

## SOLUTION BRIEF

# Modern, Powerful, and Efficient Data Lakehouse Infrastructure with Dremio and Pure Storage

## Enable Effective and Impactful Data Consumption

Developing a modern data infrastructure is a complex task, as enterprises aim to extract maximum value from their data while minimizing complicated data pipelines. Traditional data warehouses often lead to data silos, vendor lock-in, inconsistent sources of truth, and convoluted processes. With the exponential increase in data volume and variety, legacy data infrastructures struggle to scale and deliver the necessary access and performance needed for modern analytics. Data administrators face the challenge of building modern data lakes while maintaining critical legacy data pipelines. Pure Storage has partnered with Dremio, the unified data lakehouse platform, to help enterprises build a future-proof, scalable, and efficient data infrastructure.

### Unrivaled Scalability for the Iceberg Data Lakehouse

Pure Storage FlashBlade is an enterprise-grade, on-premises solution that supports the native Amazon S3 API. It offers high performance object storage and dynamic data management, enabling next-generation workflows on-premises or adjacent to your public cloud. The solution's industry-leading data management policy engine optimizes performance and durability while adhering to data locality requirements. Pure Storage is extremely scalable, supporting low-touch, nondisruptive expansions, and can store billions of objects.

Dremio helps data teams deliver faster access to their data. Deploying the Dremio Unified Lakehouse Platform on top of FlashBlade object storage maximizes the value of your enterprise's data while exercising full control of data placement, lifecycle, and outcomes.

Together, Pure Storage and Dremio empower enterprises to take full control of their data management, enabling impactful analysis and business intelligence while ensuring optimal data placement to maximize value. Pure Storage's information lifecycle management policies allow granular control over data retention, placement, and tiering to lower-cost object storage. As Pure Storage gives you unrivaled control over your data lake, Dremio facilitates simple access to the data across the entire enterprise.



#### Accelerated Queries

Pure Storage and Dremio deliver real-time data access and interactive-speed response times, ensuring swift processing and analysis of complex datasets for business insights.



#### Seamless Scalability

Pure Storage and Dremio provide seamless scalability, handling highly parallel workloads and growing data volumes efficiently, improving analytic runtimes and complex data model development.

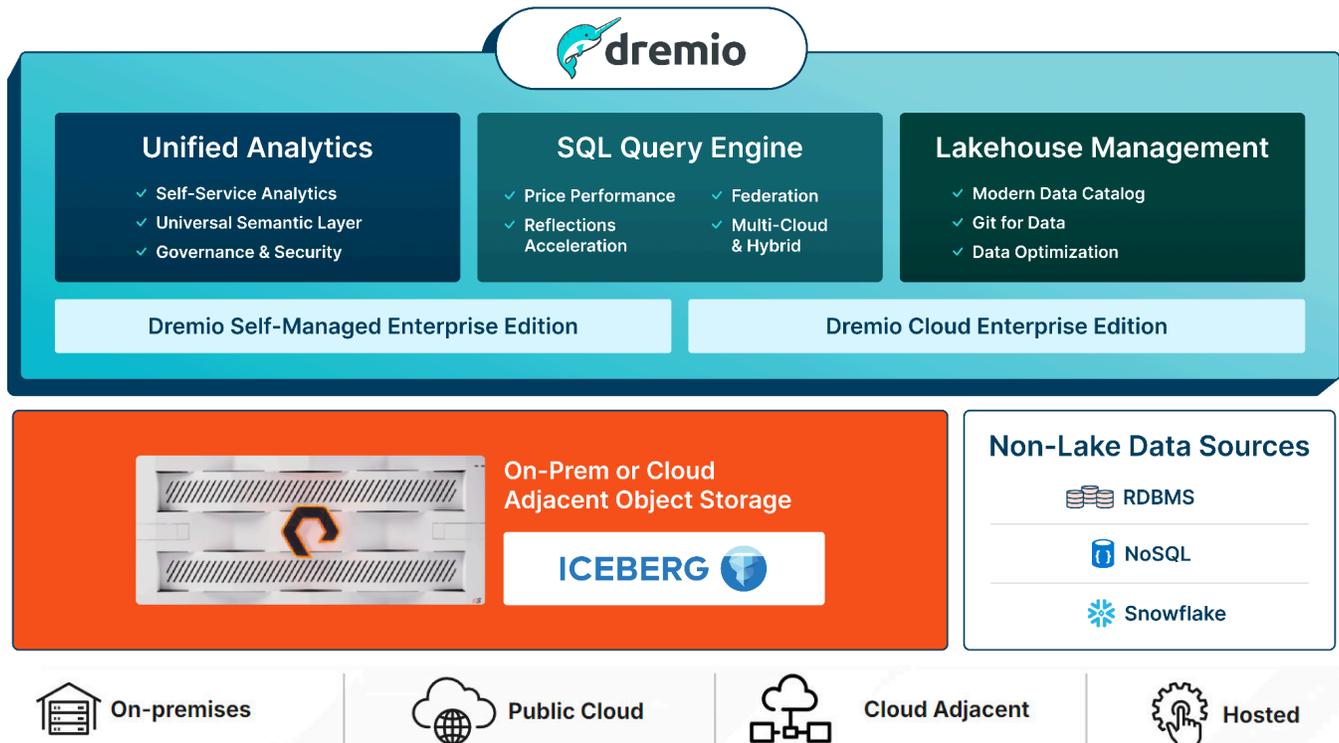


#### Operational Simplicity

Pure Storage and Dremio offer seamless tiering and independent scaling, enabling cost-effective management of massive data volumes for compliance, backtesting, and precise analysis.

## Dremio: The Unified Lakehouse Platform for Self Service Analytics and AI

The Dremio Unified Lakehouse Platform empowers you to make effective and impactful use of your enterprise's data. With Dremio's SQL query engine, data can be queried in place on Pure Storage, providing data users across the enterprise with access to the data lake. Dremio can query data from a wide range of sources, including S3 object storage, block and file storage, and relational databases like Amazon Redshift and Postgres. Dremio's semantic layer makes it easy to build and share data products over your current data infrastructure, giving you direct access to all your data during phased migrations to a modern data lake. Dremio supports platform-agnostic data formats like Parquet and Apache Iceberg, avoiding vendor lock-in and future-proofing your data infrastructure.



## Maximize the Value of Your Data

### Unmatched Performance and Capacity

Pure Storage FlashBlade offers unparalleled performance and capacity, creating a high-performance, resilient, and centralized data repository. This ensures that your data lakehouse can handle massive datasets without any loss of performance. FlashBlade's architecture supports rapid data access and processing, making it ideal for demanding analytics workloads and large-scale data consolidation.

### Lightning-Fast Queries

FlashBlade's all-flash performance, combined with Dremio's SQL Query Engine powered by Apache Arrow, delivers 4x to 1000x faster performance compared to traditional data lake engines. FlashBlade as a high-speed lakehouse storage solution is perfect for large datasets requiring real-time data analysis.

### Simplified and Resilient Lakehouses

FlashBlade's high-performance, scalable, and resilient object storage, when paired with Dremio's advanced lakehouse management features, offers a robust solution for modern data environments. FlashBlade ensures rapid access to large datasets, streamlined data operations, and enhanced data durability. Its robust architecture guarantees business continuity with high resilience and reliability, allowing for massive data consolidation without any performance degradation. This centralized repository supports continuous operations and disaster recovery, ensuring that your data is always available and protected. This integration empowers enterprises to achieve faster insights, reduce operational complexity, and maintain a highly efficient and scalable data environment, ultimately driving better business outcomes.

## Streamline Hadoop Modernization with Dremio and Pure Storage

Modernizing your Hadoop cluster with Dremio and Pure Storage simplifies management and boosts performance. Traditional Hadoop clusters face high maintenance, performance bottlenecks, scalability issues, and data silos. The Dremio and Pure Storage solution reduces operational burdens with an intuitive interface and easy-to-use all-flash storage. It enhances performance through query acceleration, enabling real-time analytics and faster data processing. The solution offers seamless scalability, allowing easy expansion of storage and compute resources. By providing a unified view of data across various sources, it breaks down data silos and facilitates comprehensive analytics and data-driven decision-making.



### Additional Resources

- Discover the power of [FlashBlade](#) for your data analytics needs.
- Take advantage of a free FlashBlade [test drive](#).
- Learn more about [Dremio](#) and take a free [Dremio test drive](#).