

## Accelerate Your High-Performance Scale-out ONTAP Environment by Load-Balancing with DataSphere

Machine learning load balances across clusters while a metadata engine places the right data, in the right place, at the right time — automatically and without application disruption.

### THE CHALLENGE

Until recently, Network Attached Storage (NAS) was rarely considered a solution for mission-critical and high-performance applications. NAS offers a rich environment for file-based management, which is often preferred by enterprise IT, but falls short when compared with the reliability and block-level performance of Storage Area Networks (SAN). With the advent of flash memory, NAS solutions now possess the scale-up performance needed to compete with SAN. However, NAS is still governed by metadata operations, often lacks leading scale-out capabilities, and is limited to proprietary clustering solutions that do not allow for scaling across vendors and generations of clusters.

While the Network File Systems (NFS) protocol brings simplicity of use and management at the file level, clients (applications and their storage stacks) must communicate with a NAS device using NFS operations to retrieve information about or gain access to a file. These commands are sequential operations that add performance overhead as they act on metadata requests or on a file. Traditional NAS performance is therefore limited by how fast it can service sequential NFS protocol requests and file activity.

### SOLUTION OVERVIEW

#### Offload metadata to dramatically accelerate performance of ONTAP FAS clusters

DataSphere is a metadata engine that leverages machine learning to separate and offload the architecturally rigid relationship between applications and where their data is stored. Through data virtualization, DataSphere abstracts the data path from the metadata path to connect different storage resources across a global namespace, effectively virtualizing data.

Offloading metadata delivers predictable, low-latency operations by guaranteeing that metadata operations do not get “stuck” in the queue behind data requests. Rather than having to wait for sequential operations to complete, DataSphere enables parallel access with the latest optimizations of the standard NFS v4.2 protocol at the application client, improving the performance for metadata and small file operations.

#### Extend and tier a scale-out an ONTAP cluster

DataSphere can easily create a powerful scale-out architecture that load balances shares and files across controllers to accelerate application performance and reduce hotspots. When configured with DataSphere, a group of FAS arrays can

### WITH PRIMARY DATA AND NETAPP, CUSTOMERS CAN:

- Load balancing across clusters
- Overcome performance bottlenecks
- Automate load balancing at the file level, non-disruptively
- Seamlessly add StorageGRID Webscale as a storage tier to your infrastructure

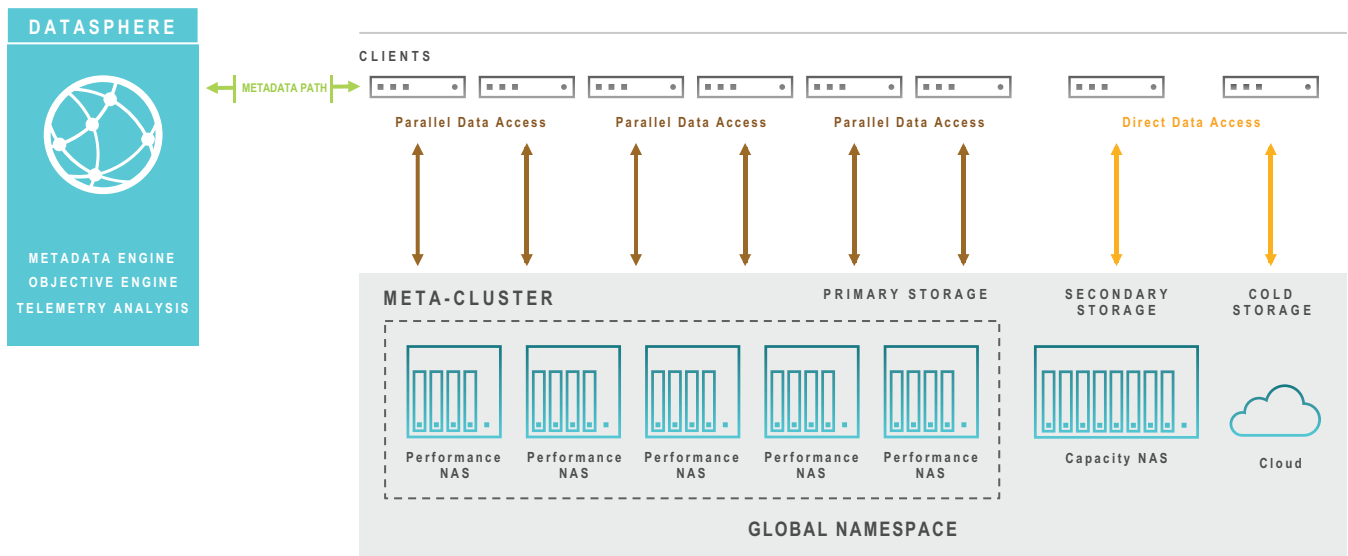
### SOLUTION COMPONENTS

#### NetApp Products

- NetApp® ONTAP
- NetApp StorageGRID
- Webscale storage appliance with E-Series storage hardware

#### Primary Data Products

- DataSphere – Machine learning metadata engine (Enterprise or LOB)
- DataSphere Extended Services – Data mover



**Figure 1:** Automatic tiering of live data within arrays, across clusters, and into S3 object storage end-points, for 7-mode and Clustered Data ONTAP.

be logically pooled together as a scale-out cluster under a global namespace. With its unique metadata engine and out-of-band architecture, DataSphere creates a comprehensive scale-out environment from a pool of FAS controllers. Using a global namespace, DataSphere presents a group of volumes as a cluster, enabling data to be load balanced across the volumes, accelerating app performance through parallel data access. This scale-out “meta-cluster” can extend into the cloud or other S3 object storage to deliver data redundancy and cost-savings at the same time (Figure 1).

**SOLUTION ARCHITECTURE OVERVIEW**

Primary Data’s DataSphere provides a storage metadata engine can be installed in physical or virtual environments, allowing files to reside anywhere within a global namespace, on any managed storage, irrespective of the storage protocol. NetApp ONTAP and StorageGRID Webscale (SGWS) provide storage solutions at various performance points that DataSphere will intelligently place data on based on the application need.

DataSphere’s objective language and DSX data mover work together to provide dynamic location placement and movement of files.

The location and placement of a file is based on user-defined objectives in combination with the storage characteristics (Size, owner, cost, age, IOPS, etc..) and user actions, such as reading a file. This enables enterprises to use SGWS to keep unused data accessible while preserving storage capacity and performance on the FAS storage for more active files. This means that SGWS can be viewed and used as a low-cost, highly reliable storage tier to the existing NAS storage. In addition to the benefits of tiering to SGWS, DataSphere enables additional savings and agility by leveraging the SGWS as a unique tier to store cold data, manage snapshots, and enable on-demand usage models.

**ABOUT PRIMARY DATA**

Primary Data develops intelligence and automation software for enterprise data management across on-premises IT infrastructure and into the cloud. Its DataSphere platform combines metadata management and machine learning to move the right data to the right place at the right time across a global namespace, automatically and without application disruption.