research

HEWLETT® PACKARD

HP | Distributed File System | Solutions

The Best Way to Achieve Global Workgroups in a Windows®-UNIX® Environment

Common Internet File System (CIFS) versus Network File System (NFS)

Rev. 2

Table of Contents

Introduction

This paper compares the two leading distributed file systems and examines how each is suited or unsuited to achieving global workgroups in a mixed UNIX®-Windows® environment. Special attention is paid to CIFS and to HP's CIFS/9000 product.

The Necessary Goal: Global Workgroups

Today, society and business are rocketing toward a dynamic and interoperable world, with business operations and workgroups in all parts of the planet needing to share information. Factors such as globalization of services and product development, extremely aggressive time-to-market goals, and increased need for information exchange among employees, suppliers, and customers have made it imperative that information transfer be fast, free, and unfettered.

Today's businesses need a computing infrastructure that allows them to meet the demands of 24 x 7 access on a global basis.

The IT Reality: Mixed UNIX and Windows

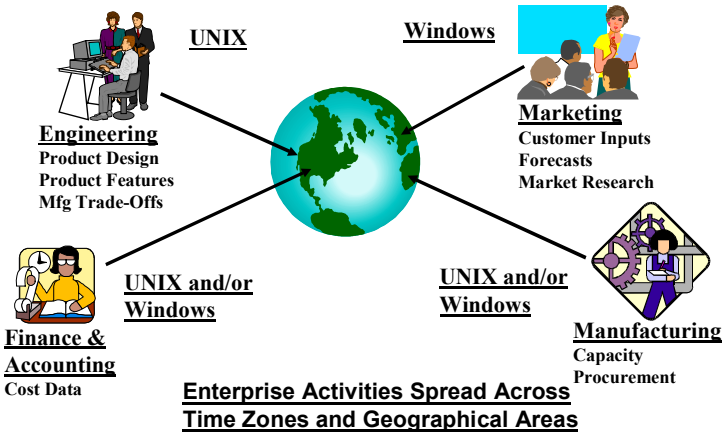
There is a potential roadblock to truly worldwide information access by all parties. It is rooted in the recent history of information technology.

Throughout business, government, and industry, IT departments have had to make a choice between operating systems. Many chose variations of UNIX, especially HP-UX, as the environment for running mission-critical applications. UNIX delivers the robust high availability and performance required for managing large databases, providing the 24 x 7 capability needed for Internet and intranet servers, and handling the host of emerging e-services.

Windows, while viewed as less robust, was plenty good enough to become the corporate workhorse for day-to-day tasks: word processing, spreadsheets, accounting, graphics, desktop publishing, e-mail, and much, much more. Windows is also widely employed in file and print serving. Easy to use and, in one form or another, ubiquitous on personal computers and corporate desktops, Windows has developed a huge installed base in all corners of the globe.

The result is that today nearly all enterprises find themselves with a mixture of Windows-UNIX servers and clients. In fact, UNIX and Windows have been operating side by side for years, and it may appear that they interoperate quite well. After all, many millions of Windows clients have remote access to UNIX servers. In fact, if you work in a corporate environment, you have probably already manipulated files on a UNIX server that appears to you as simply the H: or J: or K: drive on your Windows NT® machine.

Figure 1: Today's Worldwide Enterprises Use Both UNIX and Windows



The Problem: Back-and-Forth UNIX-Windows File Access

But within this seemingly transparent interactivity, there lurks a problem: the lack of *complete back-and-forth file system access* between Windows and UNIX. Users want this access; they want to be able to combine their Windows and UNIX systems seamlessly, with complete interoperability between the two. They want to use a variety of hardware platforms and operating systems without problems, and they want to access any file, on any system, from any location.

The Solution: Implementing a Distributed File System

The path to complete interoperability is the distributed file system—although, as we shall see, "file system" can be somewhat of a misnomer. In fact, the major reason for choosing a distributed file system is Windows-UNIX interoperability. Interoperability specifies how smoothly a user of one operating system can access files on another platform. This is *the* major issue in heterogeneous OS environments.

The Choices: NFS and CIFS

To implement a distributed file system in a Windows-UNIX environment, there are two main choices:

- **NFS**: The system administrator can install Network File System (the most common UNIX distributed file system) on every machine;

  -or-

- **CIFS**: The system administrator can install the Common Internet File System (the Windows distributed file system) on every machine.

Both of these can deliver the "back-and-forth" file system access required for global interoperability. But the performance, security, and administration are quite different.

(Another file system, Distributed File System, or DFS, also offers UNIX-Windows NT interoperability. However, this technology has not been widely adopted, and the number of users is minuscule compared to NFS and CIFS. The technology is complicated, expensive, and immature—and for these reasons has high deployment costs and an uncertain future.)

What You Get with NFS

Developed by Sun Microsystems, Network File System is the most common UNIX distributed file system. (It is nominally available on nearly every popular machine and operating system except the Macintosh* and MacOS, but in practice its penetration has been limited to the UNIX platform.) Although available for Windows, it is not usually included in distributions of Windows 95, 98, or NT. This means that Windows users who want NFS must purchase it separately and install it.

At this writing, the version of NFS most commonly in use is version 3; version 4 is in development. The NFSv4 specification is being developed by a working group of the Internet Engineering Task Force (IETF) and is in the very early stages of definition. Preliminary draft specifications indicate that NFSv4 attempts to address some of the issues facing NFSv3 in the Internet world. However, it is important to note that the NFSv4 specification is several years away from finalization and even more years away from widespread adoption.

NFS is a *de facto* standard, which means there is no guarantee that any changes to NFS or its protocol will be endorsed by Sun and other vendors. Its capabilities may vary between vendors, but later releases of NFS, including version 3 and version 4.x, are generally backward compatible with older versions.

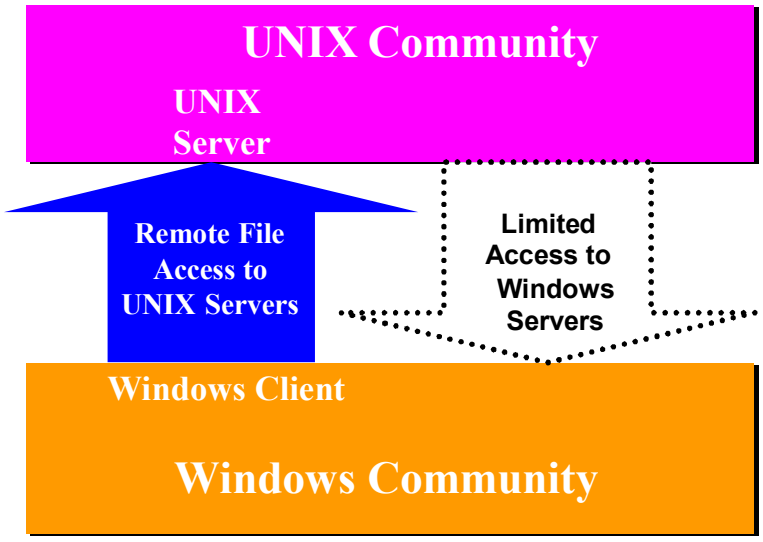
NFS and File Access

NFS includes IP support and read-only client-side caching. It offers strong integration with the Veritas file system VxFS.

* Macintosh computer is a product of Apple Computer, Inc.

When it comes to interoperability between UNIX and Windows, however, plain-vanilla NFS does not measure up. Windows clients have remote access to files on a UNIX NFS server, but users of UNIX workstations cannot access files on Windows servers.

Figure 2: With standard NFS, UNIX users have limited access to files on Windows servers



Microsoft® Windows NT Services for UNIX Add-On Pack makes it easier for customers to integrate Windows NT Workstation 4.0 and Windows NT Server 4.0 into their existing UNIX environments. It contains an NT/NFS gateway, allowing Windows NT users to access files on UNIX systems, and UNIX workstation users to access resources on Windows NT. However, this add-on pack is not widely used.

Other Characteristics of NFS

NFS is not connection-oriented. It offers good performance over local area networks (LANs), but is not good for use with the wide area networks (WANs) that form today's Internet and intranets. NFS maintains only limited state information about files, which means only limited integrity and recoverability. Moreover, NFS is generally considered to have major shortcomings in wide area connectivity, security, and high availability (file replication).

What You Get with CIFS

CIFS, or the Common Internet File System, is the Windows specification for remote file access. It is actually a file system access protocol designed for the Internet. It is not intended to replace existing file access protocols such as NFS, but to complement them. A full cross-platform implementation of CIFS comprises two products:

- **CIFS**, a standard part of every Windows system shipped recently, shipping today, and shipping in the future (Windows 95, Windows 98, NT 4.0, Windows 2000).
- **CIFS/9000**, HP's implementation of CIFS and a standard part of its HP-UX operating system.

CIFS had its beginnings in the networking protocols, sometimes called Server Message Block (SMB) protocols, that were developed in the late 1980s for PCs to share files over the then nascent Local Area Network technologies (e.g., Ethernet). SMB is the native file-sharing protocol in the Microsoft Windows 95/98, Windows NT, and OS/2 operating systems and the standard way that millions of PC users share files across corporate intranets.

CIFS is simply a renaming of SMB; and CIFS and SMB are, for all practical purposes, one and the same. (Microsoft now emphasizes the use of "CIFS," although references to "SMB" still occur.) CIFS is also widely available on UNIX, VMS, Macintosh, and other platforms.

Despite its name, CIFS is not actually a file system unto itself. More accurately, CIFS is a remote file access protocol; it provides access to files on remote systems. It sits on top of and works with the file systems of its host systems. CIFS defines both a server and a client: the CIFS client is used to access files on a CIFS server.

CIFS is not intended to replace HTTP or other standards for the World Wide Web. CIFS complements HTTP while providing more sophisticated file sharing and file transfer than older protocols such as FTP. Like NFS, CIFS is a de facto standard.

A Short History of the Common Internet File System

Late '80s	Microsoft, 3Com, and HP co-develop Server Message Block protocol for interoperability between Windows clients over emerging network technologies (Local Area Networks, Wide Area Networks).
1994	Microsoft releases NT SMB Server source code.
1994/1995	AT&T ports NT SMB Server to UNIX and makes it available to UNIX vendors as Advanced Server for UNIX (ASU). Samba UNIX Server to NT released as open source product.
1996	HP announces its Advanced Server for UNIX product, AS/9000. Microsoft renames SMB protocol as Common Internet File System and publishes specification.
2000	HP announces CIFS/9000 product for HP-UX with UNIX-client-to-NT-server capability, stronger UNIX file system integration.

About CIFS/9000

Hewlett-Packard's implementation of CIFS for UNIX is called CIFS/9000. The Common Internet File System product for HP-UX 11, CIFS/9000 comprises two products:

- CIFS/9000 Server
- CIFS/9000 Client

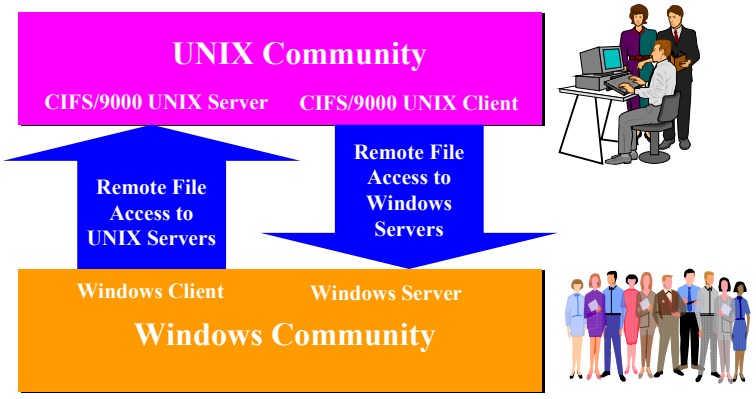
CIFS/9000 also provides a Pluggable Authentication Module (PAM) to allow HP-UX 11 users to gain access to Windows authentication servers.

CIFS/9000 is a follow-on product to AS/9000 (Advanced Server for UNIX) in which HP-UX is a file server to Windows platforms. However, it is more complete than AS/9000 in that it provides both server and client modules for both HP 9000 servers and workstations. CIFS/9000 is a no-charge product which will be automatically ignited—installed at the factory—on every HP 9000 server and workstation beginning in March of year 2000.

CIFS and File Access

CIFS is the key to high-performance cross-platform interoperability between UNIX and Windows. With CIFS/9000 installed on the HP-UX side, Windows clients running CIFS have remote access to HP-UX file servers running CIFS/9000 UNIX Server. At the same time, the UNIX client can access files on the Windows server via CIFS/9000 UNIX Client. CIFS maintains state information for easier recovery from link failures.

Figure 3: CIFS/9000 gives full file system access between UNIX and Windows



Other Characteristics of CIFS

CIFS is a remote file access protocol that is compatible with how applications already share data on local disks and network file servers. CIFS uses TCP/IP networking, a reliable, connection-oriented protocol with scalable performance over both local and wide area networks. With its ability to handle timeouts and retransmissions, TCP is the preferred protocol for WAN transfer. Thanks to its TCP implementation, CIFS provides good support for file sharing and remote file access over WANs.

CIFS is designed to enable all applications, not just Web browsers, to open and share files securely across the Internet. With CIFS, existing applications and applications for the World Wide Web can easily share data over the Internet or intranets, regardless of computer or operating system platform. CIFS utilizes the Internet's global Domain Naming Service (DNS) for scalability and is specifically optimized to support the slower-speed dial-up connections common on the Internet.

CIFS is an open, cross-platform technology based on the native file-sharing protocols built into Microsoft Windows and other popular PC operating systems, and supported on dozens of other platforms, including UNIX. It is extremely robust, preserving file integrity and allowing easy recovery, and it works well over both LANs and WANs. It also provides a path to global user authentication.

CIFS allows groups of users to work together and share documents across the Internet or within their corporate intranets. Users don't have to install new software or change the way they work. It incorporates the same high-performance, multi-user read and write operations, locking, and file-sharing semantics that are the backbone of today's sophisticated enterprise computer networks.

Detailed File System Comparison

This section includes a more detailed comparison of NFS and CIFS, with special emphasis on areas of concern to systems administrators, including:

- User authentication and file system security
- Performance
- Administration overhead and supportability

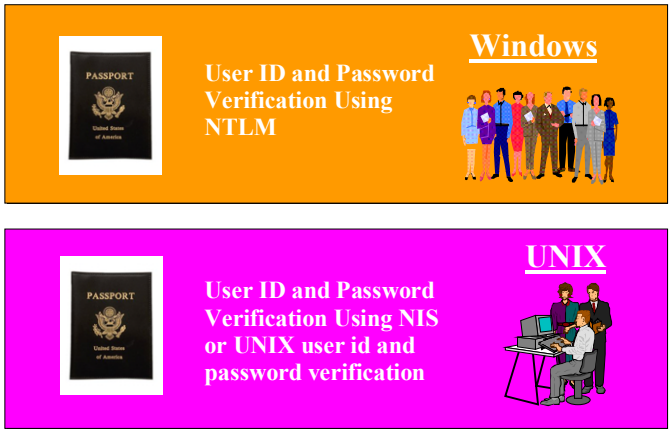
Authentication

A secure networked file system ensures that only authorized users can view and modify information. But a security problem can arise, especially when two disparate operating systems are allowed access to each other's files. A security solution is authentication, which ensures that a user who is accessing file data is indeed the intended user. Authentication means that a user ID and password are checked against databases maintained by system administration to make sure that the user ID and password combination are valid.

In an environment where UNIX and Windows coexist, there must be a way to authenticate users across both platforms. But today, UNIX-Windows authentication is fragmented because UNIX and Windows use different network authentication methods. UNIX most frequently employs Network Information Services (NIS) for user authentication, while Windows uses a CIFS client accessing an NTLM (NT LAN Manager) authentication server. With Windows 2000, support for Kerberos authentication will be added.

Because Windows and UNIX use different authentication schemes, administrators must keep two sets of passwords and user IDs. The maintenance of two different authentication methods for UNIX and Windows is a major issue for Enterprise IT organizations, who would almost universally prefer to have a single authentication method and a single password database.

Figure 4: UNIX/Windows NT user authentication is fragmented



NFS and Authentication

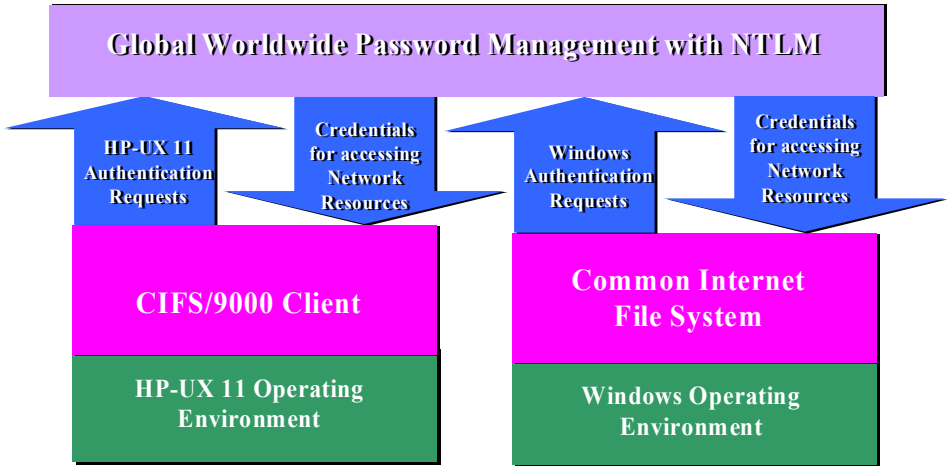
In the most common implementations of NFS, only a UNIX user ID (UID), without verification, is required as a credential to access file system data. This lack of authentication permits an intruder to easily gain access by using a valid UID within an NFS RPC (remote procedure call). Although accommodated by recent NFS versions, global authentication facilities such as Kerberos or Diffie-Hellman public key exchange are seldom used (see the glossary at the end of this white paper for a definition of Diffie-Hellman encryption).

CIFS and Authentication

CIFS/9000 is an enabling technology for common user authentication, allowing HP-UX 11 to use Windows Domain servers as the single, global authentication service. CIFS/9000 client and server software validates each user using the NTLM authentication protocols. The CIFS/9000 PAM NTLM module integrates NTLM authentication with UNIX login and password services. Together they provide HP-UX users with transparent and secure access to the same database against which Windows users are authenticated.

When Windows 2000 is introduced into this environment, CIFS/9000 will continue to provide seamless integration into Windows authentication services. As a component of the CIFS protocol, NTLM authentication is utilized by Windows 2000 and will continue to be supported into the practical future. Later, as Kerberos slowly unfolds as the primary Windows 2000 authentication method, HP-UX will be ready to fully integrate with Windows authentication via a PAM Kerberos module and a future CIFS/9000 Kerberos release. Today, CIFS/9000 provides the best possible common authentication solution.

Figure 5: Using CIFS, UNIX can access Windows user authentication services



To sum up: NFS in general lacks global authentication capability. CIFS requires that each user be authenticated by the CIFS server. It makes global user ID and password authentication possible across both UNIX and Windows, allowing tighter security and easier maintenance in both environments.

Performance

Performance is another important consideration in choosing a distributed file system, and one of the best ways to improve performance is with caching. Client caching can improve the throughput of a distributed file system by increasing the performance of remote file access, increasing the scalability of the server, and decreasing the network traffic.

In operation, data from the client is sent to a local cache. Once the client cache is filled with data, the next read/write operations are generally processed locally instead of going to the server; this avoids network traffic and reduces the utilization of server resources. When the client requests data, the server sends a chunk of data, often larger than the size request, that will be located in the client cache.

NFS and Caching

A caching tool, CacheFS manager, was implemented with NFSv3, allowing this technology to use both memory and disk client cache. However, caching is only for read access; writes are not cached but are instead handled directly by the serving host.

CacheFS typically speeds access to data by a factor of 4; however, it is not recommended for filesets with actively changing files. Furthermore, CacheFS must be specified for every directory mounted. In the case of automount, every client must have the same CacheFS directory configured, thus either increasing system administration overhead or decreasing the benefits of caching.

CIFS and Caching

CIFS provides advanced, robust client caching and includes caching of both reads and writes. As a result, CIFS offers higher automatic data replication and improved data access performance. This technology minimizes network traffic when repeatedly accessing common items such as Web graphic files, and it results in better performance—especially in the kinds of operations found in today's e-services and Web-centric applications.

Like CIFS itself, CIFS/9000 features sophisticated client caching. This brings the reliability and performance of CIFS to the Windows–HP-UX environment.

To sum up: NFS is often criticized for poor write performance. CIFS has much more aggressive caching than NFS and includes both read and write caching, resulting in better overall performance.

Deployment and Administration

Any new technology, no matter how beneficial, must justify its cost. And cost only begins with purchase; a major issue is the amount of IT resources that must be devoted to deployment, including setup and migration. Support is also an area of concern.

NFS: NFS requires a mixture of vendors to achieve a total solution. Windows users who want NFS must purchase and install it separately. This not only requires resources at implementation but also raises questions about support later.

By choosing NFS, the system administrator must install Network File System on every Windows machine. Because of the number of Windows machines in a typical enterprise, the deployment can involve *hundreds of installations*.

CIFS: CIFS is essentially free: CIFS itself comes as a standard component of all Windows platforms (Windows 95 and later), and CIFS/9000 is automatically ignited on HP-UX 11. Furthermore, there is a large base of existing Microsoft clients and servers with CIFS already installed.

A system administrator who chooses CIFS simply installs CIFS/9000 on the enterprise's HP-UX 11 machines. Because the number of UNIX machines is normally much lower than the number of Windows machines, this mean a system administrator might have to perform only *a dozen installations* in a typical computing environment.

What about migration? For end users, the migration from AS/9000 to CIFS/9000 is relatively transparent. A Windows user will see little difference between accessing an HP-UX system with AS/9000 and an HP-UX system with CIFS/9000.

CIFS/9000 is a native implementation on HP-UX, not a port from the Windows environment as is the case with AS/9000. Hence, the administration and management of CIFS/9000 are much more consistent with HP-UX administration and management.

To sum up: NFS is not the best solution from a deployment or supportability standpoint. CIFS and CIFS/9000 offer much easier and less expensive deployment, as well as a nearly seamless integration with Windows. And CIFS is essentially free.

Conclusion

For the optimum path to cross-platform file access between UNIX and Windows, the clear choice is CIFS coupled with CIFS/9000. This solution offers excellent performance and security, while reducing the overhead required for administration and deployment. NFS, while a good local solution, is too simplistic for a global enterprise. Furthermore, NFS has a fractured release history (version 2, 3, and 4 are significantly different) and high deployment costs.

CIFS is especially well-suited for enterprise infrastructures, e-services, and Internet/intranet computing because:

- CIFS has file sharing between Windows and UNIX.
- CIFS has excellent fault tolerance, with the ability to recover from link failures.
- CIFS is optimized for wide area networking links.
- CIFS is secure, with support both for anonymous transfers and for secure, authenticated access to named files.
- CIFS is integrated within the operating system (HP-UX 11 and Windows) for performance and scalability.
- CIFS is free and supportable.

In short, CIFS and CIFS/9000 form ***the*** interoperability solution for environments with a mix of UNIX and Windows platforms.

Feature Comparison Table

General Features Comparison

	NFS	CIFS Today[1]	Proposed Future CIFS[2]
Interoperability between hardware vendors	✓	✓	✓
Integration in an NFS environment	✓☺	✓	✓
Integration in an NT environment	✓	✓☺	✓
High Availability	✓	✓	✓
Read caching	✓	Some[3]	✓
Write caching		Some[3]	✓
WAN support	Soon	✓	✓
Global namespace		✓[4]	✓[4]
Non-root administration		✓	✓
Single point of administration		✓	✓

[1] Current CIFS specification, CIFS/1.0; implemented in Windows NT 4.0, Advanced Server, Samba, and by other vendors.
[2] Windows 2000 (Windows NT 5.0) functionality. Not all functionality documented in official CIFS specification. Microsoft may not document functionality. No specification exists currently.
[3] Caches one user and one file at a time using opportunistic locks.
[4] Client support for Global Namespace in CIFS specification. Server support is a Microsoft Windows NT-only proprietary technology (Distributed File System). Other CIFS servers can be leaf nodes.

Questions and Answers

Q. Is CIFS/9000 available now?

A. CIFS/9000 is scheduled to be available in the first quarter of 2000.

Q. How does CIFS/9000 compare with AS/9000?

A. CIFS/9000 provides both a Windows server and a client for HP-UX 11, which is more capability than is offered by AS/9000. Furthermore, because CIFS/9000 was developed for the UNIX environment, it is also more integrated with HP-UX. By contrast, AS/9000 was derived from Microsoft's Advanced Server for UNIX (ASU), which was originally ported from Windows to UNIX by AT&T. Because of this, ASU has suffered from UNIX integration issues. The more notable integration issues are centered around system administration and incompatibility with basic system utilities such as backup tools.

Q. Will HP continue to support AS/9000?

A. HP has no current plans to discontinue AS/9000, although we expect customers to migrate to CIFS/9000 over time. We are continuing to sell and support AS/9000 and to monitor the marketplace.

Q. Does CIFS/9000 support both Windows NT 4.0 and Windows 2000?

A. Yes. Specifically, Windows NT 4.0 uses NTLM for user authentication, and Windows 2000 will continue to support NTLM as an authentication protocol. CIFS/9000 servers can be accessed by Windows 2000 clients, and CIFS/9000 clients may access Windows 2000 servers.

Glossary

ACL

Access control list, metadata that describes which users are allowed access to file data and what type of access is granted to that data. ACLs define "access rights." In this scheme, users typically belong to "groups," and groups are given access rights as a whole. Typical types of access rights are read (list), write (modify), or create (insert.) Different file systems have varying levels of ACL support, and different file systems define different access rights. For example, DOS has only one set of rights for a file (since only one user is considered to use a DOS system). A POSIX 6-compliant file system allows multiple rights to be assigned to multiple files and directories for multiple users and multiple groups of users.

ASP

Application service provider, an e-business that essentially "rents" applications to users.

Authentication

Scheme to ensure that a user who is accessing file data is indeed the intended user. A secure networked file system uses authentication to prevent access occurring from someone pretending to be the intended user.

Authorization

Ensures that a user has access only to file system data that the user has the right to access. Just because a user is authenticated does not mean he or she should be able to read or modify any file. In the simplest form or authorization, users are given read or modify permissions to individual files and directories in a file system through the use of access control information (called an Access Control List, or ACL.)

CIFS

Common Internet File System, a specification for a file access protocol designed for the Internet.

CIFS/9000

Hewlett-Packard's implementation of CIFS for UNIX. CIFS/9000 provides both server and client modules for both HP 9000 servers and workstations.

Credential

A piece of information that identifies a user. A credential may be as simple as a number that is uniquely associated with a user (like a social security number), or it may be complicated and contain additional identifying information. A strong credential contains proof, sometimes called a verifier, that the user of the credential is indeed the actual user the credential identifies.

Diffie-Hellman

A protocol used to securely share a secret key between two users. Diffie-Hellman protocol uses a form of public key exchange to share the secret key. Diffie-Hellman is known to be susceptible to an interceptor's attack; but authenticated Diffie-Hellman Key Agreement, a later enhancement, prevents such a middle-person attack.

Encryption

Encryption ensures that data is viewable only by those who possess a secret (or private) key. Encrypted data is meaningless unless the secret key is used to decrypt the data. Encryption and decryption of data is called ciphering.

Integrity

Integrity ensures that file system data is not modified by an intruder. An intruder can not intercept a file system data packet and modify it without the network file system discovering and rejecting the tampering.

Kerberos

An authentication and authorization security system developed by MIT and the IETF working group. It is based on secret key technology and is generally easier to manage than a public key infrastructure because of its centralized design. However, Kerberos is not as scalable as a public key infrastructure.

NFS

Network File System; NFS is by far the most widely used distributed file system.

Public Key

An encryption method by which two users exchange data securely, but in one direction only. A user, who has a private key, creates a corresponding public key. This public key can be given to anyone. Anyone who wishes to send encrypted data to the user may encrypt the data using the public key. Only the user who possesses the private key can decrypt the data.

Public Key Infrastructure

Method of managing public key encryption. Although public key technology has the advantage of never exchanging decryption keys, it has the disadvantage of being difficult to manage. Some issues include distribution of public keys with proof of the key's ownership and revocation of expired or terminated keys.

Samba

An open source product that first appeared in the mid-1990s. Samba provides Windows NT server capability for UNIX systems, including most of the capabilities of Advanced Server for UNIX, with the exception of the Primary Domain Controller (PDC) and Backup Domain Controller (BDC) synchronization protocols. Unlike Advanced Server for UNIX, Samba was designed for the UNIX environment and is much more highly integrated. Although Samba is widely used, vendor support for it is not generally available.

Secret Key

Secret key, also known as symmetric-key or shared-key, encryption is a ciphering technique by which two users exchange data by encrypting and decrypting data with a shared secret key. Data is both encrypted and decrypted with the same key. The secret key must be exchanged securely (such as through the "cones of silence") since anyone knowing the secret key can decrypt the data.

SMB

Server Message Block, the file-sharing protocol at the heart of Windows networking. SMB is shared by Windows NT, Windows 95, Windows for Workgroups, and OS/2 LAN Manager. CIFS is essentially a renaming of this protocol.

For More Information

Contact any of our worldwide sales offices or HP Channel Partners (in the U.S. call 1-800-637-7740) or visit our Web site at: **www.hp.com**. Keyword CIFS.

Microsoft, Windows, and Windows NT are U.S. registered trademarks of Microsoft Corporation. UNIX is a registered trademark of the The Open Group.

Technical information in this document is subject to change without notice. HP believes information on competitors to be accurate at the time of publication, but does not guarantee its accuracy.

© Copyright Hewlett-Packard Company 2000

Printed in U.S.A.
01/00