databricks

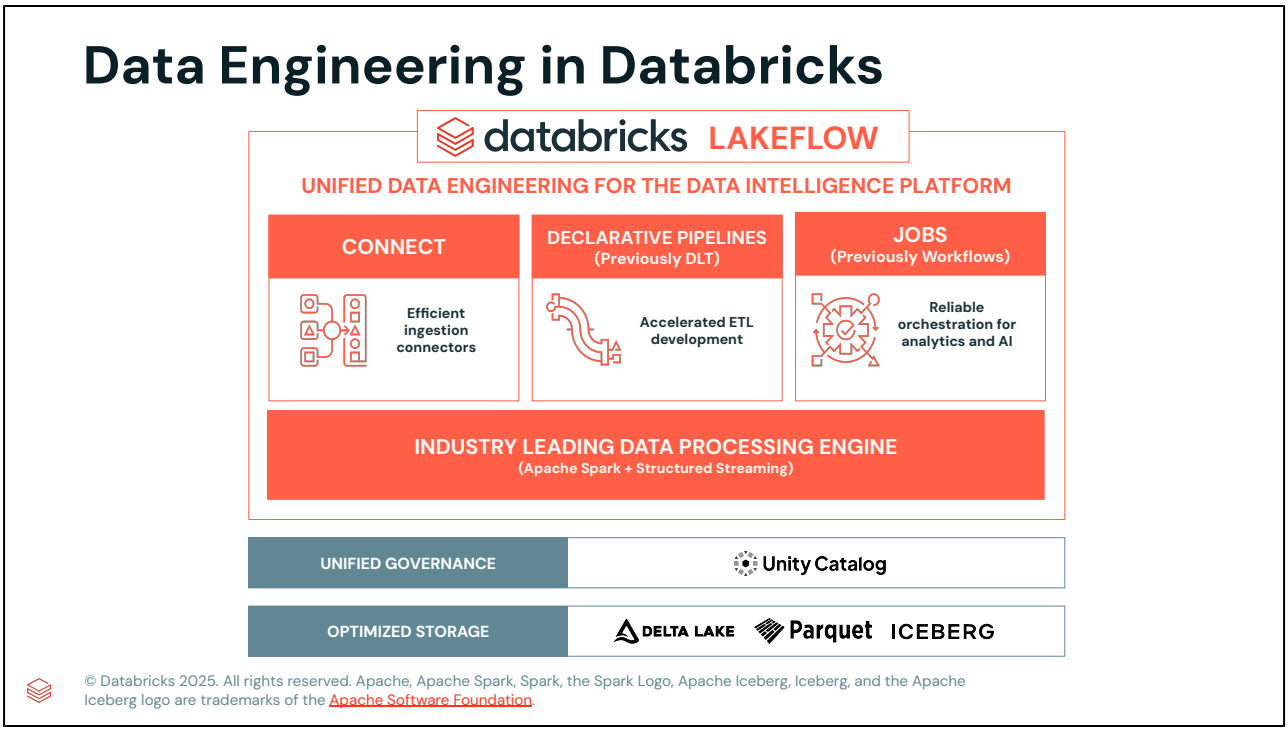
Introduction to Data Engineering in Databricks

LECTURE

Data Engineering in Databricks

© Databricks 2025. All rights reserved. Apache, Apache Spark, Spark, the Spark Logo, Apache Iceberg, Iceberg, and the Apache Iceberg logo are trademarks of the Apache Software Foundation.

In this lecture, we'll explore how Databricks streamlines data engineering by enabling seamless data ingestion and management using LakeFlow Connect and advanced storage solutions like Delta Lake, Parquet, and Iceberg.



It all begins with optimized storage using Delta Lake, Parquet, or Iceberg.

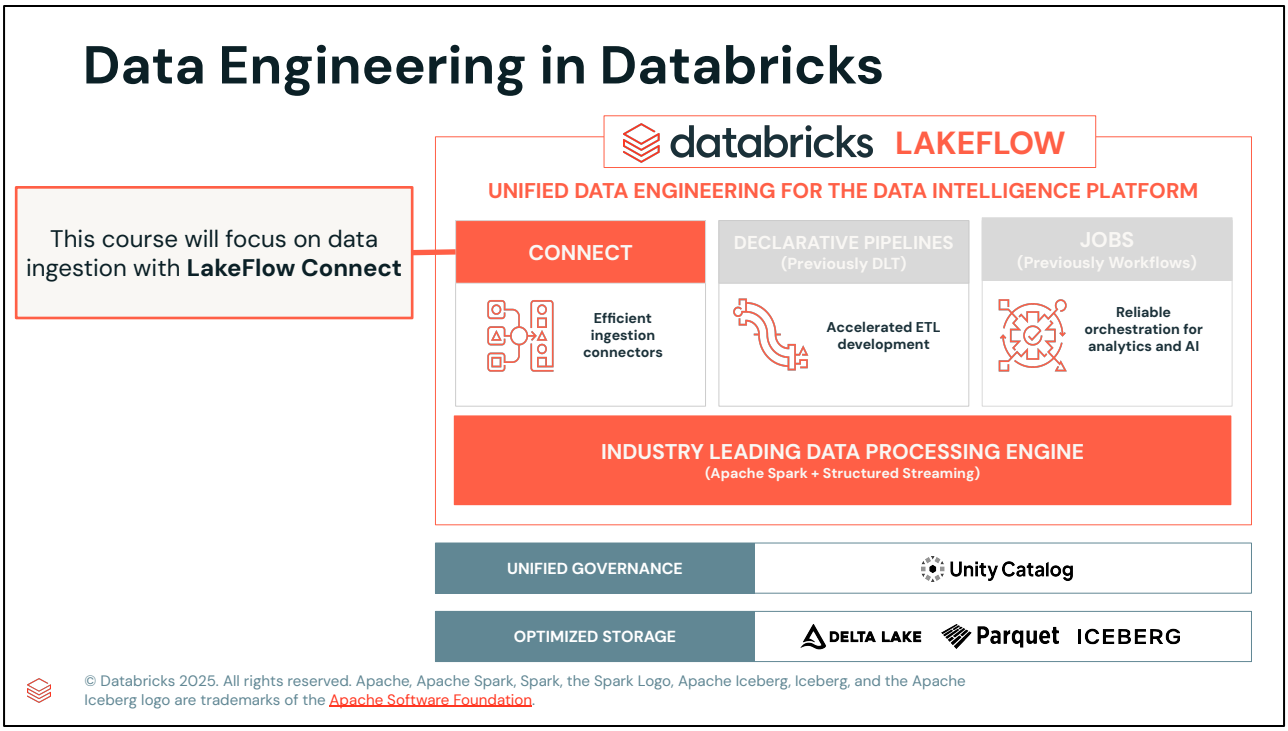
Built on top of this storage layer is unified governance with Unity Catalog. Unity Catalog is a centralized data catalog that provides access control, auditing, data lineage, quality monitoring, and data discovery across Databricks workspaces.

Databricks then offers **Lakeflow**, an end-to-end data engineering solution that empowers data engineers, software developers, SQL developers, analysts, and data scientists to deliver high-quality data for downstream analytics, AI, and operational applications. Lakeflow provides a unified platform for data ingestion, transformation, and orchestration, and includes the following components:

- **Lakeflow Connect**: A set of efficient ingestion connectors that simplify data ingestion from popular enterprise applications, databases, cloud storage, message buses, and local files.

- **Lakeflow Declarative Pipelines**: A framework for building batch and streaming data pipelines using SQL and Python, designed to accelerate ETL development.

- **Lakeflow Jobs**: A workflow automation tool for Databricks that orchestrates data processing workloads. It enables coordination of multiple tasks within complex workflows, allowing for the scheduling, optimization, and management of repeatable processes.



This course will focus on data ingestion with LakeFlow Connect into the Databricks Data Intelligence platform.

© Databricks 2025. All rights reserved. Apache, Apache Spark, Spark, the Spark Logo, Apache Iceberg, Iceberg, and the Apache Iceberg logo are trademarks of the Apache Software Foundation.

Thank you for completing this lesson and continuing your journey to develop your skills with us.