ActualTech Media

Making True Hybrid Cloud Storage Simple and Fast

Dan Kusnetzky

Data management in a hybrid cloud storage environment can be challenging and complex. Each storage environment requires its own tools. Creating a simple, drop-in approach that produces immediate improvements is quite difficult. What ways are companies addressing this challenge?

To start with, it's necessary to approach making hybrid cloud storage both simple and fast from a new and different direction. Traditional approaches that build up a custom solution that combines tools provided by different suppliers often compound the problem of complexity and may not address the issue of overall performance.

When it comes to hybrid cloud storage, network management and security appears to have a technology that can be applied in the storage segment. Using this approach, it's possible to quickly understand a sophisticated storage environment made up of many "moving parts," built on a number of different technologies, and provided by (sometimes many) different vendors.

> Traditional approaches that build up a custom solution that combines tools provided by different suppliers often compound the problem of complexity and may not address the issue of overall performance.
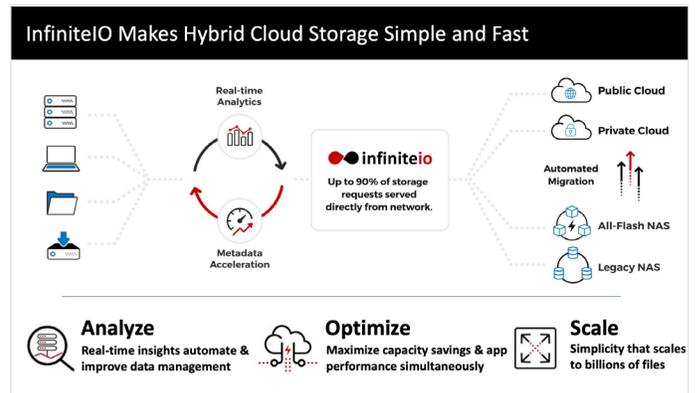

Figure 1: An overview of InfiniteIO's capabilities.

Deep packet inspection (DPI), a tool that's been utilized in network management and security, offers surprising and useful answers. DPI combined with machine learning offers a great deal of promise. This approach should make it possible for enterprises to quickly and easily build a complete understanding of complex storage environments; that allows organizations to optimize access to storage and minimize storage costs.

INFINITEIO MAKES HYBRID CLOUD STORAGE SIMPLE AND FAST

InfiniteIO was founded with the aim of making it possible for enterprises to easily and effectively address the challenges faced with hybrid storage (see **Figure 1**).

They've developed an appliance-based solution that can be installed without disrupting current operations. It's placed in front of enterprise NAS solutions and learns about the storage in use, including the hot files and their associated metadata.

Then it begins to respond to metadata requests out of its own internal DRAM on behalf of the back-end storage devices. This approach offloads serving metadata requests from the back-end filers. Depending upon the application, this can be the majority of storage requests.

InfiniteIO's solution can answer up to 90% of storage requests, responding in only tens of microseconds of latency. Studies show that 70% to 90% of the storage requests coming from the network are for file metadata. This includes data such as file name, file size, file creation date, the date the file was last updated, the list of files in a folder or directory, and who has access to the file.

By responding to these requests, InfiniteIO can offload this work from NAS systems. InfiniteIO dramatically accelerates the response to these requests.

Once in operation, InfiniteIO makes the NAS systems more efficient, and provides greater overall performance. With InfiniteIO, the majority of the storage requests are handled before they get to the NAS systems.

While InfiniteIO is in operation, it learns a great deal about the enterprise's files. It learns what files are and are not actively being used. Enterprises can then transparently tier files to more cost-effective storage.

The key here is that all storage still appears to be local to applications, regardless of storage location. This means the enterprise can free up prime Tier-1 storage and migrate cold files, while hot, i.e. active, files remain in Tier-1 storage. This processes isn't disruptive to workloads or applications, and requires no changes to the enterprise storage infrastructure.

> Once in operation, InfiniteIO makes the NAS systems more efficient, and provides greater overall performance. With InfiniteIO, the majority of the storage requests are handled before they get to the NAS storage systems.

Let's explore how InfiniteIO works. DPI allows InfiniteIO to analyze data as it travels over the network, enabling InfiniteIO to accelerate storage requests by responding directly to the majority of data requests without having to go back to the NAS systems for data. Enterprise users continue to connect to storage using the storage system's IP address.

> InfiniteIO's solution is based on a shared-nothing cluster architecture offering "redundant everything" that provides enterprise resiliency, without a single point of failure.

Many members of InfiniteIO's engineering team came from TippingPoint, a provider of network intrusion prevention systems. They used DPI to inspect IP packets as they were flowing by on the network. It was then possible to detect intruders and quarantine their network flow. This approach is commonly used in network security applications.

InfiniteIO refined this approach and focused it on inspecting every NFS storage request, in real time, to quickly understand the environment. Metadata from active files is held in memory, analyzed and used to accelerate and consolidate storage. It also enables real-time analytics and transparent tiering of files.

SCALING TO BILLIONS OF FILES WITHOUT DISRUPTION

InfiniteIO's solution is based on a shared-nothing cluster architecture offering "redundant everything" that provides enterprise resiliency, without a single point of failure.

It's possible to drop this solution into the network just like a Layer 2 network switch, without requiring the enterprise to experience downtime or any workflow changes. Once installed, it immediately begins to "crack" every NFS message, learning active files, and responding to metadata requests on behalf of the storage behind it.
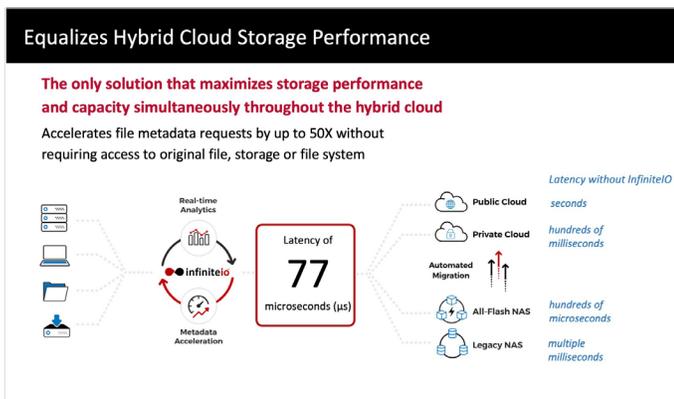
Figure 2: InfiniteIO can make your storage as fast as flash-based systems.

Then it initiates a one-time scan of the storage and grabs the metadata for all the files on the storage servers. Once complete, it can serve *all* of the metadata requests and offload this task from the servers. This makes all files "hot."

Enterprises are then able to transparently tier inactive files, and storage administrators can create storage policies that define what should be considered inactive or "cold" data.

EASY POLICY CREATION

This makes it easy to migrate all files that haven't been actively used in the last 180 days, for example. It would also be easy to select project data that's no longer in use by file name or directory, and migrate it immediately without having to wait for it to "age out."

Exceptions are also supported. For instance, it's possible to migrate all of the files that haven't been active in a specific time period to object storage, with the exception of those in specific directories or created by a specific person or group. It's easy to create multiple policies to represent the needs of individuals, groups, or business units.

> It's easy to create multiple policies to represent the needs of individuals, groups, or business units.

InfiniteIO offers a great deal of flexibility for policy management. Once the policies are created, they can be immediately executed to match inactive files and start migrating them. Or, by default, they can be run at midnight.

Another advantage is that there's never a need to rescan the storage environment to pick up metadata changes. InfiniteIO always has the current metadata for all files. Policies simply examine this metadata. All migrated files continue to look like they're local to workloads and applications, no matter where they are.

Once the original scan's been completed, InfiniteIO offloads all metadata requests. Requests for actual data are handled by the NAS server. In the case of writes, the NAS provides updated metadata that is analyzed by DPI, allowing InfiniteIO to always be up-to-date and consistent with the storage system behind it.

If a read or write request is made for data that's been migrated, InfiniteIO will simply bring that data back. It will be placed back in the original NAS system in its original location. The application has no idea that this occurred, since everything works as it did before. The requested data just shows up.

This is all accomplished without capacity-based pricing. Enterprises simply purchase an appliance once and then manage as much data as needed.

SUPER FAST PERFORMANCE

InfiniteIO is super fast. It can respond to a typical metadata request in under 100 microseconds (µs), as **Figure 2** shows. This is as fast as flash-based systems, regardless of where the file is actually stored.

InfiniteIO also protects migrated data from unnecessary recall. All ingress charges are eliminated, and the performance level is maintained for both local and migrated files for 70% to 90% of storage requests.

Although the original goal was to protect the metadata for local and migrated storage, the result is it can equalize and accelerate the performance for storage access for both on-premises and migrated files. By allowing the migration of inactive data to the cloud, enterprises receive another benefit, because backup of active files can be accomplished faster, reducing backup window length.

One InfiniteIO customer saw 30% to 40% faster backup prior to tiering or migrating their data. This is because backup applications make a surprising number of metadata requests. Backup software is constantly checking file access, determining what to back up. InfiniteIO was absorbing all of those requests and responding in microseconds. InfiniteIO's real-time analytics made it possible for this customer to understand what was happening. It was an interesting and unexpected benefit of their use of InfiniteIO.

INSIGHTS FROM REAL-TIME ANALYTICS

InfiniteIO offers a dashboard providing real-time analysis of its functions. It offers a quick overview of system health, real-time performance statistics, storage utilization and latency graphs, number of NFS calls per second, and estimated savings. See **Figure 3**.

> One InfiniteIO customer saw 30% to 40% faster backup prior to tiering or migrating their data. This is because backup applications make a surprising number of metadata requests.

InfiniteIO's machine learning capabilities keep its metadata up to date and accurate without requiring re-scans of the storage. This makes it easy to see how much data is active and how long it's been since additional data has been used.

This analysis also makes it possible for the enterprise to realize significant savings by moving inactive data to more cost-effective storage. It also reports, in aggregate, how much storage a migration will save.

FULFILLING THE PROMISE OF SIMPLE, FAST HYBRID CLOUD STORAGE

InfiniteIO lives up to its promise of making hybrid cloud storage simple and fast. Once installed, it immediately goes to work. There's no software to install, no configuration changes or special mounts required. InfiniteIO doesn't need to "stub" or use
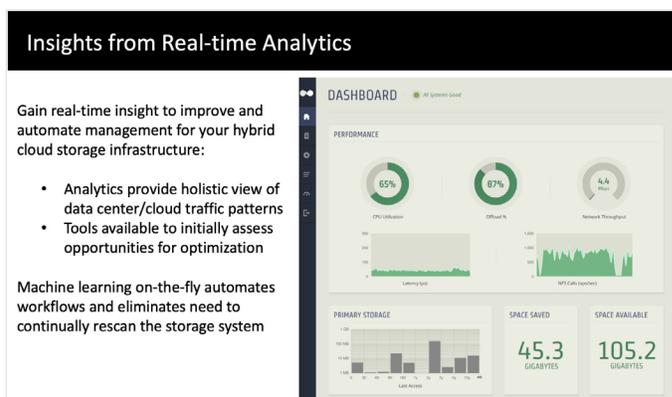

Figure 3: The InfiniteIO dashboard.

dynamic links to manage the hybrid storage environment. The file metadata is abstracted.

Furthermore, there's no need for continuous scans of the storage behind the InfiniteIO system. Inactive data is just transparently migrated to object storage and brought back on demand. If you're looking for a better way to do storage, check them out.

Our Take

Undertaking digital transformation for your business means re-thinking how you do everything. That even applies to storage.

For storage to work in the cloud era, it needs to be faster than ever, with the lowest latency possible. Moving data into and out of those clouds is both time- and resource-consuming; those are two things most IT organizations don't have in excess.

That's why you should look hard at anything that frees up resources and boosts speed. The InfiniteIO solution is elegant and simple, and should be on the short list of any company looking to speed up their storage for hybrid cloud scenarios.

IN PARTNERSHIP WITH:

infiniteio™

For more information, visit infinite.io