WHITEPAPER

# File Server Consolidation Using Data Dynamics StorageX

File server consolidation improves utilization of storage resources and simplifies manageability of file servers. Using Data Dynamics StorageX reduces the complexity of file server consolidation and enables effective enterprise-wide file storage management.



DATA DYNAMICS

# Contents

## Executive Summary

The proliferation of application files and user-generated documents is driving continued expansion of network storage requirements within enterprise organizations. A common solution to this problem has been to add more file servers and storage, a strategy that may work in the short term. But as time progresses, the cost and complexity of managing a large number of file servers outweigh the initial benefits. Issues that may arise are inefficient storage utilization and increased complexity in the ongoing management of user-to-file-share mappings and backup/restore operations.

Implementing a file server consolidation solution using Data Dynamics StorageX can address these issues and provide organizations with tools to manage file server resources efficiently. Utilizing a Global Namespace (GNS) to provide a single, organized view of all network file data, StorageX allows administrators to create automated, policy-driven data migration routines that simplify the consolidation process and optimize storage resources. GNS is an abstraction layer between the logical and the physical layers, which removes the dependency between them. This strategy solves the immediate problem and anticipates similar issues that may appear as the consolidated file servers approach capacity once again.

This paper describes a Windows file server consolidation solution using StorageX and is intended for IT architects, IT administrators, and the technical staff responsible for IT infrastructure and storage management. It explains how to:

- Create a GNS and thereby virtualize user access to network shares.
- Consolidate data to reduce the number of file servers and Network-Attached Storage (NAS) devices.

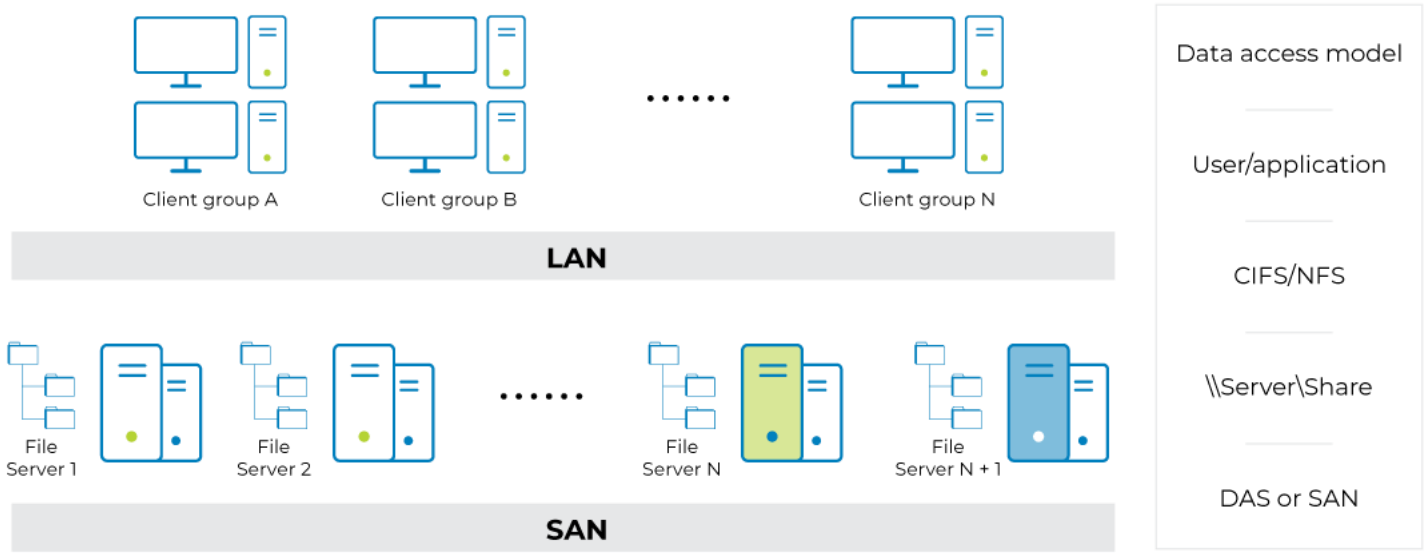
## Reference Architecture

Enterprises today have file shares with large amounts of unstructured file data, on multiple file servers or NAS devices across the IT infrastructure (in this document both file servers and NAS devices are referred to as" file servers"). However, the incentive for using file shares goes beyond simply facilitating application and file sharing among users and user groups. The IT group needs to ensure that corporate data is centrally located and that a backup policy is deployed in accordance with company data protection policies and legal requirements. The proliferation of business-critical data residing on file shares increases demand for additional file share capacity—directly increasing administration and maintenance costs.

A practical approach is to simply add new file servers when capacity is reached on the existing ones. However, since the physical placement of the files mandates the logical access scheme for file shares, this approach will need to accommodate the inflexible ties between the physical placement of files and the logical access path. Adding new file servers requires the creation of new shares on the new file servers or redistribution of shares, resulting in changes to end users mappings. This in turn adds complexity and increases cost over time as management and maintenance tasks become more time consuming. Experience demonstrates that simply adding more file servers results in inefficient resource utilization across the enterprise, while the complexity of managing the infrastructure and the user access scheme is increased considerably.

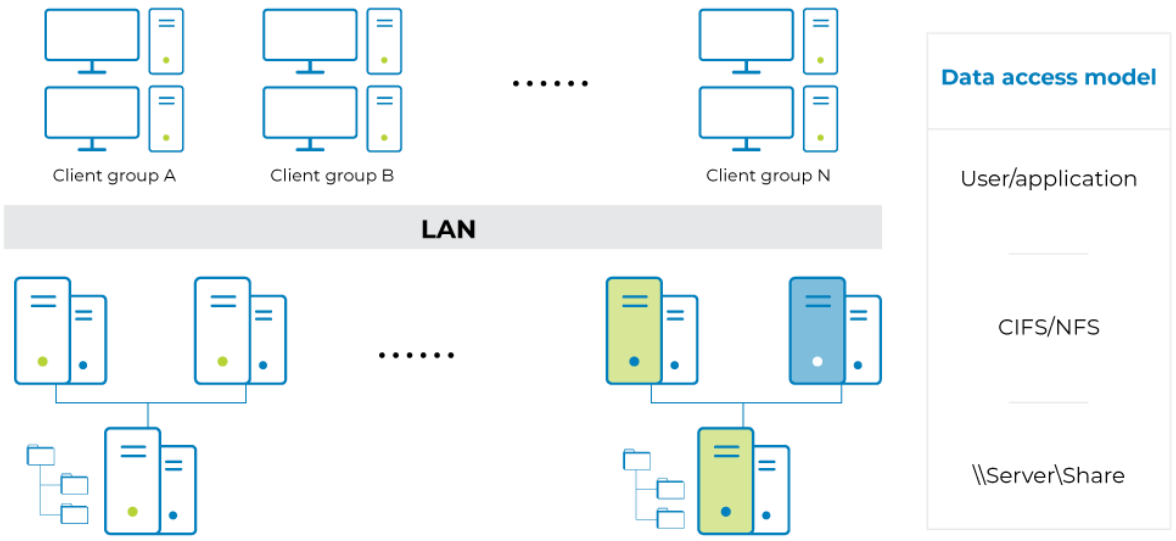
## Adding File Servers

As shown in the following figure, end users are divided into client groups (often reflecting their organizational affiliation) with their file shares mapped according to the client group and business functions. As the demand for capacity increases, new file servers are added and users are mapped to new shares on new file servers. This is a relatively easily implemented solution, but as more and more file servers are added, it becomes evident that this is a point solution that solves the immediate problem only, and as the complexity increases new problems will arise.




## Consolidating File Servers

Adding more file servers requires data relocation and remapping, accomplished by modifying and maintaining complex user logon scripts; and as a result resource utilization, such as storage capacity, floor space, power, cooling, and administration, decreases. The next step is file server consolidation, which consolidates data on fewer (scalable) large capacity file servers. The process of consolidating file servers can be quite complex and it requires a meticulous analysis prior to execution as well as a subsequent verification phase. It is necessary to ensure that file security is preserved and data access for the various user groups is intact. The consolidation requires planned downtime and, depending on the amount of data, it may be so time consuming that it cannot be performed within an acceptable (for example, a weekend) maintenance window. Consolidation can have significant impact on the production environment.

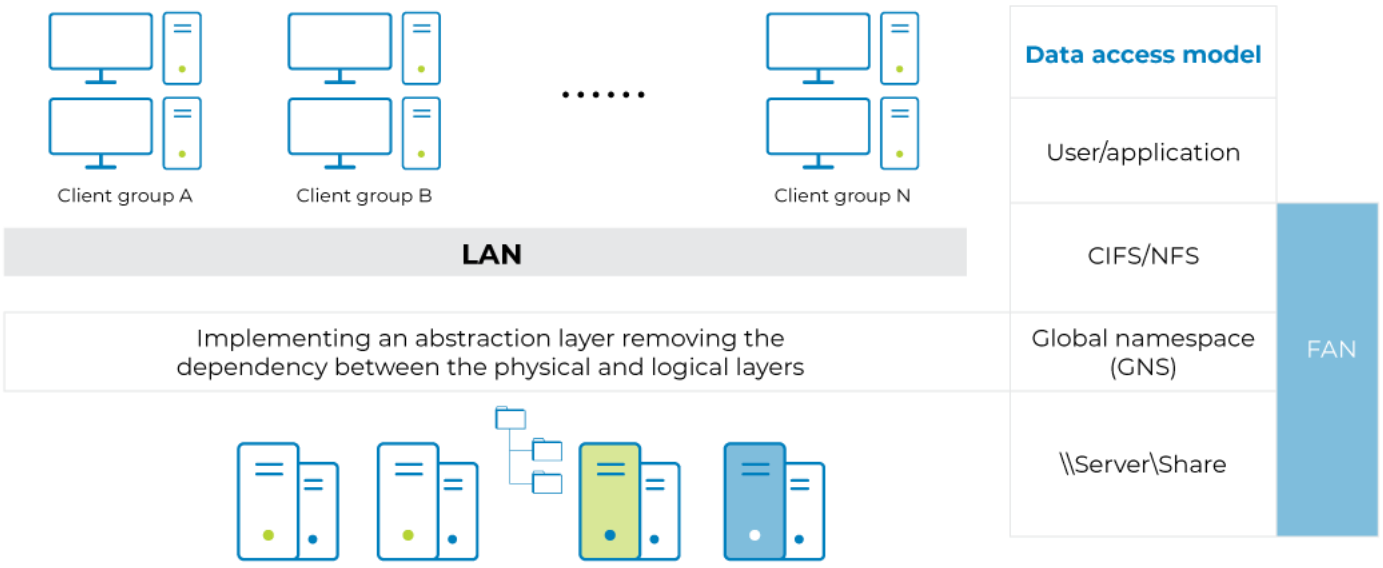


File servers are consolidated into fewer large file servers and the problem is solved temporarily; however, the inflexible relationship between the physical and logical layers persists. So when space on the consolidated file servers starts filling up, the same capacity problems will arise. Typically uneven utilization of the file servers occurs, and from time to time capacity is reached on a particular file server. A common problem is this: a bottleneck occurs when the amount of stored file data combines with the throughput from a particular file server (performing backups) makes it impossible to back up the data within the backup window. This can occur with a Storage Area Network (SAN)-attached file server on which storage capacity is scalable and continuously added, while other resources on the file server are not. Adding more and larger file servers does not solve the original problem. Even worse, more and larger file servers increases the time required to perform the next consolidation—which in turn impacts downtime in the production environment as well.

## Virtualizing File Servers

To solve the issue of periodically having to perform more and more complex file server consolidation with larger amounts of data, you must remove the dependency between the physical and the logical layer for shared file access. By doing so, the infrastructure becomes much more flexible, decreasing the need for maintenance downtime for performing changes in the physical layer, since these tasks can be performed non-disruptively and without impact to production. In addition, this must be done while still adhering to good infrastructure design practices observing Reliability, Availability, and Serviceability (RAS), scalability, and manageability principles and not precluding further growth of the infrastructure and integration of new components.

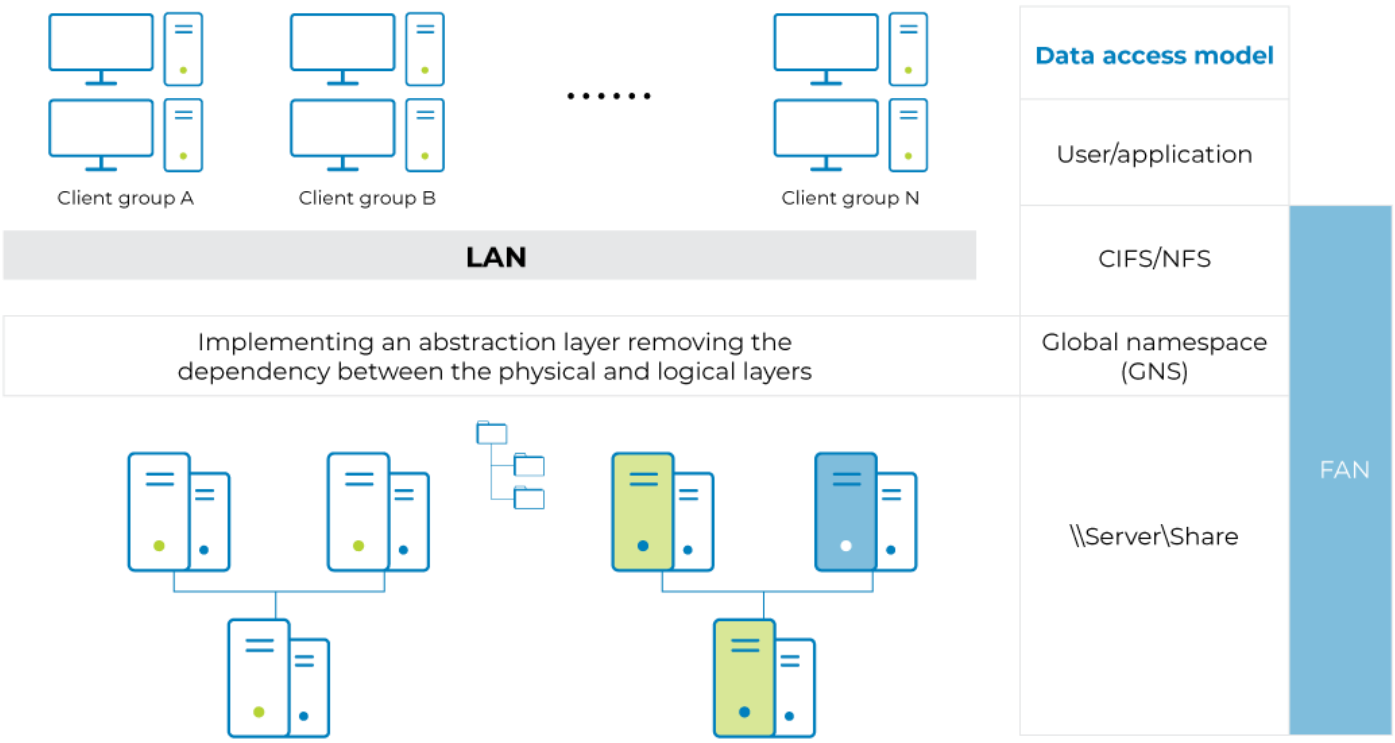
A best practice when implementing StorageX is to create a global namespace, as illustrated in the following figure. The file servers are now virtualized for the end user, and there is an enterprise-wide StorageX interface for IT administrators. A benefit of carefully designing and implementing the GNS is that the directory structure becomes intuitive to end users and helps them navigate within the shared file space.




You can architect this File Area Network (FAN) using StorageX, an integrated suite of tools that logically aggregates distributed file data across heterogeneous storage environments and across CIFS- and NFS-based file systems, while providing policies to automate data management functions.

## Consolidation in a Virtualized File Server Infrastructure

Once StorageX is deployed and the GNS is created, you can perform file server consolidation. StorageX facilitates the consolidation and relocation of file shares non-disruptively and with limited impact on production. As shown in the following figure, file servers are consolidated without changing the directory structure.




StorageX is an open, standards-based application that can be seamlessly introduced into an IT infrastructure. It uses existing file systems, which means that administrators are not required to change their network operating procedures to leverage the benefits of StorageX. It integrates into the existing network security framework and administrators can use security settings such as group permissions to automatically create and populate a GNS. No software or agents are required on the client machines accessing the namespace; and unlike many global namespace solutions, StorageX does not require the introduction of a new protocol on the network.

The remainder of this document will focus on how file server consolidation is facilitated with StorageX to perform advanced enterprise-wide file management such as policy-based file placement and file lifecycle management in the FAN.

## Guidelines

This section provides high-level guidelines for characterizing, designing, deploying, and managing file server consolidation solutions using Data Dynamics StorageX and Data Dynamics UNCUpdate.

Data Dynamics UNCUpdate provides storage administrators with a reliable tool for reporting on files that contain Universal Naming Convention (UNC) references to storage resources under consideration for migration. UNCUpdate can report on UNC entries encountered in a wide variety of files, providing administrators with the option to update UNC references to target the storage resource at its new location.

## Server Consolidation Assessment

In theory, the objectives for consolidation projects are to gain better resource utilization and to align the IT infrastructure with the long-term strategy for the data center. In reality, often there are other drivers that determine when to perform IT infrastructure changes such as data consolidation:

- To replace file servers that are end of life or coming off lease
- To optimize utilization of storage resources if you are running out of capacity
- To reduce the physical footprint and power consumption in the data center to reduce cost
- To improve storage management (fewer administrators managing larger amounts of data and a greater number of users)
- To reduce the time required for backups and avoid increasing the backup window

As the IT infrastructure grows, at some point you need to re-evaluate the current infrastructure and ensure that it is working efficiently. In addition to identifying and resolving current pain points, you should also try and address predictable future needs. Deploying StorageX can significantly reduce time in the analysis phase, enabling faster and safer server consolidation. The virtualization of file space when you implement a GNS becomes the foundation for efficient and ongoing space management supporting long-term IT infrastructure and business needs. This includes advanced file management based on defined policies, file lifecycle management, as well as the ability to leverage dynamic changes in the physical infrastructure when necessary.
Administrators should calculate the costs of maintaining and operating the current file services infrastructure—taking into account its growth rate and growing complexity—and compare that to the estimated cost of maintaining a consolidated solution with StorageX.

## File Server Consolidation Using StorageX

The primary activities to perform with StorageX in a file server consolidation solution are:

1. Design the global namespace and consolidated environment.
2. Implement the global namespace.
3. Migrate data and update the global namespace.

**NOTE:** Although it is possible to use StorageX for migration without implementing a GNS, it is strongly recommended that you implement a GNS. If your time line demands migration and consolidation be performed as soon as possible without implementing a GNS, you can use a consolidated Distributed File System (DFS) root server. However, you should subsequently implement a domain-based GNS. For more information contact your representative.

## Designing the global namespace

Design the global namespace directory structure such that it both supports the needs of the organization and is intuitive to users. Consider organizing your namespace based on departments or perhaps based on geographic locations. Often you can use the existing Active Directory (AD) Organizational Unit (OU) structure as the foundation for a GNS design. Use the StorageX reporting feature to gather data and create an overview of the file server environment. In the planning phase, an overview can help you develop the namespace design, which you can then further refine so that the logical access model meets the requirements of the organization.

**NOTE:** Administrators should be familiar with Microsoft Distributed File System (DFS) planning, requirements, and best practices before implementing a global namespace.

StorageX provides policies to automate namespace population. Administrators can use these policies to search for servers and shares that match a particular text string, and then use permission-based policies to generate namespaces based on group permissions. After the namespace is in place, policies that scan for changes in the physical resources supporting the logical namespace can be run automatically on a scheduled basis. You can configure policies to report changes or to update the logical namespace automatically to reflect the changes detected by the policy.

We recommend that you use domain-based DFS roots (Active Directory integrated) to observe High Availability (HA), and use your Domain Name System (DNS) server deployment as a guideline for where to place DFS root servers. If your environment is distributed over multiple sites, wherever there is a DNS server plan to place a DFS root server.

If your environment prohibits using domain-based DFS roots (if there are existing Windows 2000 Domain Controllers (DCs) or in a Windows NT environment), you must use a standalone DFS root. We recommend implementing the standalone root on a Microsoft Cluster Server for HA. When you use a standalone DFS root and the root server becomes unavailable, then the DFS is unavailable until the root server is available again.

## Implementing the Global Namespace

Having designed the GNS, deploy it using the StorageX Namespace Creation wizard, to structure and deploy the global namespace. While the GNS is based on Microsoft DFS it depends on the Microsoft DNS for lookup of file servers, so a consistent deployment of DNS must be present prior to implementing GNS.

## Migrate Data and Update the Global Namespace

When planning data migration, start by characterizing the existing file services environment and determining what can be consolidated. Identify available resources, evaluate storage utilization, and estimate the cost of maintaining the environment. StorageX reports are a powerful tool in the assessment phase. From a central point of administration, IT administrators can discover storage resources and display information such as disk space usage, share permissions, and age of files to determine data migration requirements.

Consider these factors for the server configuration: the new file servers must meet performance requirements and have sufficient storage capacity to consolidate data from the source servers and still have room to grow. Moving storage to a SAN environment can assist in this process. The newly deployed servers should also meet the HA requirements of the organization. This could mean server clustering or DFS root/link duplicates in a domain-based DFS as well as developing a plan for disaster recovery.

Other considerations include:

- Administrators must determine the impact to application uptime and users—and how to resolve or message downtime to the user community.
- Older versions of Windows operating systems may need to be upgraded or patched in order to support DFS.
- If source file servers are performing other functions in addition to serving file shares, these functions must be examined to determine how they may be affected by the migration.
- Once designed, the new server configuration should be brought online and verified in a test environment.
- If no issues arise during the testing, the servers can be deployed into the production environment in preparation for the data migration.

Prior to consolidation, run Data Dynamics UNCUpdate in scan mode on each individual share in order to identify which shares, directories, and file types may have embedded links that will require modification once the namespace is implemented. Data Dynamics UNCUpdate will generate a report that can be used in the document update process.

## Workflow and Examples

This section describes how to implement and leverage StorageX to perform file server consolidation—the activities to be performed and use case examples.

## Global Namespace Design

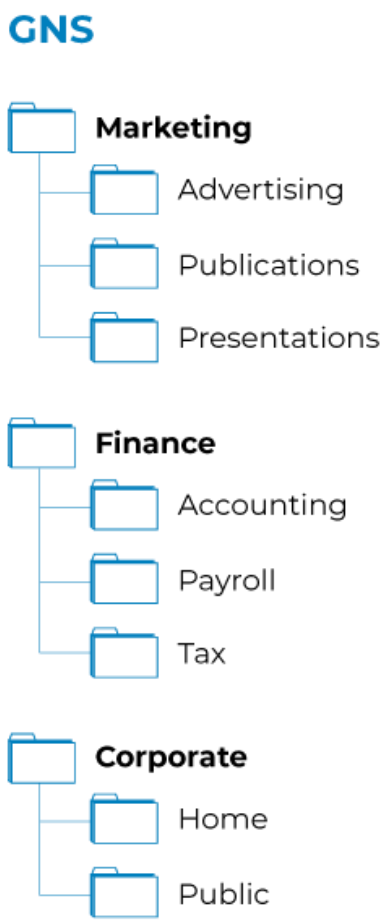
To perform the actual global namespace design, a common approach is to use the organizational structure and logically order folders belonging to a specific organizational group or function. This approach is probably not very different from how the shared file space was implemented initially. And then growth may have forced the logical groups to be split and folders located where capacity was available.

**Example 1. Designing the namespace**
In this example the GNS reflects the organizational affiliation shown below for the users in the Marketing and the Finance department. The users share mappings are:

**Marketing**
E:\ Advertising [\\FileSrv4\Advertising]
F:\ Publications [\\FileSrv2\Publications]
G:\ Presentations [\\FileSrv3\Presentations]
H:\ Home [\\FileSrv1\Home]
P:\ Public [\\FileSrv4\Public]

**Finance**
E:\ Accounting [\\FileSrv1\Accounting]
F:\ Payroll [\\FileSrv3\Payroll]
G:\ Tax [\\FileSrv1\Tax]
H:\ Home [\\FileSrv2\Home2]
P:\ Public [\\FileSrv4\Public]

The goal is to make the access scheme intuitive and clear for the user. In order to do so, the GNS design is based on the file shares used by the different user groups in the organization. As shown in the following figure, file shares used by marketing are logically grouped together, file shares used by finance are logically grouped together, and corporate shares (which apply more widely across user groups) are logically grouped together.



## Global Namespace Creation

Based on the GNS design, the GNS is now deployed in the file services infrastructure. As part of the GNS design phase, determine what type of DFS root and how many root servers to use depending on the environment.

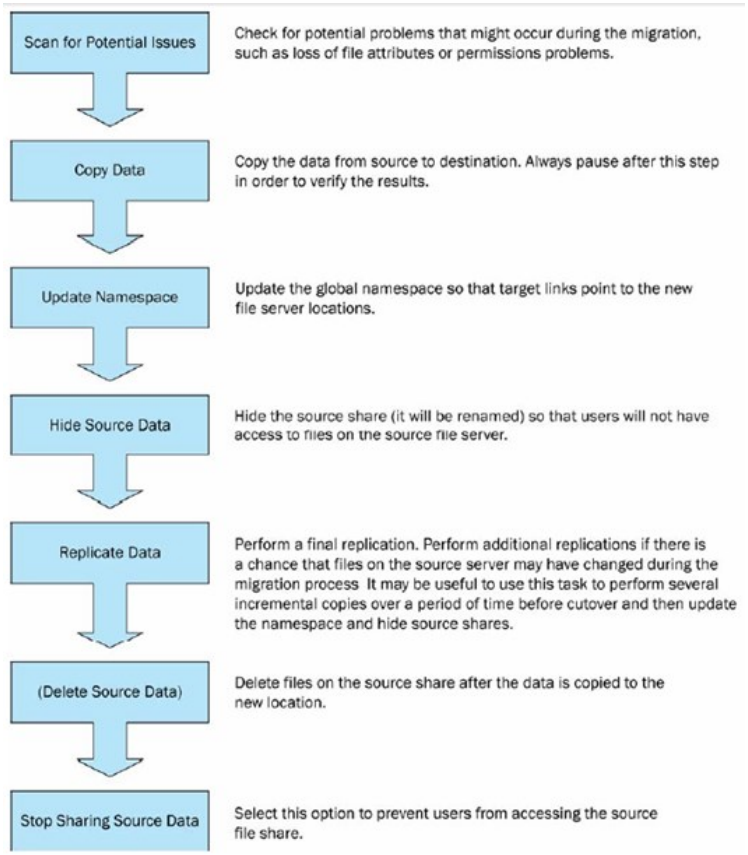
## Building the Global Namespace

Once the namespace layout and type of DFS root are determined, perform the following tasks to build the namespace:

1. Choose which servers to host the DFS root and build the namespace using the StorageX graphical interface. The DFS links should point to the original source server shares.
2. If network home directories are part of the consolidation process or if it is desirable to have users' home folders in the DFS namespace, create and execute a StorageX home folder policy. This will modify users' home folder paths in the Active Directory to point to its virtualized location.
3. Notify users about new namespace implementation. Create a StorageX namespace report and provide it to users in order to explain and clarify how the new changes may affect them.
4. Update user login scripts and drive mappings to point to the global namespace. Perform this task in a maintenance window and preferably roll it out on a department-, location-, or domain-based schedule in order to minimize risk. The next time users log in, they will be directed to the virtualized namespace paths.
5. Update any procedures or software that automatically creates embedded links in documents; for example, update spreadsheet macros so that these links use the new GNS path.
6. Run Data Dynamics UNCUpdate in modify mode against shares that may have files with embedded links. If UNCUpdate was previously run in scan mode, use the XML file created to reduce the actual runtime when updating the embedded links.

## File Migration

After the namespace has been deployed, administrators can begin the migration process. A typical migration involves the tasks shown in the following figure, all of which can be automated using StorageX.



**NOTE:** A number of replication and security options are available to optimize data transfer and control how file security is handled during the migration. See the StorageX online help for details.

StorageX uses a replication agent to migrate data from one system to another. Whenever the migration process begins, StorageX attempts to deploy the replication agent onto the machine hosting the destination target. If the agent cannot be installed on this machine, it attempts to deploy the agent onto to the machine hosting the source target. Using the source or destination machine is preferred, but if for some reason, the agent cannot be installed on either system, StorageX uses a proxy agent to copy the data from the source to the destination. If a proxy agent is not specified in the host properties window---the replication agent on the StorageX server will be used.

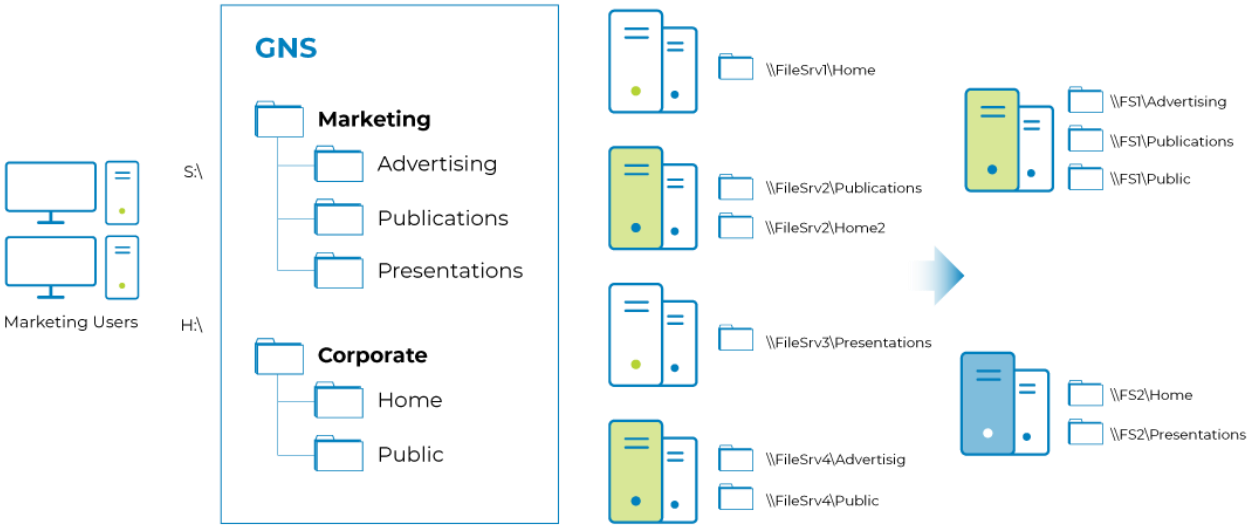
Virus scanning software can slow replication, and in some cases, can cause locked file errors during the transfer process. Disable virus scanning software to speed up migration, but do this only if you have verified that your environment is virus free and protected. You should also assess the probability that files on the source file shares are locked:

- Files such as Outlook PST files are typically locked while Outlook is running.
- MS Access files are exclusively locked when open.
- MS Word and Excel documents are locked in read-only mode to the second accessing process.

The StorageX snapshot replication option integrates with Windows 2003 VSS snapshots and NetApp file server snapshots to allow replication of locked files.

Because server name dependencies are eliminated with the implementation of GNS, it becomes possible to add, expand, and move data transparently without disrupting user access to files. The storage administrator can now perform a server consolidation without impacting users or applications. Use a StorageX migration policy to migrate data based on administration requirements, and then update the global namespace to reflect the newly consolidated environment.

**Example 2. Performing the consolidation**
In the following figure, file servers are consolidated (the Marketing group only is shown). The file shares placed on FileSrv 1 - 4 are migrated to FS1 and FS2 without impacting user file access.




It is important to emphasize that file migration can have a significant impact on applications and documents that existed prior to the implementation of the GNS. For example, if users have embedded UNC pathnames in documents or shortcuts to files that have moved to a new location, administrators must take measures to ensure that links are updated as part of the consolidation process. Use Data Dynamics UNCUpdate to automatically discover and modify shortcuts, embedded links, and UNC references user documents. Once applications and documents with embedded links point to UNC paths Through the GNS, further modification is unnecessary even after future migration.

## Creating and Executing the Migration Policy

The procedure below details the high-level tasks required to perform a Windows-to-Windows migration. A number of policy actions not described below can be altered depending on file server platform and user or corporate policy requirements. See StorageX online help for further information.

1. Create a migration policy for the file shares to be moved to the newly provisioned servers and select the migration actions to be performed. The administrator can choose to run actions one after the other and pause after each completed action, or pause only if an error occurs. A number of replication and security options are available to optimize data transfer and to control how file security is handled during the migration.
2. Create tasks to migrate source shares to the destination server. This process includes specifying the source shares UNC path and destination UNC path. Shares on the destination server can be created at this time if they do not exist already.
3. Schedule a time to run the migration or manually execute the policy to begin the process. Assuming no errors were discovered, this will create a baseline copy of the data. You can now perform additional incremental copies if you wish by restarting the migration task.
4. Schedule a time to update namespace and re-enable the migration. The policy will update the target links in the DFS namespace, hide the source share, and perform a final replication.
5. Allow enough time for end users to provide feedback. If file access problems occur, resolve them and re-enable the migration policy. Shares and data will be removed from the source file servers.

## Post-Migration Tasks

Perform a verification test to ensure that the new environment is operational, that all file data is available to users as expected, and that the desired backup policy is applied (and works) at the new file location. Once testing is complete, the source file servers can be decommissioned or repurposed for other use.

Migration is now complete. A best practice is to create a StorageX application snapshot schedule, which backs up the StorageX database and associated files to protect against corruption or loss of application functionality. Policies include disaster recovery, share monitoring, policy-based data movement, and other events scheduled and implemented with StorageX. For most sites, set this to once a day and keep the first 30 versions. After the first month, you could move to a weekly backup, which keeps 12 versions (or 3 months worth of data). In most instances it is not necessary to enable the “include reporting database information in snapshot,” option as this increase the size of the application snapshot.

**NOTE:** If the StorageX application service and/or database become unavailable, the DFS namespace managed by StorageX is not impacted.

## Procedures Overview

This section describes the most common problems encountered when carrying out file server consolidation with StorageX, as well as how to anticipate and resolve them. While they may not apply to your environment, we recommend that you review them.

## Verifying Hostname Resolution

- Always perform a hostname resolution verification test prior to implementing GNS. Because GNS relies on DNS lookup, it is important to verify that both forward and reverse lookup for all file servers are consistent.
- For clients, verify that the Dynamic Host Configuration Protocol (DHCP) scopes correctly point to the DNS servers.
- If you have special users, for example mobile users that access the network differently, you must verify name resolution from their laptops as well.

## Access Permissions for StorageX

- When using StorageX, you must give the service account permission to update the active directory for the DFS root as well as full control on DFS root share and subfolders.
- To ensure the migration of files is not complicated by, for example, users removing permissions for Local Administrators from their home directory, add the StorageX service user account to both the user groups Local Administrators, and Backup Operators groups on all file servers and the StorageX server.

## Creating the GNS

- When you create the domain-based root design, create the root first as a standalone DFS root and then do all the modifications against the standalone root. Constantly making modifications on a domain-based root adversely impacts performance. When the namespace is fully created, then back it up with StorageX, delete the standalone root, create a new domain-based root, and restore the backup to the domain-based root.

## Handling Identical Share Names

- It is common to have coincident share names in the existing file space environment. When consolidating from multiple file servers, consider how to handle identical share names in situations in which files on the shares cannot be consolidated into one directory or when file shares must be preserved and placed at different (logical) locations in the GNS.