Primary Data

NetApp®

DataSphere Intelligent Data Management and NetApp StorageGRID Webscale

Machine-learning metadata engine puts the right data in the right place, at the right time

The Challenge

To be a competitive business is to be a technology company and data is the currency that drives success. As enterprises continue down the path of digital transformation and cope with the resulting explosive growth of data, they look for more sophisticated data management tools while keeping costs in check.

Enterprise customers have more choice than ever before when selecting storage to meet their applications' needs. With different products offering a wide range of performance capabilities, protection features, and prices, customers can strategically choose the option that best matches expected data demands. However, storage and server administrators are forced to take a bottom-up approach to meeting data requirements—by purchasing and allocating storage based on expected peak performance and protection requirements. This overprovisioning of infrastructure is expensive, inefficient, and still prone to problems, because unanticipated activity can grind critical workloads to a halt.

Most enterprise data is stored on NAS file systems (NFS and SMB), but object storage is a new option that features lower costs and high-capacity density. It delivers performance suitable for colder data when accessed on-premises as data is passed through the Amazon Simple Storage Service (S3) protocol. Still, there are challenges that impede the broader adoption of an object tier. Primarily, enterprises don't want to change their applications to read and write to S3. But they also need to automate the detection of cold data and move that data on demand between file-based and object storage, without disrupting the running of applications.

So, how does an enterprise easily integrate a cost-effective object storage tier with high-performance NAS without disrupting application workloads, while increasing IT's confidence of meeting SLAs?

The Solution

NetApp and Primary Data want to close the gap between business and IT. The two companies are working together to enable IT to define data management objectives with fine granularity, improve the utility of storage infrastructure, reduce costs, improve application performance, and automatically load-balance data across the infrastructure to meet SLA requirements.

Primary Data's data management platform, DataSphere, automatically and nondisruptively moves data to the right storage to meet these objectives, so that desired service levels are always met. DataSphere employs machine-learning software to build intelligence into how an enterprise

HIGHLIGHTS

- Seamlessly add the NetApp® StorageGRID® Webscale storage solution as an on-premises storage tier.
- Move live data back from object stores without disrupting application access or reconfiguring apps to support object storage.
- Reduce costs with automatic data deduplication and compression.
- Optimize performance and costs by restoring only the files needed.
- Free up NetApp ONTAP® storage capacity by moving cold data to StorageGRID Webscale.

SOLUTION COMPONENT

NetApp Products:

- StorageGRID Webscale
- All Flash FAS systems running ONTAP software

Primary Data Products:

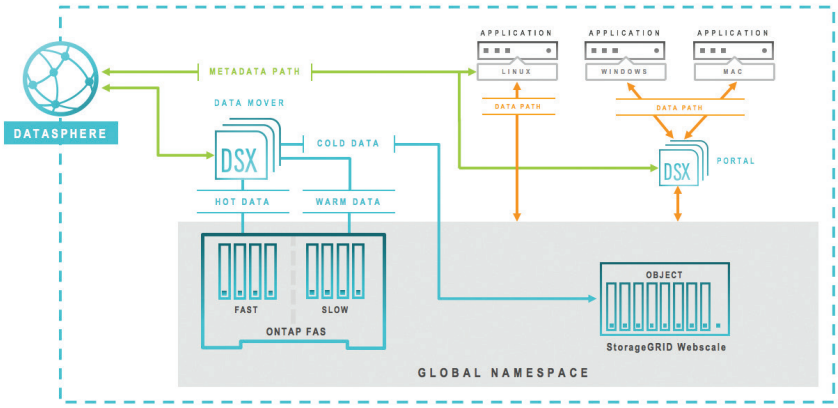
- DataSphere—machine-learning metadata engine
- DataSphere Extended Services (DSX)—Data Mover

automates data management across its IT infrastructure, both on-premises and in the cloud. While adding awareness between applications and infrastructure, DataSphere virtualizes data and creates a global namespace. DataSphere makes heterogeneous datastores such as NetApp FAS and StorageGRID Webscale simultaneously available to all applications without requiring any changes to applications.

DataSphere helps enterprises reduce the TCO for managing and storing data by defining and accounting for data movement based on user-defined price-to-performance targets. With native support for the S3 protocol, DataSphere makes it easy to use StorageGRID Webscale in a global namespace as a low-cost, highly reliable storage resource to tier cold data, manage snapshots, offer data governance, and enable on-demand usage models.

DataSphere moves live data across storage tiers without disrupting application workloads. This means that data can be archived into StorageGRID Webscale and transparently brought back to a performance tier if it becomes active again. IT can deliver cost savings with no overhead in complexity or performance impact.

Primary Data's DataSphere provides a metadata engine that 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 FAS systems and StorageGRID Webscale provide storage solutions at various performance points, and DataSphere intelligently places data on these solutions according to application need.



DATASPHERE MANAGES DATA MOBILITY FROM ONTAP TO STORAGEGRID WEBSCALE BY SEPARATING METADATA FROM THE DATA PATH.

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 storage characteristics (size, owner, cost, age, IOPS, and so on) and user actions, such as reading a file. This enables enterprises to use StorageGRID Webscale to keep unused data accessible while preserving storage capacity and performance on the FAS system for more active files. StorageGRID Webscale can therefore be used as a low-cost, highly reliable storage tier to the existing NAS storage. In addition, DataSphere offers savings and agility by using StorageGRID Webscale as a unique tier to store cold data, manage snapshots, and enable on-demand usage models.

> "With DataSphere, we were able to improve the economics of how we run our business. We get more out of our IT investments and can consistently provide higher SLAs to our customers while maintaining healthy margins."
>
> Director of IT
> Service Provider

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.

ABOUT NETAPP

NetApp is the data authority for hybrid cloud. We provide a full range of hybrid cloud data services that simplify management of applications and data across cloud and on-premises environments to accelerate digital transformation. Together with our partners, we empower global organizations to unleash the full potential of their data to expand customer touchpoints, foster greater innovation and optimize their operations. For more information, visit www.netapp.com. #DataDriven