



IT Knowledge • Business Results



Microsoft File Serving and NAS

ESG Report

By Tony Asaro
Senior Analyst

March 2006

Introduction

Microsoft is one of the leaders in file server and NAS technology, and its solutions have been embraced by companies of all sizes across most verticals. Windows Storage Server 2003 (WSS) is being offered through various OEM partners and solutions, and it provides many benefits to small, medium, large, and enormous companies. The NAS market is growing on many different levels and file storage is becoming increasingly critical to business. File storage includes a wide range of documents, including Word, Excel, PDFs, PowerPoint, scanned images, video, audio, CAD/CAM, source code, check images, x-rays, computer generated images (CGI), and photographs. These types of files are often referred to as unstructured data and current ESG research indicates that data growth in this area is exceeding that of other data types, including database and e-mail.

Figure One: File Storage



NAS storage systems have file-level awareness that allows users to access, copy, delete, recover, and search for files using tools that are independently provided by the NAS storage system apart from any host-based application. Advanced NAS solutions must provide basic functionality that makes it easy for users to find and protect their file data. In Windows Storage Server (WSS) 2003 R2, there is a native search capability that allows for the quick search and retrieval of files. Additionally, WSS supports a feature called Volume Shadow Copy Services (VSS) that has fundamentally changed the way that data is protected in Windows environments. With VSS functionality, Windows systems and applications are able to create their own protection copies of data in the background without disrupting ongoing tasks. Users are also able to recover files that have been deleted or corrupted without requiring the intervention of an IT person. And the IT department is freed from the mundane task of responding to dozens and even hundreds of requests to recover data.

One of the core advantages of WSS is that administrators who are already familiar with the day-to-day management of Microsoft systems can leverage their existing expertise to add NAS capacity in a seamless manner. Additionally, Microsoft WSS easily integrates with a myriad of tools and applications provided by Microsoft and its partners and in many cases integration requires little to no additional configuration changes. These include backup applications, anti-virus software, replication solutions, and management applications.

The Road Ahead

ESG believes that NAS will continue to increase in implementation and the rate of growth will rise due to the massive creation and retention of files of all types. Microsoft will certainly continue to be a major force in the NAS arena and is uniquely positioned to provide value to a wide spectrum of customers of all sizes, industries, and geographies. ESG has watched Microsoft from the beginning as it entered the NAS market and witnessed its increasing success and the evolution of its products and technology.

Microsoft Windows Storage Server 2003

A great deal of application data resides on direct-attached storage (DAS) or internal storage found within servers. DAS and internal storage are typically single points of failure. If they break, applications become unavailable and customers could potentially lose data. Additionally, downtime can result in losing money and negatively impact customer satisfaction.

DAS can become difficult to manage as a customer's environment grows. It can require separate management of each individual system. Due to their 1-to-1 nature, backing up DAS systems can be an increasingly cumbersome task with each new server and DAS storage system added to your environment. Creating islands of DAS systems can lead to difficulties managing capacity. Customers often overestimate capacity requirements and buy more than they need. The cost of unused capacity begins to add up as more servers are added to the environment.

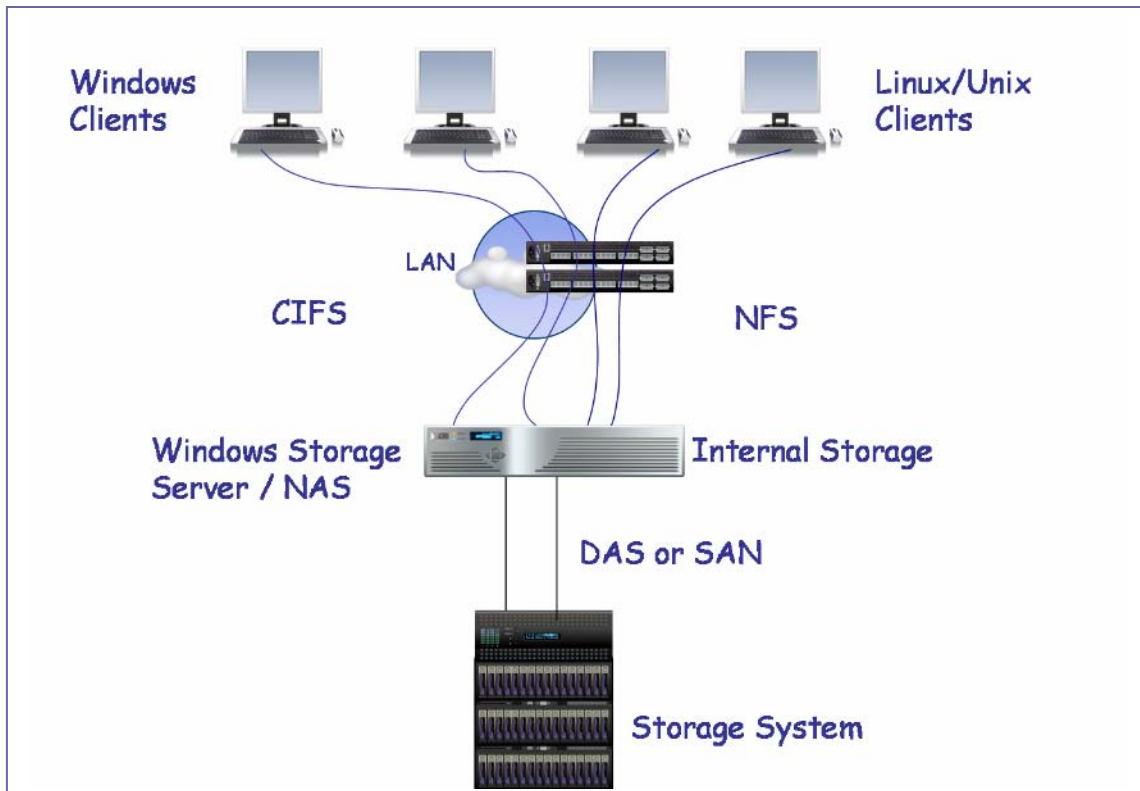
DAS is a good solution but as the number of servers grows, users will see a cross-over point where network storage becomes a better solution. It's like comparing a staircase and an elevator. It may be more practical and faster for someone to use the stairs to get to the second floor, debatable at the third floor, impractical at the fourth, and so on. If you are huffing and puffing at the fifth floor, it's time to get a NAS solution.

One of the key values of NAS solutions is the consolidation of both servers and storage. To manage explosive growth, companies need to be able to centralize the management of file services and storage. NAS solves these issues by providing high data availability and the ability to consolidate storage. NAS simplifies management, centralizes backup and data protection, and provides greater capacity utilization.

Windows Storage Server Defined

The Windows Storage Server (WSS) 2003 solution from Microsoft provides a comprehensive NAS offering that is built on the Windows Server 2003 standards-based platform. WSS 2003 communicates with client workstations over the LAN using the CIFS (Windows) and/or NFS (Linux/Unix) protocols.

Figure Two: Microsoft Windows Storage Server



WSS can be attached to a wide variety of storage systems, including internal storage and SAN solutions that utilize FC or iSCSI protocols. A number of partners provide turn-key solutions, including Dell, HP, and IBM, that integrate the Windows Storage Server and the backend storage arrays.

Microsoft Virtual Disk Service

WSS is easy to install and configure, especially for administrators who are already familiar with Windows server systems. The Microsoft Virtual Disk Service (VDS) has greatly simplified the initialization and configuration of backend storage systems. Now that leading storage system vendors support VDS functionality, it has become very easy to configure storage volumes with WSS.

The ability to expand a volume non-disruptively is very valuable. With the rapid growth of storage capacity and increasing demands for file-based storage, it is very clear that system administrators will need the ability to adjust volume sizes on demand. With VDS and a new Microsoft utility called Storage Manager for SANs (SMfS) (available in R2 versions), this can be accomplished easily and transparently. VDS and SMfS functionality is fully supported within WSS, and these built-in features can be extremely useful in creating, managing, expanding, and deleting storage volumes.

Microsoft Volume Shadow Copy Service

WSS also fully supports Microsoft Volume Shadow Copy Service (VSS) data protection functionality. The VSS feature offers two options, a full volume copy of the primary data and a differential copy. The full volume copy is essentially an exact replica of the primary data. It takes the same amount of capacity as the original primary data. Companies that create full volume copies typically use them for testing and data mining and not data protection.

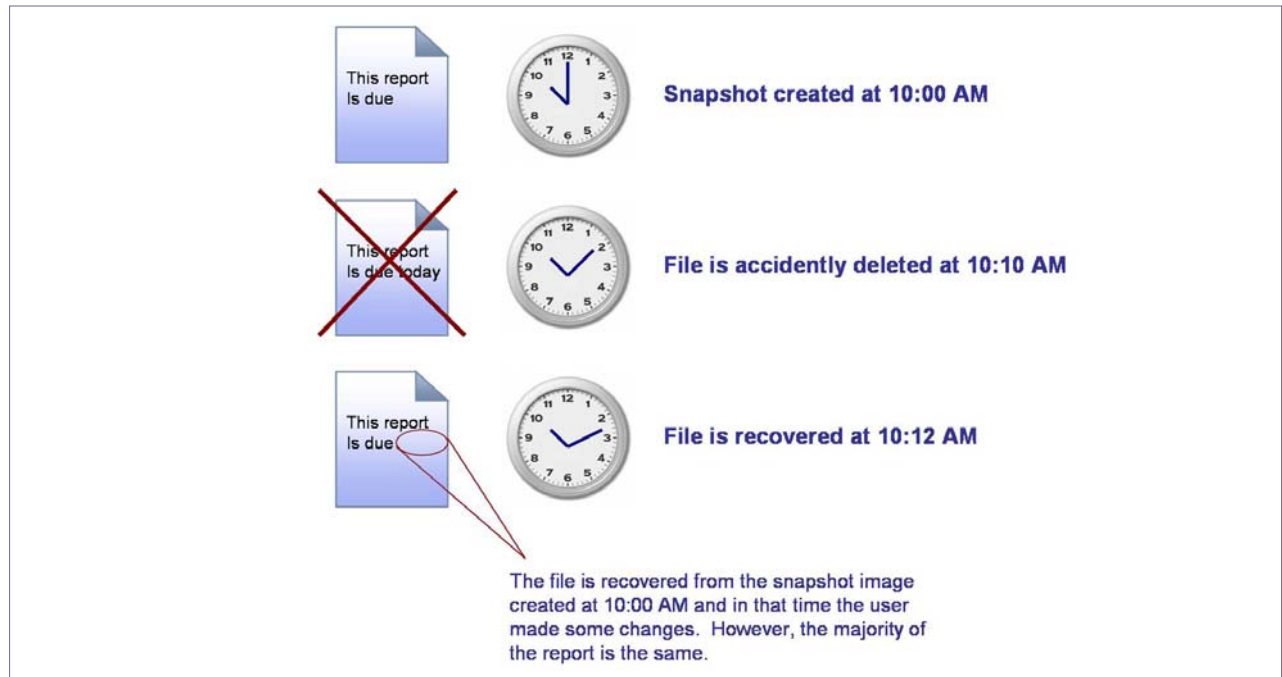
Differential copies or snapshots are logical read-only images of the primary data and only new data is stored in a “difference” area on the storage array. The snapshot data is read-only and cannot be edited, deleted, or corrupted. With VSS, snapshots can be created almost instantly without disrupting any application tasks that may be running. This functionality enables systems administrators to rapidly protect data. Since VSS is built into the OS, it is able to make snapshot copies with an absolutely minimal impact on system resources. This is important, since the storage system’s primary function is to service read/write requests from users and applications. The storage system shouldn’t have to sacrifice its hardware resources to make copies of data. VSS helps WSS to meet robust reliability and availability requirements by reducing the system resources required for snapshots to negligible amounts.

A VSS snapshot is space efficient, meaning that it does not require a large amount of capacity to protect primary data. This is very different than full volume copies, which make an exact replica of primary data. Making a full volume copy of 5 TB of primary data requires an additional 5 TB of capacity. Making a snapshot copy of 5 TB of primary data requires no additional capacity at its initial creation. In fact, making 50 snapshots of 5 TB of primary data also requires no additional capacity. The VSS snapshot will only consume capacity when *new* data is created. As a result, VSS snapshots require much less capacity than full volume copies, providing a lower cost method for protecting primary data.

The primary reason that customers use VSS snapshot copies is to protect and recover data. The following are common examples that outline the value of VSS snapshot technology:

- The system administrator has configured the storage system to make VSS snapshots once an hour on the hour.
- A user has been working on a 25 page report all week and then at 10:10 AM accidentally deletes the file. The report is due by the end of the business day.
- The user accesses the VSS snapshot copy that was created at 10:00 AM and recovers the report. The majority of the report is the same except for the new data that was created by the user in the last ten minutes. While this work has to be re-done, the other 25 pages are still intact.

Figure Three: Recovering a File using VSS



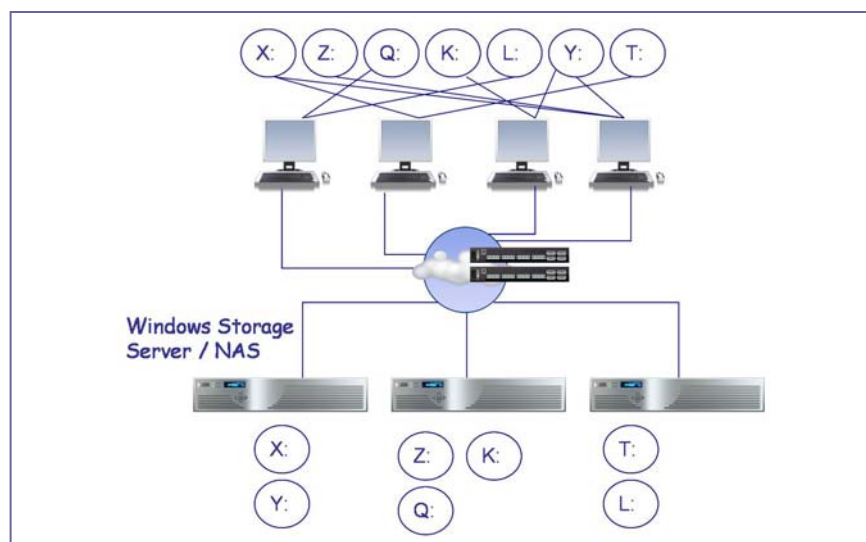
Data Protection Manager (DPM)

Microsoft Data Protection Manager (DPM) is a software application that can be installed on Windows Storage Server (WSS). The combination of DPM with WSS provides a solution that offers primary NAS for file storage and sharing, and offers data protection for other Windows servers that have internal, DAS, or even SAN-based storage. By adding DPM, the investment that customers make in Microsoft NAS can be extended to assist in protecting all of their Windows file data throughout the Enterprise.

Distributed File System (DFS)

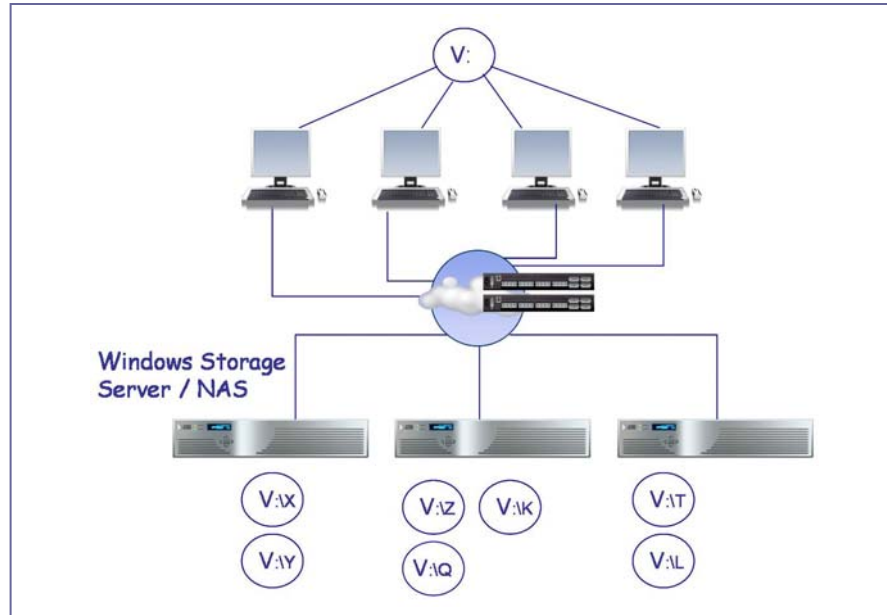
Microsoft DFS is a global name space solution that can significantly simplify management of multiple file servers and improve user workflow. Managing a large number of discrete file servers can be a cumbersome task when performing client adds, moves, and changes. Additionally, companies often have dozens, hundreds and even thousands of NAS file servers, which create complex environments for users to navigate.

Figure Four: Multiple File Server Environment



DFS creates a single aggregated file server interface that provides easier access to multiple Microsoft file server/NAS storage systems. Without DFS, system administrators and users see multiple file servers in the form of different network drive letters. With DFS, they see a single drive letter with access to all of the same files and directories.

Figure Five: DFS Global Name Space



Additionally, DFS can work in conjunction with built-in Windows replication functionality (DFS-R) to provide high data availability. Using DFS-R, the replication of files and directories between file servers within a DFS namespace can be managed and implemented transparently. DFS-R functionality allows IT administrators to publish content to remote file servers, and retrieve remote file server data for centralized backup. In the case of a file server failure within a DFS namespace, replicated data will still be available and accessible from another file server. DFS-R also supports a remote differential compression (RDC) algorithm that helps to minimize the network bandwidth that is required to replicate files. This has been proven to dramatically reduce the WAN bandwidth that is required for replication.

Active Directory

Surprisingly, many non-Windows based NAS solutions that support the CIFS are not fully integrated with Active Directory. While user authentication, permission, and access rights will still work with heterogeneous NAS solutions, these NAS systems are not seen as objects by Active Directory. The majority of companies that have Windows servers use Active Directory to centrally manage users, servers, and policies. Active Directory drives a number of critical functions and tools, including Group Policy rules. For example, Windows administrators can set Group Policy schedules for Windows Installer that automatically installs and distributes anti-virus software, server upgrades, and backup applications. Active Directory works with DFS policies to control replications and provide load balancing for performance. Security, including access controls, IPSec support, and encryption, is all managed through Active Directory for WSS NAS solutions. All of these capabilities provide powerful, secure, and comprehensive management of Microsoft NAS environments.

Customer Environment

Recently a large European car maker installed a group of non-Windows based NAS appliances in a few departments, only to find out that some key group policies were not supported. These appliances required too much manual intervention and were not cost effective. The customer replaced their appliances with Windows Storage Server to ensure seamless integration with all the group policies set in their Active Directory deployment.

Application Integration

The Windows Storage Server (WSS) 2003 and R2 version are NAS solutions with at its core a Windows server that can run multiple applications. The same WSS that supports NAS functionality and DFS also supports a wide range of antivirus software, backup applications, print server functionality, and other server tasks. This allows companies to consolidate servers. Additionally, it gives customers a choice of different applications to support the Microsoft NAS environment. For example, ESG believes that anti-virus software should be required for all NAS products. NAS solutions, like any server on the network, are vulnerable to viruses. Every NAS solution offers antivirus support, but IT administrators often prefer to implement a consistent anti-virus solution that can be easily updated across multiple servers. Fortunately, there is a wide range of anti-virus solutions for NAS customers to choose from with WSS 2003 R2. Another example is remote replication software for disaster recovery. There are a number of leading remote replication solutions that can be used with Microsoft NAS. This gives customers the freedom to choose the solution which they believe best meets the needs of their business. ESG is aware of 50 thousand remote mirroring licenses in the market today with the majority running on Microsoft Windows servers.

Another important example is that only NAS systems built with Microsoft WSS R2 can use Microsoft SharePoint natively. Any files stored on Microsoft WSS NAS systems can leverage built-in SharePoint functionality for web-based collaboration and file sharing.

File Server Resource Manager (FSRM) is a suite of tools that enables administrators to understand, control, and manage the quantity and type of data stored on their Windows servers. System administrators can configure quotas to limit the amount of data stored at the folder, share, or volume level. FSRM enables administrators to efficiently monitor existing storage resources, and it can be used in the planning and implementation of future policy changes. FSRM addresses capacity, policy and quota management.

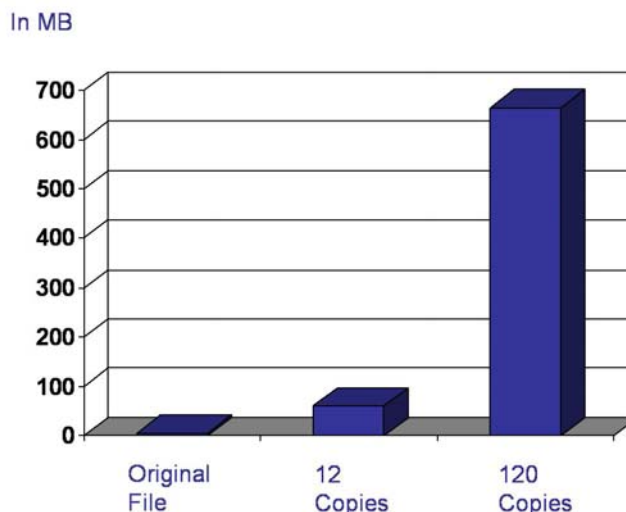
Platform Partners

Customers can choose from a number of platform partners that provide turn-key solutions, including Dell, HP and IBM. Having these leading vendors as well as a number of other OEM partners allows customers to have a wide range of options to choose from. Additionally, these vendors are well known market leaders that have customers in small, medium, and large corporations, the government and the military, as well as global reach.

Single Instance Storage

ESG has found that there are a large number of duplicate files that are stored over and over again. Duplicate files consume a great deal of storage, which creates a domino effect of inefficiency. First, it increases the cost of storing data. Consider a 5 MB file that has been e-mailed to a dozen users. If each of these 12 users decides to store that same file on a network drive, 65 MB of capacity will be consumed by that one file. If each of these users sends the same file to another 10 users and they in turn store the file on a file server/NAS system, the amount of total capacity explodes to 665 MB.

Chart One: Capacity Inefficiency



The domino effect continues at the time when backup operations are performed. The backup application must store all of the file copies that have been created on multiple file servers, requiring more time and more tapes to complete the process. The need for more tapes impacts the process of managing tape media, which in turn requires more administration time. Also, managing more tapes increases the probability of human error, which can lead to data loss.

If remote replication is being performed, the process of copying and transferring that data over the network is negatively impacted. It requires more time and bandwidth, which impacts recovery point objectives. Since so much data is being moved over the WAN, more costly bandwidth is required. Additionally, at the other end of the wire, more storage is needed for all of those copies.

Over time, the domino effect may require companies to buy new storage systems based on capacity requirements. With each new storage system there are further requirements for software licenses, maintenance agreements, additional tapes to perform backups, increased media management with a statistically greater chance of data loss, costly WAN bandwidth, and more.

Windows Storage Server 2003 R2 differentiates itself from Windows Server 2003 with built-in support for single instance storage (SIS), which physically consolidates duplicate file data but logically maintains independent file metadata through sparse links that appear to the end-user to be just like the original files. If we go back to the original 5 MB file that was ultimately distributed to 120 users and consumed 665 MB of capacity, with Microsoft NAS single instance storage, it will only require 5 MB of storage. Think about the economic impact of this capability. ESG believes that this is extremely useful functionality that goes a very long way to solve this domino effect of inefficiency.

Search

Microsoft supports Content Indexing on its Windows Storage Server 2003 R2 solution, which provides search capability that enables users to find and access files quickly. ESG believes that search is an extremely useful capability with file-based storage systems. The reason is that as the number of files increase, navigating through them all to find the file that you are looking for becomes more difficult. If you are looking for multiple files, the challenge (and frustration) escalates. And if those files were created by different applications, then the time, complexity, and chance of success are all impacted. The ability to perform search can significantly simplify the process of finding and accessing the files you need. This has a direct impact on improving business efficiency.

Microsoft Windows Server 2003

Microsoft Windows Storage Server 2003 is built from Windows Server 2003 and shares many of the same features and functions. It is important to point out that a large number of Windows Server 2003 systems are being used for file and print applications. According to Microsoft, over 19% of all Windows Server 2003 systems are being used for this purpose in new deployments.

Using Microsoft Windows Server 2003 for file services provides an easy and low cost way to create shared storage, offering much of the same value as Windows Storage Server 2003, but typically in a scaled down capacity and without some of the advanced features.

Choosing a NAS Solution

There are a number of viable NAS solutions in the market today. The evaluation process will differ based on the needs of the environment. Fundamentally, there are a number of things to ask and consider. If the majority of the environment is Windows-based with some mix of Linux/Unix, then using Microsoft NAS should be on the short-list. Moving Operating System prejudices aside, Microsoft NAS has good NFS and excellent CIFS performance. And quite frankly, CIFS performance is the harder of the two to optimize. Additionally, Microsoft has excellent price/performance, a metric that often gets missed with various benchmarks. Any environment can be over-engineered to reach a performance number but there is always an associated cost. Price/performance should always be considered and evaluated to decide what you are willing to pay to meet the practical needs of the business.

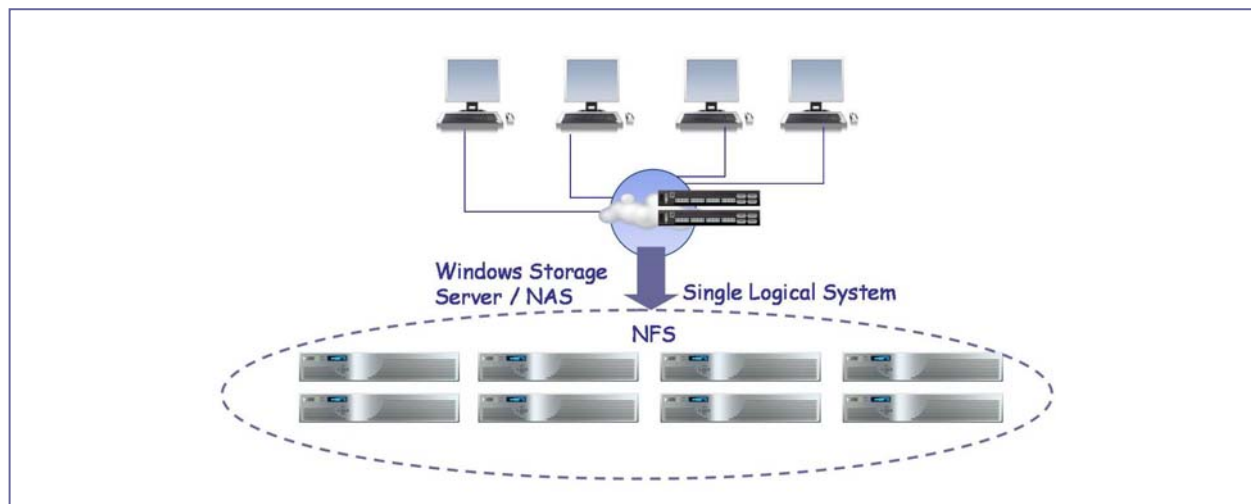
Using Microsoft NAS to store Linux and Unix data is supported and yet not widely adopted. In some cases, there may be legitimate technical reasons for not using Microsoft NAS in these environments. It is also possible that system administrators in a Linux/Unix environment may feel much more at home with the command line of a Linux/Unix-based NAS solution or another solution that uses the same nomenclature. Nevertheless, it is counterproductive to reject Microsoft NAS solutions out of hand for Linux and/or Unix just to avoid crossing Operating System lines. Microsoft NFS performance is very comparable to many Unix based appliances (see www.spec.org), and by using Windows NAS solutions, companies can effectively consolidate their file server solutions while also taking advantage of many new advanced WSS features. Certainly, companies with a mixed environment with a greater proportion of CIFS to NFS traffic should seriously consider Microsoft NAS for both.

At this point, some administrators might begin to consider the mix of CIFS and NFS traffic that makes sense, but this is not the right question to ask. As stated above, Microsoft NAS provides excellent CIFS performance and good NFS performance. ESG feels that the important issues to consider include the following:

What type of NFS performance is required? Many high performance computing environments require bleeding edge results from Linux and Unix-based applications. These environments are sometimes referred to as the “Lunatic Fringe” and require an enormous amount of resources to enable them to handle lots of large-block data streams. This is not what Microsoft NAS should be used for. However, if Linux is being used by a development group, for example, then Microsoft NAS is well-suited for this purpose. Especially since many of these same users are probably using Windows for their PCs and Microsoft NAS for their home directories.

Microsoft will support NFS scale out and can cluster up to eight WSS NAS systems. This means that customers can start with a single node Microsoft WSS NAS system and add nodes one by one up to a total of eight. ESG is a vocal proponent of clustered storage architectures because they enable linear scalability and greatly improve reliability without adding significant complexity to the management of the systems. Clustered architectures increase processing power, memory, and bandwidth. Not every NAS solution supports clustering and its performance scalability. Clustered storage provides customers an easy way to grow their environments as needed.

Figure Six: NFS Clustering



A more relevant issue than performance is price/performance. It is important to balance performance with price in order to find the right solution for your environment. Ask yourself what the right price/performance is for your applications and environment.

Performance is important to any application and environment. However, hardware advancements and protocol improvements continue to make performance less of a critical path. This then forces customers to engage in a discussion on functionality and manageability, which ESG feels are increasingly the more important factors when choosing a NAS solution.

Functionality and Manageability

Storage system vendors should not be in the business of selling capacity. They should provide solutions that allow customers to optimize the capacity that they have. Most, if not all, storage system vendors only pay lip service to this idea. However, the value of single instance storage (SIS) is both quantifiable, saving real cost that can be easily measured improving the overall efficiency of managing storage. As discussed above, the ability to retain only unique data and remove copies can drive down the overall cost of storage by a magnitude. Does your NAS solution support SIS?

Is Microsoft or Linux/Unix the dominant platform? If Windows is the dominant platform, then companies can take advantage of Microsoft NAS features, functions, integrated tools, and applications. Active Directory can be leveraged for better management. How well integrated is your NAS solution with other valuable Microsoft tools like Active Directory?

DFS creates a global name space in order to greatly simplify the management of NAS environments that have dozens, hundreds, or thousands of file systems. How well does your NAS solution integrate with DFS, DFS-replication, and management with Active Directory?

ESG feels that Index Search capability will become requisite with file system storage in the near future. If you store thousands, hundreds of thousands, millions of files or more that were created by multiple applications, how do you easily retrieve and access this data? Search greatly simplifies this process, improving efficiencies and the overall workflow process. Does your NAS solution support Search capability?

These are some of the question that ESG feels are important when choosing and evaluating a NAS solution.

ESG's View

Over the last couple of years, Microsoft has evolved its NAS solution from providing basic functionality to offering advanced features. Additionally, providing a suite of tools, including VDS, VSS, Active Directory, DFS, GPO, applications, and ISV solutions that enhance its NAS, make it even more compelling and useful to customers. The addition of Search capability and Single Instance Storage (SIS) creates an even greater leap in value. ESG is a big proponent of Search and reducing duplicate data on storage and feel that both of these capabilities will become pre-requisites for primary NAS systems over time. We feel even more strongly about this with Microsoft leading the way. While performance is always important, there are other considerations that require priority, including the ability to rapidly access data and the reduction of the domino effect of inefficiency. Having said this, Microsoft has excellent CIFS performance, good NFS performance, and with its NFS clustering, can scale performance by adding up to eight nodes into a cluster. Additionally, Microsoft's price/performance story is a very attractive one. Microsoft has shown the kind of real leadership and innovation that raises the stakes for file servers and broader NAS storage solutions as a whole.

All trademark names are property of their respective companies. Information contained in this publication has been obtained by sources the Enterprise Strategy Group (ESG) considers to be reliable but is not warranted by ESG. This publication may contain opinions of ESG, which are subject to change from time to time. This publication is copyrighted by The Enterprise Strategy Group, Inc. Any reproduction or redistribution of this publication, in whole or in part, whether in hard-copy format, electronically, or otherwise to persons not authorized to receive it, without the express consent of the Enterprise Strategy Group, Inc., is in violation of U.S. copyright law and will be subject to an action for civil damages and, if applicable, criminal prosecution. Should you have any questions, please contact ESG Client Relations at (508) 482-0188.