ActualTech Media

DATACORE

Automate Your Data Placement Across Multiple File Servers, NAS, and Object/Cloud Storage for Simplicity and Savings

Trevor Pott

CONTENTS

IN THIS PAPER

File storage is a difficult problem for many organizations. Not all files are equal, each has its own useful lifespan, importance to the organization, and regulatory considerations. Complicating matters is that these files are often scattered across multiple storage systems. This paper discusses how DataCore vFilO solves these problems, saving organizations money, and making life easier for end users and systems administrators alike.

Putting data governance standards into practice remains one of the toughest problems facing IT, whether the reason for those standards is regulatory compliance, or simple efficiency. Not all data has the same value, and the value of data changes over the course of its lifetime. DataCore vFilO solves these problems.

Guidelines for file availability, durability, responsiveness, and access control that seem simple on paper often prove impractical to institute, made more difficult by the perpetual, astonishing growth of data, particularly unstructured data. Systems administrators are constantly running out of space on multiple storage systems while seeing others go mostly unused.

> ## Making efficient use of storage means solving the file distribution problem.

From the end user's point of view, finding files across multiple storage locations can be frustrating. This is because each individual storage system has its own namespace; when files move between physical storage systems their relative position changes.

In addition to worrying about the namespace problem, administrators struggle to optimize storage efficiently, balancing the desire to have everything in the same location forever against the need to manage capacity consumption and archive old files.

By unifying your file storage infrastructure, your organization will reap significant operational benefits, but the manual process to set up appropriate levels of data protection can be tedious and time consuming, especially if the files are distributed across multiple independent storage systems. A solution that simplifies and automates this process greatly eases this headache.

The File Distribution Problem

Multiple storage systems are useful because they can put the storage closer to the end user. They're necessary to address concerns such as redundancy against failure, and they're inevitable because you can only put so much capacity or performance in a single system.

As previously stated, data has differing and mutable value, changing over the course of its lifetime, which complicates where that data should be placed, creating an incentive to change the location of the data over time. Until DataCore vFilO, however, this has not been an easy process.

The traditional IT model of having all workloads for an organization confined to a single, centralized, nearby data center is increasingly rare. Organizations have multiple storage systems, each with a different cost per bit, both for capital expenditure, and for ongoing operations. At the same time, files that are infrequently (or never again) used are often occupying space on the same storage devices as files that see frequent use, because both humans and computers look for grouped files in the same namespace.

Both end users and applications prefer to have high-throughput, low-latency access to the files they need, but this is complicated because storage systems offer differing performance, and are increasingly distributed geographically. The end-user experience is thus not only a factor of the physical capabilities of the storage system, but the user's proximity to the storage system and the layout of file shares along departmental boundaries

> ## We might not notice a few dozen milliseconds of latency when opening a file to work on, but it can be crippling to demanding applications.

Making efficient use of storage means solving the file distribution problem. Files that need to be frequently accessed should be the only ones located on expensive storage, and a copy of them needs to be geographically proximate to where they will be used. Meanwhile, files that will not be accessed again excepting backups or legal discovery purposes need to be archived off to low cost storage, such as object storage, or cloud storage.

Solving this problem manually is effectively impossible. Building one's own storage solution out of shell scripts

and strict business processes about where to place files is frustrating, but is the best most organizations have been able to do so far. There is a better way.

Solving the File Distribution Problem

Solving this problem has traditionally required kludgy workarounds. These range from making copies of files across multiple storage systems, to physically restricting both users and workloads to specific geographic areas where they will be proximate to the files upon which they need to operate. As the world transitions to a hybrid cloud IT infrastructure model, as well as a distributed work-force model, this is increasingly untenable.

User perceptions aren't the only concern regarding file distribution across devices, especially when that distribution encompasses far-flung geographies. Data that takes a long time for applications, IoT devices, and other machines to access impacts the services those machines provide. And machines are far less forgiving about delays than humans. We might not notice a few dozen milliseconds of latency when opening a file to work on, but it can be crippling to demanding applications.

Bringing all of an organization's file storage together under vFilO allows files to be placed on specific storage systems, with copies on however many storage systems are required, and with changes to files automatically replicated to other storage systems when this occur. File placement policy can be based on a number of considerations, including performance, age, frequency of use, geographic proximity to data consumers, and load leveling.

Introducing DataCore vFilO

DataCore vFilO is a next-generation file and object storage virtualization solution that controls data placement across diverse storage resources according to the data governance rules you set. This is accomplished by empowering organizations to consolidate separate filers under a single, global namespace, ensuring they achieve the maximum value from existing investments, while enabling non-disruptive modernization to meet tomorrow's needs.
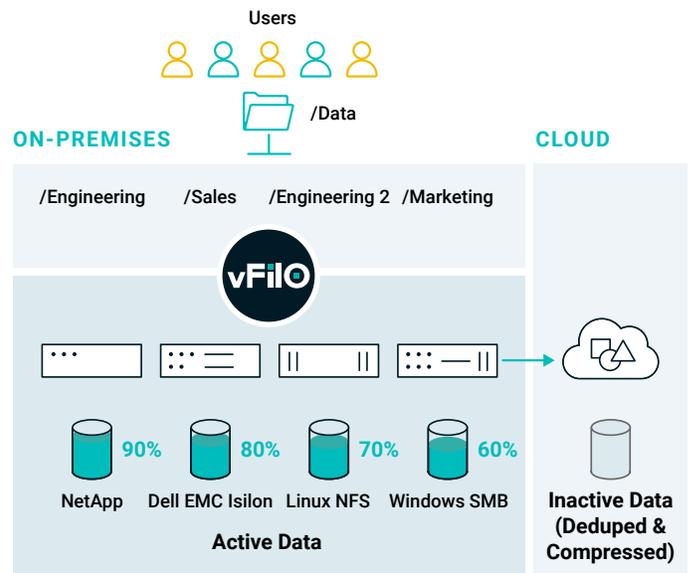

Figure 1: DataCore vFilO architecture

vFilO makes distributed file storage convenient again by combining multiple file servers and NAS systems, as well as S3-compatible object and cloud storage, into an integrated system for better visibility, flexibility, and control (see **Figure 1**). vFilO provides efficiency across all storage systems, unified search across multiple shares, a single global namespace, as well as other benefits.

vFilO can assimilate existing shares. This means there's no file migration required, and that the solution can be operational in minutes with minimal disruption. Once existing shares have been assimilated, vFilO can export them as sub directories of a single namespace, virtualizing an organization's entire file storage, and ingesting all the metadata for the files, and the shares they were organized under.

This enables all of that organization's file shares to be discoverable from a single mount point, and allows vFilO to manage file placement throughout the organization's entire file storage infrastructure.

Delivering adequate responsiveness is another traditional storage problem vFilO can significantly ameliorate. By controlling file placement across all storage infrastructure components, administrators can ensure that files requiring the best-performing physical hardware live there, while other files (such as inactive data) are moved to less costly storage, or even archived altogether.

Show Me the Money

vFilO offers benefits far beyond namespace consolidation. Simplifying your file storage infrastructure offers a number of financial advantages, both in terms of capital costs and ongoing operational expenditures.

As noted earlier, solving the problem of file distribution across an organization's entire file storage infrastructure allows administrators to create policies that place the frequently used files on the fastest storage, and less frequently used files on slower storage. A single namespace also makes discovery, searching, indexing, and tagging easier. This reduces frustration for human users, and can pay outsized dividends for applications that rely on these capabilities.

Moreover, manually spreading files out across multiple devices results in both stranded performance and stranded capacity—10% of unused capacity here and 40% of unused capacity there adds up quickly. vFilO automatically load balance your capacity in order to avoid this.

> **Having all file shares on a single virtualized file system makes any IT tasks where discovery is important significantly easier.**

Controlling data distribution is also a very direct way to reduce capital expenditure on storage. For example, holding on to inactive data is an expensive burden to carry, but regulatory concerns make it mandatory. A better solution is to archive inactive data to public cloud storage or local object storage, where it's compressed and deduplicated to reduce its footprint (and cost), and yet can be rehydrated for access on an as-needed basis, completely transparently to the data consumer.

Unifying file access on vFilO has operational benefits for IT. Having all file shares on a single virtualized file system makes any IT tasks where discovery is important significantly easier. Backups, audits, legal discovery, and archiving are all a lot less troublesome if you don't have to find every file server across the organization, analyze each one for new shares, and then add them to the appropriate maintenance routine.

Similarly, vFilO makes everyday implementation of file services less burdensome. It represents a single target for IT policies ranging from information security to storage quotas. And you don't have to be concerned with—or build workarounds for—the implementation nuances of individual storage systems. vFilO handles all of that for IT teams, presenting a single file storage interface against which all IT policies can be enacted. With vFilO physical file servers, NASes, object storage systems, and even public cloud storage all become little more than different physical locations where bits and bytes rest.

> **DataCore vFilO helps you achieve significant cost savings on your data storage investments.**

Last, but certainly not least, vFilO's in-place assimilation capabilities allow it to be easily deployed within a brownfield environment, which not only allows vFilO to be operational minutes after initial deployment, it also dramatically simplifies the IT burdens associated with ongoing mergers and acquisitions, a normal part of corporate life that many IT teams dread.

Take the Pain Out of File Storage

Managing file storage remains a significant problem for IT teams. Data governance is a challenge complicated by geographic distances, regulatory concerns, and the natural diversity of equipment that's the result of normal organizational growth. vFilO takes the pain out of file storage for end users, applications, and the IT teams that support them.

DataCore vFilO helps you achieve significant cost savings on your data storage investments. By controlling where and when to move data based on cost and business objectives, vFilO lets you leverage lower-cost storage alternatives and gain financial flexibility. Try the DataCore cost savings calculator to see how much money you can save for your organization: www.datacore.com/products/vfilo/savings-calculator.