

Group Logic White Paper | October 2009

The Best of Both Worlds

Sharing Mac Files on Windows Servers

A Technical Best Practices White Paper



About This Document

Apple's Mac OS X computer platform has seen its market share grow rapidly in Enterprise environments recently. This has been driven by the popularity of innovative products like the iPhone and the MacBook Air, as well as Mac OS X's security and its ability to run virtualized instances of Windows applications. As a result, Macs are entering the Enterprise through senior level executives or formal corporate pilot projects.

Laura DiDio from industry consultancy Information Technology Intelligence Corporation released a survey in late 2008 that showed 68% of some 700 companies polled planned to allow their end-users to deploy Macs as their work systems in the next 12 months. Separately, a recent survey by the Enterprise Desktop Alliance showed that 91% of IT administrators indicated that integration and management parity between Macs and PCs were major issues of importance to their organization.

This whitepaper discusses the issues involved in sharing files between Macs and Windows servers and compares the options for successful integration. It is intended for systems administrators, technical evaluators and decision-makers who are considering upgrading or purchasing Group Logic's ExtremeZ-IP solution for the first time.

Overview

While Apple's Xserve provides a pure Mac file server solution, most organizations use Windows file servers. Existing investments in Windows technology and training, heavy usage, a high volume of users, virtualization and cluster environments often mean that changing a company's IT strategy is neither economically nor politically feasible. In some situations, such as a creative department within a large corporation, the Mac population is relatively small and must be able to work efficiently with the storage platform used by the majority. To further complicate matters, as Mac OS X begins to penetrate new corporations and educational institutions, administrators must understand and resolve the issues involved in integrating these Macs within the existing Windows file server infrastructure.

Mac users must find an efficient, reliable, and compatible method for sharing files with these servers. This white paper will discuss the issues involved in sharing files between Macs and Windows servers and compare the options for successful integration.

Three methods are widely used and readily available:

- The Mac OS X SMB client uses the native file sharing protocol for Windows. The SMB client is built into Mac OS X.
- Services for Macintosh (SFM), provided by Microsoft in older versions of Windows, uses Apple Filing Protocol (AFP), the native file sharing protocol of the Mac.
- ExtremeZ-IP™, developed by Group Logic, is file sharing software that also uses AFP to store and access files on Windows file servers.

Mac Files

An explanation of how the Mac works with files is key to understanding the issues that arise when connecting Macs to Windows file servers.

- The Mac File Structure
- In the Mac environment, a single file includes four or more parts:
- The data fork – the actual data in the file.
- The resource fork, which can include important information about fonts, settings, and configuration.
- Finder information, which includes metadata such as which application created the file and what kind of file it is.

Extended attributes, which are one or more “named streams” of data. Extended attributes were introduced Mac OS X 10.4 (Tiger) and are used extensively in 10.5 (Leopard) and 10.6 (Snow Leopard).

The Mac file structure plays an important role in creating the user-friendly Mac experience. For example, the Mac uses the Finder information to recognize which application should be used to edit a file so the user does not have to add file extensions such as .doc or .xls. The Mac uses extended attributes to track Finder comments and to mark a file as a potential security threat after it has been downloaded from the Internet.

A reliable file sharing solution must provide for this file structure.

Mac File Naming Conventions

In the Mac environment, nearly all characters are valid for use in file names. Mac files can have diverse names, including such characters as ?, /, *, or characters outside the Roman alphabet. Mac OS X file names are represented using Unicode, a standard that can encode all written languages including those from many European nations, as well as complex Asian languages such as Japanese and Chinese. Storing Mac files with these characters in their names on a Windows file server presents a challenge. While the Microsoft Windows' NTFS file system stores file names in Unicode, many Windows applications (including Windows Explorer) do not recognize or display these characters properly. A reliable file sharing solution must provide flexibility and balance in providing file name compatibility for both Mac and Windows users.

Mac Behavior and Windows Server Performance

In creating the Mac user experience, the Finder and other system components can put heavy demand on a file server. For example, when the user displays a folder that contains files, the Mac enumerates the contents of the folder(s) and collects file system metadata for every file it finds and displays those in the Mac Finder. The Mac will sometimes enumerate the same folder(s) multiple times and often actually opens and reads parts of the files in order to collect the Finder information and display previews, icons, and other information. The result is a significant number of operations generated against a server which can have a negative impact on performance. This problem grows linearly as the number of users increases.

Searching for files presents another performance challenge. The Mac frequently performs searches against the server looking for a particular file, either at the request of the user or through a system

process. Since the Windows server doesn't support or understand how the Mac searches through files, the Mac will look at every file on the Windows system. This operation might be acceptable if the server has few files. However, if many users are conducting simultaneous searches or if you have many files, these searches can have a negative impact on system performance. An efficient and reliable file searching solution must address these issues.

Apple and Microsoft File Protocols

The Mac can use two major protocols - AFP and SMB¹ - to talk to Windows file servers. AFP fully supports the Macintosh because it was designed by Apple and has been enhanced to support the new features of Mac OS X. SMB is designed for Windows file sharing but does not support the full semantics of the Mac OS X and UNIX file systems, an important source of incompatibilities. The UNIX community has proposed extensions to SMB to address this incompatibility, but they have not yet been adopted by Microsoft for Windows.

File Sharing Options

Three file sharing solutions are widely used and readily available: Microsoft's Services for Macintosh (SFM), Mac OS X's built-in SMB Client, and Group Logic's ExtremeZ-IP. The following sections evaluate how each solution addresses the Mac / Windows compatibility challenge.

Services for Macintosh

SFM, based on the Apple Filing Protocol, has been built into the Windows server since Windows NT 3.1. However, Microsoft has not maintained SFM and it has been removed completely from Windows Server 2008. More importantly, SFM only supports version 2.2 of AFP, which is four revisions behind what is currently used in the latest versions of Mac OS X. SFM includes longstanding bugs and also doesn't support Microsoft Clustering or Microsoft's Distributed File System (DFS).

Because Microsoft has not enhanced or maintained this software, many Mac OS X features do not work with SFM. For example, SFM does not support file names greater than 31 characters, files greater than 2 GB in size, long passwords, Bonjour file server discovery, automatic reconnect, Microsoft DFS or fast file searching.

SFM does not fully support Active Directory, which is Microsoft's Directory Services and security architecture for central authentication and authorization services. For example, SFM cannot provide Kerberos-based single sign-on authentication. Using single sign-on, a user can authenticate once and gain access to the resources of multiple software systems.

SFM does not support the full range of Unicode characters that can be created by the Mac. This limitation prevents users from copying files that use these characters in a file name to the server. SFM's support for Unicode is limited to common Mac characters such as /, *, and ?.

To deal with the four parts in the Mac file structure described previously, SFM stores the extra information as NTFS alternative data streams. Alternate data streams are a standard method of associating additional components of information with a file. In SFM, if a Windows user or some other system process moves the Mac file, the extra information stays with the file.

¹ Both AFP and SMB run on top of TCP/IP. AFP should not be confused with the legacy AppleTalk protocol.

For users with Mac OS X 10.6 Snow Leopard, their ability to find and connect to SFM file shares is limited. Mac OS X 10.6 no longer supports the AppleTalk protocol, which is used exclusively by SFM to allow Macs to browse the network to look for file servers. As a result, Mac OS X 10.6 users need to manually enter the address of each file server to connect.

SFM does not address the issue of Mac behavior and its impact on server performance. When the SFM service starts, it builds an index of the entire file system, an operation that often means long delays before Macs can access the file server, sometimes from 15 minutes to hours. The lack of availability of the file server to Mac users translates directly into reduced productivity and frustration for end-users who cannot complete their work.

SFM does not support new features from Microsoft or Apple, including Microsoft Cluster, DFS, Apple's Network Spotlight full content searching, and Apple's Time Machine client backup capability.

Since Microsoft provides minimal maintenance support and has entirely dropped SFM from Windows 2008, SFM is not a viable option for providing Macs access to Windows file servers.

Mac OS X SMB Client

This option is based on Microsoft's SMB protocol and is built into Mac OS X. The SMB protocol was designed to support Windows file sharing. The SMB client makes the Mac look like a Windows client, but the Mac has to make compromises because it is acting like Windows and some of the core features of Mac OS X don't map well to this protocol. For example, recent innovations in Mac OS X 10.5 Leopard included Time Machine backups and Network Spotlight, and are only available over the AFP protocol.

Of critical importance, each version of Mac OS X has changed the SMB client's behavior with respect to key Mac functionality. As a result, there are incompatibilities and file format differences between each major release that can be a source of data corruption, user frustration and help desk calls. ExtremeZ-IP (discussed below) is the only solution which provides full compatibility between Macs and Windows servers regardless of version of Mac OS X or Windows that is being used.

Mac File Structure

The SMB client's approach to supporting the unique Mac file structure presents challenges, specifically how the SMB client stores the data fork, resource fork, Finder information and extended attributes of each file. With Mac OS X 10.6, the SMB client uses the same format as ExtremeZ-IP and SFM, namely the extra file components are stored in NTFS alternative data streams.

With Mac OS X 10.5 and earlier, the approach is different, which can cause significant incompatibilities, especially in mixed Mac OS X 10.5 and 10.6 environments. In Mac OS X 10.5 and earlier, when the SMB client transfers Mac files to the Windows environment, it separates the data file from its metadata and creates hidden "._" (a.k.a. "dot underscore") files that hold the Mac resource fork and Finder information in a format known as AppleDouble. This transfer method is one of the central difficulties for Mac users.

Mac OS 9 used AFP, not the SMB client, and files transferred with AFP (or Mac OS X 10.6) are stored differently. AFP servers and Mac OS X 10.6 SMB clients use NTFS alternative data streams and do not create "._" files during a transfer. This means legacy data, as well as new 10.6-created content on a Windows file server, is in this format. The Mac OS X SMB client prior to 10.6 does not recognize this format. When the Mac OS X SMB client goes to the server to look for data, it looks for the "._" files and,

when it doesn't find them, assumes they do not exist. Valuable metadata are lost and files lose their association with applications. This is a major issue for companies with legacy data and for mixed Mac OS 9 and Mac OS X environments.

The hidden “.” files confuse Windows users and sometimes result in deleted and moved data files and the loss of critical data. Because these files are hidden, they are easily disconnected and lost from the main data file by accident when files are moved, renamed, or archived. The result is data loss for the Mac client, including the loss of the association of a file to the application that created it. In addition, in situations where a user may lock or change permissions on a “.” file, the data file may not open, further confusing the Mac client.

The following screenshot shows how Mac files and their associated “.” files appear when placed on a Windows server with the Mac OS X SMB Client:

Name	Size	Type	Date Modified
._Ducky.tif	96 KB	Microsoft Office Do...	3/15/2007 10:39 AM
Ducky.tif	294 KB	Microsoft Office Do...	10/12/2003 2:02 PM
._Slogan.eps	79 KB	EPS File	3/15/2007 10:40 AM
Slogan.eps	1,154 KB	EPS File	2/16/2005 1:12 PM
._Helvetica.dfont	1 KB	DFONT File	3/15/2007 10:40 AM
Helvetica.dfont	969 KB	DFONT File	1/13/2006 4:59 PM

File Naming

Another source of difficulties with the SMB client is that the full range of Mac file naming conventions are not available for all users. Before version 10.5 of Mac OS X was released, users could not copy files with names containing standard Mac characters to the file server using SMB. Legacy files with names that contained these characters are not available to Mac clients who use OS version 10.4 and earlier. In Mac OS X 10.5 and later, SMB supports file name characters in the same way that SFM and ExtremeZ-IP support them. However, this new capability introduces significant compatibility problems for companies that use the SMB client in mixed 10.4 and 10.5 environments. If the SMB protocol is used, Mac OS X 10.4 clients may not have access to files created with Mac OS X 10.5.

The SMB Client in Mac OS X 10.5 and later doesn't provide the flexibility to control what kind of file name characters can be copied to the file server. With full support for Unicode, Mac OS X 10.5 clients can copy files to the server whose names cannot be seen or accessed by many Windows applications, which causes complications and end-user frustration when files cannot be shared across platforms.

Performance

Since the SMB protocol was designed for Windows and not Mac clients, some Mac behaviors, when executed over the SMB protocol, can have significant negative performance implications on a file server. Since the Mac is acting like a Windows client, the server does not optimize its behavior for the Mac client. As a result, the high level of enumerations, metadata retrieval, and building of previews by Mac software translates to additional I/O operations on the server, which result in performance penalties for all users.

Search

The SMB protocol has no built in support for searching for files, so when a Mac client searches with SMB it must iterate through the entire file system. This is not only a lengthy operation for the end-user, but consumes valuable server resources throughout the process. (See page 9 for more information on the impact of the lack of optimizations).

Additionally, Apple's innovative Network Spotlight full content searching feature is not supported over SMB, as Network Spotlight is only available over the AFP protocol.

Microsoft DFS-N

Microsoft's DFS-N creates a "virtual" file namespace that gives Administrators better control over their file server infrastructure. DFS is increasingly being deployed to enable flexible file server configuration changes, improved reliability / disaster recovery scenarios, as well as optimized file access over WAN links. The Mac OS X SMB client provides no support for Microsoft's Distributed File System Namespaces (DFS-N).

Backup

Mac OS X's Time Machine backup solution provides seamless backup and versioning of files to a network file server. This reduces the need for Enterprises to deploy new client backup infrastructure for Mac clients, and also enables Mac users to self-recover files without calling the help desk. However, Time Machine is not supported over the SMB protocol.

ExtremeZ-IP

ExtremeZ-IP, an AFP-based file server developed by Group Logic, resides on the Windows server and replaces SFM. ExtremeZ-IP resolves the file sharing problems inherent in both SFM and the SMB client and fully supports all versions of Mac OS, from Mac OS 9 through Mac OS X. ExtremeZ-IP does not require installation of client software on the Mac.

All Mac OS X features are fully supported by ExtremeZ-IP. Active Directory and single sign-on, automatic reconnect, long passwords, Time Machine backups, fast file searching (including full content searching with Network Spotlight), and DFS are all available with ExtremeZ-IP. Group Logic continually maintains, tests, and updates ExtremeZ-IP and ensures that it meets quality assurance standards and has no long-standing bugs.

File Naming

To support the diverse file naming options available to a Mac, ExtremeZ-IP maps special Mac characters to Unicode in the same format as SFM so Windows applications that support Unicode can also support these file name characters. Unlike SFM, ExtremeZ-IP also supports the broader set of Unicode characters that Mac OS X can create.

ExtremeZ-IP also provides flexibility in how these special Mac characters are handled. With ExtremeZ-IP, administrators can define and enforce file name policies on their servers. Policy enforcement results in better use of disk space by keeping files off the server that don't belong there, keeping files that break naming policies out of important workflows, and delivering improved compatibility among Macs, Windows, and other platforms all with minimal Administrator intervention.

ExtremeZ-IP's approach to handling file naming is compatible across all versions of Mac OS X, from 10.4 through 10.6. These unique capabilities allow end-users to find their files on the server without opening help desk tickets and with minimal impact on server performance.

Mac File Structure

With ExtremeZ-IP, the Mac file structure components—the resource fork, Finder information, and extended attributes—are stored as alternative data streams associated with the file. From the perspective of a Windows user or system application, the file appears as a single file. So if the file is moved, these alternate data streams which contain the Mac-specific data go with the file even if the application doesn't see them. Since Windows users don't see Mac-related information, they cannot destroy it and, since ExtremeZ-IP and SFM use the same format, legacy data on the file servers are compatible

ExtremeZ-IP's approach to handling the Mac file structure is compatible across all versions of Mac OS, from Mac OS 9 through Mac OS X 10.6.

Microsoft DFS-N

ExtremeZ-IP enables Macs to be integrated into environments leveraging DFS Namespaces as well as DFS Replication. Mac users can navigate the DFS namespace using an intuitive helper application or directly through the Mac Finder. Which option is used depends on the user's configuration and workflow. Also, ExtremeZ-IP fully supports DFS site costing for WAN configurations, as well as DFS home directories. See the related White Paper "Integrating Macs with Microsoft DFS" on Group Logic's web site: <http://www.grouplogic.com/resource-center/pdfs/Integrating-Macs-w-Microsoft-DFS-A-Technical-Best-Practices-White-Paper.pdf>.

Backup

ExtremeZ-IP also supports Apple's innovative Time Machine capability, which provides self-service backup for Mac clients over the AFP protocol. Users' documents are automatically backed up to the Windows file server, and users can selectively restore previous versions directly from their desktop, without help desk intervention. ExtremeZ-IP also provides options to set quotas on the amount of backup space used for individual users and enables Macs to be backed up with the existing Windows backup infrastructure.

Search

ExtremeZ-IP enables Macs to leverage Microsoft's Windows Search 4.0 to conduct full content searches directly from the Mac Finder. ExtremeZ-IP acts as a bridge between Apple's Network Spotlight and Microsoft's Search software. Additionally, ExtremeZ-IP provides fast file name searching via a built-in index. These unique capabilities allow end-users to find their files on the server without opening help desk tickets and with minimal impact on server performance.

Active Directory Compatibility with ExtremeZ-IP

ExtremeZ-IP fully supports Active Directory including Kerberos single sign-on for Mac OS X. Kerberos is the standard authentication for Active Directory. With Kerberos and ExtremeZ-IP, the Mac user can log into a Mac, get a Kerberos ticket, and then access all servers in the domain. ExtremeZ-IP honors all password policies—expiration, complexity rules on client change, as well as other rules for changing passwords. Therefore, password policies are consistent among all client platforms. In addition, ExtremeZ-IP can warn the user a configurable number of days before his or her password is going to expire.

Mac and UNIX permissions are traditionally narrower in scope than Active Directory permissions. Common permissions are owner, group, everyone, read, write, and execute. Active Directory provides many more possibilities for permission groups. Using ExtremeZ-IP, users can take advantage of these additional security capabilities. ExtremeZ-IP maps Active Directory permissions to effective Mac permissions to ensure security. ExtremeZ-IP also provides flexible options for simulating UNIX permissions to Mac clients for optimal application compatibility.

Special challenges exist when Macs are configured to use network-based home directories. When the Mac client tries to access a home directory on a large network, the system enumerates all of the home-directory folders on the server. This process can consume a considerable amount of time as the server has to look through the entire list to find the user's home folder. ExtremeZ-IP includes two optimizations in a feature we call "Access Based Enumerations" (ABE) that mitigate this problem. The first optimization uses the user's profile and lists only folders in that profile. In the second optimization, for users that do not want to use Active Directory, the Mac equates the user name with the home directory and, when the user logs on, displays only the folder that matches his or her user name.

Performance Caching in ExtremeZ-IP

ExtremeZ-IP is tuned to handle the performance demands of Mac clients efficiently and quickly.

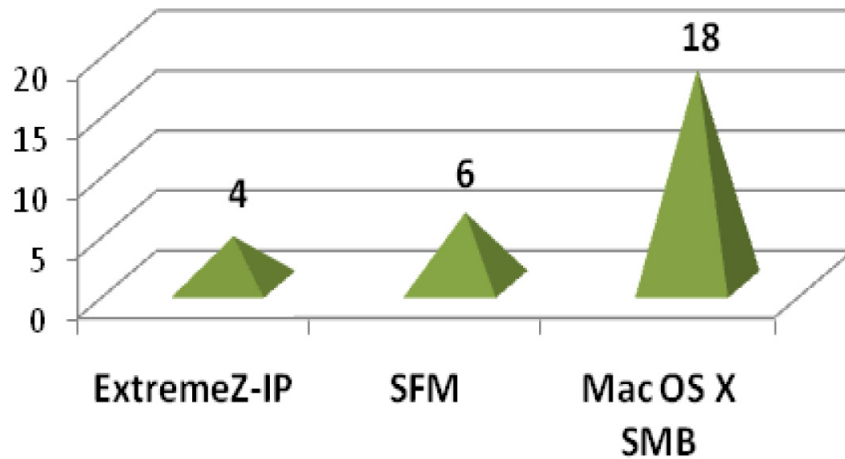
In order to create its friendly user experience, the Mac sends many requests to the file server which can put a heavy load on the system. Group Logic has studied the request patterns of the Mac and introduced caching and pre-fetching technology that predicts what the Mac client is doing and fulfills requests without accessing the file system directly. This technology significantly reduces latency on common enumerations and file metadata requests. More importantly, by keeping the load off the file system, ExtremeZ-IP reduces the use of server resources and frees up processing power and I/O for other operations.

ExtremeZ-IP uses sophisticated algorithms to cache file metadata and other key information. This caching improves performance and user-perceived response time. ExtremeZ-IP maintains an index that includes information about the files it is sharing and the file system structure. As Mac users access files and folders on the server, ExtremeZ-IP dynamically manages a cache called the "node table cache," which contains the information about the files most recently accessed by the Mac client.

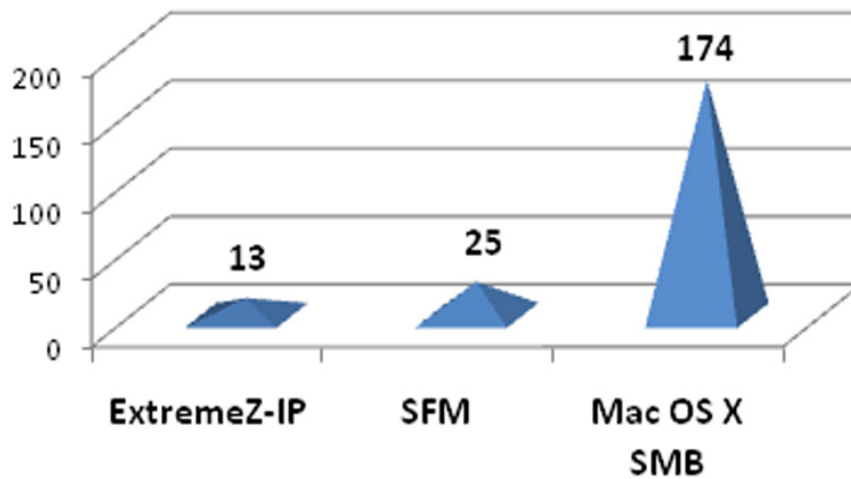
The information ExtremeZ-IP caches includes traditional metadata such as modification dates but also Finder type and creator information, custom icons, file sizes, offspring count, user access rights. ExtremeZ-IP also continually monitors changes in security and updates its cached information accordingly.

The following charts compare the time to display the contents of a folder in the Finder for ExtremeZ-IP with those for SFM and Mac OS X SMB client. These charts show the performance of a single user accessing the server. The impact scales dramatically when multiple users are involved, even if fewer files are in the folder.

Seconds For Finder to Display Folder with 10,000 Files

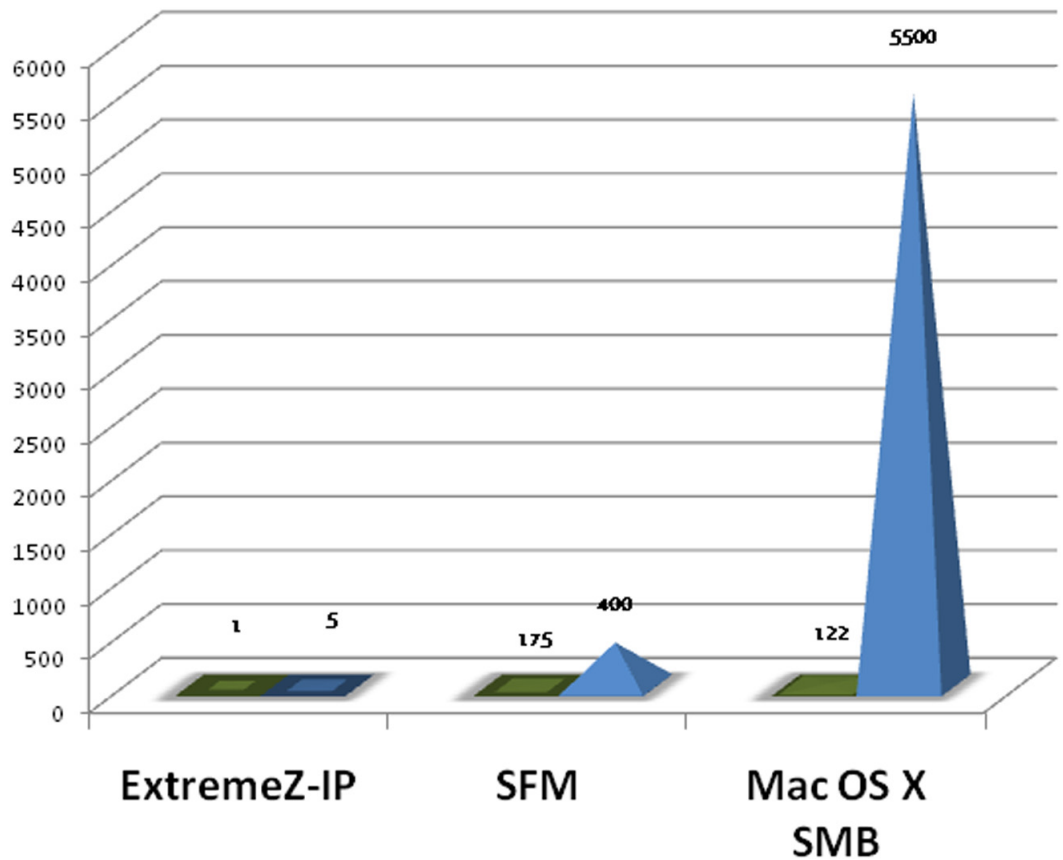


Seconds for Finder to Display Folder with 35,000 Files



The chart below shows the time required for the same three file sharing options to fulfill a request by a Mac client to search for a file with a particular name, which is a very common operation on a Mac. Because ExtremeZ-IP maintains an index to fulfill these requests, it can respond very quickly without having an impact on server performance. On the other hand, the SMB protocol has no support for server-side searching, so the Mac client must traverse the entire file system across the network to look for the file. This lack of support can have a significant impact on server performance.

Time to Search 225K and 1M Files (in seconds)



Conclusion

When it comes to connecting Macs to external file sharing resources, many users and IT professionals discover that while the Mac has basic support for Windows file sharing protocols, something more is needed. This is what we call Compatibility without Compromise: maintain the innovative Mac user experience and functionality; conform to the organization's IT policies and best practices; and improve end-user and network performance and scalability.

Group Logic's ExtremeZ-IP file server meets and exceeds this need. ExtremeZ-IP is a Windows-based software package that implements the Mac's native AFP file sharing protocol, but provides IT professionals the ability to integrate the Macs into the Windows-based IT infrastructure properly and efficiently.

For end-users, ExtremeZ-IP provides support for the features of Mac OS X such as long file names, large files, full content searching, automatic backups, as well as rock-solid stability, performance and lack of data corruption required to get the job done.

For the administrators, ExtremeZ-IP provides support for IT policies such as single file storage of Mac information, Active Directory, Microsoft DFS, single sign-on, caching to keep Mac-intensive file operations off the server, and flexible file naming policies. This translates to measurable business value in the form of increased productivity and IT infrastructure efficiencies, ease of integration and significantly reduced volume of help desk calls.

ExtremeZ-IP comes with the backing of a vendor who understands the issues surrounding Mac / Windows integration and provides timely support and fixes that keep pace with the innovations from both Apple and Microsoft. That is why over 3,000 companies worldwide have trusted ExtremeZ-IP for more than 10 years to provide the best integration between Macs and Windows file sharing infrastructures.