How a Global Top 10 Financial Institution Reduced Compute Wastage and Simplified Analytics

Implemented Denodo's Embedded MPP engine to optimize cost and efficiency, redefining analytical consumption through a shared, governed architecture

Background

This leading global financial institution, consistently ranked among the world's top-10, operates across multiple divisions, each responsible for producing curated, high-quality datasets to support enterprise-wide reporting, governance, and compliance. Within this structure, the enterprise governance and shared services function supports business lines such as Finance, Risk Control, Compliance, Legal, IT, HR, and Operations, working together to ensure transparency, regulatory adherence, and enterprise efficiency across the organization.

To deliver consolidated insights, this enterprise function consumes data produced by other divisions, most of which use Databricks to prepare curated Delta tables stored in Azure Data Lake Storage (ADLS).
For years, the organization relied on Denodo as its logical data access layer, providing unified visibility and governed access across distributed data.

However, as data volumes and adoption grew, this model began to strain. Each department needed its own Databricks cluster to query Delta data – a safeguard for governance that now introduced cost and operational complexity. This tension led the enterprise function to evaluate the recently launched Denodo Embedded MPP engine to consolidate compute within the governed layer, creating a shared, secure, and cost-efficient analytical environment across the enterprise function.

Challenges

While the distributed Databricks model empowered each division to manage its own transformation pipelines, it created an operational burden for the enterprise governance and shared services function, which depended on cross-divisional data for consolidated analytics and reporting.

Although Databricks is a highly capable compute engine for data engineering, transformation, and advanced analytics, where its Spark-based architecture delivers significant power and flexibility, using the same environment for lightweight and ad hoc analytical consumption proved inefficient. Each tenant maintained its own Databricks cluster to query Delta data stored in ADLS. This ensured isolation and compliance but resulted in a highly redundant setup, with dozens of clusters consuming compute resources for simple dashboard and reporting workloads.

Over time, maintaining and supporting each cluster, along with the pipelines that connected them through Denodo, became increasingly costly. The architecture provided strong infrastructure isolation across tenants but lacked operational efficiency, similar to using a high-performance engine for everyday city driving – powerful, but impractical for day-to-day demands. The business needed a more cost-effective model that could retain the same level of governance while optimizing compute usage and simplifying maintenance across its data landscape.

By centralizing compute but federating control, the enterprise function proved that governance need not constrain innovation, it can scale it.

The Solution

This global institution's enterprise function designed a rigorous proof of concept to test whether Denodo's Embedded MPP engine, a Presto based distributed compute framework integrated into Denodo, could serve as a shared analytical compute layer without sacrificing performance or governance. The team recreated production-grade workloads across both Databricks and Denodo MPP, using Delta format data stored in Azure Data Lake Storage.

Three workload types were tested: simple filters, complex analytical queries with joins and window functions, and full table scans. Each was executed 100 times to validate consistency. The results showed comparable performance, with Denodo MPP delivering near-identical query times to Databricks for dashboard and reporting workloads.

Beyond performance, Denodo MPP removed Databricks' cold start latency for dashboards, eliminated the cost of maintaining multiple underused clusters, and simplified cross source joins by allowing a single query plan to span both cloud and on-premise data.

Armed with evidence, the organization rearchitected its analytics layer:

- A single Denodo MPP cluster now supports all tenants.
- Logical isolation and compliance are maintained through Denodo's policy-based access and tagging framework.
- Databricks continues to serve as the high-performance platform for data preparation, transformation, and machine learning.

The outcome is a balanced architecture that aligns compute with purpose, using Denodo MPP for governed, cost efficient analytical consumption, and Databricks for advanced data engineering and analytics.

For this financial institution, the shift was not just about technology optimization: it redefined how data governance and performance can coexist within the same operational model.

Key Benefits

The introduction of Denodo's Embedded MPP engine enabled the business to simplify analytics operations while maintaining the strict governance and performance standards expected of a global financial institution.

Compute Efficiency: A single shared Denodo MPP cluster replaced multiple Databricks clusters, significantly reducing compute waste and operational overhead.

Cost Optimization: They identified up to USD 150,000 in annual Databricks compute savings tied to Denodo-related workloads, excluding additional efficiencies gained from reduced maintenance, support, and pipeline provisioning.

Governance by Design: Logical tenant isolation and policy-based access ensured security and compliance without physical infrastructure duplication.

Simplified Operations: The enterprise governance and shared services function can now manage one scalable compute layer, while Databricks remains dedicated to high-performance data engineering and machine learning.