Updated on 23/07/2025

Sign up

Vertica Analytics Platform Training
2 days (14 hours)

Presentation

Master Vertica Analytics Platform in its expert dimension with this comprehensive training course, designed for data engineers, architects and analysts wishing to exploit the full potential of this high-performance data warehouse.

The course begins with a deep dive into Vertica's MPP architecture, columnar storage, projections, segmentation, replication and high availability mechanisms. You'll learn how to size and configure your clusters to guarantee performance and resilience.

You'll learn how to efficiently ingest massive volumes of data, in real time or by batch, via COPY, Kafka, Spark, or ETLs like Talend. Optimized modeling (partitioning, projections) becomes an essential performance lever.

You'll master advanced SQL querying, analytical functions, query profiling, as well as compression and tuning with Database Designer. BI integration, APIs, Python and native machine learning will enable you to leverage your data on a massive scale.

As with all our training courses, this one will be presented with the latest Vertica Analytics Platform updates.

Objectives

- Understand Vertica's advanced architecture and how it works in MPP.
- Design, model, segment and optimize high-performance data schemas in Vertica
- Master strategies for ingesting massive data via COPY, Kafka, Spark or ETL
- Be able to secure, govern and trace access, processing and audits in critical environments
- Know how to industrialize Vertica processing and deployment with scripts and CI/CD automation.

- Leverage Vertica's native analytics and machine learning engine for predictive analytics

Target audience

- Data Engineers
- Data architects

Prerequisites

- Basic knowledge of SQL language

Vertica Analytics Platform training program Introduction to

Vertica

- Vertica architecture
- Use cases
- Comparison with other data warehouses
- Key components of the Vertica platform

Architecture and basic concepts

- Distributed columnar architecture
- Columnar vs. row storage
- MPP model
- Fundamental concepts
- Projections
- ROS/WOS
- Segmentation and partitioning
- Sharding and data replication
- High availability architecture

Installation and configuration

- System and network requirements
- Cluster installation
- Single node vs. multi-node
- Initial configuration
- Cluster setup
- Security and firewall
- Management interface

Data ingestion

- Ingestion methods
- COPY, INSERT, EXPORT
- COPY LOCAL vs COPY DIRECT
- Mass loading
- Error handling
- Optimizing ingestion
- Integration with other tools
- Kafka, Spark, Hadoop
- ETL connectors
- Real-time and continuous ingestion

Data modeling

- Schemas and databases
- Table creation
- Supported data types
- Constraints and indexes
- Projections
- Manual vs. automatic creation
- Optimizing projections for queries
- Partitioning and segmentation

SQL querying in Vertica

- SQL compatibility
- Standard and specific functions
- Advanced analytical functions
- Windows, aggregations, groupings
- Temporal queries
- Queries with joins and subqueries
- Distributed queries on multiple nodes

Performance optimization

- Execution plan analysis
- EXPLAIN and PROFILE
- Projection tuning
- Using Database Designer
- Data compression
- Resource management
- Workload Management
- Performance monitoring
- Vertica Management Console
- Log tables and system views

Security and access management

- Authentication and authorization
- Roles, users, groups
- Data security
- Encryption at rest and in transit
- Audits and monitoring
- Access logs
- Securing sensitive requests

Maintenance and administration

- Rebalancing and unused projections
- Optimization of projections and segments
- Backup and restore
- Incremental backup
- VBR
- Version upgrade and scheduled maintenance

Visualization and BI

- Connection with BI tools
- Tableau, Power BI, Looker, Qlik
- Integration via ODBC/JDBC
- Creating dashboards and analytical queries
- Visualization case studies with Vertica

Advanced integrations

- Integration with Hadoop / Hive
- Integration with Spark
- Integration with Python / R (Machine Learning)
- Vertica REST API and SDK

Machine Learning with Vertica

- Introduction to Vertica ML
- Native models
- Creating training sets
- Model training
- Evaluation and prediction
- Integration with Jupyter notebooks

Deployment and DevOps

- CI/CD with Vertica SQL scripts
- Schema versioning
- Automation with Bash / Python scripts
- Use of containers and Kubernetes deployment

Companies concerned

This course is aimed at both individuals and companies, large or small, wishing to train their teams in a new advanced computer technology, or to acquire specific business knowledge or modern methods.

Positioning on entry to training

Positioning at the start of training complies with Qualiopi quality criteria. As soon as registration is finalized, the learner receives a self-assessment questionnaire which enables us to assess his or her estimated level of proficiency in different types of technology, as well as his or her expectations and personal objectives for the forthcoming course, within the limits imposed by the selected format. This questionnaire also enables us to anticipate any connection or security difficulties within the company (intra-company or virtual classroom) which could be problematic for the follow-up and smooth running of the training session.

Teaching methods

Practical training: 60% hands-on, 40% theory. Training material distributed in digital format

to all participants.

Organization

The course alternates theoretical input from the trainer, supported by examples, with brainstorming sessions and group work.

Validation

At the end of the session, a multiple-choice questionnaire verifies the correct acquisition of skills.

Certification

A certificate will be awarded to each trainee who has completed the entire course.