



A High AI-Q[™]
Company

Unifying Retail Data at Scale: A Cloud-Native Data Modernization

A full-scale data modernization and migration, unifying fragmented enterprise data across SAP HANA, BigQuery, and multiple operational systems into a single Databricks-based data lake.

Overview

QBurst consolidated scattered enterprise data into a unified Databricks data lake, using internal accelerators to reduce the manual engineering effort typically required for large-scale data unification.

- Unified previously fragmented data from SAP HANA, BigQuery, and enterprise tools into a single Databricks-based lake.
- Digitized workflows reduced manual effort by 40%.
- Architecture seamlessly supports 10% annual store growth across regional footprints.



Client Profile

Leading American discount closeout retailer offering deeply discounted, overstocked, and closeout products from name-brand and private-label suppliers. Experiencing rapid expansion and relying on a technology-driven retail model, the company focuses on operational efficiency, scalability, and personalized in-store experiences through data and digital transformation.

Fragmented Data and Scalability Constraints

Data scattered across disconnected systems limited the client's ability to see and act on operations in real time.

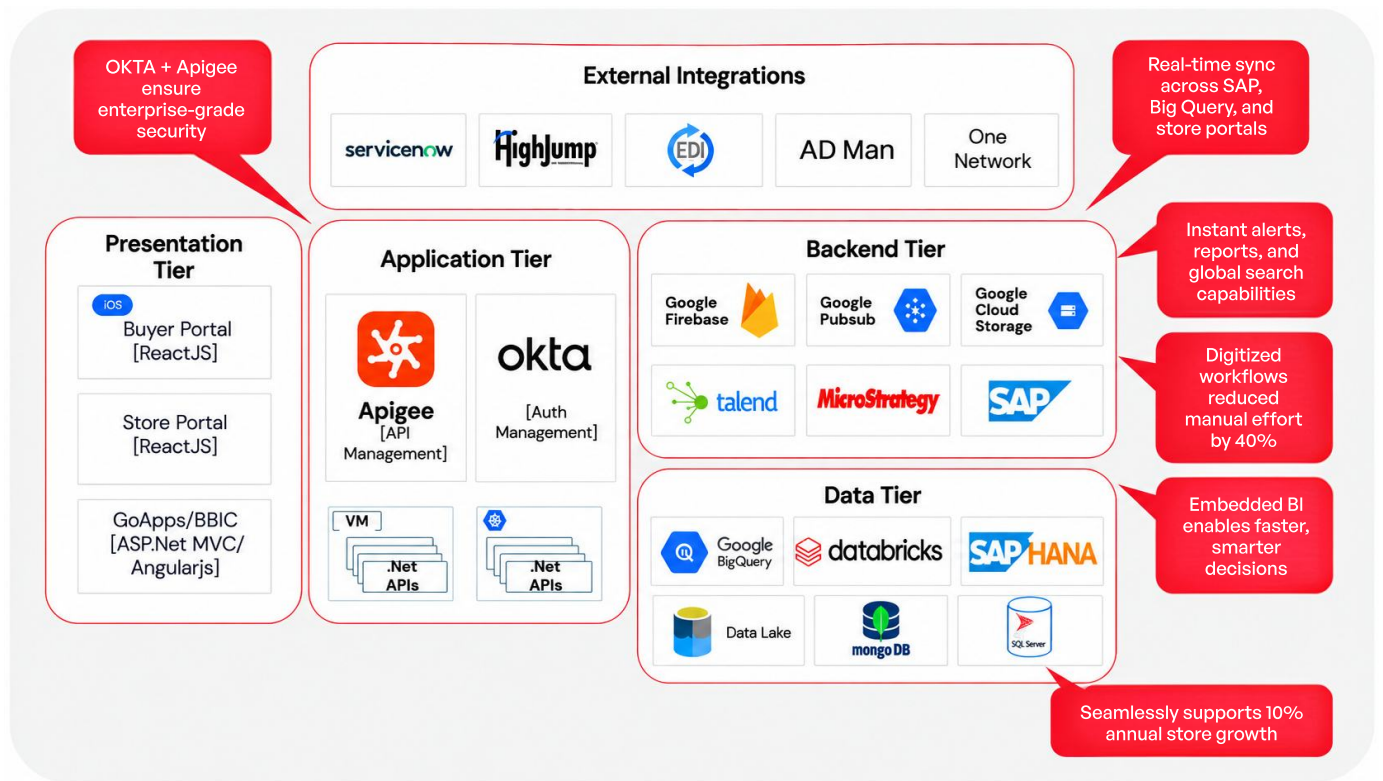
- **Data Fragmentation:** SAP HANA, BigQuery, and multiple enterprise tools each held partial, disconnected views of operational data, with no single source of truth.

- **Manual Effort Overhead:** Reconciling data across systems manually consumed significant operational effort across teams.
- **Scalability Pressure:** Rapid expansion required a data architecture capable of scaling 10% annually without performance bottlenecks.
- **Limited Real-Time Insight:** Lack of instant analytics, real-time alerts, and global search capabilities delayed operational decision-making.
- **Seasonal Load Variability:** The business's seasonal nature created extreme fluctuations in system load that legacy infrastructure could not accommodate.

Data Modernization and Migration to a Unified Databricks Lake

We executed a full data migration and modernization initiative, consolidating the client's fragmented data landscape into a single Databricks-based data lake, using QBurst's internal accelerators to reduce the manual effort typically required to migrate and reconcile data at this scale.

- **Accelerated Data Unification:** Used internal migration accelerators to extract, transform, and consolidate data from BigQuery and connected enterprise tools into a single Databricks data lake, reducing manual data engineering effort.
- **Unified Data Architecture:** Structured the consolidated data within Databricks to serve as a single governed source of truth, replacing the fragmented, tool-specific views that previously existed.
- **Real-Time Synchronization:** Established real-time data sync between core systems and the unified lake, enabling instant alerts, reporting, and global search capabilities.
- **Embedded BI & Analytics:** Connected the unified data lake to downstream BI tooling (MicroStrategy) for embedded analytics and operational decision support.
- **Supporting Application Layer:** Buyer and Store Portals, along with GoApps/BBIC, provide the interfaces through which regional teams access unified data insights, secured via Okta and API-managed through Apigee.



End-to-End Enterprise System Architecture

Technical Highlights

- **Data Consolidation Layer:** Internal accelerators automated extraction and transformation of data from BigQuery, SAP HANA, MongoDB, and SQL Server sources into Databricks.
- **Cloud Infrastructure:** Built on Google Cloud, using Firebase and PubSub for real-time event handling and Cloud Storage for supporting data assets.
- **Access & Governance:** Implemented Unity Catalog for governance and data cataloging with automated information capture for new jobs and data assets.
- **Consumption Layer:** MicroStrategy delivered embedded BI and reporting directly against the unified Databricks data lake.

Impact: Unified Data and Operational Efficiency

- **Data Modernization Outcome:** Migrated and consolidated fragmented data from SAP HANA, BigQuery, and enterprise tools into a single governed Databricks data lake.
- **Reduced Manual Effort:** Digitized workflows and automated data consolidation reduced manual effort by 40% across operations.
- **Scalable Growth:** Architecture supports 10% annual store growth without performance degradation.
- **Real-Time Decision-Making:** Embedded BI, instant alerts, and global search now run against a single source of truth instead of fragmented systems.