Bigquery on GCP

What is Bigquery?

Google Bigquery is a Cloud Datawarehouse powered by Google which is Serverless, highly scalable, and cost-effective designed for making data driven business decisions quickly. It offers both the batch and streaming insertion capabilities and is integrated with Tensorflow as well to perform machine learning using sql like dialects.

Bigquery Features

- Upload data and run SQL.
- Fully managed.
- Separate storage and computing.
- Compute scales with usage, without cluster resizing.
- Thousands of cores are utilized per query.
- Deployed across multiple data centers by default, with multiple factors of replication to optimize maximum data durability and service uptime.
- Stream millions of rows per second for real-time analysis.
- Terabytes of data analysed in seconds.
- Storage can scale to Petabytes.

Datatypes Supported in Bigquery

- Array
- Bool
- Bytes
- Date and Datetime
- Geography
- Numeric – Integer, Float, Decimal, BigdecimalS
- String
- Struct
- Time and Timestamp

Bigquery Interaction Modes

Cloud Console

- Run queries and examine results.
- Manage databases and tables.
- Save queries and share across the organization for re-use.
- Detailed Query history.

Data Studio - Quick visualization on top of Bigquery data.

API - To interact with bigquery programmatically.

Bigquery Limits

- Concurrent rate limit for on-demand, interactive queries — 50.
- Daily query size limit — Unlimited by default.
- Query execution time limit — 6 hours.
- Maximum number of tables referenced per query — 1,000.
- Daily destination table update limit — 1,500 updates per table per day
- Maximum number of standard SQL query parameters — 10,000
- Maximum columns in a table, query result, or view definition — 10,000
- Concurrent rate limit for on-demand, interactive queries against Cloud Big table external data sources — 4.