DIGGIBYTE

Building a Unified Databricks Lakehouse

Modernizing Synapse + SQL Managed Instance into a unified Databricks Lakehouse with Lakebase as the serving layer.

CLIENT INFO:

EUREKA FORBES

Client:
Eureka Forbes

Use Case:
Platform Modernization & Data Engineering Transformation

OUTCOME MAP ALIGNMENT

Azure Synapse → Diggibyte → Databricks

10 TB Structured data in scope

300 Stored procedures to be modernized

30 Power BI reports retained with data-source connection updates

A scalable Databricks Lakehouse foundation modernizes legacy SQL-based processing, transforms data engineering workloads, and supports improved performance, scalability, governance and future readiness.

DIGGIBYTE

THE CHALLENGE

Modernizing a fragmented data architecture

Eureka Forbes needs to transition from a fragmented Synapse + SQL Managed Instance architecture to a scalable, unified Databricks Lakehouse while maintaining continuity across existing analytics, downstream systems and ingestion pipelines.

MIGRATION COMPLEXITY

- Fragmented platform architecture: The existing environment spans Synapse and SQL Managed Instance, creating a need for platform modernization.
- Large-scale data transformation: ~10 TB of structured data and associated transformations need to be migrated and modernized.
- Legacy SQL processing: ~300 stored procedures need to be converted into scalable, maintainable Databricks-native pipelines using SQL, PySpark and Workflows.

BUSINESS REQUIREMENTS

- Power BI continuity: ~30 existing Power BI reports need to be retained without redesign or migration to AI/BI platforms.
- Downstream system continuity: Data flow to WebEngage, Microsoft Dynamics 365 and internal web applications needs to remain uninterrupted.
- ADF ingestion continuity: Existing ADF ingestion pipelines need to continue without disruption.

BUSINESS OBJECTIVE

Establish a scalable, unified Databricks Lakehouse that modernizes data engineering, improves performance and scalability, maintains Power BI continuity, and minimizes disruption to business operations.

DIGGIBYTE

THE SOLUTION

A scalable, unified Databricks Lakehouse

The modernization transitions legacy SQL-based processing to distributed, high-performance Databricks data pipelines and establishes a Medallion Architecture to improve data quality, governance and reusability.

WHAT WE DID		HOW WE DID
01	**Platform Modernization**	Transitioned from the fragmented Synapse + SQL Managed Instance architecture to a scalable, unified Databricks Lakehouse and established Databricks Lakebase as the primary serving layer.
02	**Data Engineering Transformation**	Migrated and modernized ~10 TB of structured data and associated transformations, converting ~300 stored procedures into Databricks-native SQL, PySpark and Workflows.
03	**Medallion Architecture**	Implemented a Bronze → Silver → Gold Medallion Architecture to improve data quality, governance and reusability.
04	**Analytics & BI Continuity**	Retained ~30 existing Power BI reports with only data-source connection updates, preserving report structure, visualizations, business logic/calculations and user experience.
05	**Performance & Scalability**	Improved query performance and workload scalability using Databricks, optimized compute and storage utilization compared to the Synapse 500 DWU model, and enabled elastic scaling for future data growth.
06	**Risk-Minimized Transition**	Executed a phased migration with comprehensive validation, controlled cutover strategy and 4-week hypercare support to ensure minimal disruption to business operations.

DIGGIBYTE

IMPACT & OUTCOME

Modernized data platform with analytics and business continuity

The modernization provides a unified Databricks Lakehouse foundation while maintaining continuity of existing Power BI reports, downstream integrations and ingestion pipelines.

BUILD ON DATABRICKS

~10 TB	~300	~30
Structured data in scope	Stored procedures to be modernized	Power BI reports retained

MODERNIZATION SCOPE

1. **~10 TB of structured data in scope for migration and modernization.**
2. **~300 stored procedures converted into scalable, maintainable Databricks-native pipelines.**
3. **~30 existing Power BI reports retained with data-source connection updates.**

SUSTAINABLE VALUE

1. **Unified Databricks Lakehouse foundation supports future data growth and analytics needs.**
2. **Elastic scaling and optimized compute and storage utilization support performance and cost optimization.**
3. **Foundation enables future capabilities including advanced AI/ML use cases, real-time analytics and intelligent data products.**

REQUIRED BEFORE EXTERNAL SUBMISSION

Prior written approval from Diggibyte is required before the submitted customer story can be used for any public marketing purposes.

Contact Us: www.diggibyte.com sales@diggibyte.com