

Cloud Modernization of Enterprise Data Platform

Replacing legacy reporting systems with a Databricks and Delta Lake Medallion architecture to overcome processing latency and enable cloud scalability.

Overview

QBurst collaborated with the client's in-house team to modernize their data ecosystem from on-premises infrastructure to a cloud-based reporting platform. Utilizing a Databricks and Delta Lake Medallion architecture alongside Tableau, the platform streamlines batch ingestion and delivers low-latency enterprise analytics.

- Eliminated performance limitations of legacy tools (SSRS, SAP BO) by transitioning to a unified cloud consumption tier.
- Reduced manual reporting overhead for IT teams while enabling cloud scalability and improved data accessibility.



Client Profile

Based in the US, the client is a prominent academic hospital chain specializing in medical research, treatment, and clinical education.

Challenges: Fragmented Data and Legacy Bottlenecks

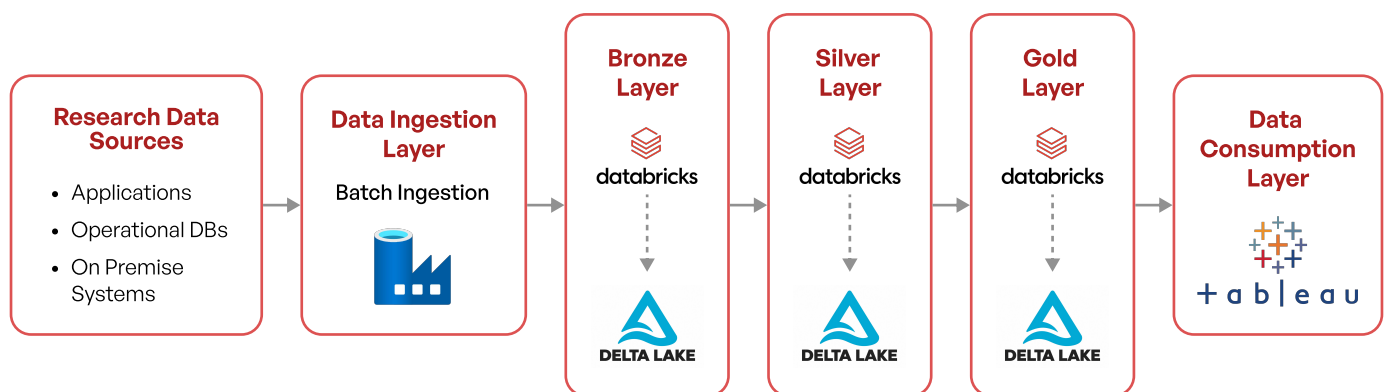
Performance limitations and high latency caused by legacy reporting systems, specifically SSRS and SAP BO.

- Fragmented data across legacy on-premises systems, operational databases, and research applications.
- Excessive manual reporting effort required from IT teams to extract and synthesize operational and research insights.
- Lack of cloud scalability and remote accessibility across clinical and research data workflows.

QBurst Solution: Medallion Lakehouse Architecture

We partnered with the client's in-house engineering team to design and deploy a cloud-based reporting architecture. The platform consolidates disparate research and operational data sources into a structured Databricks and Delta Lake pipeline before delivering insights via modern visualization tools.

- **Data Ingestion Layer:** Implemented automated batch ingestion to collect raw data from on-premises systems, operational databases, and research applications.
- **Bronze Layer (Raw Data):** Ingested raw datasets directly into Delta Lake managed by Databricks for reliable, scalable storage.
- **Silver Layer (Refined Data):** Cleaned and structured incoming data within Databricks to create validated datasets for intermediate analysis.
- **Gold Layer (Business Aggregates):** Processed refined data into high-performance business aggregates using Databricks and Delta Lake to power reporting.
- **Data Consumption Layer:** Replaced legacy SSRS and SAP BO tools with Tableau for streamlined, accessible data consumption across teams.



Modern Data Platform

Technical Highlights

- **Migration Accelerator:** Reusable modernization framework to accelerate the migration of data and transformation.
- **Lakehouse Foundation:** Delta Lake as the data lakehouse for ad hoc exploratory analytics and Silver layer preparation.
- **Processing Pipelines:** Databricks pipelines for data transformation and Gold layer implementation.
- **Metadata Governance:** Unity Catalog for enterprise data governance and cataloging.

Impact: Enhanced Performance and Reduced IT Effort

- **Legacy Elimination:** Successfully transitioned away from rigid SSRS and SAP BO reporting setups.
- **Reduced Manual Overhead:** Automated batch pipelines and transformed data layers significantly decreased manual effort for IT teams, driving over 55% increase in team productivity.
- **Latency Reduction:** Reduced data processing latency by 60% compressing query load times from over 45 minutes down to under five minutes.
- **Cloud Scalability:** Delivered a modern, highly accessible cloud infrastructure capable of scaling alongside enterprise growth.