CLOUDERA

TRAINING
DATASHEET

DENG-257

Trino for Enterprise Data Federation: Unified Analytics with Cloudera

Course Overview

Training Details

- **Format:** Instructor-led
- **Level:** Intermediate
- **Duration:** 2 Days
- **Platform:** Cloudera Data Warehouse

Core Areas

- Distributed SQL & Trino Architecture
- Enterprise Data Federation
- Cross-System Analytics
- Performance Optimization
- Production & Deployment

Hands-On Focus

- Configure Trino warehouse
- Execute cross-system queries
- Design federation architecture
- Optimize & troubleshoot queries

Capabilities You Will Build

- Design **enterprise-grade data federation architectures** across systems
- Execute and optimize **high-performance distributed SQL queries**
- Configure and manage **Trino connectors for cross-system access**
- Diagnose and resolve **query bottlenecks and performance issues**
- Build **scalable, production-ready federation solutions**

About This Training

Modern enterprises struggle with fragmented data spread across data lakes, warehouses, and operational systems. Traditional ETL pipelines introduce latency, increase costs, and slow decision-making.

This training enables organizations to adopt **real-time data federation using Trino**, allowing teams to query distributed data sources instantly — without duplication or movement.

Participants will learn how to build **enterprise-grade federated analytics architectures** that reduce data engineering overhead, accelerate insights, and unlock business value from existing data ecosystems.

Business Value & Outcomes

By the end of this training, organizations can:

- **Eliminate ETL bottlenecks and reduce pipeline complexity** through query-in-place federation
- **Accelerate time-to-insight** from hours/days to seconds with real-time distributed querying
- **Reduce data movement and infrastructure costs** by avoiding unnecessary data duplication
- **Enable real-time, cross-system analytics** across data lakes, databases, and cloud platforms
- **Improve governance, security, and team productivity** through unified and controlled data access

Enterprise Use Case

Real-Time Fraud Detection (Banking)

- Analyze transactions and customer data across multiple systems in real time using Trino-based federation — without data movement.
- Outcome: Faster fraud detection, unified insights, and reduced infrastructure cost

DENG-257

Trino for Enterprise Data Federation: Unified Analytics with Cloudera

Day 1 – Distributed SQL & Federation Foundations

Module	Key Concepts Covered	Hands-On Activities	Business Value Delivered
Introduction to Trino in Enterprise	Distributed SQL architecture, Trino vs Hive/Impala, enterprise use cases	Launch Trino Virtual Warehouse, execute queries, observe execution stages	Understand how to **replace traditional ETL-heavy analytics with real-time federation**
Architecture & Query Execution Deep Dive	Coordinator/worker nodes, query lifecycle, memory & spill handling	Analyze query plans, monitor execution metrics	Enable teams to **diagnose performance issues and optimize query execution**
Enterprise Data Federation & Virtualization	Data virtualization, cross-domain access, federation vs replication	Design multi-domain architecture, map domains, plan joins	Reduce **data duplication costs and enable unified analytics across systems**
Multi-Catalog & Cross-Domain Queries	Catalog configuration, cross-catalog joins, pushdown behavior	Configure catalogs, execute cross-domain queries, analyze data movement	Enable **real-time analytics across distributed data sources without data movement**

VERSION:
2.0.0

DENG-257

Trino for Enterprise Data Federation: Unified Analytics with Cloudera

Day 2 – Performance & Enterprise Deployment

Module	Key Concepts Covered	Hands-On Activities	Business Value Delivered
Advanced Query Optimization	Predicate pushdown, dynamic filtering, join strategies, cost-based optimization	Tune distributed queries, compare performance scenarios	Improve **query performance and reduce infrastructure cost**
Monitoring & Troubleshooting	Query metrics, bottleneck detection, memory & spill tuning	Diagnose slow queries, apply tuning improvements	Reduce **downtime and accelerate issue resolution**
Workload Management & Resource Control	Resource groups, concurrency, query prioritization, multi-tenancy	Configure resource groups, simulate workloads	Ensure **stable performance in multi-user enterprise environments**
Enterprise Architecture & Production Best Practices	Hybrid federation, scaling, HA, deployment strategies	Design enterprise architecture, map workloads	Build **production-ready, scalable federation systems**

VERSION:
2.0.0