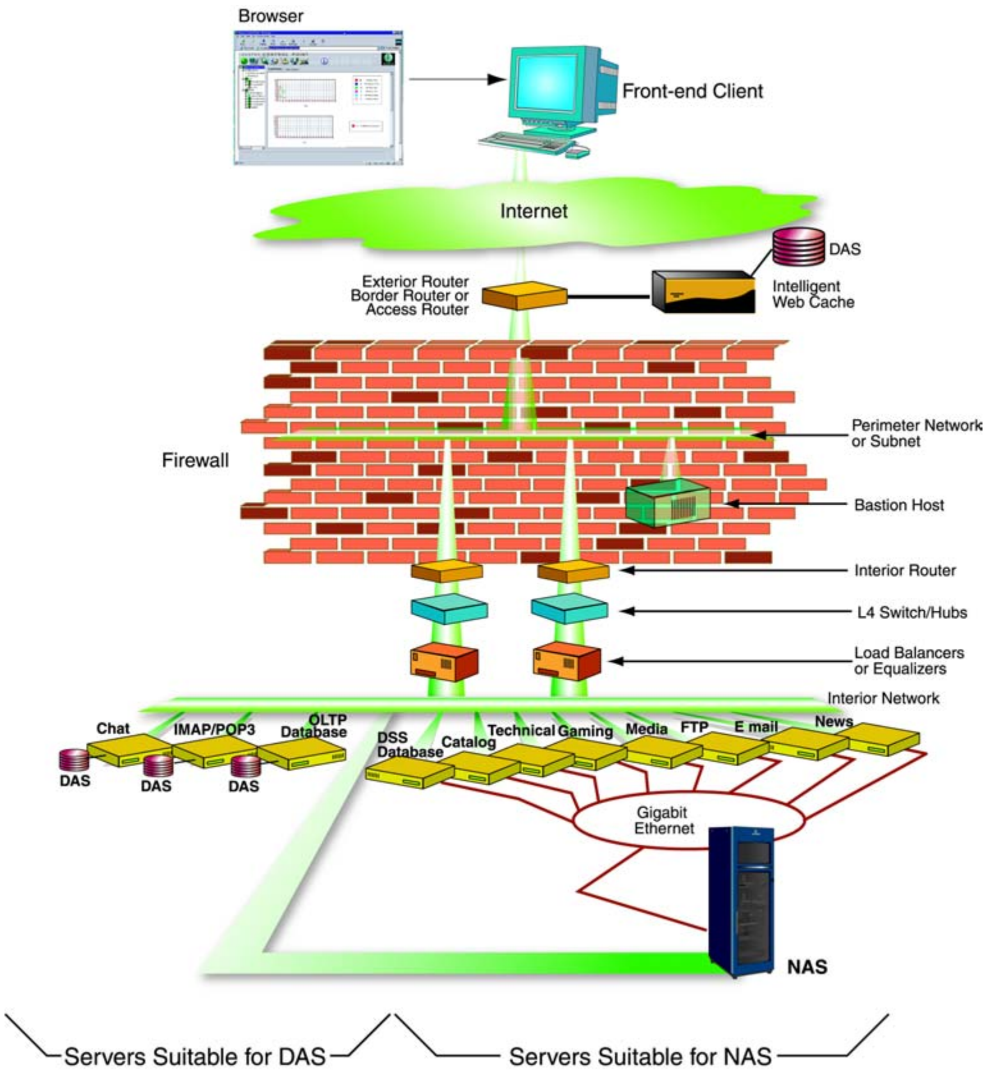


Figure 16 – A multi-level functionally specialized hierarchical website architecture provides an optimal balance of 24x7, flexibility, speed and manageability .

Challenges faced in optimizing website architectures

There are many website architecture design challenges faced by B2B and B2C Internet technology enterprises. When a "fat O/S" model Direct Attached Storage (DAS) is used for large scale Web server farms in a flat cluster – like topology, the same problems experienced by corporations with many distributed general purpose servers are encountered. (See **Figure 12**.) This is because an individual Web server is asked to do more than it is designed to do.

Flat two-tier architectures ask the servers to process both IP requests, file system requests and manage storage.

Servers are designed to be compute intensive, yet this flat two-tier architecture asks the servers to process both IP requests, file system requests and manage storage for whatever domain names are associated with that particular server. Reliability is decreased as each server's working set increases to perform both IP requests and filesystem processing.

There is no single point of disk backup, and an inefficient network based backup scheme is often used. Disk utilization is inefficient since disk space cannot be allocated as needed and files must often be replicated. Load balancing techniques are less easily applied since there is no clean way to divorce file system processing from IP-request processing.

Implementing a multi-level functionally specialized website architecture solves these problems.

Implementing a multilevel functionally specialized website architecture with devices optimized to do only one job well can solve these problems. This architecture balances computing tasks in a way that enables an optimal balance of 24x7, flexibility, speed and manageability in a cost effective manner not achieved by other designs.

The back-end NAS servers functionally balance workloads and improve overall website performance.

- Border Routers route traffic to and from the Internet.
- Specialized front-end Intelligent Web Caches, such as the product series of Cache Flow Inc., handle many IP requests in front of the firewall and relieve unnecessary I/O activity from the firewall, L4 switches, servers and storage on the back-end networks.
- Routers and bastion hosts on the subnet in the firewall provide security.
- Switches, hubs and load balancers provide redundancy and route traffic.
- Servers with Direct Attached Storage (DAS) may support heavily transaction oriented I/O workloads particularly those having a large percentage of writes.
- Optimized "thin O/S" Web servers improve reliability and availability by running only the minimum code necessary to accomplish the specific function of Web serving for non-transaction oriented workloads where I/Os tend to be large and data is referenced repeatedly.
- Back-end NAS servers consolidate data, improve availability, reduce backup windows and improve network and system manageability. The back-end NAS servers are optimized for moving raw data between disks and networks, and remove the file system's workload from the front-end servers to functionally balance overall website performance. The back-end NAS server provides a common access point and disk pool for Web server applications such as HTML pages, CGI scripts, mailspool files and news.

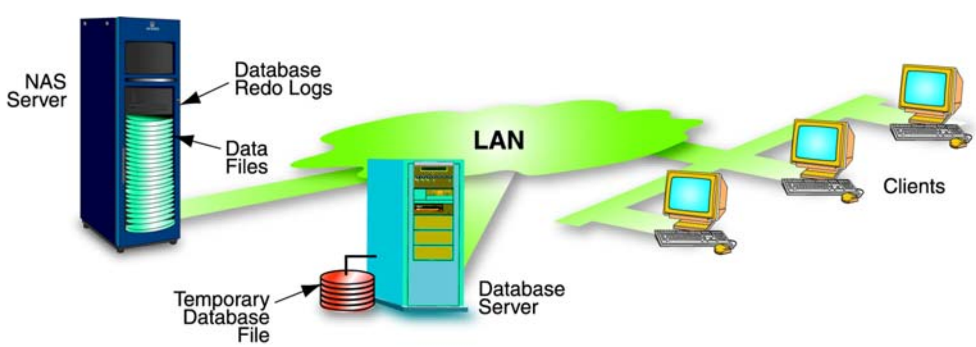
Decision Support (DSS) applications

Many companies are striving to gain a sharp competitive edge by building decision support applications, like data warehousing, data marts and data mining. In these applications, companies collect large amounts of data, then conduct statistical analysis to identify important trends. NAS products are well suited for decision support applications because the massive amounts of source data required is often stored outside the database server and periodically read into the data mart for off-line queries.

In some cases it is not necessary to populate the data mart's disks with data but to leave the original data on NAS and read it in over the network instead of over an I/O channel. Therefore, NAS can save hours in preparing the data mart with refreshed data for the DSS (query) process.

In reading the original data directly from NAS, a typical data warehouse or data mart query involves a process known as a *join.* In a *join,* two database files are read into the server memory and the data from the two files is combined into a larger file. Since the files are often much larger than the database server's memory, the *join* is done on pieces of each table, creating temporary database files. These temporary files are sometimes stored on the database server's disks and then linked together to create the final *joined* database file. Putting these temporary database files on the server's DAS disks (**Figure 17**) can reduce the additional network traffic that occurs from sending them back to the NAS server where the source data resides. A further discussion of how NAS devices are deployed for DSS and other relational database applications is available in an Oracle white paper prepared by Auspex, available at http://www.auspex.com.

Figure 17 – Decision Support (DSS) application with database files on NAS and temporary database files on the database server's DAS.



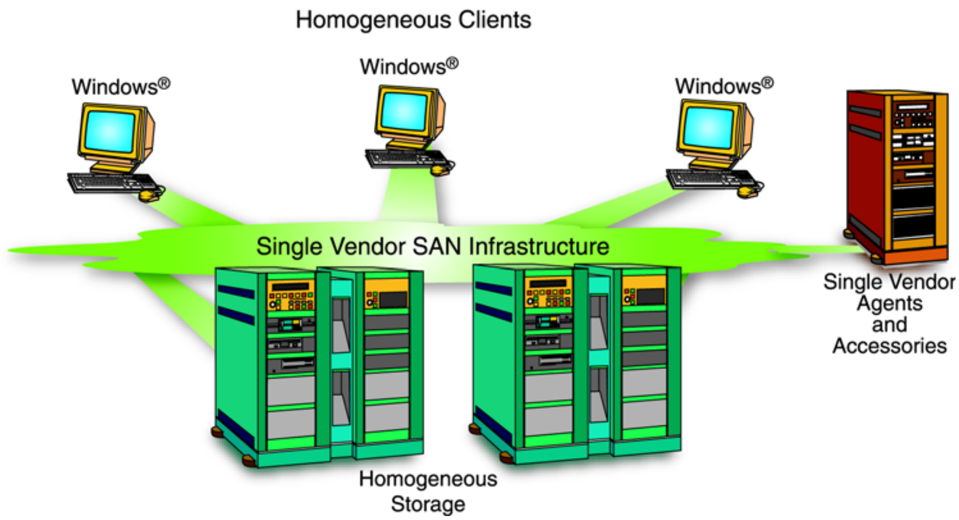


Figure 18 – Most SANs deployed today are single vendor systems

FC-SAN applications

When considering an FC-SAN application it is important to note that a standardized and open FC-SAN architecture is still a vision. Early adopters of SAN technology have implemented proprietary SAN applications. For example, EMC markets a proprietary SAN known as Enterprise Storage Networks (ESN), while Compaq markets its own proprietary SAN known as Enterprise Network Storage Architecture (or ENSA).

Applications for centralized management benefits

The primary SAN vision is in the creation of a common storage pool from which *any application* has access through the SAN to *any storage device.* Unfortunately, data sharing and heterogeneous platform interoperability are not available today in SAN, as they are on NAS. This means that SANs today are largely deployed in order to centrally manage large amounts of data. To reduce administration costs, SAN storage is divided into "pools" (based on operating system type). Disk drives supporting different operating systems are housed in the same cabinet, but assigned to different operating systems, with no UNIX® and Windows® data sharing without the use of gateways or meta data controller schemes such as Tivoli's SANergy™. This highlights the difference between SAN storage sharing (or Type I subsystem partitioning) and NAS Universal Data Sharing (UDS) as shown in **Table 2**.

Many storage and networking experts today recommend that the first implementations of SAN be done with homogeneous servers and homogeneous storage devices (see **Figure 18**). *A homogeneous SAN deployment would be appropriate with a proprietary SAN architecture* such as offered by EMC® or Compaq® but the enterprise becomes "locked in" to one vendor.

It is important to note that in *proprietary* SAN implementations it is possible to share volumes assigned to heterogeneous hosts within one type of storage subsystem.
For example, volumes assigned to Sun® servers can be shared in a Symmetrix® subsystem with volumes assigned to HP® servers. In this scenario, the Sun® volumes are accessible only to Sun® servers, and HP® volumes are accessible only to HP® servers. This is called *subsystem partitioning* and is not the *true data sharing* delivered by NAS. Regardless, while subsystem partitioning is sometimes called "data sharing" by some vendor sales people, it is more correctly identified as "resource cabinet sharing" or heterogeneous disk drive co-location. IBM® has established a convenient way of thinking of data sharing as shown in **Table 2**, with Level III (as with NAS) being the most desirable.

Table 2. IBM® definitions of three types of data sharing

Heterogenous Platform Data Sharing	Definition
IBM® Type I data sharing	Subsystem partitioning - Where disks assigned to heterogeneous servers share the same cabinet (SAN).
IBM® Type II data sharing	Data copy sharing - Where data is copied to a second disk and then accessed by a heterogeneous server.
IBM® Type III data sharing	True data-sharing - Concurrent reads and writes from heterogeneous servers to a common disk (NAS).

Applications that do not require true data sharing (Level III)

IT organizations often want fully open SANs where heterogeneous clients can access heterogeneous storage devices to share data on a common volume with concurrent read and write access. This is known as *true data sharing*—and it is not available on SAN today because it requires more intelligence on the storage system than is available from current SAN systems.

Since no SAN standards exist for SAN data sharing, it is prudent to use NAS for this purpose.

Due to a lack of standards in the SAN market, IT professionals should consider SAN only in applications that do not require true heterogeneous data sharing. Accepted network data sharing standards, such as NFS and CIFS, are being considered for custom SAN file systems to provide the intelligence for true interoperability between heterogeneous platforms and storage. However, this effort is far from completion and is one of the areas where SAN still falls behind NAS architecture.

Applications where security risks are well managed or low

It is usually best to implement early SAN deployment with low security applications; deployments should be conducted with applications that are not critical from a performance point of view

Many NAS vendors implement completely secure NFS and CIFS integrated locking schemes. By comparison, its Fibre Channel (FC) or Fibre Channel Protocol for SCSI (FCP) implementations expose SAN in five categories of security attacks that can allow data to be stolen or destroyed on a physically secure Fibre Channel fabric.

Please note that the network security terms used in **Table 3** are defined in the Glossary of Terms at the end of this report. See also the Fibre Channel Industry Association's home page at http://www.fibrechannel.com.

Proprietary SAN vendors may eventually implement schemes to compensate for these vulnerabilities in the Fibre Channel and Fibre Channel Protocol for SCSI specifications. In the meantime, IT professionals should determine whether the SAN architecture they are considering has a work-around to compensate for these security weaknesses. Zoning is one such potential work-around for building security into SANs. However, zoning is not standard, it tends to be too coarse for storage and too static for clients, and it defeats the basic objectives of SAN. Otherwise, *it is usually best to implement early SAN deployment with low security applications.*

SAN Security Risks that stem from FC/FCP and allow data to be stolen or destroyed	FC/FCP SANs	Networks/ NAS
Node Name/Port Name "spoofing" at Port Login time	Yes	No
Source Port ID "spoofing" on data-less FCP commands	Yes	No
"Snooping" and "spoofing" of Fibre Channel Arbitrated Loop (FC-AL)	Yes	No
"Snooping" and "spoofing" after fabric reconfiguration	Yes	No
"Denial of service" (DoS) attacks can be made in User mode	Yes	No

Table 3 – Security risks in Fibre Channel (FC) and Fibre Channel Protocol for SCSI (FCP)

Applications where congestion control can be managed to avoid performance bottlenecks

Along with excellent security, networks have sophisticated congestion control that is proven and standardized. Congestion can occur in a network or SAN when one link or node receives more traffic than it can handle. Congestion can cause performance bottlenecks in the data path over the SAN between application and disk.

Neither Fibre Channel nor FCP provides a good built-in method of congestion control. Some Fibre Channel switch vendors have implemented congestion control outside of the Fibre Channel standard, but FCP, which is based on SCSI, does not control congestion. (See **Figure 19**).

As with the five categories of FC and/or FCP security risks for SANs (See **Table 3**), IT professionals should consider whether the SAN vendors have successfully developed offsets to the congestion problems of FC or FCP. If not, early SAN deployments or initial production testing should be conducted with *applications that are not performance-critical* or where data flow can be accurately specified and planned in advance.

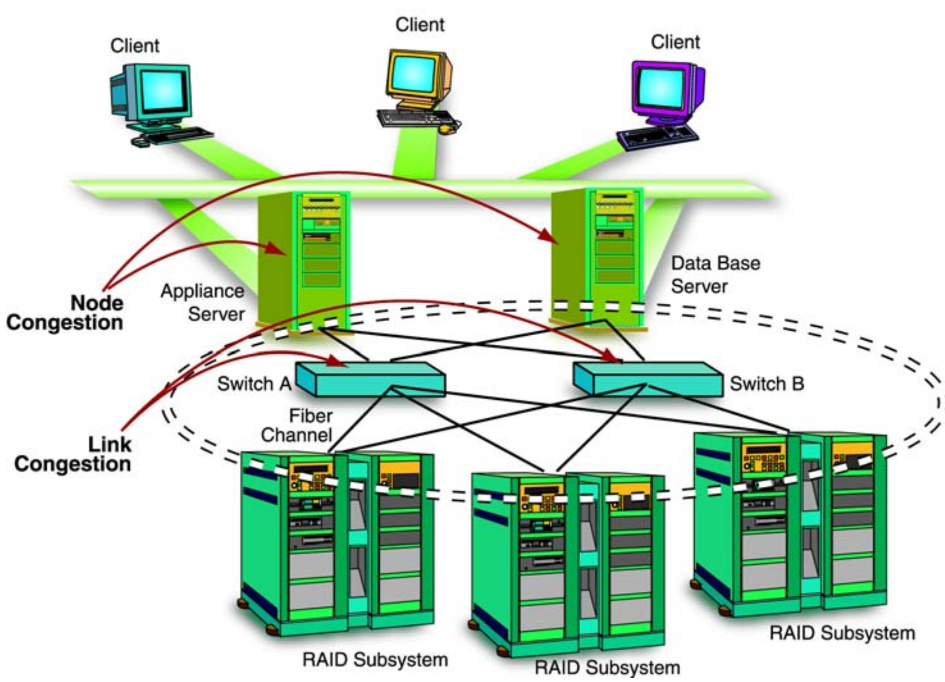


Figure 19 – Link and node congestion in a SAN

Solving FC-SAN problems

The security, congestion and heterogeneous data sharing problems of current FC-SAN deployment may eventually be solved and standards may eventually be developed. Meanwhile, early SAN implementations will likely implement work-arounds. However, these work-arounds will probably be IT manpower intense and often require vendor intervention. For example, EMC's SAN™ has an elaborate method of entering World Wide Names (64-bit permission schemes) at initialization of Symmetrix® in a SAN™. EMC® service engineers must be called to make changes in this configuration. Some issues may not even be inhibitors to early SAN adoption, depending on whether the deployed applications are critical to an enterprise. The biggest problem with SAN™, however, is that enterprises want the promised low-cost, high-availability benefits of FC-SAN now, while FC-SAN probably won't be able to deliver until standards are established.

Given the weak standards history of the storage industry and the strict standards adherence of the networking industry, it is no surprise that NAS is in production for 48% of the large enterprises surveyed by ITCentrix Inc.®, while only 7% of enterprises have implemented SAN™ in production.

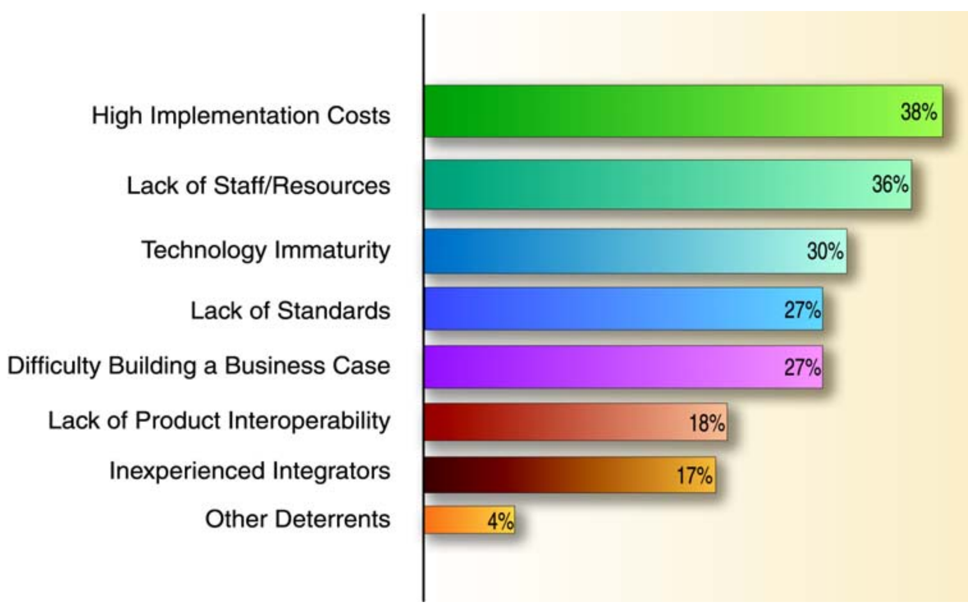
In conclusion, SAN™ can today deliver centralized management benefits like storage resource pooling and LAN-free backup. However, today's SAN™ deployments require significant implementation and installation costs with high IT management overhead. The enterprise that might benefit from SAN today is large, has significant financial and human resources, and can see a future competitive advantage in being an early-adopter of FC-SAN today.

What to expect from a FC-SAN implementation

Four out of ten respondents mentioned the high cost of FC-SAN implementation as a deterrent to FC-SAN implementation.

A good barometer as to what might be expected from a FC-SAN deployment can be gleaned from data collected in mid-2000 and reported in ComputerWorld magazine. **Figure 20** lists a tabulation of responses from 100 IT professionals considering FC-SAN implementation or testing. It is important to note that about 4 out of 10 respondents mentioned the high cost of FC-SAN implementation as a deterrent to FC-SAN implementation. Although great strides are now being made in FC-SAN management software and interoperability of FC switches is expected in the next few years, "Lack of Staff" to implement and manage FC-SANs was mentioned as the second most frequent FC-SAN deterrent in this survey.

Figure 20 – Deterrents to FC-SAN implementation.



iSCSI or E-SAN applications

iSCSI is an emerging industry standard protocol for delivering SCSI commands over IP.

There are no known E-SAN deployments at this time since iSCSI HBAs and NICs are just becoming available. The type of applications that will be suitable for E-SANs will likely be the same as those for FC-SAN unless the SAN file-sharing problem is solved. However, security and congestion control issues are likely to be solved with E-SAN since Ethernet technology is being driven by the networking industry. Ethernet solved security and congestion control problems years ago and can be expected to solve them for Gig-E and 10Gig-E.

New applications for remote E-SANs may well be enabled for disaster recovery planning since Ethernet uses the same transport (IP) and network (TCP) protocols used for the Internet. This would render "encapsulation schemes" such as FC/IP unnecessary. Additionally there is much speculation as to whether TCP/IP stack processing will be a performance problem for iSCSI and E-SAN technology. Early IETF subcommittee drafts of the iSCSI spec assumed that HBAs with hardware to outboard the stack processing from the local server would be required. But there may be implementations at the low end for E-SANs that do not use these specialized NICs.[4] If the major expected volume of Gig-E and 10Gig-E networking equipment and iSCSI NICs occurs, it may well drive the price of iSCSI HBAs down to a commodity level thereby making deployment of E-SANs more cost effective than FC-SANs. In any case, despite the security, performance and lower cost potential advantages of iSCSI and E-SAN technology, it is not here today. An assessment of whether iSCSI or other storage over IP technologies will displace or cap the growth of FC-SANs in the future is discussed further in **Chapter 5** as we look into the future.

[4] Fibre Channel HBAs offload the host of network processing, which is one reason they are significantly more expensive than commodity Ethernet NICs.

Unifying NAS and SAN for Enterprise Storage

4

In spite of vendor hype from EMC, MTI and others claiming to offer integrated NAS and SAN architectures today, it is not possible to deploy an *optimally integrated* NAS and SAN architecture at this time. There are significant technical challenges to integrating block oriented (SAN) and file oriented (NAS) storage access methods. As we discussed in **Chapters 1 and 2,** blocks are *physical* data concepts referring to the organization of data on a disk drive in terms of sectors and tracks. Files are *logical* data representations (understandable only to a file system) and are made up of multiple blocks, sectors or tracks. A disk drive has no knowledge of files. The translation between the logical file data and the physical organization of this data into blocks on a disk drive is the job of the file system. File systems are usually embedded in a client or server's operating system (local file system) but remote file systems over a network (NFS or CIFS) have become standard for offloading *block retrieval* processing using redirector software[5] .

It is not possible to deploy a fully integrated NAS and SAN architecture at this time.

The technical challenges

There are two categories of questions that need to be resolved before NAS and SAN architectures can be integrated into a common architecture.

Two categories of questions need to be resolved before NAS and SAN architectures can be unified.

- How will the rivalry between FC and IP be resolved as the underlying transport mechanism for SANs of the future? Will TCP/IP (Ethernet) become prevalent for E-SANs or will other schemes evolve to prominence?
- How and where will a NAS like remote file system be implemented for SAN? What characteristics will/should this file system have regarding large scale multilingual (Type III) file sharing?

Neither set of questions has an easy answer since there are many vendors in both storage and networking that have vested interests in either preserving or changing the status quo of today. We believe that the future trends to Universal Data Sharing (i.e., true data sharing or Type III data sharing) and continued advances in Ethernet networking technology are safe bets for the future. This is because a bilingual (UNIX® and Windows®) file system is absolutely needed for SAN data to be shared, and the potential of Ethernet technology to minimize cost of reconfiguration for eventual NAS and SAN convergence is significant. Ethernet is highly likely to continue its high performance and cost effective evolution as is scalable Universal Data Sharing (UDS).

Neither set of questions has an easy answer.

Ethernet is likely to continue its high performance and cost effective evolution as is large scale multilingual true data sharing such as offered by Auspex.

Chapter 4 discusses the potential of future storage networking protocols (FC and IP) and the different directions that an integrated NAS and SAN architecture can take. **Chapter 5** discusses the more difficult issue of what characteristics should a unified NAS and file system have relative to UDS.

[5] Appendix A discusses this process in detail.

How will the rivalry between FC and IP be resolved?

The first important topic to examine is the likelihood that E-SANs displace FC-SANs.

How can NAS and SAN integration be resolved without first resolving the question of what type of SAN will be the basis for this integration? Well, it really can't. So the first important topic to examine is the likelihood that E-SANs displace FC-SANs as the primary basis for storage networking in the future. Ethernet networking technology has been playing "catch up ball" compared to Fibre Channel in terms of data transmission rates. Fibre Channel was originally conceived as a separate networking protocol optimized for large block storage data with a larger packet size than Ethernet (IP transport) and therefore lower networking overhead. However, the Fibre Channel committee made a fundamental mistake in planning next generation transmission rates at 2 Gigabits per second instead of an order of magnitude increase as was the planned case for 10 Gigabit Ethernet (10Gig-E). Although some vendors such as EMC are implementing this 2 Gigabit per second standard, efforts are currently underway to increase this transmission rate to 10 gigabits per second. This is illustrated in **Figure 21**.

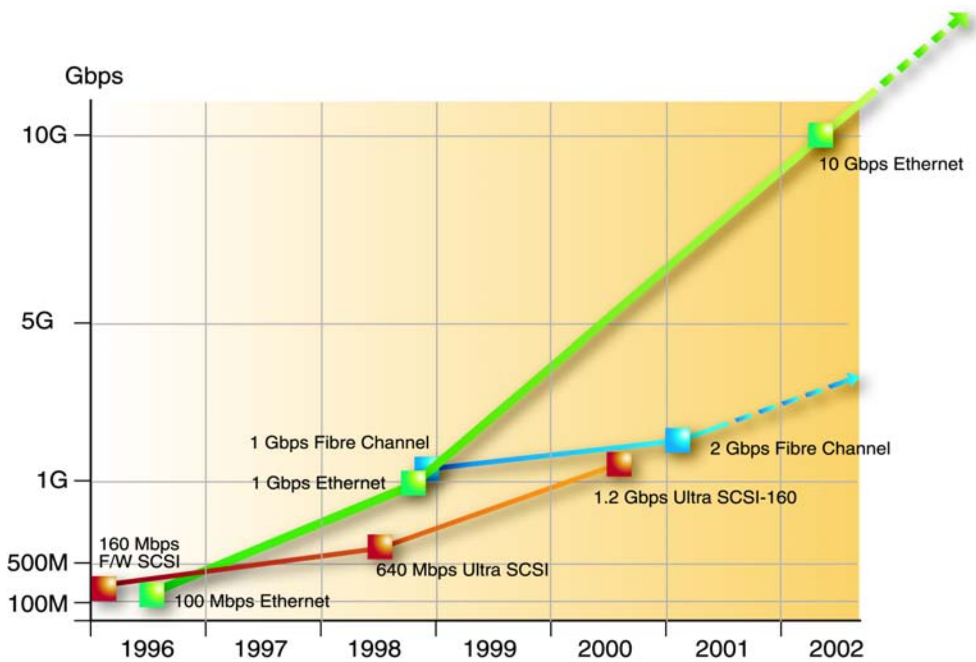


Figure 21 – Transmission rates of Ethernet and Fibre Channel.

The argument for IP Storage (iSCSI[6]) vs. Fibre Channel

As can be seen earlier in **Figure 2** and the **Title Page** of this report, iSCSI (SCSI over IP) and Fibre Channel are similar in that they both send packets containing SCSI commands over lower-level network protocols from the *Source* to the *Destination* computer in the network. In the case of Fibre Channel (FC), these packets are called SCSI-FCPs, which are sent over the lower level FC protocols[7]. In the case of iSCSI these packets are called iSCSI PDUs.

There are various arguments concerning whether the newer specification (iSCSI) will displace the older specification (Fibre Channel). Primary arguments against Fibre Channel are:

[6] There are flavors of IP storage protocols such as SoIP (See glossary of terms) being proposed other than iSCSI to the IETF. However, iSCSI by far has the most support from vendors and will be the primary storage, over IP technology, discussed in this report.

[7] Fibre Channel has defined a specification for each layer of networking. See Appendix C.

1) The specification is not complete for security and congestion control.
2) All FC implementations are proprietary whereas Gigabit Ethernet will be open.
3) Data transmission rates of 10GigE will eclipse FC in the near future.
4) 10GigE potentially will use existing network management skills and not require the expense of new staff.
5) Expensive Fibre Channel network analyzers need not be purchased.
6) Ethernet networking equipment will enjoy lower pricing than FC due to mass volume production expected from the networking industry.

Conversely, the major arguments against iSCSI are:

1) There are no deployed iSCSI SANs today.
2) Major investments in SAN management tools will be required before E-SANs become viable.
3) The specification will take forever in the IETF due to competing storage vendor interests.
4) When standards do become finalized there will be many exceptions (like the SCSI standard), which will result in the lack of a true usable standard.
5) Processing of the TCP stack is CPU intensive and will require offloading stack processing from the client or server to specialized iSCSI HBAs in a manner similar to the way FCP handles stack processing on FC HBAs. If this is the case, iSCSI HBAs will be just as expensive as FC HBAs are today.

Although these arguments and counter arguments ricochet throughout the storage industry, no one doubts that the network industry's Gigabit Ethernet initiatives will continue with evolutionary determination. This is why we believe that expected continued advances in Ethernet networking technology is a safe bet for the future. Whether and when SAN technology is based on this leading IP storage technology (iSCSI) depends on resolving the iSCSI issues discussed above. If the first iSCSI HBA implementations do not outboard stack processing, early IP storage efforts may begin at the low end due to the intensive processing requirements of TCP intensive workloads.

How and where will "NAS-like" remote file systems be implemented for SAN?

FC-SAN information is basically partitioned storage.

Whether SANs based on Fibre Channel (FC-SAN) or Ethernet (E-SAN) UNIX® and Windows® information cannot be shared as in the case of a universal IBM Type III (**Table 2**) data sharing architecture, such as the Auspex NetServer™ family of products. FC-SAN information is basically partitioned storage (IBM® Type I data sharing – **Table 2**). Sharing on FC-SANs is only accomplished:

1) Through methods that involve multiple data images and data migration and conversion software (Type II data sharing – **Table 2**)
2) Through low performance client or server emulation software (**Appendix B**)
3) Through software gateways where a common repository of metadata is checked before access (metadata controllers such as Tivoli's SANergy).

All three of these methods are less than direct and do not meet the ideal objective of Universal Data Sharing (UDS) that allows *any* user attached to *any* network to read and write to *any* file anytime, *any*where from a *single data image*. This high standard of multilingual true data sharing has already been achieved by the Auspex NetServer design and is the single important topic discussed in **Chapter 5**. We believe that the evolution

Figure 22 – FC-SAN information cannot be shared.

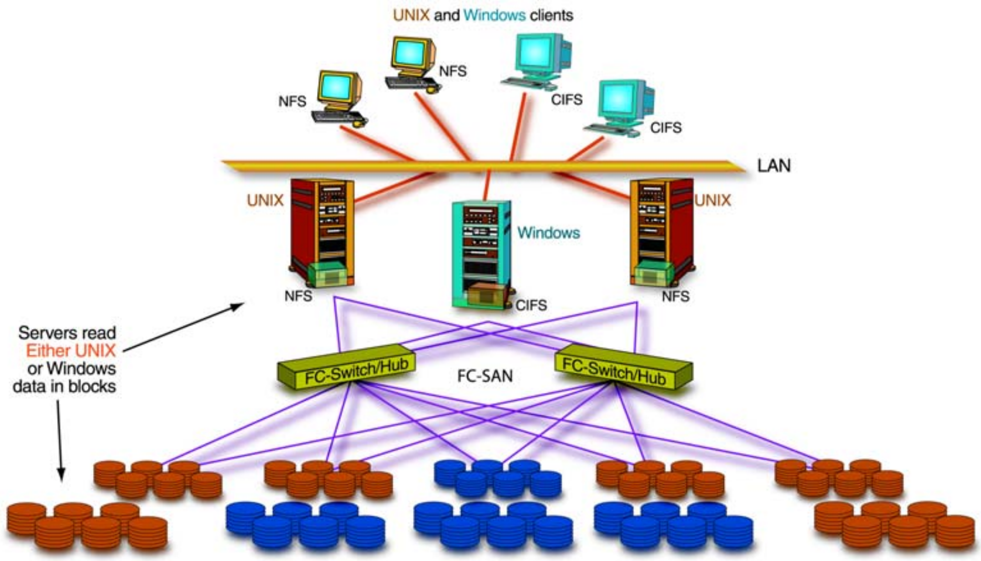


Figure 23 – E-SAN information cannot be shared.

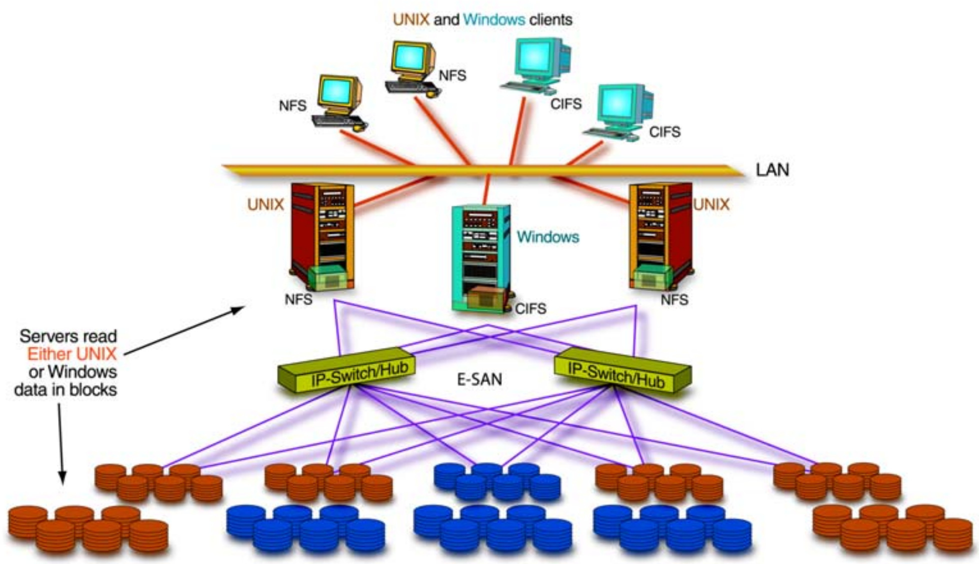
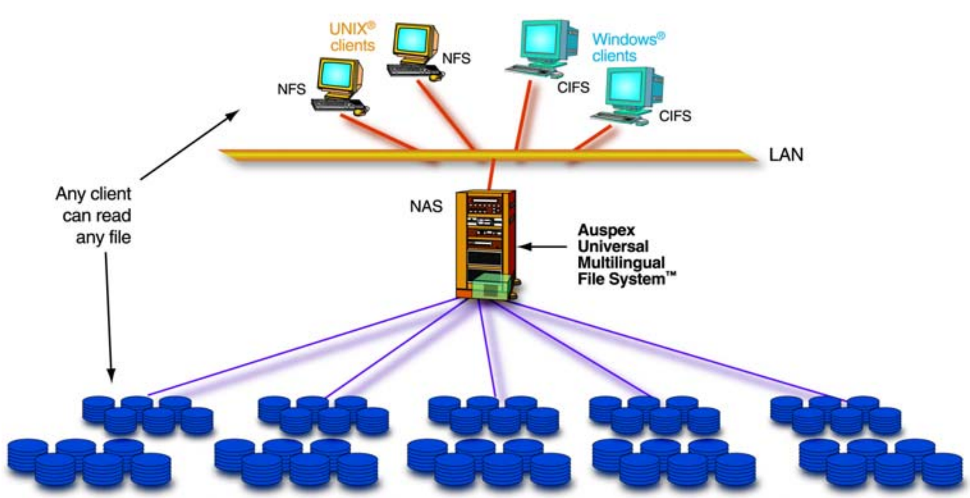


Figure 24 – NAS information can be shared.



toward the Auspex standard of UDS by the rest of the industry will eventually occur and that it is therefore a safe bet to deploy to Auspex today when planning for eventual NAS and SAN integrated architectures.

The evolution of data sharing by the rest of the industry will eventually reach the Auspex standard.

What will a unified NAS and SAN architecture look like?

There is no escaping the fact that a true multi-lingual networked file system is a fundamental aspect of an integrated NAS and SAN architectural design. This is why Auspex has taken the first step toward a unified NAS and SAN architecture with the May 2001 announcement of the NetServer3000™ product family. Specifically Auspex implemented Fibre Channel connections behind its file server nodes (NAS heads) so that much larger amounts of storage (68TB) can be managed by its efficient, scalable, expandable I/O Node design with fault tolerant server failover option. This model is known as the *Vertically Unified NAS and SAN Model* and is shown in **Figure 26**. The Auspex highly scaleable parallel architecture makes it possible to expand storage capacity by adding more I/O Nodes (than the current three) in the future along with faster file server interconnects (like the Auspex SCI bus) to accommodate capacities up to a petabyte (petabyte = 1024 terabytes) and beyond! This can only be accomplished with a parallel expandable architecture as is unique for Auspex in the NAS market at this time.

A true multi-lingual networked file system is fundamental to integrating NAS and SAN.

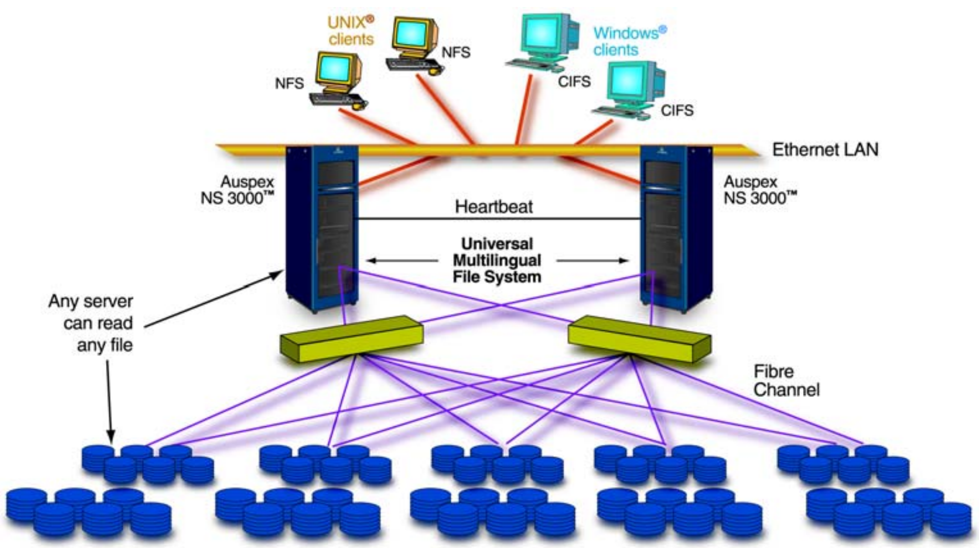


Figure 25 – The Auspex vertically unified model.

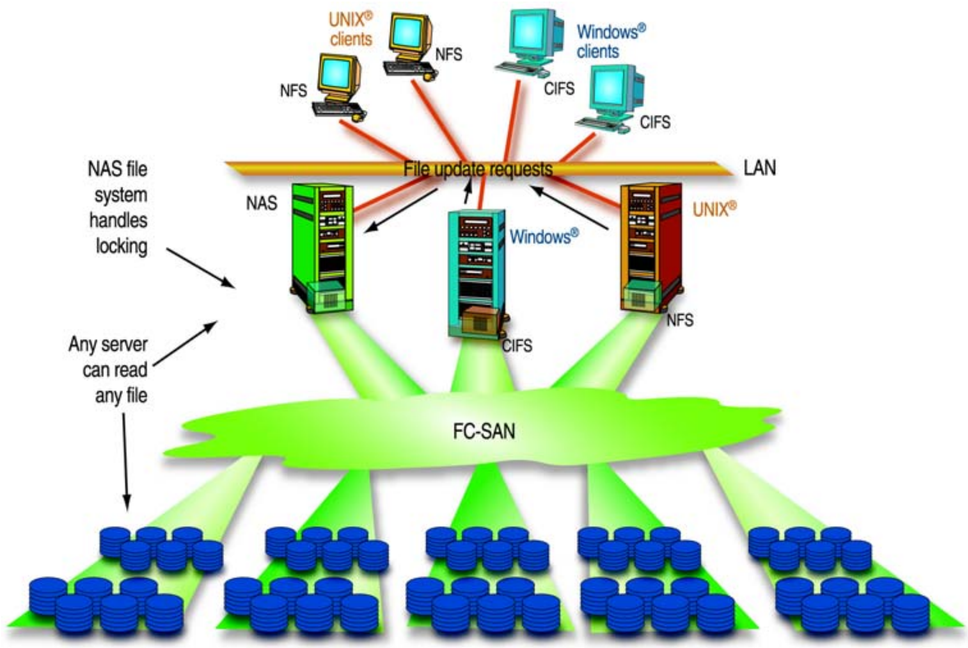


Figure 26 – Horizontally unified model NAS and FC-SAN.

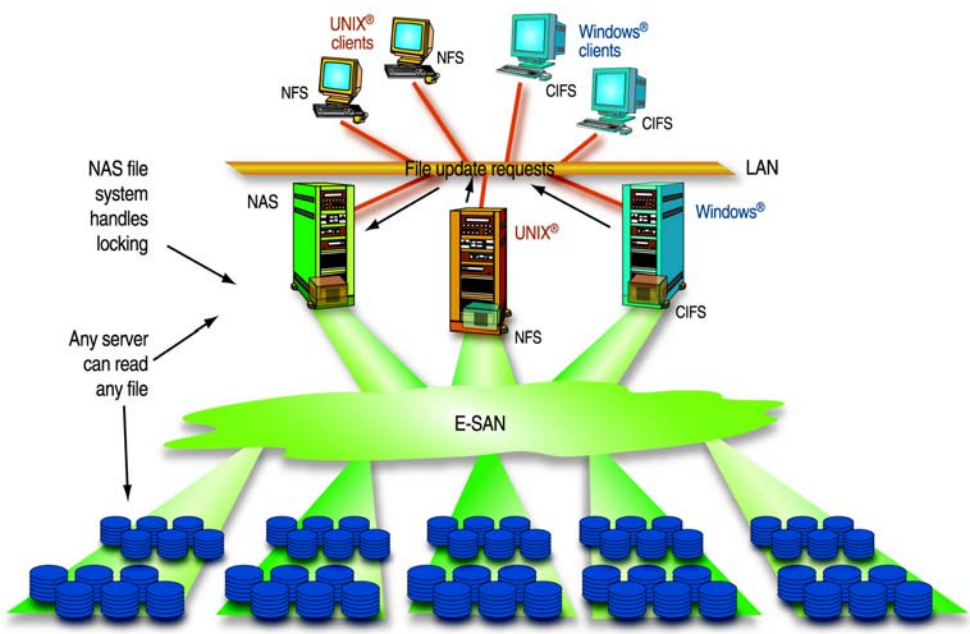


Figure 27 – Horizontally unified model NAS and E-SAN.

Other unified NAS and SAN architectures are known as the *Horizontally Unified NAS and SAN Model*. These architectures involve the use of software gateways where file metadata is checked either on the NAS device or on a separate metadata controller (SANergy™) and block data returned to UNIX® and Windows® servers through direct attachment or over the network. EMC's HighRoad™ file system software product is a move in this direction. But EMC's Celerra™ product does not implement universal file sharing so Type III data sharing is limited to the amount of storage that can be controlled by a single Data Mover (EMC's word for the file processing node or NAS head). The ideal requirements for a true multilingual file system with large-scale UDS are discussed in more detail in **Chapter 5**. Horizontal unification may occur for either FC-SAN as shown in **Figure 26** or for E-SANs (iSCSI) as shown in **Figure 27**.

Selecting the Right NAS Product to Prepare for NAS and SAN Integration

5

A review of what we have discussed thus far can be summarized as follows:

We have presented a review of the three major architectures available today for enterprise storage and discussed appropriate applications for each. A brief review of the key topics presented thus far is appropriate:

1) Analysts predict that NAS and SAN will converge in the future.
2) You cannot deploy a fully integrated NAS and SAN architecture today.
3) There is uncertainty as to whether Fibre Channel (FC) or Internet Protocol (IP) will serve as the transport for future NAS and SAN integration although either is possible
4) SANs cannot easily provide Universal Data Sharing (Type III – See **Table 2**) for UNIX® and Windows® data without a common file system.
5) SANs primarily provide infrastructure benefits to the IT department.
6) NAS is the only solution that provides business benefits to non-IT departments and these benefits occur primarily through fully lock protected UNIX® and Windows® data sharing from a single data image with scalable read and write capability.
7) Many NAS data sharing applications involve very large scale capacity requirements (e.g., Streaming Media, PDM, CASE/SCM).
8) Development of a common file system for SANs is likely to take many years and be impeded by both standards committees of rival vendors and difficult technology hurdles.
9) Captive NAS devices are trending toward the separation of the storage array from the NAS server (NAS head).
10) Client and / or server emulation software for UNIX® and Windows® data sharing are less for data sharing than optimally designed "on board" NAS file systems.

Selecting the most scaleable UNIX® and Windows® file sharing solution is the best strategy

Selecting the most scaleable UNIX® and Windows® file sharing solution is the best strategy.

The above argument strongly suggests that an enterprise would be foolish to wait to deploy the business benefits of *NAS unique* UNIX® and Windows® UDS. An enterprise should also remain cautious concerning the uncertainty of future SAN transport protocols (FC vs. IP). For these reasons, Auspex believes that selecting the most scaleable UNIX® and Windows® file sharing solution is the best strategy for an enterprise to pursue when planning for eventual NAS and SAN integration.

Since no one believes that future NAS and SAN integration will be complete without scaleable Type III UNIX® and Windows® data sharing (NAS has it and SANs do not), it makes sense that this feature will come from NAS and not SAN.

Large scale Type III data sharing is THE critical characteristic for future NAS and SAN integration.

Only the Auspex NetServer product family has been specifically designed for NAS from the beginning.

Only Auspex provides "scalable" Universal Data Sharing (UDS).

For these reasons this chapter will examine ONE single question:

- Which of the major NAS product offerings provide fully lock protected[8] Type III UNIX® and Windows® Universal Data Sharing (UDS) from a single data image[9] for high capacity scalability?

If any user is to have full read/write access to any information, anytime, anywhere from a single data image, then large scale Type III UDS is THE critical characteristic for future NAS and SAN integration and one that can be selected today.

A complete discussion of the architectures of the four major approaches and considerably different approaches to NAS can be found in the first edition of the ***Auspex Storage Architecture Guide,*** which is available from your Auspex representative. The four alternatives reviewed in this report are:

1. General purpose file servers from Sun®, HP®, Compaq® and others
2. File Servers from Network Appliance®
3. The Celerra® File Server and Symmetrix® Storage System from EMC®
4. The Auspex NetServer™ product family

It will be seen from this report that only the Auspex NetServer™ product family has been specifically designed for NAS from the beginning and that competitive approaches were not.

The many meanings of 'Scalable"

The word 'scalable" is the most misused and misunderstood word in the storage industry today.

Definition #1 – "Scalable" can refer to the ability to easily add more *capacity.*

Definition #2 – "Scalable" can refer to the ability to linearly increase *performance.*

Definition #3 – "Scalable" can refer to the ability *to both add capacity AND increase performance linearly.*

When an NAS vendor, other than Auspex, refers to their products as "scalable," they are referring to one of these three meanings of the word.

Auspex NetServers were designed with a more strict definition of "Scalable"

The Auspex NetServer product family, however, was the only NAS product line designed from the outset with a more strict definition of "scalable" that includes large scale UDS or Type III "true data sharing."

Definition #4 – The Auspex definition of "scalable" refers to the ability to add capacity while increasing performance linearly and providing *large scale multi-lingual Universal Data Sharing (UDS) to any user on any network with securely locked read AND write privileges from a single data image.*

[8] Full lock protection means the honoring of both UNIX® and Windows® permission schemes with load balancing so that any user has equal access to files from a performance point of view. i.e., no users are "second class citizens."

[9] The importance of data sharing from a single data image is. In the area of higher storage system utilization and in the cost advantage of staff being able to manage larger amounts of data per person.

If *any* user is to have full read/write access from *any* network to *any* information, *any*time, *any*where from a *single data image*, then any file must be available to *any* user (for reading and updating writing), even as system capacity scales up to 68TB of data and beyond. Unless the storage system can provide this capability, the entire concept of easy and cost effective management of large amounts of data is not possible. With the Auspex NS3000™, performance actually increases as system capacity is scaled and additional I/O nodes are added.

Competitive NAS solutions are not scalable to large capacities in terms of read/write Universal Data Sharing (UDS)

Only an architecture where processing power is scalable, like the Auspex NetServer™ product family, can scalability be provided to this strict definition. The reason is simple, namely that a single file server can only do so much work, otherwise performance degrades. Since the Auspex NetServer™ parallel architecture can be scaled up in terms of the number of file servers[4] (in addition to the number of disk drives), this objective is uniquely accomplished. To put this capability in perspective, **Figure 28** shows the Network Appliance, EMC and Auspex products drawn proportionately to show the maximum scalable size of each design.

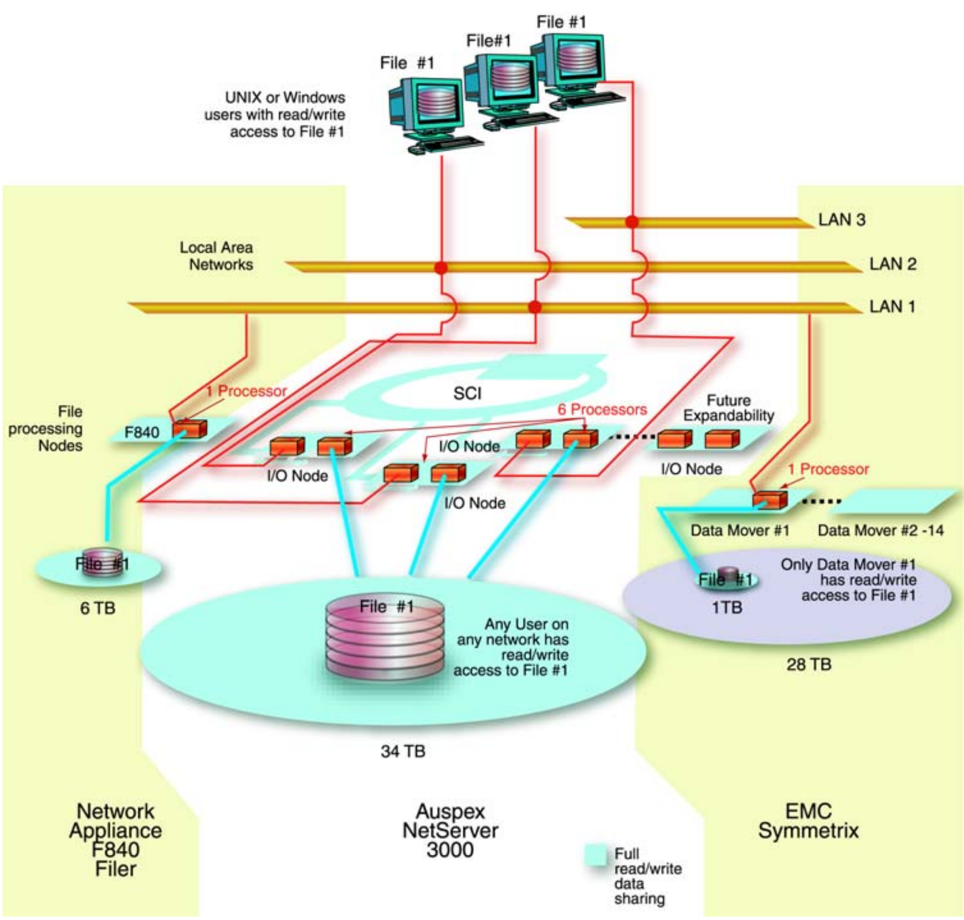


Figure 28 – Only Auspex NetServers™ allow any user attached to any network to share any file anytime, anywhere from a single data image.

[10] The Auspex NS3000™ I/O Node IIs (individual file servers) exchange information on a private high speed SCI bus using a custom message passing scheme to allow large scale multi-lingual Universal Data Sharing between Nodes.

Network Appliance F840 Series and Auspex NetServer 3000™ design comparison

The F840 single processor design scales to 6TB for Type III data sharing.

The Network Appliance F840 file server architecture uses a custom hardware and software design that has remained almost unchanged since its introduction in 1994. The product line was originally designed for a low-end, low-cost market that used non-mission critical data. While optimized for UNIX® and Windows NT® Type III file sharing, Network Appliance F840 file server has a major design problem, however, when it comes to adding both capacity while increasing performance linearly and providing Type III data sharing to any user on any network. The design is based on a single processor that does everything, and one processor can only handle so much storage. In the case of Network Appliance that number is 6TB making the product suitable for small NAS applications within a department or workgroup. The single processor serially manages all system activities over a single system bus. This single processor workload includes file system processing, network processing, RAID management and system management processing. By comparison, an Auspex NetServer 3000™ uses multiple processors in line with the company's design philosophy of *functional specificity* where software and hardware work in parallel for greater efficiency. Whereas the single F840 processor handles many tasks, the Auspex design not only distributes these tasks to separate network processors, separate file system processors, separate service processors, but also offloads RAID management to powerful dedicated hardware RAID processors in each of three file server nodes as shown in **Figure 28**. In addition, the NS3000™ parallel design allows scalable UDS to 38TB whereas the F840 is dead-ended at 6TB. The Auspex separate service processor is not shown in **Figure 28**.

EMC Celerra™/Symmetrix™ and Auspex NetServer 3000™ design comparison

To change file assignment requires unwanted downtime and complexity.

Celerra's file sharing scalability is effectively limited to the amount of storage that can be handled by a single processor.

Similar to the Network Appliance F840 design, EMC's Celerra® has a single processor for each of 2-14 file processing nodes (Data Movers) that "front end" a Symmetrix® storage system. Unlike Auspex, EMC *did not* design Celerra® for scalable UDS past 1TB. This is because Celerra® can not be scaled for file server processing power for Type III data sharing as the NS3000 can. At system initialization a Celerra® Data Mover (DM) is permanently assigned ownership of a given so that the assigned data mover has both read and write access to that file. To change file assignment requires unwanted downtime and complexity. Since there is no cooperation between Data Movers (DMs), due to lack of a scalable parallel architectural design, Celerra® has the limitation that only users on networks attached to DM #1 can update (write to), File #1 as shown in **Figure 28**. Other files assigned to Data Movers 2-14 have only "read access" to File #1. This design means that Celerra's Type III UNIX® and Windows® file sharing scalability is limited to the amount of storage that can be handled by a single processor. Although Celerra® is specified to have a total (non-shareable) capacity of 28 TB of data, scalability for Type III UNIX® and Windows® file sharing is limited to a much lower number of terabytes. This number is realistically the same as the 6TB limit specified by Network Appliance in their F840 design as shown in **Figure 28.**

When it comes to UDS, Celerra® should be looked at more as partitioned storage. It is not an optimal design for allowing any user to have full read/write access from any network to any information, anytime, anywhere from a single data image since UDS dead-ends at 1TB, as opposed to 38TB for Auspex.

If Celerra® were to offer parallel scalable processing architecture among all Data Movers, then full and scalable Type III UNIX® and Windows® file sharing could be achieved. This, however, is not the case and will have to be addressed by EMC® engineers if Celerra® is to become optimized for Type III data sharing. Today, Celerra® can be thought as a collection of single processor file server nodes sharing the same storage cabinet. The Auspex NetServer 3000™ on the other hand, allows full concurrent read and write access to all files on the system by all users. Another way to picture Celerra®, is that it lacks a communication bus and file system coherency software between file server nodes and a customized message-passing scheme like the Auspex NetServer 3000™.

Other NAS products and the Auspex Net Server 3000™ design comparison

There are many other products in the NAS market but all are based on single processor designs in each file system node. These products all have the same Type III UNIX® and Windows® file sharing scalability limitations as the F840 and Celerra® since none are based on a large-scale parallel and functionally specific architecture such as the Auspex product family. For a thorough discussion of data sharing and other important features of the Auspex product family, a separate report titled the ***Auspex NS3000 Product Guide,*** is available at www.auspex.com or from your Auspex representative.

Summary and Conclusions

6

This report has discussed the three major choices in storage architecture available today and appropriate applications for each: Direct Attached Storage (DAS), Network Attached Storage (NAS) and Storage Area Networks (FC-SANs and E-SANs). We have seen that NAS architectures deliver business benefits to both IT and non-IT departments and that the benefits of SANs are primarily for the IT department.

The status of future NAS and SAN integration

We have also reviewed the issues regarding NAS and SAN integration. It is apparent that today's Fibre Channel SANs (FC-SANs) have no apparent way to implement high performance and scalable Type III UNIX® and Windows® file sharing in the foreseeable future without agreement on a universal file system such as has been implemented by NAS vendors. In fact FC-SANs may be eclipsed by SAN architectures based on the IP networking transport protocol. Most likely this transport will be Ethernet (TCP/IP) although we view this issue with great uncertainty. We have called this potential architecture Ethernet SAN or E-SAN to easily differentiate it from SANs based on Fibre Channel (FC-SAN). In fact NAS currently uses IP-based transport, and this makes E-SANs and NAS a more logical integration point than FC-SAN and NAS.

Only hybrid variations of NAS and SAN integration can be deployed today

In spite of vendor hype from EMC®, MTI® and others claiming to offer integrated NAS and SAN architectures today, it is not possible to deploy an *optimally integrated* NAS and SAN architecture today. What exists at the moment is proprietary hybrids with shortcomings compared to an ideal solution. Most analysts project that NAS and SAN will eventually become integrated in the future, and we have examined the ideal characteristics of such an approach.

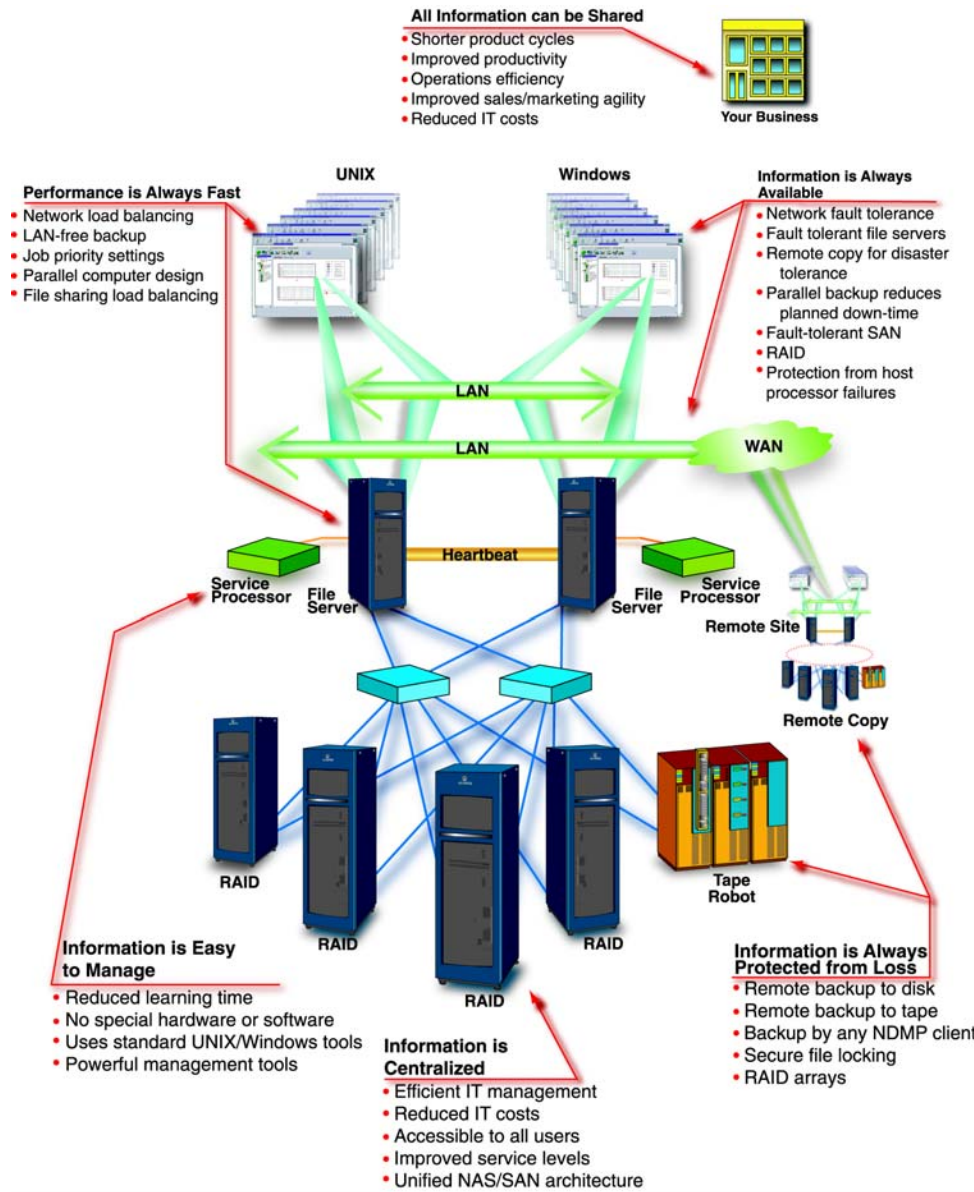
Auspex has taken the first steps toward true NAS and SAN integration

The Auspex NetServer 3000™ product family has taken the first steps toward such integration by implementing Fibre Channel connections between the RAID array and the NAS head. The Auspex NAS head parallel and expandable architecture makes Auspex the only candidate among NAS vendors to manage very large amounts of storage with its existing file system design and cooperative scheme for file sharing and file coherency maintenance between file server nodes.

Only NAS provides non-IT department business benefits

In light of the demonstrated business advantages of many applications for NAS-based UNIX® and Windows® Type III data sharing, it is proposed that an enterprise can safely deploy a NAS architecture with optimal Type III data sharing characteristics. This will allow the enterprise to not wait to benefit from these considerable business benefits.

The business and infrastructure benefits of the Auspex NS3000™ Series.



Conclusions regarding an effective short-term storage architecture strategy

We propose five effective short-term storage architecture strategies as follows:

1) **Enterprises should deploy "best of breed" NAS solutions today** if there are business benefits to justify the investment. "Best of Breed" is defined to mean the ability of a NAS Head to allow any user attached to any network to share any file anytime, anywhere from a single data image.

2) **Enterprises should evaluate new NAS procurements carefully against a tight definition of scalability.** Make your vendors discuss this topic and their plans for the future. Scalability is a much maligned and misused vendor term. The tight and appropriate definition of "scalable" refers to the ability to add both capacity while increasing performance linearly and providing Universal Data Sharing to any user on any network on a *large scale basis*.

3) **Enterprises with an existing proprietary implementation of SAN should be cautious** in expanding this investment given the FC vs. IP uncertainty and the low probability that FC-SANs will solve the Type III data sharing, security, congestion control and distance limitation problems inherent in today's FC-SAN solutions.

4) **Enterprises that have not yet implemented a SAN architecture should wait if at all possible** until the FC-SAN vs. E-SAN picture clarifies further. Testing of E-SANs would be appropriate for these enterprises since E-SANs can be expected to hold security, congestion control and long distance transport efficiency advantages over FC-SANs.

5) **Future advances in Scalable Universal Data Sharing (UDS) and Ethernet are both safe bets** for technology in the future to minimize cost of reconfiguration. Ethernet is highly likely to continue its high performance and cost effective evolution, as is scalable UDS.

NAS Offloads Work from the Local Client or Server to the Remote Server[11]

A

File system I/O is responsible for as much as 25% of a local client or server processor workload depending on the application suite that is running. The table below illustrates how the file system block retrieval process is offloaded from the local client or server by a network attached storage server. Although the process is discussed for the Microsoft Windows NT® operation system, a similar process occurs for UNIX® except the I/O is redirected using the Networked File System (NFS) instead of Server Message Block (SMB) protocol. The I/O process begins when an application passes file requests to the ***Environmental subsystem*** which is a user-mode protected process that Windows NT® runs and supports applications native to different operating systems environments. Examples of these subsystems are the Win32® subsystem and the OS/2 subsystem.

The Environmental subsystem then issues file requests to the Windows NT® ***Executive.*** The Executive is a collection of kernel-mode modules that provide basic operating system services to the environment subsystems. The Executive implements a message-passing facility called a *Local Procedure Call* (LPC) facility. It works very much like the Remote Procedure Call (RPC) facility used for network processing.

The ***Object Manager*** is the part of the Windows NT® Executive that provides uniform rules for retention, naming and security of objects such as files[12] . Before a process can manipulate a Windows NT® object (e.g., file), it must first acquire a handle to the object (e.g., file handle). An object handle includes access control information and a pointer to the object itself. All object handles are created through the Object Manager. In this way, the same routines that are used to create a file handle can be used to create a handle for another type of object.

Other modules in the Executive manage virtual memory, i.e., ***Virtual Memory Manager,*** and processes, i.e., the ***Process Manager.*** These and the other executive services of the Executive are involved in the process of determining whether an I/O needs to be executed *to* or whether it can be read *from* cache memory without the need to perform a disk I/O.

[11] The source of the material in Appendix A is the Microsoft Windows NT® Developers Guide.

[12] The Object Manager manages the global directory for Windows NT® and tracks the creation and use of objects by any process. This namespace is used to access all named objects that are contained in the local computer environment. Some of the objects that can have names include the following:

- Directory objects
- Object type objects
- Symbolic link objects
- Semaphore and event objects
- Process and thread objects
- Section and segment objects
- Port objects
- File objects

The Environmental subsystem then sends either synchronous or asynchronous I/O requests to the ***I/O Manager*** that gets control of the I/O process and constructs one or more structures called ***I/O request packets***. It then passes these packets to the appropriate ***file system driver*** for local or remote file system execution.

TheWindows NT ***Executive*** contains a top thin layer known as Executive services that serves as the interface between the user-mode environment subsystems and kernel mode. The ***I/O Manager*** is the part of the Executive that manages all input and output that needs to be conducted if the file request cannot be satisfied from memory. The I/O Manager supports all ***file system drivers, hardware device drivers and network drivers*** and provides a heterogeneous environment for them. This uniform interface allows the I/O Manager to communicate with all drivers in the same way, without any knowledge of how the devices they control actually work. The I/O Manager also includes ***driver support routines*** specifically designed for file system drivers, for hardware device drivers and for network drivers. It allows multiple file systems and devices to be active at the same time while being addressed through a formal interface.

In the case of requests for data from a remote server, a client *Redirector* is invoked. A client redirector refers to the software required to access over a network. Whenever a client workstation requests file or print services from a server, the client redirector software issues special commands understood by the server software. Windows NT® uses a client redirector language called SMB (Server Message Blocks) to communicate between clients and servers.

For an application to retrieve a file from a local computer therefore, Windows NT® calls one of the local file system drivers, fastfat.sys or ntfs.sys to get those files off the local hard disk. But if an application retrieves a file from another computer over the network, the operating system (specifically, the I/O Manager) calls the redirector (using SMB protocol) to get that file. It’s called the redirector because the request for the file (that is built on the local client by the remote File System driver) is “redirected” away from the local hard disk and toward a network drive. The following table shows how block retrieval (Steps 17 and 18) is offloaded from the local server so that this activity.

I/O Processing Activity		Client	Network Server
1.	File request to environmental subsystem	Yes	No
2.	File request passed to NT executive	Yes	No
3.	Object Manager Processing using LPC	Yes	No
4.	Obtaining File Handle	Yes	No
5.	Checking security of access to object	Yes	No
6.	Virtual Memory Processing	Yes	No
7.	Process Manager Processing	Yes	No
8.	Is I/O in cache or on disk	Yes	No
9.	Is I/O local or remote (local or remote FS)	Yes	No
10.	Construction of I/O request packets	Yes	No
11.	Packets passed to local or remote FS driver	Yes	No
12.	Invocation of redirector on client (for remote I/O)	Yes	No
13.	Packetizing of request and client NIC processing	Yes	No
14.	Client TCP/IP processing	Yes	No
15.	Low level network protocol (frame) processing	Yes	Yes

16.	Network Server TCP/IP processing	No	Yes
17.	**Remote File processing (block retrieval)**	**No**	**Yes**
18.	**Update of remote directory (directory coherency)**	**No**	**Yes**
19.	Network Server TCP/IP processing	No	Yes
20.	Update of local directory (directory coherency)	Yes	No
21.	TCP/IP processing of returning packets	Yes	No
22.	Passing packets to remote FS driver on client	Yes	No
23.	Passing packets to environmental subsystem	Yes	No

The Inefficiencies of Emulation Software in Data Sharing

B

Separate servers make data sharing difficult

Sharing data is difficult with general purpose file systems.

Maintaining separate Windows® and UNIX® servers, with separate sets of clients, is the path of least resistance. However, this approach makes it difficult for UNIX® and Windows® systems to share files and system management facilities such as backup. Users moving from UNIX® to NT® are likely to be dissatisfied with performance, availability and reliability.

The biggest disadvantage of the approach is that sharing files between different types of clients may require a difficult-to-manage backup scheme resulting in additional administration cost for administrators to learn. In addition, UNIX® system management facilities like backup are available only for UNIX® users and Windows® backup facilities are available only for Windows® users. If Windows® runs on typical PC servers, users moving from UNIX® to Windows® may become less productive and less satisfied due to the lower availability and performance of PC servers. This approach avoids the complexities involved in evaluation and implementing the other three approaches for administrative personnel. However, the cost of fragmenting the work environment and lost user productivity is often far greater than the initial investment in a superior architecture, such as the NS3000™.

Client-based emulation uses processing power inefficiently

Client-based emulation can be an administrator's nightmare.

Client-based emulation implements an NFS protocol stack on a PC client or a CIFS protocol stack on a UNIX® client. For example, SunSoft® PC-NFS implements an NFS protocol stack on a PC client. This approach has the advantage of requiring no changes at the server by system administration personnel and any problems with the product affect only clients using the emulation software. Technically savvy users can often manage the solution themselves. For this situation or when users only access files on a foreign system on an occasional basis, client-based emulation can be an appropriate solution to mixed UNIX® and Windows® environments.

However, these products tend to be relatively slow because of the extra work that the client processor must do to emulate the foreign protocol. They can also create stability problems on the client. To allow two-way file sharing, two different products are needed: one to allow Windows® clients to access UNIX® systems and another to allow UNIX® clients to access Windows® systems. There is also a major cost associated with client-based emulation that is the administration overhead of keeping all the clients current with new releases or getting the software installed on all new clients. Furthermore, since it is quite difficult to mask fundamental differences in file systems, client-emulation is usually imperfect. Perhaps the biggest drawback to client-based emulation occurs when the user cannot solve all installation, configuration and administration problems placing a burden on system administrators. Finally, if there are large numbers of clients, total costs to the organization can be high.

Restrictions of server-based emulation

Server-based emulation suboptimizes performance.

Server-based emulation implements foreign protocol conversion software on a server, for instance, CIFS on a UNIX® server or NFS on a Windows® server. An example, this is the

Samba® suite of freeware components that implements a CIFS protocol stack on a UNIX® server. TotalNET Advanced Server (TAS) from Syntax is bundled with Sun's Netra 150® server product and is an example of a commercial server-based emulation product.

Server-based emulation is generally better than client-based emulation in terms of availability, performance and manageability since no special software is required on the client machine. Since servers are usually more powerful than clients, performance tends to be better. Availability tends to be better than client-based emulation since servers tend to be more tightly controlled, monitored and configured. In addition, management difficulties that do occur are confined to servers and not spread over an entire client population. Finally, server-based emulation is likely to be better because the product's central location is more strategic.

Server-based emulation products, however, execute as user-level processes as opposed to running in the UNIX® kernel. This is not the most efficient way to tune a protocol on a server, and performance is less than kernel-based software since it takes far more instructions to accomplish the same amount of work. Like client-based emulation, server-based emulation is usually imperfect due to the difficulties of emulating facilities like file locking and security on a system that has different features and is a fundamentally different machine. It is particularly difficult to support Windows® users on UNIX®, since Windows® is more flexible and offers more options. In essence clients that use emulation are still "second-class citizens" when compared to native clients. Although server-based emulation is a step up from client-based emulation and is appropriate for more users and provides more intensive file access of the foreign system, its performance limitations make it less than ideal for large numbers of users or even moderate numbers of users with high-intensity application.

Advantages of bilingual Network Attached Storage

Bilingual file servers are the best choice.

Bilingual file servers such as the NS3000™ are typically the most appropriate solution for high-intensity mixed UNIX® and Windows® environments. They become more attractive as the amount of data increases and as the number of clients requiring access to both UNIX® and Windows® files increases. Technically, the bilingual file server is greatly superior to both of the emulation-based approaches. Performance and reliability are likely to be much better, since protocol stacks are part of the kernel, not user-level add-ons. Furthermore there are no "second-class citizens" since a bilingual server treats CIFS and NFS as peers even under heavy concurrent loads on both protocols. Bilingual servers provide the best structure for integrated management of file locking and can make sure that NFS users cannot violate CIFS locks. The bilingual file server also provides "best of breed" administration and management. For example, UNIX® backup tools can be used for all files while Windows® administrators can manage the system using standard administrative tools. Because of the scalability, manageability and reliability of this approach, the total cost of ownership over time is less than other approaches especially for high-intensity applications.

Although acquisition costs may be higher for a bilingual file server such as the NS3000™, total cost of ownership (TCO) is lower over the life of the product due to:

- centralized backup
- higher productivity
- data integrity
- lower administration

A Taxonomy of Network Protocols

C

There are many network protocols in use today. Understanding where a particular protocol resides relative to the OSI (Open Systems Interconnect) model is shown in **Table 4**. Efficient communications between a *Source* and *Destination* computer in a network depend on the protocol for each level of the networking process so that the services provided are appropriate for the type of data being transferred. An example of this would be the difference between UDP and TCP. UDP is used in streaming media (audio and video) broadcasting and drops frames in the event of a problem.

Missing some pixels in a video broadcast is less important than having an interruption in the video broadcast. TCP, however, continues to re-send the frames until they are successfully delivered. Auspex file servers support both TCP and UDP natively in hardware as part of the company's advantage in Streaming Media Architecture and data flow efficiency as discussed in **Chapter 3** of this report.

Level or Layer	Name	Examples of network applications and protocols
7	Application Layer	In the Internet world common applications include Telnet, FTP, HTTP, NTTP, SHTTP, IRC, NFS, RPC, CIFS, SMTP and DNS
6	Presentation Layer	EBCDIC, ASCII, XDR
5	Session Layer	SEP, sockets
4	Transport Layer	TCP, FCP, SPX, STP, UDP, iSCSI, IPFC
3	Network Layer	IP, IPX, CLNP, FC
2	Data Link Layer	PPP, SLIP, FC, LAP protocols i.e., MLP, LAPD, LAPM, LAPF, FC
1	Physical Layer	10Base T, 100Base T, 1000Base T, 10,000Base T, ATM, ESCON, FDDI, FICON, HIPPI, RS 232, RS 449, FC

Table 4 – Today's network protocols as they relate to the OSI model.

Glossary

100BaseT
See fast Ethernet

ATM
Asynchronous Transfer Mode. A suite of network protocols providing low-level services spanning local- and wide-area networks. ATM is intended to provide the switching and multiplexing services necessary to carry voice, data and video and multimedia traffic using fixed 53-byte cells. Standards are being defined to allow ATM to emulate traditional LANs (LANE).

b
Abbreviation for bit (e.g., 10 Mb/s Ethernet).

B
Abbreviation for byte (e.g., 120-GB total capacity).

CAE
Computer Aided Engineering.

CAM
Computer Aided Manufacturing.

CE
Concurrent Engineering refers to the parallel process of product design, test fixture development and manufacturing planning to accelerate product life cycles.

CIFS
Common Internet File System. A statefull, connection-oriented, network file-sharing protocol developed by IBM and Microsoft as part of LAN Manager. CIFS is the native file sharing protocol for systems running Windows for Workgroups, Windows95 and Windows NT. Sometimes referred to as SMB.

Daemon
The word daemon comes from Greek mythology and is today used to refer to a software routine that runs in the background of an application program. Daemons were guardian spirits or ghosts and in UNIX software terms, daemons respond to requests from other processes across a network. On Windows NT, daemons are called services.

DataXpress
Communication among the NS2000's multiple hardware processors and software processes are handled by DataXpress, a low-overhead message-passing kernel executing on each processor.

Denial of Service (or DoS) attack
DoS refers to a general category of network security attacks that are designed to bring the network to its knees by flooding it with useless traffic. DoS attacks can take an application down but do not compromise data integrity.

ECAD
Electrical Computer Aided Design

Ethernet
A Network protocol developed by Xerox in cooperation with Digital Equipment and Intel in 1976. Ethernet supports data transmission rates varying from 10 kilobits/sec (10BaseT) to 10 megabits/sec (10Gig E). The Ethernet specification formed the basis of the IEEE 802.3 specification.

Fast Ethernet or 100BaseT
Defined by the IEEE 802.3 committee, provides a 100 Mb/s standard that is compatible with existing 10BaseT installations, preserving the CSMA/CD media access control (MAC) protocol.

FC
An acronym for Fibre Channel.

FC/IP
An acronym for Fibre Channel over IP, a "tunneling" network protocol for sending FC packets over IP networks.

FCP
A Fibre Channel Protocol that encapsulates SCSI commands.

FDDI
Fiber Distributed Data Interface. A standard for local area networks that typically uses fiber-optic media capable of data rates up to 100 megabits/second over distances up to 100 km. An FDDI network is a token-based logical ring, and is often constructed as a pair of counter-rotating redundant rings (called dual-attachment mode) for reliability. Ethernet, in contrast, is a bus-based, non-token, 10-megabits/second network standard.

Fibre Channel is an ANSI standard designed to provide high-speed data transfers between workstations, servers, desktop computers and peripherals. Fibre channel makes use of a circuit/packet switched topology capable of providing multiple simultaneous point-to-point connections between devices. The technology has gained interest as a channel for the attachment of storage devices, but has limited popularity as high-speed networks interconnect. Fibre channel can be deployed in point-to-point, arbitrated loop (**FC-AL**), or switched topologies. Fibre channel **nodes** log in with each other and the switch to exchange operating information on attributes and characteristics. This information includes **port names** and **port IDs** and is used to establish interoperability parameters.

Fibre Channel Protocol
An ANSI standard covering Fibre Channel protocol for SCSI.

FMP or Functional Multiprocessing

The term Auspex uses for its unique distributed parallel processing NS3000™ architecture. Each NS3000™ I/O node is based on an Asymmetric Multiprocessing design with two processors and a unique real time OS called the DataXpress kernel. Each processor simultaneously and efficiently executes different functions in the network file serving process. One processor handles network processing and the other processor handles File and Storage Processing. A Service Processor Node is based on the traditional general-purpose single CPU computer running the general purpose Solaris® OS and is used primarily for system management activity. Up to three I/O nodes and one service processor node are connected by a Scalable Coherent Interface (SCI). System software consists of a unique custom messaging system that enables efficient network and storage processing on the I/O nodes and efficient system and data management on the host node. The FMP architecture improves system availability compared to other approaches by isolating the I/O nodes from unplanned outages of the general purpose OS (Solaris®), and I/O processing can continue even in the event that the host node is down. This architecture provides for the Auspex unique large scale UDS advantage among NAS vendors. See also SMP, parallel processing, SCI.

Functional specificity

An Auspex design philosophy where software, hardware and network elements perform specific tasks in parallel for optimal efficiency and scalability.

Gigabit Ethernet

A standard of the IEEE 802.3 committee which provides a mechanism for conveying Ethernet format packets at gigabit speeds. The goals of the gigabit Ethernet include: preserve the CSMA/CD access method with support for 1 repeater, use the 802.3 frame format, provide simple forwarding between Ethernet, fast Ethernet and gigabit Ethernet, support both fiber and copper (if possible), and accommodate the proposed standard for flow control. At the time of this writing it appears that Fibre channel will be adopted to provide the physical layer for the first implementations of gigabit Ethernet.

HBA

Host Bus adapters for Fibre Channel or iSCSI connect the network to the client or server system and sometimes offload network protocol processing (FCP or TCP respectively) from the *Source* computer to the network interface card (NIC) or HBA.

i-man®

An internet-centric software product used by manufacturing enterprises to manage product content for collaborative commerce.

IETF

The Internet Engineering Task Force is the standards body controlling standards for the Internet.

IP

The IP (Internet Protocol) is the protocol for routing packets on the Internet and other TCP/IP-based networks.

iSCSI

A standard being finalized by an IETF (Internet Engineering Task Force) standards committee for the sending of SCSI commands over IP networks.

LADDIS

An acronym formed by names of the group (Legato®, Auspex®, Data General®, Digital Equipment Corporation®, Interphase® and Sun®) that developed and popularized SPEC's vendor-neutral NFS server benchmark of the same name. See SPEC.

LAN
Local area networks or LANs are networks of computers that are geographically close together; this usually means within the same building.

LFS
Local File System. A file system type developed by Auspex and used in the NetServer for file-system communication between the network processors and the file processor and between the host processor and the file processor. LFS provides local file operations similar to NFS remote operations, but without the protocol processing overhead. See also VFS, MAC Media Access Control.

MCAD
Mechanical Computer Aided Design

MCAD
Mechanical Computer Aided Design is a term used to denote the software and hardware used by mechanical design engineers to generate an electronic model/simulation of mechanical parts and assemblies.

MRP
Material Resource Planning.

NFS
Network File System. NFS is an ONC application-layer protocol for peer-to-peer, distributed file system communication. NFS allows a remote file system (often located on a file server) to be mounted transparently by client workstations. The client cannot perceive any functional difference in service between remote and local file systems (with trivial exceptions). NFS is the most popular ONC service, has been licensed to over 300 computer system vendors, runs on an estimated 10 million nodes and is a de facto UNIX® standard. See also VFS, ONC and NFSv3.**NetOS**
The operating system of the Auspex NetServer™ product family.

NFSv3
NFS version 3. References to NFS generally imply NFS version 2 protocol. NFS version 3 is an update to the NFS protocol. Significant among the many changes made for NFSv3 are the adoption of a safe asynchronous write protocol and the use of block sizes up to 64 KB. Other protocol changes are intended to improve the overall network and client efficiency and provide improved support for client-side caching.

NFS ops/s
NFS operations per second. Typical NFS operations include: lookup, read, write, getattr, readlink, readdir, create, remove, setattr and statfs.

NIS/NIS
Network Information Service. This is ONC's general name-binding and name-resolution protocol and service.

Node
See FibreChannel.

ONC
Open Network Computing. The trade name for the suite of standard IP-based network services—including RPC, XDR and NFS—promulgated by Sun Microsystems®.

Operating System

The operating system is the most important software program that runs on a computer. The Operating System (OS) performs basic tasks such as recognizing input from a keyboard, sending output to the display screen, keeping track of files and directories on the disk and controlling peripheral devices such as disk drive and printers or a mouse. The OS acts as a traffic cop and schedules the various programs that the computer executes. The OS is also responsible for security, ensuring that unauthorized users do not access the system. Operating systems can be classified as follows: 1) Multi-user – allows two or more users to run programs at the same time. 2) Multi-processing – supports running a program on more than one CPU. 3) Multi-tasking – allows more than one program to run concurrently. 4) Multi-threading – allows different parts of a single program to run concurrently. 5) Real Time – Usually a stripped down OS that responds to input instantly.

OSI

An acronym for Open Systems Interconnect.

Parallel processing

When a single computer simultaneously uses more than one CPU to execute a program. Ideally parallel processing makes a program run faster because there are more CPUs running it. In practice, it is often difficult to divide a program so that separate CPUs can execute different portions without interfering with each other. Among NAS vendors, only the Auspex NS3000™ effectively overcomes this problem by designing each I/O node with two processors each performing separate portions of the network file-serving task. In addition, the NS3000 links multiple I/O nodes together by a highly efficient Scaleable Coherent Interface (SCI) interconnect, which allows the multiple nodes to act as one system. See also Functional Multiprocessing (FMP).

PDM

Product Data Management is a term that denotes the systems and methods that provide an electronically integrated structure for all types of information. It can be used to define, manufacture and support how products are stored, managed and controlled.

Port / Port ID

See FibreChannel.

RPC

Remote Procedure Call. An RPC is an (almost) transparent subroutine call between two computers in a distributed system. ONC RPC is a Sun-defined session-layer protocol for peer-to-peer RPC communication between ONC hosts. ONC RPC underlies NFS.

RAID

Redundant Array of Independent Disks. RAID is used to increase the reliability of disk arrays by providing redundancy either through complete duplication of the data (RAID 1, i.e., mirroring) or through construction of parity data for each data stripe in the array (RAID 3, 4, 5). RAID 5, which distributes parity information across all disks in an array, is among the most popular means of providing parity RAID since it avoids the bottlenecks of a single parity disk.

SCI

Scaleable Coherent Interface is an ANSI standard (#1596-1992) that is the modern equivalent of a processor-memory-I/O bus and a Local Area Network combined and made parallel to support distributed multiprocessing. The SCI interconnect has very high bandwidth, very low latency and a scaleable architecture. This allows building large high performance systems and is used by Convex/HP supercomputers, Sun Clusters, Sequent, Auspex and others. Network latency has been measured at 150 times less than previous network connections for efficient and fast communication between computer nodes.

SCSI

Small Computer System Interface. An intelligent bus-level interface that defines a standard I/O bus and a set of high-level I/O commands. Each SCSI device has an intelligent SCSI controller built into it. SCSI is used for local data communication between a host CPU and an attached SCSI bus that contains intelligent peripheral devices such as disks, tapes, scanners and printers. There are currently many flavors of SCSI defined by different bus widths and clock speeds. The seven major variations of SCSI are SCSI 1, SCSI 2 (Fast / Narrow), SCSI 2 (Fast / Wide), Ultra SCSI (Fast / Narrow), Ultra SCSI (Fast / Wide) – also called SCSI 3, Ultra 2 SCSI (Narrow), Ultra 2 SCSI Wide.

SMB

Server Message Block protocol. See CIFS.

SMP

Symmetric Multi-Processing. A computer architecture in which processing tasks are executed in parallel on multiple, identical, general-purpose CPUs that share a common memory. SMP computer systems usually have modified operating systems that can themselves execute concurrently. The SMP architecture offers high computational throughput, but not necessarily high I/O throughput. See FMP.

SNMP

Simple Network Management Protocol. SNMP is a protocol used for communication between simple, server-resident SNMP agents that respond to network administration requests from simple-to-sophisticated SNMP manager tools running on remote workstations.

Snooping

When an unauthorized user reads private data sent to another person. This may lead to spoofing.

SoIP

SCSI over IP or SoIP is an effort announced by Nishan in June 2000 that flows SCSI on top of UDP and IP instead of SCSI over TCP and IP as in iSCSI. (See also iSCSI). SOIP may, however, be merged with iSCSI. See also Appendix C for a discussion of the differences between UDP and TCP.

Solaris 2.x

Sun's UNIX® operating system based on System V release 4.

SPARC

Scalable Processor Architecture. SPARC International's specification for the Reduced-Instruction-Set-Computer (RISC) CPUs found in systems sold by Sun Microsystems®, Auspex, etc.

SPEC

Standard Performance Evaluation Corporation. A nonprofit corporation of vendors' technical representatives that develops and certifies accurate, vendor-neutral, computer-system benchmarks. As an example, popular SPEC CPU benchmark metrics include SPECint, SPECfp and the now obsolete SPECmarks. See also LADDIS.

SPECnfs

ops/s. A measure of NFS performance standardized by SPEC. This unit of measure is often used interchangeably with SPECNFS_A93 ops/s. The A93 suffix indicates the first of what may evolve into a series of workloads, each corresponding to different LADDIS variations simulating the loads and traffic patterns of application environments like ECAD, MCAD, imaging, etc. The current version is SFS97 and incorporates NFSv3 testing.

Spoofing

In networking the term is used to describe a variety of ways in which hardware and software can be fooled or spoofed. "IP spoofing," for example, involves trickery that makes a message appear as if it came from an authorized IP address (authorized user).

SSP

Storage solution providers (SSPs) provide outsourced storage capacity services to the enterprise for Internet or intranet application.

UDS

Universal Data Sharing is the single most important feature for NAS and SAN integration. UDS refers to what IBM® calls "true data sharing," where any user can read and write to any file stored as a single data image.

UFS

UNIX® File System. UFS is the standard file system type in the BSD 4.3 kernel. See VFS.

WAN

A Wide Area Networks or WANs is a network of computers that are geographically dispersed and connected by radio waves, telephone lines or satellites.

Zoning

In a SAN environment this is a workaround to security problems with the Fibre Channel specification whereby data pools are assigned to a specific server. This defeats the basic premise of SAN whereby "any application" can have access to "any data."

$12.95

Corporate Headquarters
2800 Scott Boulevard
Santa Clara, CA 95050
USA
408.566.2000 **phone**
800.735.3177 **fax**

http://www.auspex.com **Web**

Japan
Auspex K.K.
ATT Shinkan, 6K
2-11-7 Akasaka
Minato-ku
Tokyo 107-0052
Japan

+81 3 3586 1751 **phone**
+81 3 3586 1778 **fax**

United Kingdom
Auspex Systems, Ltd
Portland House
Aldermaston Park
Church Road
ALDERMASTON
Reading RG7 4 HR
United Kingdom

+44 1189407300 **phone**
+44 1189407333 **fax**

France
Auspex Systems, SA
40, Boulevard Henri Cedex
France

+33 1 42 04 9150 **phone**
+33 1 42 04 9151 **fax**

Germany
Auspex Systems, GmbH
Ingolstaedter Stasse 22
80807 Muenchen
Germany

+49 89 350151-0 **phone**
+49 89 350151-50 **fax**

Auspex Asia Pacific
#06-02 International Building
360 Orchard Road
Singapore
238869

+65 732 9770 **phone**

Auspex Worldwide

www.auspex.com

AUSPEX