

AGENTIC LAKEHOUSE ON NETAPP STORAGEGRID WITH DREMIO



Overview: The Foundation for Enterprise AI

Modern enterprises are pushing aggressively to operationalize Agentic AI and machine learning, but progress is often constrained by legacy data architectures. While object storage holds the most complete and up-to-date enterprise data, it is rarely optimized for the performance, concurrency, and semantic consistency required by AI-driven analytics and model training. The result is slow iteration, duplicated data pipelines, and fragmented intelligence across teams.

The joint solution from Dremio and NetApp StorageGRID addresses this challenge with a high-performance Agentic Lakehouse architecture. By combining the massive scale, durability, and efficiency of StorageGRID with Dremio's high-performance query engine, open catalog, and semantic and AI-native layers, enterprises gain direct, governed access to data where it lives. This open, modern foundation accelerates data-to-insight, supports high-velocity AI and analytics workloads, and removes the need for proprietary warehouses and complex ETL pipelines.

The Integrated Solution: An Agentic Lakehouse for StorageGRID

This powerful combination creates a cloud-like lakehouse experience on-premises, optimized for the demands of high-concurrency analytics workflows.

NetApp StorageGRID: The Scalable and Durable Data Foundation

NetApp StorageGRID is the trusted, software-defined object storage platform, providing the cost-effective, durable, and scalable foundation required for petabyte-scale AI datasets.

Dremio: The High-Performance Agentic Lakehouse

Dremio is the data lakehouse platform that transforms StorageGRID into a high-performance system for BI, SQL analytics, and demanding AI/ML workloads.

Key Advantages of the Integrated Architecture

1. High-Speed Analytics and Feature Engineering

Dremio's Data Reflections and Apache Arrow Engine are the key to bypassing object storage latency, delivering sub-second, interactive query performance directly on NetApp StorageGRID. This accelerated access is vital for feature engineering, the process of preparing data for ML models allowing data scientists to rapidly iterate without waiting hours for data movement or query execution.

2. Semantic Consistency for Enterprise AI

The Dremio Semantic Layer provides a governed, unified view of all data within StorageGRID. This ensures that the features, metrics, and definitions used for model training and business reporting are consistent, minimizing errors and improving the trustworthiness of AI outputs and business intelligence.

3. Maximum Cost-Efficiency and Openness

The solution leverages the extreme cost-efficiency of NetApp StorageGRID for petabyte-scale data storage combined with Dremio's open architecture. This avoids proprietary data warehouse licensing fees and significantly reduces the total cost of ownership (TCO) while offering flexibility for future data initiatives.

KEY BENEFITS

- **Exabyte-Scale Architecture:** Effortlessly scales to accommodate massive datasets for large-scale AI model training.
- **Multi-Site Durability:** Ensures data integrity and availability with flexible protection across geographically dispersed sites.
- **S3 API Compatibility:** Provides a seamless interface for Dremio and other cloud-native tools, simplifying integration.
- **Policy-Driven Tiering:** Automates data lifecycle management, optimizing costs by moving cold data to lower-cost tiers while maintaining fast access.
- **AI Layer:** Enhances data interaction, improving performance and reducing manual effort.
- **Semantic Layer:** Ensures data accuracy and consistency with standardized definitions.
- **Reflections & Query Engine:** Pre-aggregates and optimizes data, delivering up to 100× acceleration and low-latency analytics.

Use Cases with Dremio and NetApp StorageGRID

Query and Understand All My Object Data Instantly

- Run fast SQL directly on data in StorageGRID
- No ETL, no data copies, no new warehouse
- Same definitions and metrics for every team

Power AI Models with Fresh, Trusted Features

- Serve features directly from object storage
- Same logic for training and inference
- No separate feature store to manage

Automatically Catalog and Maintain My Lakehouse Tables

- Discover and organize tables automatically
- Handle table maintenance and optimization for me
- Keep schemas and data consistent as things change

One Analytics and AI Layer Across Hybrid Environments

- Query data across on-prem and distributed sites
- Central governance with local performance
- No vendor lock-in or forced cloud move

Analyze Logs and Operational Data at Scale

- Query massive log and event datasets in place
- Build dashboards and troubleshoot faster
- No special-purpose analytics stack required





Next Steps

To learn more about how Dremio and NetApp StorageGRID can accelerate your AI and data analytics journey, please connect with a solutions expert or review the following resources:

[Dremio Agentic Lakehouse](#)

[NetApp StorageGRID](#)



[Contact Us](#)

About NetApp

For more than three decades, NetApp has helped the world's leading organizations navigate change – from the rise of enterprise storage to the intelligent era defined by data and AI. Today, NetApp is the Intelligent Data Infrastructure company, helping customers turn data into a catalyst for innovation, resilience, and growth. At the heart of that infrastructure is the NetApp data platform – the unified, enterprise-grade, intelligent foundation that connects, protects, and activates data across every cloud, workload, and environment. Built on the proven power of NetApp ONTAP, our leading data management software and OS, and enhanced by automation through the AI Data Engine and AFX, it delivers observability, resilience, and intelligence at scale. Disaggregated by design, the NetApp data platform separates storage, services, and control so enterprises can modernize faster, scale efficiently, and innovate without lock-in. As the only enterprise storage platform natively embedded in the world's largest clouds, it gives organizations the freedom to run any workload anywhere with consistent performance, governance, and protection. With NetApp, data is always ready – ready to defend against threats, ready to power AI, and ready to drive the next breakthrough. That's why the world's most forward-thinking enterprises trust NetApp to turn intelligence into advantage.

Learn more at www.netapp.com or follow us on [X](#), [LinkedIn](#), [Facebook](#), and [Instagram](#).

©2026 NetApp, Inc. All Rights Reserved. NETAPP, the NETAPP logo, and the marks listed at www.netapp.com/TM are trademarks of NetApp, Inc. Other company and product names may be trademarks of their respective owners. SB-4496-0326

