



Virtual File Manager™ Executive Overview

Enabling Comprehensive Network Data Management Services

September 2006 | TR-3513

Abstract

This document provides an overview of how Virtual File Manager can be used to deploy a Global Namespace to deliver comprehensive network data management services. The Virtual File Manager applications are described, and information is provided on how the Virtual File Manager Global Namespace differs from other namespace solutions available today.

Table of Contents

1. Introduction	3
2. What Is a Global Namespace?	4
3. Global Namespace Benefits	4
4. Global Namespace Applications	4
4.1 Nondisruptive Data Migration and Consolidation	5
4.2 Cost-Effective Business Continuity	5
4.3 Data Lifecycle Management	6
4.4 Storage Optimization	6
4.5 Consolidated Network Data Management	7
4.6 Remote Site Data Management	7
4.7 Data Classification and Reporting	8
5. Virtual File Manager Global Namespace Differentiators	8
5.1 Data Directory Services	8
5.2 Out-of-Band, Software -Based Solution	9
5.3 Standards-Based Platform	9
5.4 Not a Proprietary File System	9
5.5 Platform for Unlimited Scalability	10
5.6 Complete Namespace Managability	10
5.7 Simple to Install and Use	10
6. Summary	10

1. Introduction

As the industry's leading network data management platform, Virtual File Manager offers an integrated set of applications that logically aggregate distributed file storage. Virtual File Manager provides administrators with policies that automate data and storage services, such as consolidated network data management, data migration and consolidation, business continuity, storage optimization, data lifecycle management, remote site data management, and data classification and reporting (Figure 1).

At the core of Virtual File Manager is the Global Namespace, a logical representation of file system and storage devices that creates a unified view of data distributed across heterogeneous storage platforms. The combination of Virtual File Manager policies with the Global Namespace provides administrators with the ability to perform storage management tasks with no downtime and no disruption to users.

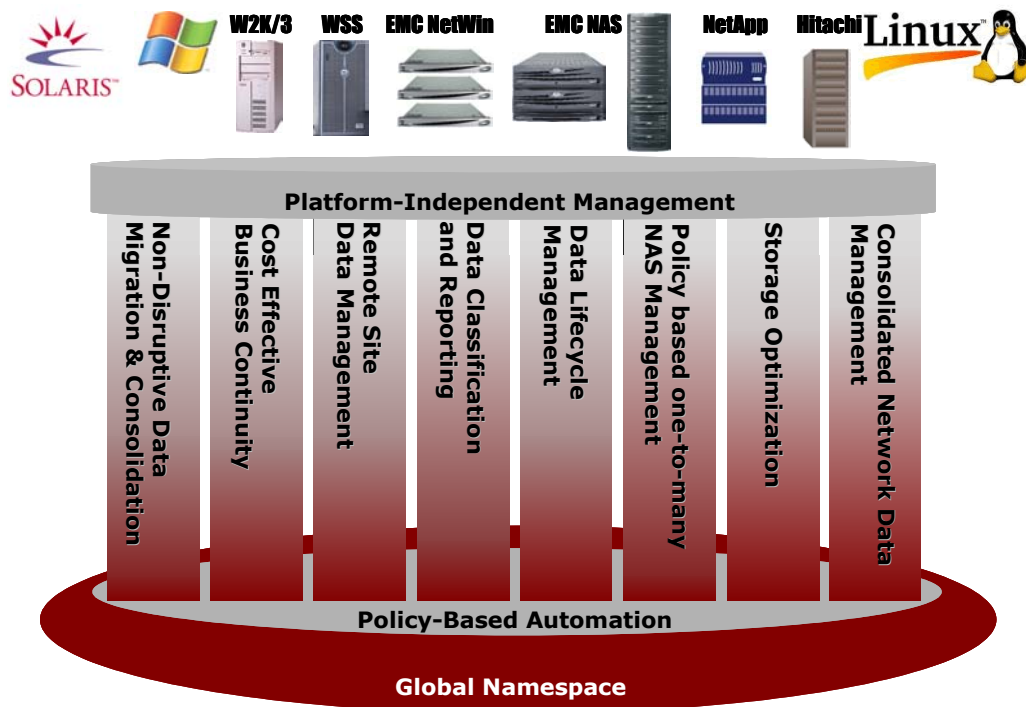


Figure 1) The Virtual File Manager Global Namespace overview.

2. What Is a Global Namespace?

Global Namespace does for files what DNS does for networking—it provides a directory service. This directory service can deliver location-independent services to users and applications across multiple, heterogeneous, distributed file systems. Once implemented, a Global Namespace enables users to access files in a logical, location-independent way, similar to accessing Web pages on the Internet. For example, when you type `www.yahoo.com`, you don't know the IP address for Yahoo, and you don't care. Similarly, with a Global Namespace in place, the user accesses information using Windows® Explorer, and the location of files is transparent.

With a Global Namespace, the administrator can distribute files in a way that optimizes performance and capacity utilization; clients then access the files via the namespace. When storage is added or consolidated and files are moved or renamed, clients are automatically redirected to the new location, often unaware that the data was moved. This approach dramatically simplifies network data management, because it permanently eliminates the need for desktop reconfiguration, drive letter remapping, and modification of login scripts when storage is reconfigured. A Global Namespace can scale to support an entire corporate environment.

3. Global Namespace Benefits

- Administrators can expand, move, rebalance, and reconfigure storage without affecting how users view and access it, and without downtime.
- Data management and data movement require far less physical administration, and these tasks are performed in less than half the time they took before.
- Administrators can manage data on heterogeneous, geographically distributed storage devices through a single console.
- A Global Namespace provides a platform for unlimited storage scalability—administrators can use a Global Namespace to aggregate multiple file systems and manage them collectively.
- Data changes are automatically updated in the namespace, and require no client reconfiguration or downtime.
- Data management and data movement are performed “behind the veil” of the namespace, and are transparent to users.
- Data can be organized and presented to users in a way that makes sense to them, regardless of how and where the data is stored.
- Users access all their files through a single drive letter, allowing them to continue accessing files in a familiar way.

4. Global Namespace Applications

Once the logical layer is in place, many data management tasks that were either impossible or cumbersome to achieve become simple. Administrators can add, move, rebalance, and reconfigure physical storage without affecting how users view and access it. This section discusses some of the key applications that a Global Namespace created and deployed using Virtual File Manager can deliver.

4.1 Nondisruptive Data Migration and Consolidation

Data migration and consolidation is no longer a point-in-time IT project for administrators. Rapid changes in hardware, technology, and capacity demands are forcing enterprises to implement ongoing data migration strategies to optimize available storage assets. However, consolidation presents many challenges for IT administrators. These challenges include minimizing client downtime, supporting ongoing data migration practices with minimal IT resources, consolidating data in a heterogeneous environment, and maximizing ROI by effectively utilizing storage resources across the enterprise.

Using a Global Namespace to facilitate data migration can result in significant time and cost savings. Enterprises that need to migrate from Microsoft® Windows NT4 to Windows 2000/2003, Novell to Windows, Linux® to Solaris®, or Windows or UNIX® file storage to NAS should consider deploying a Global Namespace to facilitate the migration. A Global Namespace creates a virtual pool of storage that enables administrators to readily identify migration requirements across platforms. In addition, once-manual processes—such as share creation, copy, and delete—can be automated to reduce errors. When used as a migration and consolidation tool, Virtual File Manager delivers many benefits in addition to Global Namespace.

Virtual File Manager Advantage

- Administrators can seamlessly migrate and consolidate data “behind the veil” of the namespace, without interrupting user access.
- A simple, intuitive GUI makes complex consolidation jobs a “drag-and-drop” operation.
- Data movement tasks can be scheduled and automated using Virtual File Manager policies.
- Virtual File Manager enables administrators to complete migration jobs faster (takes less than 50 % time).
- Virtual File Manager aggregates storage across heterogeneous, geographically distributed file systems, enabling administrators to perform data migration across devices, regardless of vendor or underlying file systems.

4.2 Cost-Effective Business Continuity

Business continuity planning has come to encompass much more than disaster recovery. Enterprises are faced with network outages, both planned and unplanned, for a variety of reasons. A number of products on the market today address the need to replicate data in preparation for a network outage, but very few address the need to fail over users quickly and seamlessly. In addition to ensuring user failover, administrators must find ways to centralize business continuity management in heterogeneous, distributed environments while keeping IT costs at a minimum.

Global Namespace provides cost-effective failover across geographically distributed sites, while giving administrators a centralized way to manage the global failover process. Because administrators can manage storage on a logical rather than a physical basis with a Global Namespace, the failover process can be completely transparent to users. Virtual File Manager provides administrators a simple yet robust platform for implementing business continuity practices, and adds significant benefits to the Global Namespace solution.

Virtual File Manager Advantage

- Instantly fail over users across different devices and/or locations
- One-step near-line storage integration for replication and failover
- Automated, policy-based setup and execution of business continuity processes

- Thousands of users can be failed over in one automatic step
- Simple, intuitive GUI makes business continuity planning and execution simple for administrators

4.3 Data Lifecycle Management

A look at the management requirements of reference data effectively illustrates data lifecycle management requirements. Reference data is information that is fixed, but that needs to be kept available for regulatory, informational, or business purposes. Examples of reference data include historical financial statements and patients' medical imaging files. This type of data produces very large file sizes, or a multitude of smaller files that take up a lot of space. For example, a single medical image file, such as an MRI, generates 20,000 individual files. These files are required to be online and available at all times, which can be costly to maintain on primary, high-performance storage systems.

Implementing a tiered storage architecture is an efficient approach to meeting data lifecycle management needs. A Global Namespace enables the implementation of a tiered storage architecture by transparently facilitating the movement of reference data from expensive primary storage onto less-expensive secondary storage based on administrator-defined criteria. In the example of medical imaging files, the hospital IT administrator could use the Global Namespace to automatically and transparently migrate medical imaging files to secondary storage 30 days after the patient's initial visit. Moving data to secondary storage not only saves hardware acquisition costs, it also enables administrators to match backup policies and investment with the business value of data. Using Virtual File Manager to create and implement a tiered storage architecture, administrators can create policies to automatically migrate files based on criteria such as age, usage, and size, among others.

Virtual File Manager Advantage

- User access to data is not affected by migrating data to other tiers of storage, because data migration is transparent and the Global Namespace is automatically updated when data is moved.
- Comprehensive data classification and reporting capabilities enable administrators to automatically classify data based on any file, volume, or directory attribute.
- Administrators can create one simple policy to drive data movement and retention practices.
- Lifecycle history reports are generated automatically, allowing administrators to determine where and when data was moved.

4.4 Storage Optimization

The explosive increase in unstructured data has led to highly distributed, unbalanced growth of servers and storage devices in most IT infrastructures. As a result, it is estimated that approximately 65% to 75% of the total storage capacity in most enterprises is not utilized. Administrators face the challenge of identifying underutilized storage devices in complex environments and load-balancing capacity across all available resources.

With a Global Namespace, administrators can create logical "pools" of storage that align with business operations, then manage and optimize storage resources assigned to the namespace. Administrators can centrally view and manage enterprisewide capacity utilization using a Global Namespace and transparently and nondisruptively balance capacity across multiple devices. In this way, Virtual File Manager provides administrators with a powerful tool for optimizing capacity utilization across the enterprise.

Virtual File Manager Advantage

- Automated, policy-based setup and execution of load balancing allows unattended storage optimization.
- Users remain unaware of data migration activities, because the Global Namespace is automatically updated to reflect data migrations transparently.
- Multiple policy settings allow administrators flexibility in choosing maximum and minimum capacity thresholds for storage devices.

4.5 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

4.6 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 the data 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%

4.7 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.

5. 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.

5.1 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.

5.2 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.

5.3 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.

5.4 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.

5.5 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.

5.6 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 a snapshot copy 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.

5.7 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.

6. 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>.



© 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. Linux is a registered trademark of Linus Torvalds. Microsoft and Windows are registered trademarks of Microsoft Corporation. Solaris is a trademark of Sun Microsystems, Inc. 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.