

Offload Your Inactive Data to the Cloud

Trevor Pott

CONTENTS

Taking Advantage of Cloud and Object Storage.....	2
The Data Placement Problem.....	2
Balancing Storage Needs.....	3
Intent-Based Data Placement.....	3
Make the Most of Your Storage.....	4

IN THIS PAPER

Are you looking to realize the maximum efficiency from your file storage infrastructure? This paper explores how vFilo helps transform how unstructured data is handled through file storage virtualization and transparent archival of stale data to the cloud.

Maintaining and scaling out multiple file servers and NAS systems quickly gets expensive. How much you spend to store data should be proportional to its value at the time, but no one has the time to go through each file or folder to identify what can be deleted, moved, or archived. DataCore vFilO solves this problem—without compromise.

The cost-effectiveness of storage can be measured in terms of utilization, performance efficiency, and placement optimization. Getting the biggest bang for your buck is not only about using what you have as efficiently as possible, it's also about making sure you're only putting on fast, expensive storage the data that absolutely needs to be there.

DataCore vFilO is a file and object storage virtualization solution that can determine where files should be placed based on age, last access, or other attributes of a file's metadata. Since these characteristics change over time, vFilO regularly checks these parameters to decide if the file should be moved to a different class of storage, including the cloud, that's most appropriate.

This process is completely transparent to end users. While the file may have been moved from one storage device to another, it still appears in the same folder it's always been in, on the same share. Files that have been moved, even those that have been archived to cloud storage, remain available to users, as if they'd never been relocated.

vFilO provides this functionality automatically, and it's a practical solution to a number of efficiency and cost problems common to file storage. DataCore vFilO frees up expensive storage space in your infrastructure, allowing for both active data growth and keeping redundant copies of data for higher levels of protection, while achieving significant cost savings.

Taking Advantage of Cloud and Object Storage

File lifecycles are reasonably predictable. Most files are written to several times in rapid succession during creation and initial revision, and then never touched again. These files do little more than occupy space, but have to be kept around for legal reasons, or because they're

required by an application to function, even if they're only very rarely accessed.

With vFilO, individual files that meet certain criteria are moved to different classes of storage. Most organizations have at least one fast file storage device, (high throughput, low latency), which is typically quite expensive. On the other end of the spectrum are public cloud storage and on-premises object storage: These are typically much slower, but also far less expensive.

DataCore vFilO frees up expensive storage space in your infrastructure, allowing for both active data growth and keeping redundant copies of data for higher levels of protection, while achieving significant cost savings.

Fast file storage makes users and applications happy. Cheap object storage makes the finance department happy. With vFilO it's possible to take advantage of both without compromising the user experience, breaking the budget, or breaking any applications.

The Data Placement Problem

For various reasons—personal preference, company policy, how an application is built, or legal/regulatory issues—files are often confined to specific directories, in specific shares. Because of this, their storage systems quickly fill up regardless of how often the files are used or how important they are.

Traditionally, when an organization starts running out of room on one department's storage system, it'll start using space on another department's storage system. Sometimes that's an officially IT-sanctioned activity and sometimes it's not, but it usually causes problems. Applications need to know exactly where the files they operate on are located, so you can't simply move those files around. And humans get annoyed if data isn't where

they expect it to be, which can turn pragmatic IT decisions into an office politics problem.

Determining what data is “hot” (in active use) and what’s not can’t be done manually. Even small organizations have millions of files, and staffing such a position would be a nightmare. Manually reviewing file usage statistics to determine where they should be placed is a legitimate candidate for the worst desk job ever.

Much of the difficulty stems from the fact that file location doesn’t tend to reflect either usage patterns (hot, warm, cold), or the relative importance of a file. An end user’s home directory or an application’s resource folder might contain thousands or even millions of files, with a significant percentage not having been read or written to in years.

So how do you solve the data placement problem? Traditional approaches to tiering and archiving data were cumbersome. They typically involved removing data from existing file structures, which meant trying to find and restore anything that had been archived to cheaper storage, which was painful.

With DataCore vFileO, data remains accessible in its original location no matter what tier of storage it’s on, and this ensures that organizations are able to make the most efficient use of their existing storage infrastructure by eliminating the problem of “stranded capacity.” With all storage under vFileO management, storage system free space becomes a resource that can be made available to the entire organization, without breaking applications or changing where people look for files.

Balancing Storage Needs

The quality of storage products varies greatly. Throughput, latency, and write resiliency all vary depending on the type of workload a storage product was designed for. This variability in performance characteristics is accompanied by a corresponding (and often quite dramatic) differentiation in price.

Thus, it’s generally not considered the best of plans to put inactive data that no one (person or machine) is using onto the best storage in your organization. No company wants to spend too much on irrelevant data

that rapidly accumulates, but actually accomplishing any form of file tiering or distribution across multiple file storage systems has traditionally been complicated. All the more difficult when different models of filers and NAS devices are involved.

How do you migrate a file to a cheaper storage system and still retrieve it when you have to, without breaking all those familiar folder/directory hierarchies we organize our files into? Not knowing where a file has gone is an annoying productivity drain to a human, but it will break applications.

Not knowing where a file has gone is an annoying productivity drain to a human, but it will break applications.

With vFileO, however, files will be individually moved to different storage systems based on criteria determined by the organization. This eliminates the need for traditional rigid hardware assignments that cause large imbalances in space consumption and response. Marketing no longer needs to experience sluggish response because they got stuck on a slow file server, while finance is less concerned with performance because they landed on a much faster system.

Not all storage is created equal, nor is the cost of maintaining that storage equal, but with DataCore vFileO organizations can manage storage capacity and performance independently of one another, and independently of departmental silos within their institution. If you need better file server performance, you add another fast file storage system to the pool. If you need more capacity, you can add more storage, or archive more data off to the cloud. But first, vFileO makes more room on the existing faster storage by migrating out files that don’t need the speed.

Intent-Based Data Placement

DataCore vFileO virtualizes separate file shares across the organization under a single global namespace by assimilating and pooling the individual file storage systems.

This untethers the namespace (where files appear to exist, from a file and share structure standpoint) from the physical storage on which the data exists. This saves organizations money by allowing less frequently used data to drift to inexpensive storage systems. And vFilo offers administrators a number of mechanisms to determine which files end up where.

Intent-based data placement allows data placement of individual files within a given folder structure to different storage devices while still presenting a single folder structure to the end user or application.

For some organizations, simply allowing vFilo to make data placement decisions based on usage would, in and of itself, be a significant improvement. Even without any further refinement, this can reduce both stranded performance and stranded capacity across entire organizations.

But for many organizations, storage decisions aren't that simple. Like data, not all workloads are created equal. Making different storage classes (storage systems with different price/performance characteristics) available is part of the answer, but both capacity and performance crunches do happen—especially in organizations with strong seasonal usage patterns—and it's important for administrators to have other mechanisms for controlling data placement.

DataCore vFilo allows administrators to perform intent-based data placement. This means they can get more performance for workloads that need it without significant budget impacts. Administrators can designate certain data as requiring more resiliency than other data, ensuring that multiple copies are kept throughout the storage pool, and can even make snapshots of data available for use by dev and test, or auditors, quickly and easily without impacting production data.

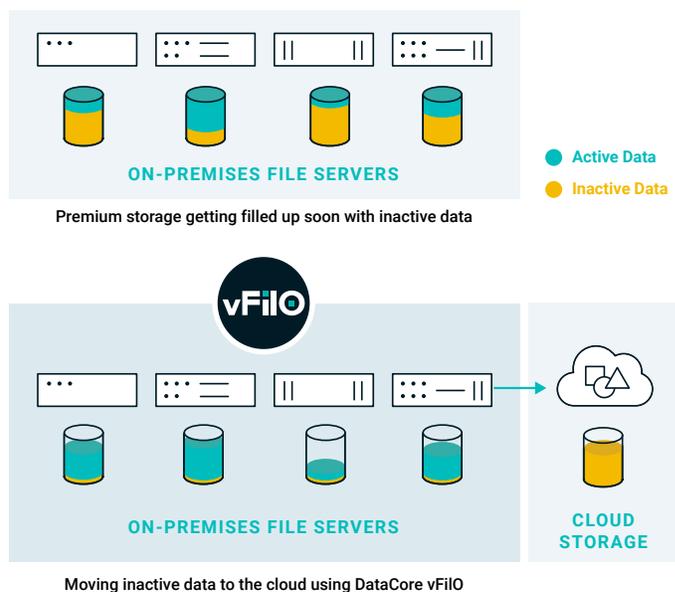


Figure 1: File storage with and without DataCore vFilo

Intent-based data placement also allows system administrators to geolocate copies of data to make access easier for distributed workforces and workloads. Certain storage classes (such as cloud) are great targets for deduplication and compression, reducing costs even further.

Most importantly, intent-based data placement allows data placement of individual files within a given folder structure to different storage devices while still presenting a single folder structure to the end user or application. Frequently used files can be placed on faster storage devices, while infrequently used files are moved to archival storage in the public cloud, or to an on-premises object storage system (see **Figure 1**).

This capability allows organizations to take full advantage of their fast file storage by ensuring it's only used for files that need the best possible performance. It also allows organizations to offload unused or infrequently used files to storage systems (like the public cloud) where capacity can be rented on an as-needed basis, avoiding significant capital expense.

Make the Most of Your Storage

DataCore vFilo empowers organizations to realize the maximum efficiency from their file storage infrastructure.

More than that, though, vFilO helps organizations transform how they handle unstructured data.

When VMware virtualized compute, it decoupled the workload from the physical server. What went into the rack stopped mattering except in edge cases, and infrastructure teams evolved beyond either obsessively matching hardware to each workload, or deliberately (and sometimes massively) overbuying in order to accommodate potential future growth.

vFilO brings the same revolution to file storage. It moves system administrators beyond having to execute an eternal space/performance/namespace balancing act or simply massively overprovisioning everything. With DataCore vFilO file system administrators can join their other infrastructure colleagues in governing their domain with policies, profiles, automation, and orchestration, without bothering their users.

Use DataCore vFilO to migrate inactive and cold data from costly NAS devices on-premises to affordable cloud/object storage. Automate data movement based on policy-driven business objectives and maximize the value of your storage investments. Try vFilO in your environment today. Register for a free test drive at www.datacore.com/products/vfilo.