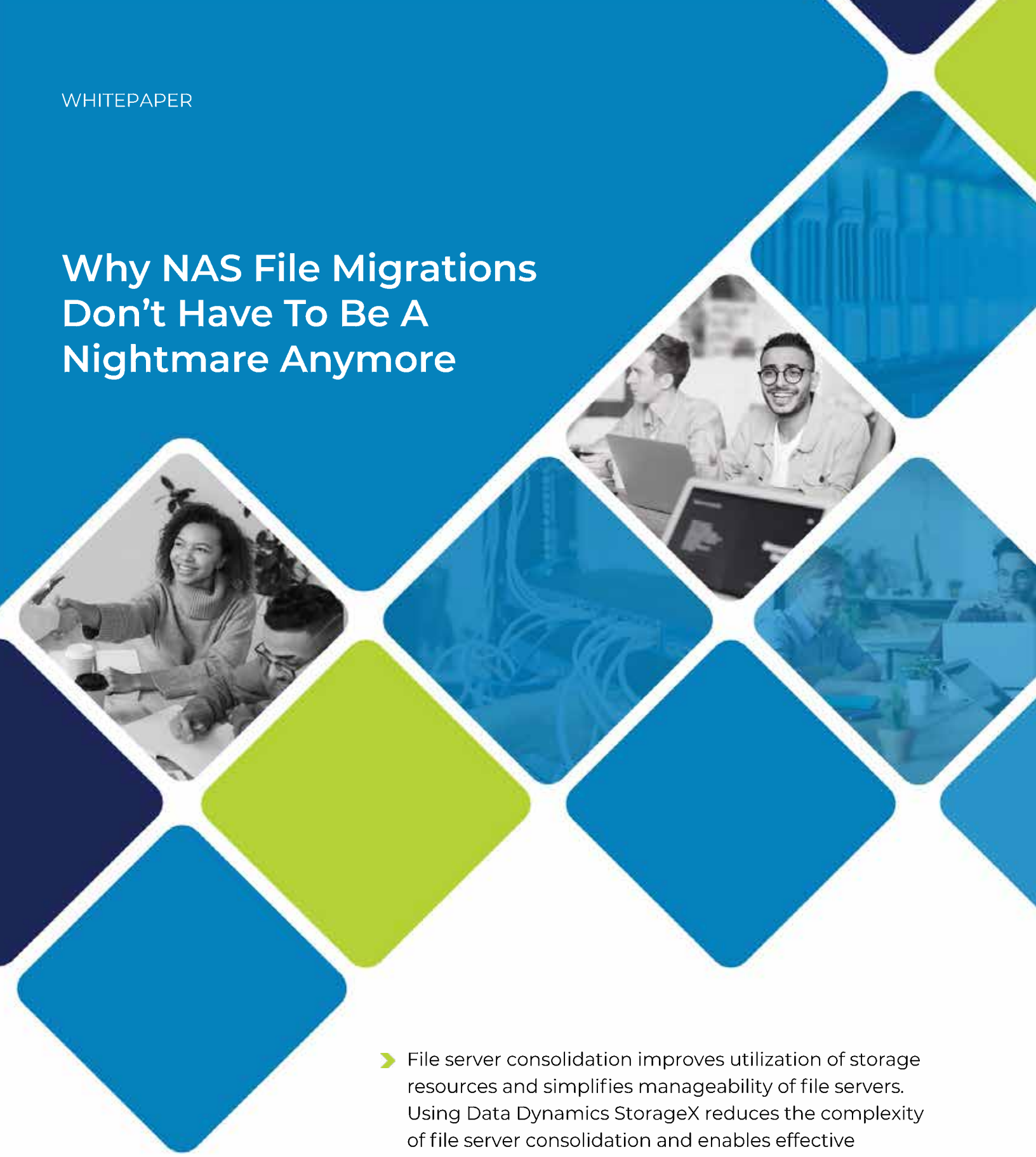
WHITEPAPER

Why NAS File Migrations Don't Have To Be A Nightmare Anymore



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

Introduction

File migration and mobility is rapidly escalating. There are always new and seemingly endless requirements to move files between systems, locations, applications, and users. New projects come up even before old ones complete.

The three most crucial purposes for file migration and mobility are tech refresh, file tiering, and NAS or filer consolidation.

Tech refresh

Tech refresh occurs when older file data storage is being retired and replaced by newer systems. The data must be migrated from the old to the new system. It also occurs when upgrading the file data storage or when changing vendors.

File Tiering

File tiering is utilized more frequently today than ever before. It's used to move files as they age to storage that aligns file value to the file storage cost. Files will likely be moved multiple times over their life times. From hot to warm to cool and eventually cold file storage. That mobility takes place both on- and off- premise and frequently now to public cloud storage. File tiering is also used for archiving, data protection, business continuity, disaster recovery, and compliance (regulatory as well as organizational).

NAS or Filer File Storage Consolidation

File storage consolidation has become much more important as data center real estate has become increasingly expensive. Organizations tend to consolidate with the same or different vendors on a new system. Commonly many smaller file servers are consolidated on significantly fewer larger NAS or Object Stores. Key motivations behind file storage consolidation are cost, simplification of management, and better control.

With file migration and mobility growing at such a rapid pace IT pros might expect that process to have become faster, simpler, easier, and more automated. But here's the rub, generally it hasn't. File migration and mobility is anything but simple. Per ongoing research, approximately 84% of all file migration/mobility projects fail based on being over budget, over time, or both. Surprisingly those failure rates are consistent regardless of whether or not the project is well staffed by internal IT or external professional services.

This paper takes a deeper dive into the file migration and mobility problems, tolls, processes, root causes, common workarounds, why those workarounds tend to fail, and how Data Dynamics is solving the problem.

Table of Contents

The File Migration/Mobility Quandary

Why is file migration and mobility still so complicated? It comes down to the tools and processes utilized.

Ineffective Limited File Migration/Mobility Tools

The two most utilized file migration/mobility tools are open source RoboCopy and rsync. The good news about these tools is that they do not have a licensing cost. The bad news is that there's a very good reason for that. These tools provide simply file mobility for NFS and/or SMB (previously known as CIFS), period. All of the rest of the processes required for file migration and mobility are notably missing. The major processes that are sadly missing include:

- Discovery.
- Planning.
- Provisioning (Vol, FS, Qtree creation, share creation/cloning, export creation/cloning).
- File copy incrementals (moving a file may not take much time but moving an entire filer or NAS does meaning files will be changing during the process).
- File copy management (multi-threading, execution, scheduling, server management, resource throttling, reporting, SID mapping, permissions transferred).
- File attributes transform.
- Project management.
- Infrastructure reporting.
- Continuous project status reporting.
- Cutover.
- Estimation (scheduling, DFS namespace update, automounter update, cleanup, stop source share, stop source export).
- Rollback.
- Service and support.

Those missing processes don't just go away because the tools don't provide for any of them. These manually labor-intensive tasks become the responsibility of the IT staff. And even though there are a few RoboCopy and rsync variations that vendors have wrapped with selected scripts to do some of the heavy lifting, the vast majority of the tasks are still left undone.

Material Consequences

Human labor must fill the gap when software tools fail to provide the required processes. Sweat equity versus software.

Spreadsheet Management

Managing migration by spreadsheet requires extensive amounts of coordination, cooperation, and communication among disparate IT administrators typically managed via a communal spreadsheet. Several different administrators manually update the spreadsheet making it inconvenient and therefore seldom up-to-date, timely, or accurate. Spreadsheet management leads to task delays, duplication, and errors...lots of errors. The most common non-rational reason given for continuing spreadsheet management: "We have always done it this way."

Scripting

Most IT pros attempt to fill the gap via scripting. Many are actually quite good at writing scripts; however, scripting tends to be quite perilous for several good reasons.

Rare is the script that goes through true quality assurance (QA) testing that commercial software goes through. Most scripts go through simple set of tests and subsequently break down under load.

More rare is the script that is regularly re-tested as systems, microcode, infrastructure, and/or software changes. Make no mistake these environmental factors are changing all the time.

Script bug fix processes tend to be ad-hoc with nominal to no regularly scheduled releases. Script documentation is erratic at best. When there is script documentation, it often does not reflect changes that have occurred to the script for bug fixes or updates. The scripts usually have to be rewritten for three reasons.

1. The script does not work.
2. The original author leaves the job responsible for that script and the documentation is non-existent or poor.
3. There is a new project and the script is inflexible because the variables for the new project are different.

Scripting can cause more problems then it solves. Problems such as increased errors, "do overs", extensive trouble shooting to find root causes, and more. Unless the IT organization is in the business of software development or has a highly disciplined software coding process, scripting is not a very good answer to file migration and mobility.

Moving Target

As the notorious Mr. Murphy so profoundly noted, what can go wrong, will go wrong. Never has that been truer than file migration/mobility. What's scoped, defined, and sized up front will change. The project is constantly evolving. Files are constantly in flux. Shares, files, and size of each will all increase. Target share will also change. All of this results in more time, resources, and budget consumed.

Organizational Risk: Inadequate Post File Migration/Mobility Testing And Validation

The project is not over once the files have been migrated/moved. There's more to be manually done. There's application as well as user access and permission checking to their shares and files; SID (secure identifier) validation; and the correction of the inevitable human errors. Access, permission, and SID rarely automatically transfer with the files. Someone has to add them back in or the shares and files are not adequately secured. This is a laborious time sink that has to be done.

Yet it is too commonly overlooked. When access, permission, and SIDs are not migrated/moved with the shares and files, they become accessible by the wrong people or those shares and files become inaccessible. Neither is acceptable because they increase application outages, lost business, lost productivity, and cost.

Severe Budget And Time Busting Root Causes

All of these file migration/mobility problems amalgamate into that 84% project failure rate. This failure rate is directly correlated to excessive professional services; longer than expected file storage system and capacity time overlap; excessive errors requiring "re-dos" or repeated file migration/mobility for the same shares and files multiple times; system errors from process breakdowns or mistakes; lost IT and user productivity; lost revenue; and the misallocation of storage, server, network, and infrastructure assets.

File migration/mobility administrators need to have extensive file migration/mobility expertise to be effective. That expertise requires training. Training is costly while making the file migration/mobility administrators more valuable and subsequently more expensive as greater expertise leads to greater compensation. It additionally increases turnover as other organizations recruit them further escalating hiring and training costs.

Make no mistake, each one of these issues are a burden to IT organizations and will bust budgets individually let alone collectively. Take the example of capacity overlap. There will be twice the systems and twice the capacity required for the length of the file migration/mobility project. Every hour, day, week, month that the project drags on requires paying for both systems simultaneously. Moreover, the system being replaced is typically off of warranty. Maintenance on that system is extraordinarily high once that happens and it has to be paid or the data is at risk in the event of some catastrophic system failure. Time equals cost here and it is the one asset that can never be recovered.

Ultimately all of these problems and issues cause so many headaches that the job is often referred to as the worst in IT.

Most Common Workarounds

The four most common workarounds to these onerous file migration/mobility issues are: file storage to file storage replication (FS2FS); server-to-server (S2S) replication; hypervisor-to-hypervisor (H2H) VM file storage replication; and file backup.

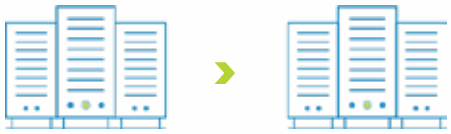
File Storage To File Storage (FS2FS) Replication

FS2FS replication has gained popularity in recent years as the go to technology to minimize the problems just illustrated. It works by taking an instant snapshot utilizing either redirect on-write (ROW) or thin provisioned copy on-write (COW) snapshots. These snapshots consume nominal amounts of actual storage capacity. The snapshots are then replicated or vaulted to the target file storage system. Unfortunately, there are several issues FS2FS that severely complicates IT organizations efforts in file migration/mobility. First and foremost, FS2FS doesn't work across different vendors and far too often fails to work with different systems from the same vendor. Additionally, too many FS2FS replications do not transfer access, permissions, or SIDS. Many fail to account for file incrementals that occur over time. This creates duplicate variations of the same file that wastes capacity on the target file storage. Deduplication reduces some but not all of that additional consumption.

Server-To-Server (S2S) Replication

S2S has been around for many years. It essentially duplicates the writes from one server immediately to another. Over the years S2S has evolved to include time stamping of every write to enable point-in-time rollbacks. It is ideal as a low recovery time data protection option. Access, permission, and SID are no problem since it is replicating between the servers and not the storage. But S2S has its significant caveats as well. It requires quite expensive duplication of server plus infrastructure hardware and software. The S2S software must be installed on every physical and virtual host. Implementing, coordinating, operating, maintaining, and managing S2S is far from trivial. Although the complexity is less than utilization of traditional file migration/mobility tools, it is still highly labor -intensive. It is best suited for changing from one hypervisor to another for data protection where the recovery point objective must be as close to zero time as possible. S2S is a very costly option.

Hypervisor-To-Hypervisor (H2H) VM File Storage Replication



H2H file migration and mobility is activated, controlled, and managed from the hypervisor kernel. It's only been available for the past several years. H2H can be quite effective for some hypervisor users with some explicit serious caveats. As of this writing, H2H currently is only available from VMware (called Storage Vmotion™) and only works with vSphere. It does not work with any other hypervisor or non-virtualized server hosts. Storage Vmotion is a relatively slow file migration, works on a per datastore basis, and is quite pricey. H2H is essentially a partial file migration and mobility tool, doesn't do anything for the non-vSphere systems file migration/mobility, and is much too costly.

File Backup

File backup is often called the poor man's file migration/mobility. The files are being backed up anyway so why not just recover them in the new storage system? File backup is neither a good file migration nor file mobility vehicle. It can do it, but files typically have to be recovered first (although not always). File backup also tends to have a relatively long recovery point objective or RPO (timeframe between backups) making the file typically 24 hours or more out of date. It is a slow process. Access, permissions, and SID are no sure thing. The file recovery process is lengthy, complicated, labor-intensive (again), and error prone. Interestingly enough, many file backup products use rsync as their file mobility engine, and carry forward much of its baggage.

There's a reason file backup is called the poor man's file migration.

There Needs To Be A Better Way To Accomplish File Migration/Mobility

The market requires a simple automated way to migrate and move file data to and from heterogeneous, homogeneous, and disparate file storage systems including object as well as cloud storage. And that file migration/mobility needs to be based on policies that enable it to be ongoing. That is exactly what Data Dynamics is endeavoring to deliver with their latest iterations of StorageX.

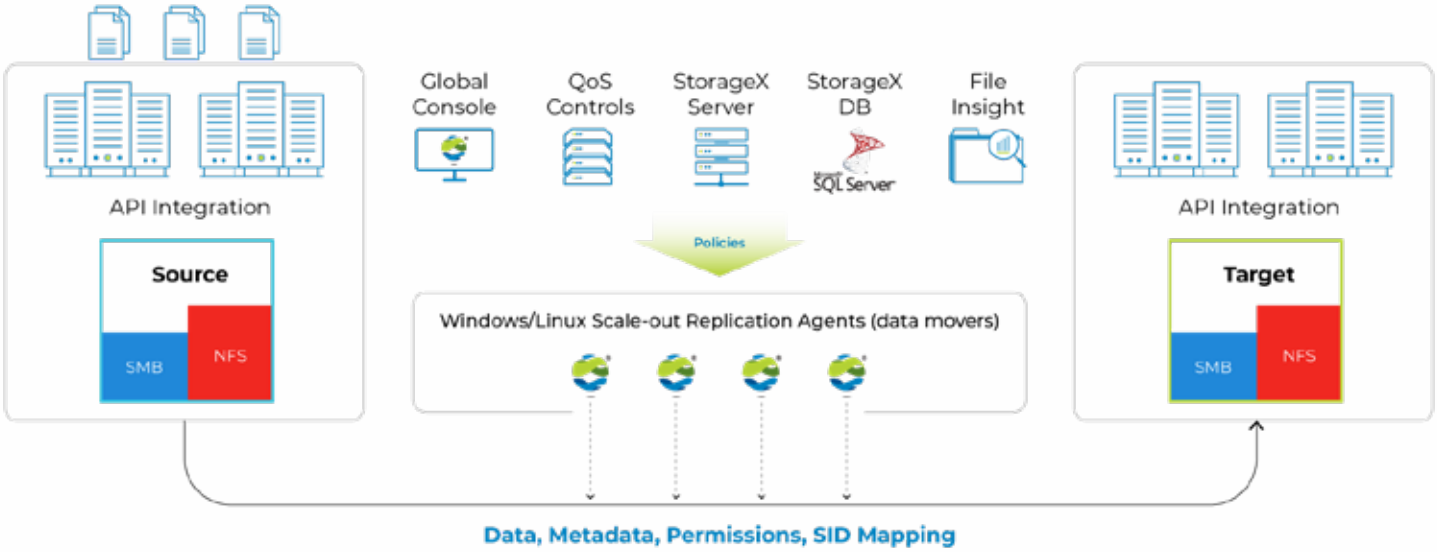
How Data Dynamics StorageX Eliminates File Migration/Mobility Nightmares

The Data Dynamics StorageX software is architected specifically to speed and simplify file migration/mobility. It eliminates file migration/mobility complexity via policy based automation and management. StorageX removes the requirements for admin expertise, scripting, or spreadsheet management. It does this by integrating and automating discovery, policy based mobility rules engines, security, with organic data movers and APIs. As a result StorageX works with the vast majority of systems including:

- NetApp Data Ontap (DOT)
- NetApp Clustered Data Ontap (ONTAP9)
- DELL-EMC Isilon OneFS
- DELL-EMC VNX
- DELL-EMC Unity
- Hitachi Virtual Storage Platform G Series
- Hitachi NAS Platform
- Standard NFS storage
- Standard SMB (CIFS) storage
- Microsoft DFS
- Object storage
- Cloud storage

What Is StorageX

StorageX is a smart out-of-band file migration/mobility software system that runs on a physical or virtual windows host. The software system comes with the StorageX Server™, StorageX Global Console™, QoS controls, StorageX DB (Microsoft SQL Server™), StorageX File Insight™ analytics plus Windows and Linux scale-out replication agents (the data movers).

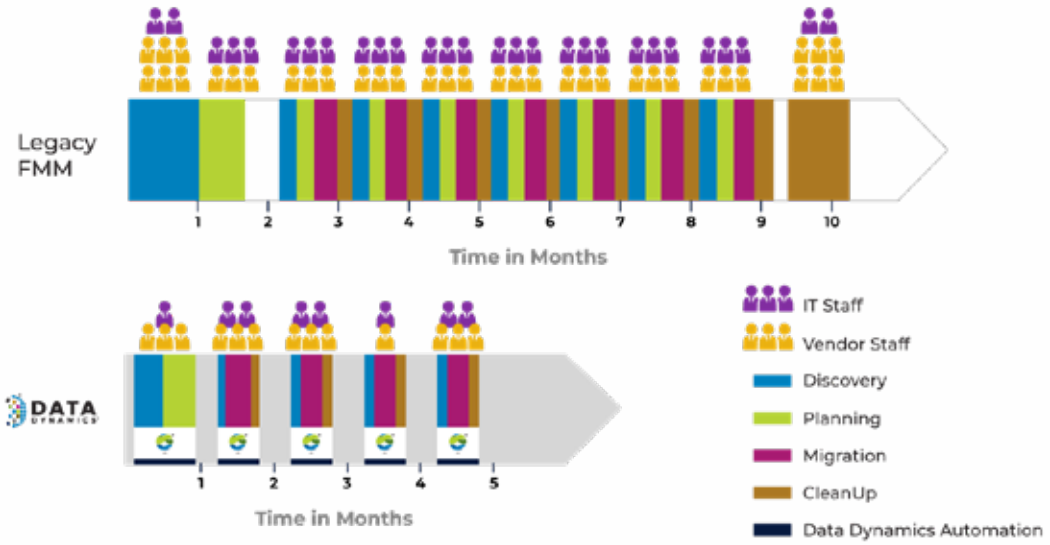


Key Differences Between StorageX And RoboCopy or rsync

It's all about automation. StorageX automates the vast majority of file migration/mobility processes.

Migration/Mobility Tasks	Robocopy	rsync	StorageX
Discovery	NA	NA	√
Planning	NA	NA	√
Provisioning	NA	NA	√
Vol, FS, Qtree Creation	NA	NA	√
Share Creation/Cloning	NA	NA	√
Export Creation/Cloning	NA	NA	√
Protocol: CIFS/SMB	√	NA	√
Protocol: NFS	NA	√	√
File Copy Incrementals	Manual	Manual	Automated
File Copy Mgmt	Manual	Manual	Automated
Multi-threading	Manual	Manual	Automated
Execution	Manual	Manual	Automated
Scheduling	Manual	Manual	Automated
Server Mgmt	Manual	Manual	Automated
Resource Throttling	Manual	Manual	Automated
Reporting	Manual	Manual	Automated
SID Mapping	NA	NA	√
Permissions transferred	NA	NA	√
File Attributes Transform	√	√	√
Project Mgmt	NA	NA	√
Infrastructure Reporting	NA	NA	√
Project Status Reporting	NA	NA	√
Cutover	NA	NA	√
Estimation & Planning	NA	NA	√
Scheduling	NA	NA	√
DFS Namespace Update	NA	NA	√
Automounter Update	NA	NA	√
Cleanup	NA	NA	√
Stop Source Share	NA	NA	√
Stop Source Export	NA	NA	√
Rollback	NA	NA	√
Service & Support	NA	NA	√

RoboCopy, rsync, and software that wrap these tools, do not. Illustrating the file migration/mobility process differences in time and resources for a tech refresh or consolidation is a real eye opener. Because StorageX completes the file migration/mobility project so much faster it actually ends up moving less total data. This is because there are fewer file changes or incrementals that have to be moved over the shorter timeframe.



Those differences add up quickly. According to Data Dynamics, StorageX will take on average 50% less time to complete a project, 50% less cost, and deliver 200% to 500% greater productivity. Typical manual migrations average about 200 TB per full time equivalent (FTE) person per year. StorageX migrations range from 500 TB to 1 PB per FTE person per year. This means far shorter transition time, shorter cutovers, faster decommissioning of old storage, faster time to value for new storage, and much reduced labor costs from increased automation as well as much greater productivity.

Are these assertions valid? Based on real-world production implementations, these assertions are conservative. They do not include the substantial savings from risk reduction, error recovery costs, unplanned downtime reduction, and their accompanying lost revenues. Those savings alone will frequently dwarf all StorageX costs.

StorageX Cost Savings

Even when those cost savings are ignored, the work, time, and financial savings are enormous. Data Dynamics has built a comparative cost savings calculator based on hundreds of real world implementations. It fairly accurately estimates what those savings will be based on several factors including:

- File data size to be moved
- Manual file migration/mobility burdened labor cost.
  - Heavy lifting skills of planning, scripting, executing, troubleshooting, error root cause analysis, and rework.
- StorageX file migration/mobility burdened labor cost.
  - Primarily project management and architecture skills because of the automation.
- Coordination effort to complete the project (high, medium, low).
  - High when there's significant cutover and reporting planning and coordination required for approximately 1000s of cutover.
  - Medium when cutover and reporting coordination is in the ~100s.
  - Low when much of the data can be copied or moved with little coordination with others, or namespace technologies used.
- Estimated number of full time equivalent personnel available for both the manual and StorageX file migrations/mobility project.
- Max file data transfer rate (MB/s) limited by infrastructure.
- Weekly file data change rate.
- Projected time in months to complete the file migration/mobility project for both the manual and StorageX methodologies.
- StorageX licensing costs.

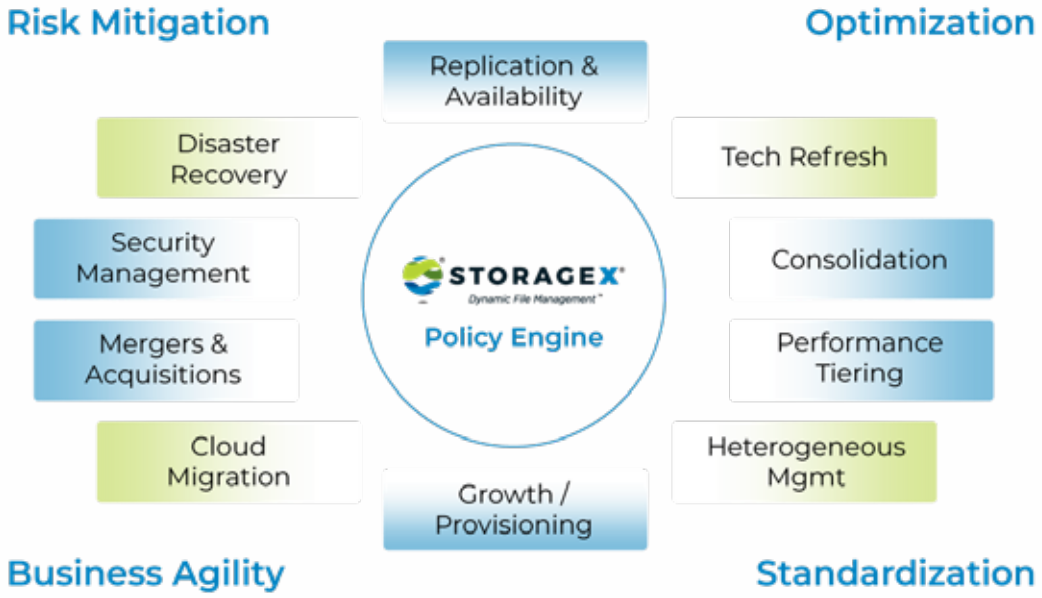
Based on some real world examples, the StorageX savings are quite impressive.

StorageX Reduction	File Migration Project			
	20TB	100TB	1PB	3PB
Cost	**49%**	**61%**	**62%**	**67%**
Time	**86%**	**73%**	**68%**	**68%**
File Data Moved	**5%**	**5%**	**16%**	**39%**

Those are huge savings and that's just the direct cost savings. The Data Dynamics savings calculator also calculates compares the cost of error recovery and business impact. As previously stated those savings are quite a bit greater. Lower cost, greater availability and reliability, and more certainty are exactly what Enterprises require. But these are just file migration projects. There are quite a few other obstinate IT file mobility problem categories solved by StorageX.

Four Other Major File Mobility Problem Categories Solved By StorageX

StorageX is an excellent file migration tool. It is also an excellent general file mobility tool for file optimization, standardization, business agility, and risk mitigation. StorageX solves as many if not more IT problems for file mobility as it does for file migration.

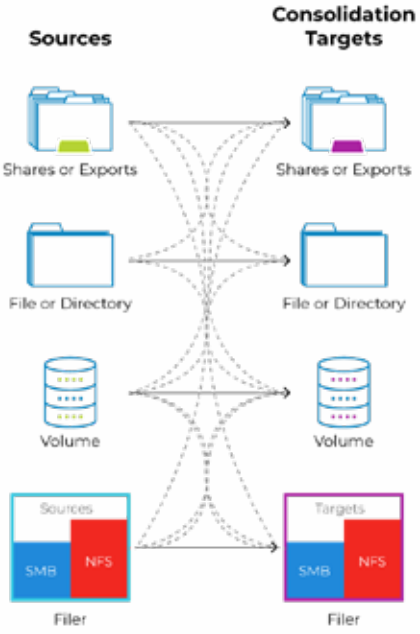


Tech Refresh & Consolidation File Mobility

File storage tech refresh is the ongoing curse of IT. It doesn't matter whether the current file storage has run out of gas (not enough performance and/or capacity), the warranty/maintenance is up and the re-upping is prohibitively expensive, or there is an ongoing effort to consolidate file storage, file storage tech refresh requires file migration/ mobility to be successful.

StorageX will move, based on policy, file data from one or many file storage systems (file servers or NAS) to one or more other file storage systems. Those file storage systems can be multi-vendor, multi-system, homogeneous, or heterogeneous. It supports any SMB to any SMB and any NFS to any NFS. The built-in StorageX API support also allows automatic target selection to match the source or enable IT administrators to determine the shape of the target destination.

When consolidating file storage, StorageX will consolidate at just about any level including shares, exports, volumes, Qtrees, Tree Quotas, or any directories. It automatically provisions and replicates the file data migrating them to pre-created shares.

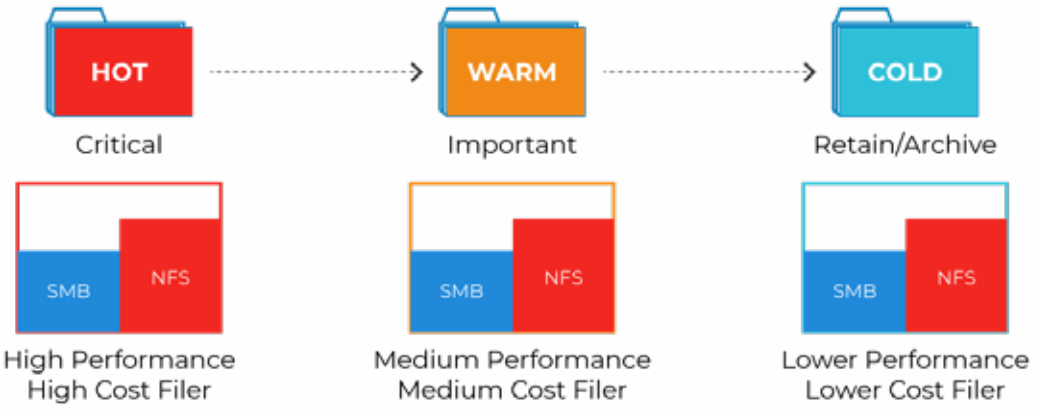


Tiered Performance File Mobility

Storage vendors have talked about storage tiering for years. Except, when they talk about storage tiering it is primarily within their storage system with different classes of drives. It is rarely between different storage systems. When it is between different storage systems, it's between specific storage systems from the same vendor. Some systems today actually tier to cloud storage. But generally storage tiering of this type has not changed much since the days of "hierarchical storage management" (HSM). It is stub based. This means when the files are accessed, or more specifically the stubs are accessed, the files are recalled to their original storage tier. In fact, the files cannot be accessed except be recalling them to the original storage tier. That means file access no matter where it is tiered is through the storage system where the files originally landed or the upgraded storage system from the same vendor.

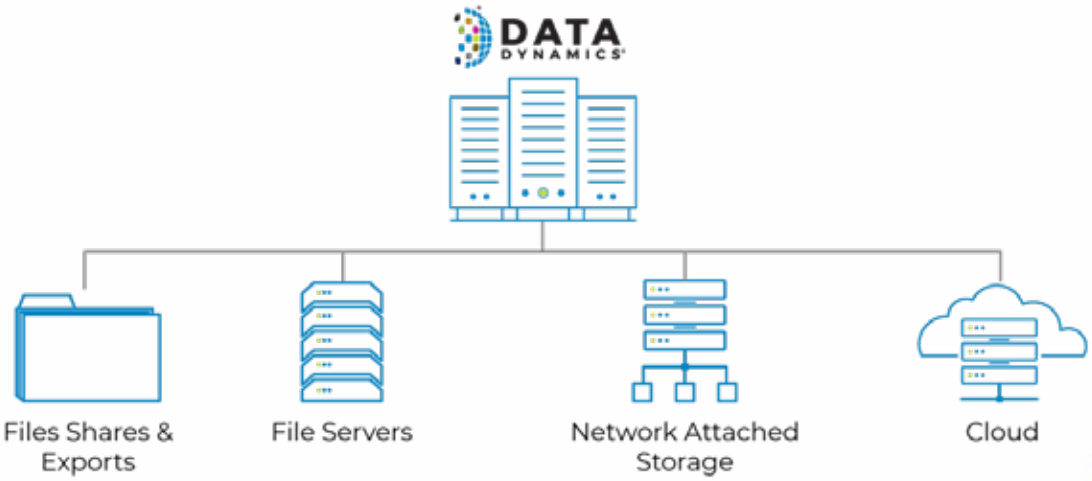
They cannot be acquired natively from the cloud storage tier. This can cause unnecessary duplicates of files in different tiers because when the file is recalled it is not deleted from the secondary or cloud tier. It is duplicated again when that file is moved based on policy back to the secondary or cloud tier. Consuming more capacity. Public cloud storage charges when copying the files out creating more cost.
Storage admins have dreamed for years to seamlessly tier files between disparate, heterogeneous file storage systems based on the file's value and the cost of the storage and to be able to access those files directly without having to move them again. It makes no sense keeping infrequently accessed or low value files on high cost high performance storage. Doing so consumes precious primary storage, storage admin time, and data protection admin time. It also increases organizational risk when a disaster occurs. Disaster recovery recovers all file data on the primary storage system regardless of value or access frequency. Recovering non-essential data lengthens recovery time. Lengthy recovery times are frequently the difference between those businesses that survive and those that do not. Per Gartner, 95% of businesses that have not recovered their systems and data within weeks are out of business within two years. Unfortunately, file storage tiering between different systems have always been deemed too difficult and impractical.

StorageX makes multi-storage system file storage tiering simple and practical. It empowers IT to tier and move files to different cost and performance classes of file storage based on policy. Files can be moved to different shares or exports based on name, age, and path criteria. StorageX will tier between different file storage, NAS, Object, and/or Cloud storage. This reduces total cost of storage, infrastructure, administration, operations, and management. And best of all, once the files are tiered, they are accessed natively from that tier. No stubs, original storage systems, gateways, or other storage systems to go through, just simple native access.



Heterogeneous File Storage Management

StorageX additionally streamlines file storage management by allowing multi-vendor file storage resources to be handled from a single console. It offers file share, storage resource container, and export discovery, in addition to provisioning, creation, and cloning. StorageX monitors and manages the file storage resources in addition to the file mobility. This allows file systems to be moved or split for growth. It also enables load balancing and rebalancing of the different file storage systems based on SLAs.



File Storage Growth And Provisioning

StorageX can also be utilized as a unified methodology for both NAS file storage provisioning and migration project provisioning. It does this through automated creation of volumes and qtress, shares and exports, and DFS links across heterogeneous storage vendors such as NetApp, Dell-EMC, Microsoft, Amazon, and others.

File Mobility To/From Cloud Storage

StorageX has the capability of moving file data to and from cloud storage including Amazon S3, Amazon Glacier, Microsoft Azure BLOB storage, and OpenStack Swift. That file mobility can be automated and/or ad hoc. Files moved to the cloud will have their compliance data transformed as well.

Mergers, Acquisitions, And Divestitures

StorageX facilitates and automates the necessary file movement and mobility required when companies merge, or divest. It maintains and manages file security. For acquisitions it transforms file systems and permissions from the old security model to the new security model. StorageX enhances SMB security by mapping SIDs to their files. For divestitures StorageX copies files out and then isolates their access to only those who have permission in the divested company.

File Mobility Security Management

Security management tends to be one of the stickiest issues for file mobility. StorageX is specifically designed to manage all of it making file mobility security management stress-free.

Copy Security
- Copy file permissions and ACL
- Clone share permissions
- Clone Export policies

Remove Access
- Delete share permissions
- Delete export policies

Re-permission
- Copy files but not permissions
- Copied files inherit new permissions

Remove invalid SIDs

SID Mapping

File Migration/Mobility DR

Disaster recovery, business continuity, or disaster tolerance varies by every file storage system. Each has its own process and methodology with no two systems or vendors alike. StorageX can manage and homogenize these processes. It will monitor and initiate failover from the primary file storage system to the designated secondary file storage system. And that secondary system can be from a completely different vendor or different system from the same vendor. Failovers and failbacks are based on user defined policy triggers.

StorageX additionally leverages NetApp SnapMirror replication. It will create, initiate, and break the replication, initiate failover as well as failback, or utilize the StorageX replication to copy the files from the primary to secondary system target.

For Microsoft DFS, StorageX will failover and/or failback the DFS link from the primary to secondary system target after it stops replicating.

File Migration/Mobility Replication Management

Maintaining business continuity (BC), data protection (DP), and disaster recovery (DR) while transitioning from one system to another can be overwhelmingly difficult. An example of this is moving from NetApp ONTAP7 to NetApp ONTAP9. There is no simple way to get this done and as a result, BC, DP, and DR readiness suffers or is non-existent during the very lengthy upgrade process.

StorageX solves this problem. In the NetApp example, it maintains that BC, DP, DR readiness while drastically shortening the upgrade process as well. It has become the most effective methodology for upgrading from ONTAP7 to ONTAP9 currently available today.

Summary and Conclusion

File migration and mobility has been a complicated, difficult, laborious, painstaking, nightmarish process. Many administrators avoid it at all costs. Data Dynamics with StorageX has changed file migration into an automated simple process that opens up a world of possibilities.

Data Dynamics, a global leader in enterprise data management, stands at the forefront of the industry-wide shift towards Digital Trust & Data Democracy. Trusted by 300+ organizations, including 25% of the Fortune 20, the company is recognized for its commitment to creating a transparent, unified, and empowered data ecosystem. Whether addressing data risk, privacy, sovereignty, optimization, sustainability, or facilitating seamless, policy-driven data migration across hybrid and multi-cloud environments, the company is ushering in a new era where data ownership, control, & actionability reside with the data owners.

DATA DYNAMICS

Contact Sales | Book a Demo