Solution Brief

StorageGRID for Internet of Things

Gain insights in real time from vast amounts of machine-to-machine data

Key Customer Benefits

Reduce Complexity of Data Repositories

- Store data globally and access it locally with a true global namespace.
- Implement erasure coding and remote copy policies.
- Manage policies and monitor storage with a single pane of glass.

Optimize Workflows

- Confirm that data is in the right place at the right time.
- Enable applications to access content directly with Amazon S3 cloud interface.

Minimize Cost of "Cradle-to-Grave" Data Management

- Improve disaster recovery economics with multisite dispersion.
- Automatically migrate aging data to tape and the cloud.
- Instantly detect and replace failing objects.
- Deploy sites that are built on software-defined storage or appliances.

The Challenge

While every mention of Internet of Things (IoT) seems to start with the game of who can cite the highest possible number of connected IoT devices, the focal point should really be how to work with the vast amounts of new data IoT devices are generating. One well-known car company built up about 14 petabytes of storage over several years. But in just a few short months, its driverless car program created an additional 4.5 petabytes!

This explosive data growth will make an already challenging IT problem – organizing the data to make it valuable – even greater. Advanced data management and organization is essential in effectively assimilating and using the flood of data that will come from IoT. However, given the scale and scope of this data, we need new thinking and approaches beyond traditional solutions. Even if your organization seems to be keeping abreast of managing growing IoT customer data, growth rates are so high that you could quickly outpace your infrastructure. That means you have to think about implementing the right solution today to have a viable infrastructure for the long term.

The Solution

The age of IoT requires a fundamental change to older, traditional data storage and management approaches. NetApp has developed a solution that bridges current and future requirements, allowing your organization to create value from IoT without storage bottlenecks or limitations. The answer is NetApp® StorageGRID® object storage. Unlike file systems, object stores group files and their metadata into objects that can be coherently accessed by different data management systems, file delivery systems, and other production and distribution applications.

Object stores are designed to support multiple storage nodes across Internet-connected sites. With the intelligent policy engine of StorageGRID, you can choose to either erasure-code your object across sites for geo-resiliency or copy your object to remote sites to minimize WAN latency and cost.

As content ages, policies can automatically move content to tape or the cloud. This capability eliminates the need to move or delete thousands or millions of files manually, and it improves the economics of your IoT data repositories.

NetApp®

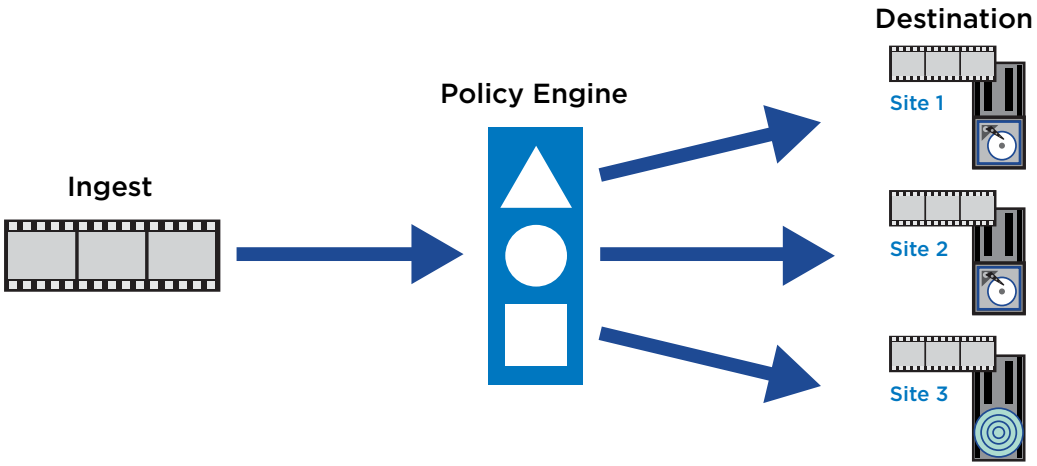


Figure 1) StorageGRID policy engine.
Content workflow and information lifecycle management (ILM) decisions drive content placement.

Reduce Complexity of Content Repositories

Whether an object is stored on a local or a remote node or is erasure-coded across several nodes within the object store, it is accessed from a single namespace. With massive scalability across this namespace, object stores free your media applications from the task of finding and moving files between storage volumes.

StorageGRID enables you to establish highly granular, flexible data management policies that determine how data is stored and protected. When developing and enforcing policies, the object store examines a wide range of characteristics and needs, including performance, durability, availability, geographic location, longevity, and cost.

Optimize Workflows

StorageGRID is a data management system. It provides sophisticated policy-driven functionality to support your asset management systems. Policies can be enforced at ingest, at rest, upon read, after metadata update, at object disposition request, or after ILM policy changes. Policies can initiate object replication, erasure coding, caching, tiering, and archiving to tape or the cloud.

This storage management, monitoring, and migration functionality delivers a better experience for all the workflows related to IoT. While asset management systems support specific workflows within an enterprise, an object store confirms that the data that those systems need is in the right place at the right time.

"As we tune the solution, customers see quite a remarkable impact on their network performance and storage performance."

Andrew Sjoquist
Founder of ASE IT

By erasure-coding, moving, or copying objects to sites as production and distribution access demands, the object store relieves your media asset management systems, users, and administrators from manually performing these tasks. When an object is stored, it is seen and accessed as one object by all clients, regardless of where it is or how many copies exist.

Erasure Coding

Erasure coding is a method of data protection in which data is broken into fragments, is expanded and encoded with redundant data pieces, and is stored across different locations or storage media. Erasure coding is used in object stores instead of traditional RAID because it reduces the time and overhead that are required to reconstruct data. In an object store with geo-dispersed erasure coding, data can be retrieved from a subset of the sites where the original object was dispersed. This feature allows object access even during temporary or permanent failure of one or more sites.

- Cloud integrated: mirroring, notification, search
- Software-defined Object Storage
- Geo-redundant and Multi-Site
- Policy-based to simplify management at scale
- S3, Swift, NFS and CIFS access
- Durable, low-cost cloud storage for Active Archive

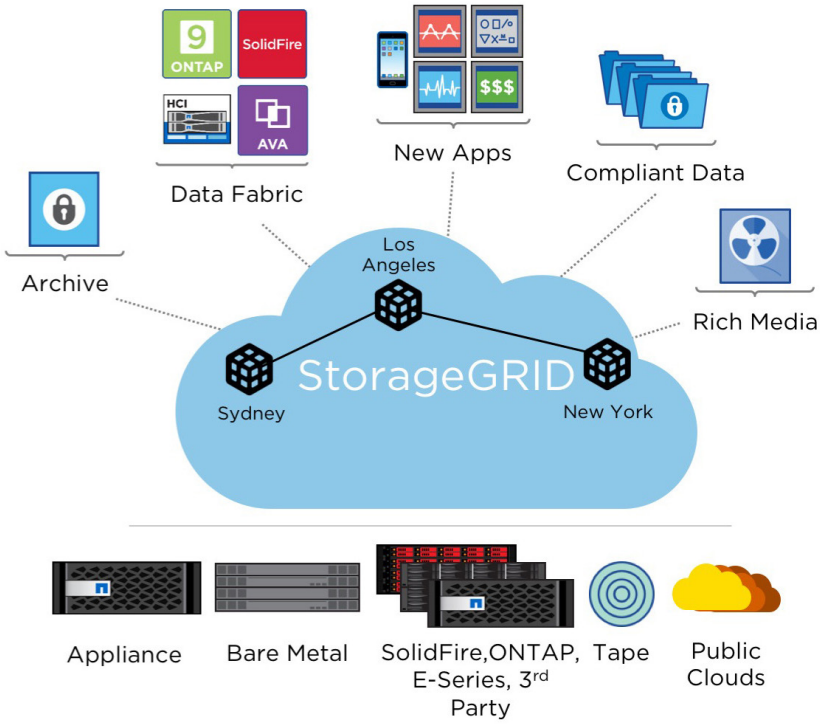


Figure 2) StorageGRID
Enable your data to be in the right place at the right time with the right performance and cost.

Minimize Cost of Cradle-to-Grave Data Management

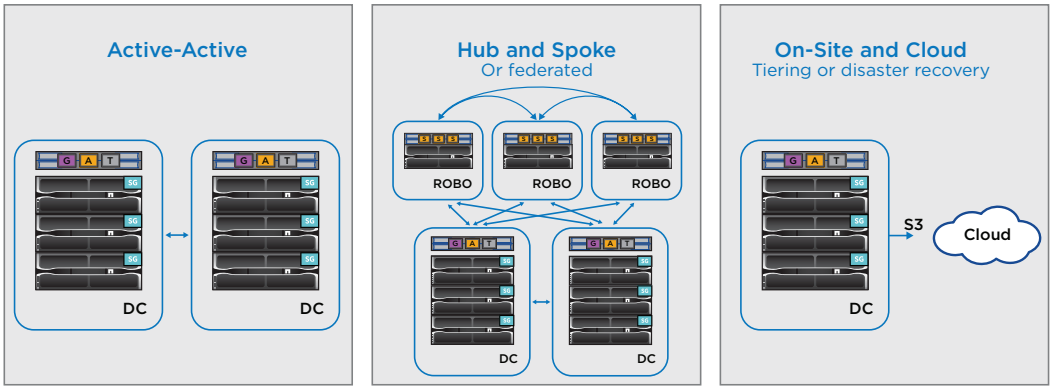
Providing ubiquitous content access often requires storing copies of the same asset in multiple places. This approach not only increases complexity, it also requires additional storage capacity. With StorageGRID, you can minimize the capacity needed while optimizing for data resiliency, data access, and distribution requirements through policy-driven automation.

The policies that you set to move and copy objects can also include deletion criteria. For example, copies of objects can be automatically deleted after set periods of inactivity, reducing capacity requirements.

The value of content changes over time, as does the cost of storing it. High-value content requires high-performance storage to keep it readily available. As content ages and is accessed less often, policies can automatically reduce replication and move that content to less expensive disk arrays, tape, or public cloud infrastructures. Management and maintenance of file copies, migrations, and deletions are driven by the StorageGRID policy engine.

The content in your repositories will live longer than the media that it is stored on, and StorageGRID will greatly reduce your migration challenges and costs in the years to come. The policy engine will drive the process of moving millions of objects from aging media to newer, lower-cost, higher-capacity media in the future.

As an 11th-generation object store with over 20 years of production deployment in the most demanding industries, StorageGRID is the platform that you can trust with your growing IoT data.



3) Flexible support for multisite resiliency strategies.
You can enable your chosen resiliency topology at the object storage level.

FEATURE	DESCRIPTION
Cloud Integration	• Cloud integrated object storage solution, with metadata search integration • S3 compatible platform, with support for self-service, allowing developers and end users to use object storage efficiently
Interfaces	• RESTful HTTP APIs include Amazon Simple Storage Service (S3) and OpenStack Swift • Standard network protocols through a NAS bridge include NFS and CIFS
Scalability	• Billions of objects • 100's of petabytes
Data integrity	• Creates a digital fingerprint as data is ingested • Offers multiple interlocking layers of integrity protection, including hashes, checksums, and authentications • Provides data object integrity verification on ingest, retrieval, replication, and migration, and at rest; suspect objects are automatically regenerated • Provides geo-distributed erasure coding for cost-effective data integrity protection across sites
Data availability	• Fault-tolerant architecture supports nondisruptive operations, upgrades, and infrastructure refreshes • Load balancing automatically distributes workloads during normal operations and failures • NetApp AutoSupport® technology automatically alerts NetApp Support engineers for proactive issue resolution • Node-level erasure coding further improves single-node availability (with NetApp E-Series Dynamic Disk Pools)
Deployment options	• Software-defined storage running on VMware and Docker deployments and managing either NetApp or third-party storage arrays • Three NetApp appliances with node-level erasure coding, Dynamic Disk Pools, and AutoSupport, providing 5 nines of availability at the storage-array level • For more on deployment options, visit www.netapp.com/storagegrid

Table 1) StorageGRID key technical features.

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