



Data Warehousing with BigQuery: Storage Design, Query Optimization, and Administration

Course description

In this comprehensive course, you will delve into the intricacies of designing, optimizing, and administering a data warehouse using Google Cloud's BigQuery. You will gain a deep understanding of BigQuery's serverless architecture, columnar storage, and massively parallel processing capabilities. The course will cover best practices for structuring data with partitioning and clustering to enhance query performance and reduce costs. You will learn advanced techniques for writing efficient SQL queries, analyzing query execution plans, and troubleshooting performance bottlenecks. Furthermore, you will explore critical administration aspects, including security with Identity and Access Management (IAM), cost management strategies, and robust monitoring. By the end of this course, you will be equipped to build, optimize, and manage scalable and cost-effective data warehousing solutions with BigQuery.

- Course level: Advanced
- Duration: 2 days

Activities

This course includes in-depth presentations, architectural deep-dives, hands-on SQL query labs, BigQuery configuration demonstrations, and practical case studies.

Course objectives

In this course, you will learn to:

- Understand BigQuery's architecture and its core features for modern data warehousing.
- Design optimal BigQuery storage schemas using advanced partitioning and clustering strategies.
- Apply sophisticated SQL techniques and BigQuery-specific best practices for query performance.
- Analyze and interpret BigQuery query execution plans to pinpoint and resolve performance bottlenecks.
- Implement robust security and access control mechanisms, including IAM, row-level, and column-level security.
- Manage BigQuery costs effectively through pricing model understanding, monitoring, and optimization strategies.
- Administer BigQuery projects, datasets, tables, and jobs efficiently for operational excellence.

Intended audience

This course is intended for data professionals responsible for designing, developing, and managing data warehousing solutions on Google Cloud:

- Data Engineers
- Data Architects
- Cloud Architects
- Senior Data Analysts
- Database Administrators
- DevOps Engineers managing data platforms

Prerequisites

We recommend that attendees of this course have:

- Intermediate SQL knowledge
- Familiarity with cloud computing concepts (preferably Google Cloud Platform)
- Basic understanding of data warehousing principles

Course outline

Module 1: Introduction to BigQuery for Data Warehousing

- BigQuery's serverless, petabyte-scale architecture and core capabilities
- BigQuery's role in the modern data analytics stack
- Key BigQuery concepts: Projects, Datasets, Tables, Views, Jobs
- Understanding BigQuery's query processing and storage mechanisms
- Demo: Setting up a BigQuery project and loading initial datasets

Module 2: BigQuery Storage Design and Best Practices

- Overview of BigQuery table types: standard, external, materialized views
- Partitioning: Time-unit partitioned tables (day, hour, month, year), integer-range partitioning, and ingestion-time partitioning
- Clustering: Concepts, benefits, selecting appropriate clustering columns, and combining with partitioning
- Schema design considerations: Denormalization vs. normalization in BigQuery, choosing optimal data types
- Hands-on Lab: Designing and implementing partitioned and clustered tables for various use cases

Module 3: Advanced Query Optimization Techniques

- Understanding BigQuery's query execution model and slot allocation
- SQL best practices for BigQuery: Avoiding SELECT *, efficient JOIN strategies, filtering early, and predicate pushdown
- Leveraging wildcard tables and efficient query patterns for large datasets
- Optimizing User-Defined Functions (UDFs) and complex subqueries
- Introduction to Materialized Views for query acceleration
- Hands-on Lab: Optimizing inefficient queries and applying best practices to reduce query cost and latency

Module 4: BigQuery Query Performance Analysis and Troubleshooting

- Interpreting query execution plans: Stages, shuffles, reads, and writes
- Identifying common performance bottlenecks: Skewed data, hot spots, excessive data scans, memory spills
- Using INFORMATION_SCHEMA and jobs.get API for detailed query analysis and metadata extraction

- Troubleshooting slot contention and managing concurrent query workloads
- Hands-on Lab: Diagnosing and resolving query performance issues using BigQuery UI, Cloud Monitoring, and INFORMATION_SCHEMA

Module 5: BigQuery Administration and Security

- Identity and Access Management (IAM) for BigQuery: Roles, permissions (project, dataset, table, and column-level access)
- Data security: Encryption at rest (default and Customer-Managed Encryption Keys - CMEK) and in transit
- Data governance: Implementing Authorized Views, Row-Level Security, and Data Masking
- Networking considerations: Private Google Access and VPC Service Controls for enhanced security
- Hands-on Lab: Configuring IAM roles, implementing row-level security, and setting up authorized views

Module 6: Cost Management and Monitoring in BigQuery

- Understanding BigQuery pricing models: On-demand vs. flat-rate (reservations)
- Strategies for cost optimization: Query limits, data expiration policies, efficient storage use, and billing alerts
- Monitoring BigQuery usage and performance with Cloud Monitoring and custom dashboards
- Auditing BigQuery activities: Leveraging Cloud Logging and Cloud Audit Logs for compliance and operational insights
- Hands-on Lab: Setting up cost alerts, analyzing billing reports, and implementing data lifecycle management

Module 7: Advanced BigQuery Features and Integrations

- Introduction to BigQuery ML: In-database machine learning for predictive analytics
- BigQuery Data Transfer Service: Automating data ingestion from various sources
- BigQuery BI Engine: Accelerating business intelligence dashboards and interactive queries
- Integration with other GCP services: Cloud Storage, Dataflow, Dataproc, Looker, Data Studio
- Working with external data sources and federated queries (Cloud Storage, Cloud SQL, Spanner, etc.)

- Hands-on Lab: Building a simple BigQuery ML model and performing federated queries against external data

Module 8: Course Wrap-up

- Recap of BigQuery data warehousing best practices for storage, query optimization, and administration
- Roadmap for continuous optimization, cost control, and operational excellence in BigQuery
- Additional resources for further learning, BigQuery documentation, and community engagement
- Q&A and discussion: Addressing specific challenges and use cases

Your course overview ends here.
But your cloud journey doesn't!

Unlock advanced training and certification pathways with **CloudThat**.



About CloudThat

Established in 2012, CloudThat is an award-winning company and the first in India to offer cloud training and consulting services for individuals and enterprises worldwide. Recently, it won Google Cloud's New Training Partner of the Year Award for 2025, becoming the first company in the world in 2025 to hold awards from all three major cloud giants: AWS, Microsoft, and Google. CloudThat notably won consecutive AWS Training Partner of the Year (APJ) awards in 2023 and 2024 and the Microsoft Training Services Partner of the Year Award in 2024, bringing its total award count to an impressive 12 awards in the last 8 years.

CloudThat specializes in Cloud Migration, Data Platforms, DevOps, IoT, and cutting-edge technologies like Gen AI & AI/ML. As a Microsoft Solutions Partner, AWS Premier Tier Training Partner, Google Cloud Platform Partner, and collaborator with leading organizations like HPE, Oracle, Nvidia and Databricks, CloudThat has trained over 850,000 professionals across 600+ cloud certifications, empowering students and professionals worldwide to advance their skills and careers.

To know more about our cloud certification training, email at sales@cloudthat.com or call us at [+91-8880002200](tel:+91-8880002200) or [+1-8555588830](tel:+1-8555588830) for those from outside India.