

Trino for Enterprise Data Federation: Unified Analytics with Cloudera

OEM: Cloudera • Duration: 2 Days (16 hrs) • Code: DENG-257

COURSE MODULES & TOPICS

Day 1 – Distributed SQL & Federation Foundations

Introduction to Trino in Enterprise

- Distributed SQL architecture, Trino vs Hive/Impala, enterprise use cases
- Hands-On: Launch Trino Virtual Warehouse, execute queries, observe execution stages

Architecture & Query Execution Deep Dive

- Coordinator/worker nodes, query lifecycle, memory & spill handling
- Hands-On: Analyze query plans, monitor execution metrics

Enterprise Data Federation & Virtualization

- Data virtualization, cross-domain access, federation vs replication
- Hands-On: Design multi-domain architecture, map domains, plan joins

Multi-Catalog & Cross-Domain Queries

- Catalog configuration, cross-catalog joins, pushdown behavior
- Hands-On: Configure catalogs, execute cross-domain queries, analyze data movement

Day 2 – Performance & Enterprise Deployment

Advanced Query Optimization

- Predicate pushdown, dynamic filtering, join strategies, cost-based optimization
- Hands-On: Tune distributed queries, compare performance scenarios

Monitoring & Troubleshooting

- Query metrics, bottleneck detection, memory & spill tuning
- Hands-On: Diagnose slow queries, apply tuning improvements

Workload Management & Resource Control

- Resource groups, concurrency, query prioritization, multi-tenancy
- Hands-On: Configure resource groups, simulate workloads

Enterprise Architecture & Production Best Practices

- Hybrid federation, scaling, HA, deployment strategies
- Hands-On: Design enterprise architecture, map workloads