

CloudPandith

Enterprise Data Engineering with Databricks: Delta Lake, DLT & Lakehouse Architecture

Real-Time Projects | Interview Ready Training

www.cloudpandith.com | 8904424822 | cloudpandith@gmail.com

🌀 SECTION 1: Advanced Delta Lake Architecture & Internals

Goal: Deep internal understanding of Delta Lake engine

Lecture 1: Delta Lake Project Introduction | Managed vs External Tables | Delta Utils & Constraints

Lecture 2: Delta Lake Architecture (Part 1) | Internal Working Mechanism

Lecture 3: Delta Lake Architecture (Part 2) | Deletion Vectors Explained

🌀 SECTION 2: Advanced Delta Lake Features

Goal: Master enterprise-grade Delta capabilities

Lecture 4: Schema Enforcement vs Schema Evolution | Time Travel | Vacuum

Lecture 5: Deep Clone vs Shallow Clone | Backup & Restore Delta Tables

Lecture 6: Change Data Feed (CDF) | Structured Streaming Integration

Lecture 7: OPTIMIZE Command | Auto Compaction Automation

🌀 SECTION 3: Delta Performance Optimization

Goal: Improve storage + query performance

Lecture 8: Disk Cache Optimization | Column-Level Encryption & Decryption

Lecture 9: Column Re-ordering in Delta Tables | Parquet to Delta Conversion

www.cloudpandith.com | 8904424822 | cloudpandith@gmail.com

● SECTION 4: Streaming & Real-Time Processing

Goal: Build real-time ingestion pipelines

Lecture 10: Spark Streaming in Databricks | Continuous vs Triggered Processing

Lecture 11: Auto Loader | Incremental File Processing | Notification Mode

● SECTION 5: Unity Catalog & Governance

Goal: Enterprise data governance model

Lecture 12: Unity Catalog Introduction | Workspace Migration Strategy

Lecture 13: Managed vs External Tables in Unity Catalog

● SECTION 6: Delta Live Tables (DLT) – Core Concepts

Goal: Declarative pipeline development

Lecture 14: Introduction to Delta Live Tables (DLT) | Bronze, Silver, Gold Architecture

Lecture 15: Create First DLT Pipeline | Pipeline UI Explained

Lecture 16: Incremental Load & Table Updates in DLT

● SECTION 7: Advanced DLT Concepts

Goal: Production-grade pipeline design

Lecture 17: Streaming Tables, Views & Auto Loader Integration in DLT

Lecture 18: Expectations in DLT | WARN, FAIL, DROP Rules | Event Logs

Lecture 19: Parameterization in DLT | SCD Type 1 & 2 | Late Arriving Data

Lecture 20: Full Refresh Pipelines | Truncate & Load Strategy

Lecture 21: Triggering Pipelines on File Arrival

● SECTION 8: Spark Declarative & Lakeflow Pipelines

Goal: Modern pipeline frameworks

Lecture 22: Spark Declarative Pipelines (SDP) | delta_sink Concept

Lecture 23: DLT to Lakeflow Migration | Modern Pipeline Architecture

● SECTION 9: Pipeline Scheduling & Orchestration

Goal: Production deployment of pipelines

Lecture 24: Databricks Pipeline Scheduling | Job vs Interactive Clusters

Lecture 25: External Table Creation in SQL Data Warehouse

● SECTION 10: Advanced Auto Loader & Ingestion Patterns

Goal: Scalable ingestion design

Lecture 26: Auto Loader Deep Dive | Incremental Load Patterns

Lecture 27: Liquid Clustering Integration with Auto Loader

● SECTION 11: Capstone Performance & Optimization

Goal: Real-world performance tuning

Lecture 28: Delta Performance Optimization Strategies (End-to-End)

● SECTION 12: Advanced Interview Preparation

Goal: Crack senior data engineering roles

Lecture 29: Spark Performance Optimization Interview Questions

Lecture 30: Broadcast Join, AQE & Delta Optimization Scenarios

Lecture 31: Real-world Databricks Architecture Interview Questions

Lecture 32: End-to-End Delta + Streaming + DLT Scenario Questions