databricks
Data Lakehouse

Overview

Enterprises today, are challenged by the number of platforms and applications deployed to manage their data related initiatives. If they are not leveraging a modern data architecture on the cloud, they are likely to:

Paying too much to manage and maintain legacy data platforms

Falling behind on productivity

Struggling to stitch together multiple point solutions

Missing new business value

Databricks is the Data + AI Company
As the world's first and only lake house platform in the cloud, Databricks is the only cloud-native vendor named as a leader in 2021 Gartner Magic Quadrant in both database management system (DBMS) and data science and machine learning platforms (DSML) with 5000+ customers across globe.

The Databricks Lakehouse Platform is a simple platform to unify all your data, analytics, and AI workloads. It combines the best elements of data lakes and data warehouses to deliver data management and performance with low-cost, flexible object stores offered by data lakes.

databricks
Lakehouse Platform

| Data Warehousing | Data Engineering | Data Streaming | Data Science and ML |

Unity Catalog
Fine-grained governance for data and AI

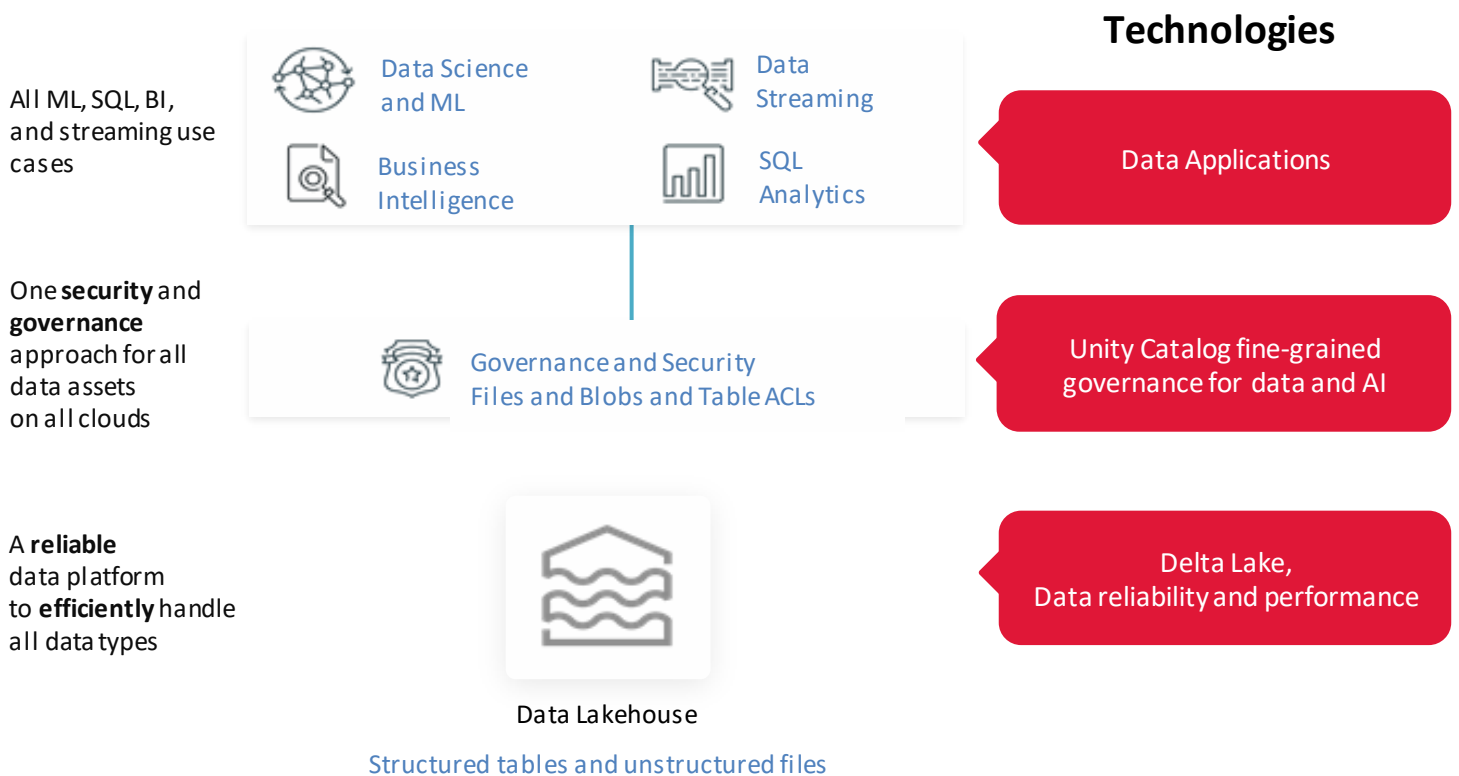### Delta Lake
Data reliability and performance

Collaborations that Create NXT Possibilities

Tech Mahindra's ability to deliver enterprise-wide digital transformation programs, helping customers derive competitive advantage by generating actionable insights from their data. Databricks's Lakehouse Platform depicts a new data architecture paradigm for enterprises and enhanced skills with Databricks integration will be positioned well with enterprises for new use cases and strategic data and analytics projects. Together with Tech Mahindra and Databrick's expertise, we can provide a turnkey solution to solve the most complex digital transformation projects.

Our Solution

The platform is a modern data platform and an industry leading lakehouse solution. Databricks provides an end-to-end platform that serves our customers more effectively than other data platform options. It is available on all major cloud services (AWS, Azure, and GCP) and accelerates time-to-value from all data, for all users and use cases.

Technologies

All ML, SQL, BI, and streaming use cases

Data Science and ML

Data Streaming

Business Intelligence

SQL Analytics

Data Applications

One **security** and **governance** approach for all data assets on all clouds

Governance and Security
Files and Blobs and Table ACLs

Unity Catalog fine-grained governance for data and AI

A **reliable** data platform to **efficiently** handle all data types

Delta Lake,
Data reliability and performance

Data Lakehouse

Structured tables and unstructured files

Simple

The platform simplifies your data architecture by unifying your data, analytics, and AI workloads on one common platform. Delta Lake forms the foundation of this cost-effective, highly scalable lakehouse by providing the reliability, performance and governance required to support all types of data workloads.

Open

As the platform is built on top of the world's most successful open-source data projects, the lakehouse platform unifies your data ecosystem with open standards and formats. This will avoid proprietary data formats and the existing investments can be leveraged with the broad ecosystem of technology and consulting partners.

Collaborative

The platform offers native collaboration capabilities and unifies the data teams to collaborate across the entire data and AI workflow.

The data engineers to analysts to data scientists — can now collaborate across all your workloads, accelerating your journey to become truly data-driven.

Success Stories

1. Fraud Analytics and Machine Learning Framework and Infrastructure for a Leading Telecom Company in US

Challenges

- High response times to fraud ML model execution on-premises
- Real time stream processing of features is required to be executed with stringent SLA

Solution

- Establish AIaaS for fraud analytics and watch tower ML Ops
- Modernization of model runtime environment on Azure
- Replace Redis with Databricks feature engineering for faster stream processing
- Telemetry and aiOpX integration to automate monitoring of lags/backpressure, count, data quality etc.

Benefits

- **Revenue** through early fraud detection
- Cost **savings** by replacing Redis
- **Faster** response times
- **Scalable** framework to integrate and operationalize more ML models

2. Modernization of Mobility Churn Analytics for a Leading Telecom Company in the US

Challenges

- Analyze customer behavior to predict churn and increase customer retention

Solution

- Leverage consumer gold data sets for predictive analysis
- Create ML feature table to summarize consumer attributes
- Label the data and train a LGBM model to predict churn and port scores
- Use Spark to build and productionize data and ML pipeline to create features, score and publish results to clients

Benefits

- High **computational efficiency**
- Model **performance** improvements
- **AI modernization** in cloud

About Tech Mahindra

Tech Mahindra offers innovative and customer-centric digital experiences, enabling enterprises, associates, and the society to RISE. We are a USD 6 billion organization with over 158,000 professionals across 90 countries helping 1262 global customers, including Fortune 500 companies. We are focused on leveraging next-generation technologies including data analytics, 5G, blockchain, cybersecurity, artificial intelligence, and more, to enable end to end digital transformation for global customers.

About Databricks

Databricks is the Data and AI company. More than 5,000 organizations worldwide — including Comcast, Condé Nast, H&M, and over 40% of the Fortune 500 — rely on the Databricks lakehouse platform to unify their data, analytics, and AI. Databricks is headquartered in San Francisco, with 20+ offices around the globe. Founded by the original creators of Apache Spark™, Delta Lake and MLflow, Databricks is on a mission to help data teams solve the world's toughest problems.