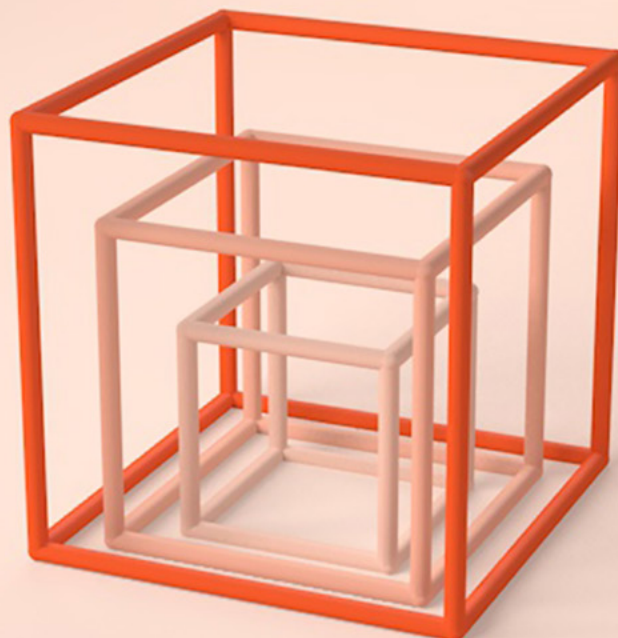


Build your data lake on NetApp StorageGRID and Snowflake



A flexible and easy-to-use data warehouse platform for your object storage.

Not long ago, enterprises had to purchase expensive, specialized hardware to build a data warehouse in their data center. Snowflake solved this problem by building a cloud-native platform for enterprises to use. The Snowflake Data Cloud is software as a service that runs on AWS, Google Cloud, and Microsoft Azure infrastructure, eliminating the need to select and manage hardware and software in a data center.

Snowflake decouples compute and storage, so you can use and pay for them separately, and you can optimize compute, storage, and services resources to reduce costs. Snowflake also allows seamless data sharing through an easy-to-use UI. By providing a simple, flexible data warehouse platform, Snowflake helps organizations get the most value out of their data.

Today's challenge with moving data to the cloud

Over time, an organization builds up a huge amount of data from various sources into its data lake, which might be stored across public and private clouds. Enterprises typically can't move all their data to a public cloud for security, governance, and compliance reasons, or because the dataset is too large. Snowflake is a cloud-only platform, so to harness its capabilities, organizations wanted to use Snowflake to augment their data lake rather than entirely moving their data into Snowflake cloud, getting the best of both the public-cloud and private-cloud worlds.

To provide enterprises with this flexibility, Snowflake introduced support for External Tables, which let you make datasets accessible without moving them into the Snowflake cloud. With External Tables, users can query data stored in files in external storage as if it were inside Snowflake itself. External Tables let organizations use Snowflake services without compromising data sovereignty.

Enabling Snowflake External Tables on StorageGRID

StorageGRID is an enterprise-grade, on-premises object storage solution that supports the native Simple Storage Service (S3) API. StorageGRID is software defined, which means that you can run it on different platforms—bare metal, VMware-based environments, or NetApp purpose-built appliances—and mix platforms in a grid. In a single namespace, StorageGRID can scale up to 16 data centers. StorageGRID offers massive S3 object storage and dynamic data management, enabling you to run next-generation workflows on premises, alongside your public cloud.

By integrating with StorageGRID, Snowflake lets you add S3 on-premises object storage as a data source while keeping the ease of use and governance that Snowflake provides. Enabling Snowflake on StorageGRID is as simple as providing Snowflake with access credentials for the S3 bucket that contains your data files. When the file is staged, users can query data stored in files in StorageGRID as if it were in a Snowflake database.

Key Features

- Augment your NetApp® StorageGRID® object storage with Snowflake services.
- Keep control of your data locality.
- Easily share data across geo locations through StorageGRID multitenancy.
- Take advantage of the StorageGRID information lifecycle management (ILM) engine to optimize resources for your active and inactive datasets.
- Scale StorageGRID performance and capacity without downtime.

StorageGRID lets you manage the data lifecycle for different workloads within a single grid and scale with low-touch operations. These abilities help enterprises run their data lake efficiently and cost-effectively. Because StorageGRID can connect to the major cloud services, you have the flexibility to build data architectures that best suit your needs.

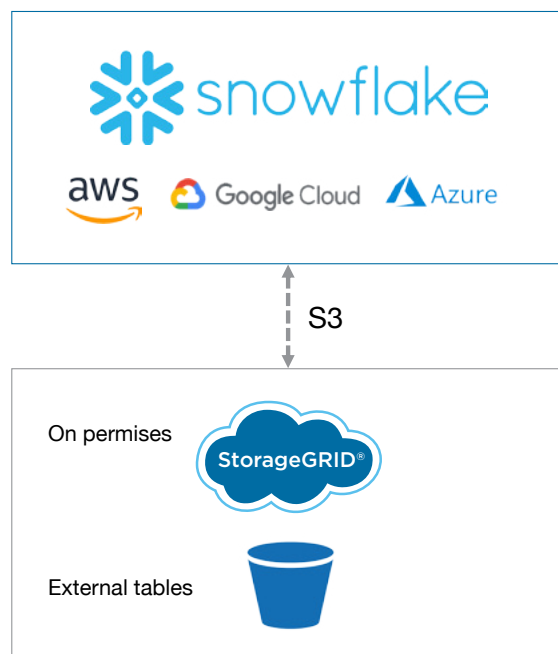


Figure 1: Using Snowflake External Tables to gain value out of your data on StorageGRID.