



Optimizing and Managing File Storage in Windows® Environments

A Powerful Solution Based on Microsoft® DFS and Virtual File Manager™

September 2006 | TR-3511

Abstract

The Microsoft Distributed File System (DFS) is a storage service that can help solve many complex problems in Windows environments. Organizations that deploy DFS can achieve immediate and significant file management benefits. In addition to making it easier to manage files, DFS provides a core set of capabilities on which powerful enterprise storage management solutions can be built.

Virtual File Manager is a solution for managing distributed file storage in Windows environments. Built on DFS, Virtual File Manager enables the integrated management of logical and physical storage elements, making it the most comprehensive Windows storage management solution available.

Table of Contents

| | |
|---|---|
| 1. Introduction | 3 |
| 2. What is a Global Namespace? | 3 |
| 3. Laying the Foundation for File Storage Management with DFS | 3 |
| 4. Virtual File Manager Global Namespace Differentiators | 7 |
| 5. Summary | 9 |

1. Introduction

Organizations in every industry are experiencing explosive growth in their requirements for file storage capacity. The proliferation of application data and user-generated documents—such as presentations, spreadsheets, graphics, and scanned documents—is driving continued expansion of network storage requirements. As a result, Windows storage configurations are growing in both size and complexity. This growth and complexity creates challenges for storage administrators as well as users. And it is a direct cause of increased storage management costs, as well as suboptimal storage utilization and data availability.

2. Enterprise Storage Management Challenges

Storage requirements are outpacing most organizations' ability to manage them efficiently. There are four key challenges relative to the growing requirements for file storage:

- **Explosive growth in enterprise storage**
Many factors are contributing to an enormous increase in network storage requirements, including the pervasiveness of wide area networking and users' ability to create and duplicate huge volumes of data.
- **Underutilization of distributed storage**
According to a Gartner Group study, only 30% to 40% of storage in distributed environments is utilized, which is low compared to 80% storage utilization in mainframes.
- **Complexity in distributed storage management**
Another Gartner Group study found that storage and storage-related management consumes as much as 75% of corporate IT budgets, and that for every dollar spent on storage hardware, much more is spent on administrative costs. Given the fact that storage requirements are increasing as much as 100% annually, organizations cannot afford to continue "throwing hardware" at the problem or hiring more administrators to handle the additional workload.
- **Achieving cost-effective high availability for distributed storage**
Organizations face a daunting task trying to make data that is spread throughout the enterprise highly available. Doing so requires a well-planned, well-designed network storage architecture that can be centrally managed. Until now, it has not been possible to centrally manage distributed storage in Windows environments, so organizations have found it very difficult to create and manage distributed file storage environments.

DFS and Virtual File Manager provide a means for addressing and solving many of these storage management problems. Based on DFS, Virtual File Manager is the most powerful automated storage management software for Windows environments. It can significantly enhance storage availability and scalability while reducing the cost and complexity of storage management.

3. Laying the Foundation for File Storage Management with DFS

Historically, with the Universal Naming Convention (UNC), a user or application was required to specify a physical server and share in order to access file information (that is, the user or application had to specify `\\Server\Share\Path\Filename`). A UNC is typically mapped to a drive letter, where `x:` might be mapped to `\\Server\Share`. Users must know the physical name and location of the machines and shares they want to access. And users must map to many drive letters (`d:`, `e:`, `f:`, `g:`, `h:`, etc.) to access information stored on

different servers. As a result, users can be overwhelmed by complicated share names and the number of places where data can be stored.

This storage management paradigm creates major problems for both users and administrators. The administrator must forever maintain the server name once it is published to users. Changing a \\servername\sharename requires the administrator to notify all users (perform “desktop touches”) so that their machines can be reconfigured to access the renamed server and shares. With this approach, administrators have very little flexibility in the way they manage users and storage. Machine name dependencies lead to inefficient administration and underutilized storage, both of which increase costs and reduce the effectiveness of storage management. However, DFS services in Windows Server 2003 and Windows 2000 provide a significant change to the traditional storage management paradigm.

3.1 A Platform for Distributed File Management

Windows Server 2003 and Windows 2000 include DFS, a storage service that can help solve many storage management challenges. DFS enables the creation of a logical file system that can unify multiple physical file systems. As a result, organizations can use it to improve storage-related functionality, reduce costs, and provide more comprehensive storage management.

DFS is also a strategic file server virtualization infrastructure on which software vendors can build value-added storage management solutions. DFS consists of a client and a server component. The client component is included with all Windows clients, and allows the client to make requests to the DFS server. The server component is included with Windows NT, Windows 2000, and Windows Server 2003. The DFS server component receives a client request and redirects or refers it to a physical target, similarly to the way a browser receives a DNS call and refers the client to a Web site.

DFS does for servers and shares what file systems do for hard disks. File systems provide uniform named access to collections of sectors on disks. Similarly, DFS provides a uniform naming convention and mapping for collections of servers, shares, and files. Another way of thinking about it is that DFS is to file storage what DNS is to networking or Active Directory is to users and computers.

3.2 A Storage Solutions Development Platform

Because DFS is a component of Microsoft Server, vendors can easily build storage solutions based on open standards with DFS at the foundation. Vendors such as Brocade significantly leverage the DFS platform to create solutions that address critical storage issues such as disaster recovery, data migration, server consolidation, storage reconfiguration, and storage optimization.

3.3 The Compelling Case for Deploying DFS

DFS lays the foundation for enterprise file storage management by providing the following core set of capabilities.

- **Eliminates machine name dependencies**
DFS removes the “once-published, forever-maintained” requirement for server names, which creates tremendous flexibility for administrators to add or move files and users without having to touch or reconfigure the desktop. Removing machine name dependencies enables the creation of a logical storage layer, which in turn enables the creation of strategic enterprise file storage solutions.

- **Enables a logical view of physical storage**
DFS separates the logical and physical aspects of storage to enable the creation of a logical layer. This means that administrators can create logical views of physical storage that match what users and applications want to see instead of how and where files are physically stored.
- **Protects investments in Windows software**
Because DFS is included in all Windows clients and every Windows server product since Windows NT 4.0, organizations can take advantage of DFS without having to purchase, load, or deploy additional operating system software.
- **Protects investments in storage hardware**
DFS supports file storage across multiple storage types—NAS, DAS, server-attached SAN—from various vendors. This means that organizations can use DFS to aggregate and increase the utilization of their existing storage devices, which can significantly reduce hardware costs.
- **Enables the development of enterprise storage solutions that are easy to deploy**
With DFS as the foundation, vendors can develop storage solutions that reside above the operating system and that do not require kernel-level software. This means that organizations can quickly and easily deploy these new storage solutions, because they do not have to retest applications or load new system-level software.

3.4 Virtual File Manager Adds Value to DFS

DFS is an immensely valuable service, but it is not a complete solution. Virtual File Manager adds value to DFS in the following ways:

- Provides a complete file virtualization solution that integrates management of the logical and physical layers
- Provides comprehensive DFS management
- Makes DFS enterprise ready by enabling its use in complex, large-scale Windows environments
- Uses DFS to deliver a storage management solution

3.5 Consolidated Network Data Management

Most enterprises today buy different types of storage from multiple vendors to meet specific business needs. A typical enterprise IT environment contains storage devices from multiple vendors, including Dell, HP, IBM, Network Appliance, and EMC, each with its own proprietary device management tools. Administrators face the challenge of finding ways to centrally manage data in such a mixed environment, and often resort to rudimentary tools or home-grown scripts. Some of the key challenges that administrators face include moving data across CIFS- and NFS-based file systems and efficiently managing storage and data with multiple GUIs.

A Global Namespace provides administrators with a single location to view and centrally manage all unstructured network data in their environment. Virtual File Manager is the only product on the market that delivers comprehensive network data management services that are device independent. Using Virtual File Manager, administrators can move and manage data across different types of storage (DAS, SAN, and NAS), from different vendors, and across NFS- and CIFS-based file systems.

Virtual File Manager Advantage

- Centralized management of network data residing on DAS, SAN, and NAS devices anywhere in the enterprise
- Management of data residing on Windows, Linux®, UNIX®, and NAS storage systems
- One simple, intuitive GUI to perform network data management tasks for all unstructured data in the environment

3.6 Remote Site Data Management

Managing data across multiple branch offices presents IT administrators with the challenge of finding ways to maintain data availability while keeping administrative costs down. In addition, the cost of replicating data for disaster recovery purposes can be quite high in an organization that has many locations. Maintaining a tape backup infrastructure at each remote site is inefficient and costly, and replicating data across a WAN for centralized backup is difficult and time consuming.

A Global Namespace provides a cost-effective solution for managing geographically distributed data because administrators have a single view of data across multiple locations and can manage it as a single entity. Data can be replicated to a central location for backup, eliminating the need to maintain tape backup infrastructure at each location.

Virtual File Manager Advantage

- Automated, policy-based replication across heterogeneous devices and multiple locations
- Ability to schedule many-to-one replications to copy data from multiple remote sites to one central device for centralized backup
- Ability to schedule hundreds of replication jobs simultaneously
- Patented Byte-Level File Differencing Replication (BFDR) copies only byte-level changes to files, reducing WAN bandwidth consumption during replication by up to 90%

3.7 Data Classification and Reporting

Data classification and reporting is a critical first step in the implementation of any successful data lifecycle management strategy. Without a clear understanding of the value of data to the organization, administrators cannot put into practice policies intended to reduce IT costs by aligning storage practices with business priorities. Today, the process of classifying information is largely manual and requires numerous hours. In order to obtain detailed department-level information, administrators must resort to manually summarizing and collating multiple reports from multiple sources.

With a Global Namespace, administrators have a logical way of grouping data for classification. Virtual File Manager takes data classification a step further and provides robust reporting capabilities that allow administrators to prepare for major IT projects. Virtual File Manager uniquely enables IT administrators to classify data based on a number of business-relevant categories, including department, location, project, user group, file age, file size, and last access time, among others. Once the classification policies are completed, administrators can run customizable reports to meet their specific needs for determining data migration requirements.

Virtual File Manager Advantage

- In-depth data classification is based on multiple parameters.
- Reporting through Virtual File Manager is automated and policy-based.
- Virtual File Manager delivers a comprehensive list of standard reports and allows administrators the flexibility to run customized reports on a number of key criteria.
- Virtual File Manager supports agentless and/or agent-based reporting.
- Virtual File Manager reports can be used as the foundation for implementing data lifecycle management strategies.

4. Virtual File Manager Global Namespace Differentiators

Global Namespace has become a buzzword in the storage industry, and many vendors are touting this capability. However, there are some features of the Virtual File Manager implementation that enterprises should consider as key requirements in deploying a Global Namespace.

4.1 Data Directory Services

Similarly to the way that LDAP delivers user authentication services and DNS delivers network-related information services, a Global Namespace provides directory services for networked storage. The Virtual File Manager Global Namespace is a central repository for storage-related information and functions, and provides administrators with a single location to manage network data.

One way to conceptualize the data directory services provided by a Global Namespace is to think about the Yellow Pages. Like the Yellow Pages, a Global Namespace is a directory. In the same way that a listing can appear in multiple sections of the Yellow Pages, a Global Namespace can provide multiple, customized views of the same data set. The Global Namespace concept is often used in the context of clustered file systems. However, a Global Namespace delivered via a clustered file system is like the White Pages. Only one listing can appear for a particular entry in the White Pages, and in clustered file systems, only one view can be provided for a given data set.

4.2 Out-of-Band, Software-Based Solution

Similarly to the way that DNS provides the physical-to-logical translation for Web addresses, enabling clients to talk directly to the Web site, the Global Namespace facilitates the physical-to-logical translation for storage. Because Virtual File Manager does not reside in the data path, it doesn't introduce any performance or latency issues when used to deploy a Global Namespace.

Unlike hardware-based network data management solutions, Virtual File Manager does not require administrators to install and manage any new devices. In addition, high availability of the Global Namespace can be achieved in a cost-effective way, because there is no requirement to purchase additional hardware to fail over the namespace. Finally, Virtual File Manager delivers comprehensive data classification and reporting capabilities that are integrated with its other data management applications, such as consolidation and business continuity.

4.3 Standards-Based Platform

Virtual File Manager is an open, standards-based software platform that can be seamlessly and nondisruptively introduced into an IT infrastructure. The software runs on any industry-standard server running Microsoft Windows 2000 or later, and does not require the deployment of a new hardware device. Because it is not in the data path, Virtual File Manager does not introduce any performance or latency issues when used to create and manage the Global Namespace. It uses the existing file system, which

means that administrators aren't required to change their network operating procedures to enjoy the benefits of a Virtual File Manager Global Namespace.

It also integrates with the existing network security framework, and administrators can even utilize security settings such as group permissions to automatically create and populate a Global Namespace. No software or agents are required to be installed on the client machines accessing the namespace. Unlike many Global Namespace solutions, Virtual File Manager does not require the introduction of a new protocol on the network.

4.4 Not a Proprietary File System

Using Virtual File Manager to implement a Global Namespace does not require any modifications to the existing network infrastructure because it simply resides on top of the existing file system. Therefore, enterprises can continue to benefit from the advantages inherent in their existing file systems (WAFL®, NTFS, VxFS), such as performance, journalizing, point-in-time recovery, encryption, compression, and security.

Virtual File Manager does not require any changes to network operations, such as snapshot and backup processes. It utilizes time-tested CIFS and NFS network protocols that are offered by existing file system vendors (Microsoft, Network Appliance, and EMC), which is a major advantage over solutions that require the introduction of a new protocol. Utilizing the underlying file system enables the Virtual File Manager Global Namespace to deliver significant benefits over aggregation solutions that use a proprietary file system.

4.5 Platform for Unlimited Scalability

There is no limit to the scalability of a Global Namespace implemented with Virtual File Manager. Administrators can use the Virtual File Manager Global Namespace to aggregate multiple file systems and manage them as a single entity. This enables administrators to overcome the scalability limitations of individual file systems and to manage tens of thousands of directories and trees through a single namespace.

4.6 Complete Namespace Manageability

Virtual File Manager provides administrators with a comprehensive tool for Global Namespace creation and management. It enables administrators to create and manage multiple namespaces through a single, intuitive console and provides a means to manage both the namespace and the underlying file system from the same console. Administrators can also dynamically populate the Global Namespace based on existing shares/export naming conventions or based on security in the enterprise.

Virtual File Manager also provides administrators the ability to monitor, scale, increase availability, audit, back up, restore, and make snapshot copies of the Global Namespace. This set of features delivers complete manageability of the Global Namespace and enables administrators to scale their deployments from a simple Global Namespace to a complex group of functional, enterprisewide namespaces.

4.7 Simple to Install and Use

Deploying a Global Namespace using Virtual File Manager is a simple process. It takes only minutes to install the software, and less than an hour to create and populate a Global Namespace. Using Virtual File Manager, an administrator can deploy a namespace and begin receiving its benefits the same day. Administrators can even use permissions established via Active Directory to automatically create and manage Global Namespaces, thereby applying the established network security framework to the new namespace.

5. Summary

Virtual File Manager builds on the Global Namespace foundation to deliver a comprehensive set of integrated network data management applications that enable enterprises to significantly reduce the cost of adding, changing, and managing storage. With a Global Namespace in place, administrators can move and manage data in a logical rather than a physical way, and are provided with a long-term solution to the growing problem of storage device proliferation. Implementing a Global Namespace enables enterprises to build a scalable, reliable storage infrastructure, dramatically simplifying data management.

For more information, visit <http://www.netapp.com/products/software/vfm.html>.



www.netapp.com

© 2006 Network Appliance, Inc. All rights reserved. Specifications subject to change without notice. NetApp, the Network Appliance logo, and WAFL are registered trademarks and Network Appliance and Virtual File Manager are trademarks of Network Appliance, Inc. in the U.S. and other countries. Microsoft and Windows are registered trademarks of Microsoft Corporation. Linux is a registered trademark of Linus Torvalds. UNIX is a registered trademark of The Open Group. All other brands or products are trademarks or registered trademarks of their respective holders and should be treated as such.