Google BigQuery

Google Cloud Platform

Google Cloud Platform

Agenda

1. BigQuery Overview
2. BigQuery Resources
3. Accessing BigQuery
4. Lab

BigQuery

- BigQuery provides a **service** for near real time interactive analysis of **massive** datasets (hundreds of TBs)
- Based on a **columnar** structure for high performance
- Query using a **SQL-like** syntax
- Only pay for storage, processing used
- **Zero** administration for performance and scale
- Supports **open** standards

“Nothing comes remotely close to the sheer power of Google BigQuery. It made large-scale data collection and crunching possible with little effort, which has translated to a significant business advantage.”

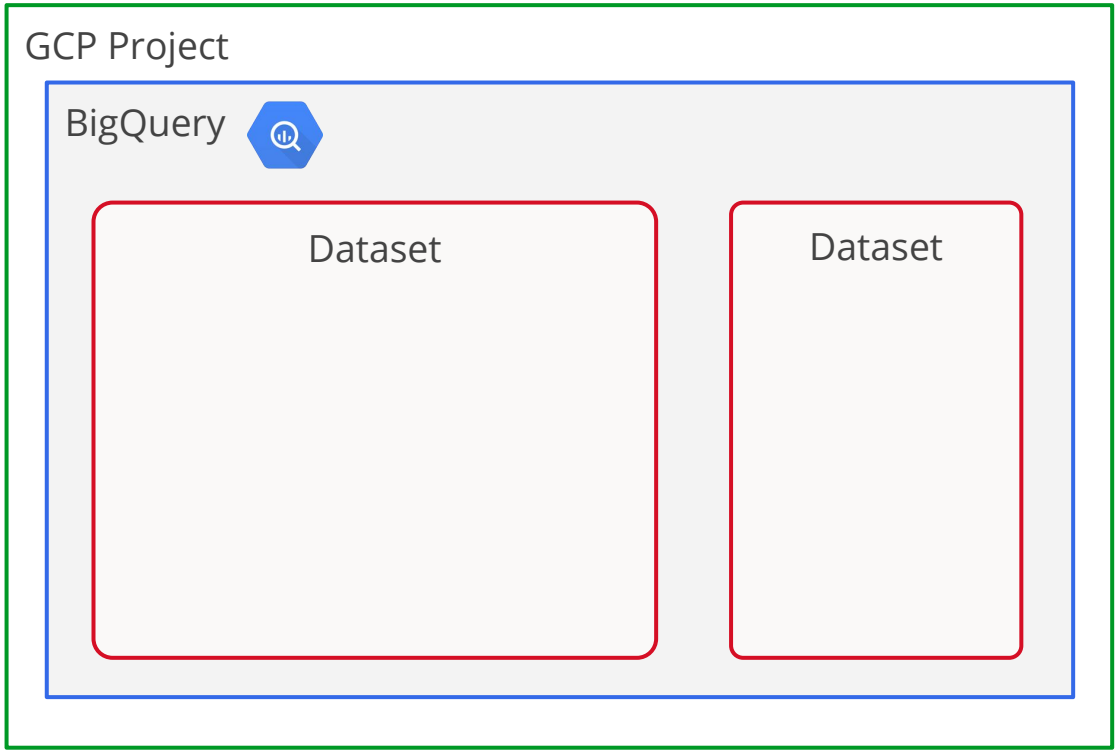
-Pradeep Kumar, Technical Architect

Agenda

1. BigQuery Overview
2. BigQuery Resources
3. Accessing BigQuery
4. Lab

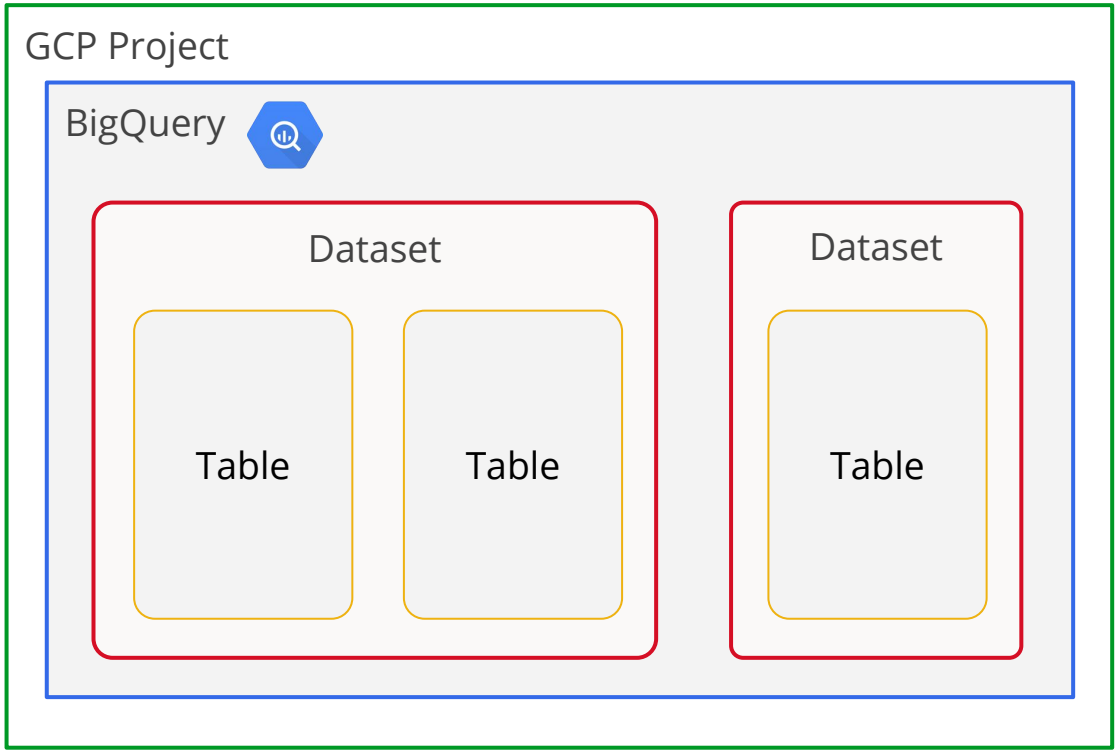
BigQuery Resources (1 of 2)

- Projects can host one or more **datasets**
- Access to data is typically controlled using Access Control Lists (ACLs) at the dataset level



BigQuery Resources (2 of 2)

- **Tables** contain data in BigQuery
- Used to define a schema for the data
- BigQuery also supports views (or virtual tables) defined by a SQL query



Sample Query

project: publicdata
dataset: samples
table: natality

```
SELECT
  weight_pounds, state, year, gestation_weeks
FROM
  publicdata:samples.natality
ORDER BY weight_pounds DESC LIMIT 10;
```

Pro Tip!

'natality' is one of a number of ready-made public data sample tables that are great for experimenting when getting started with BigQuery.

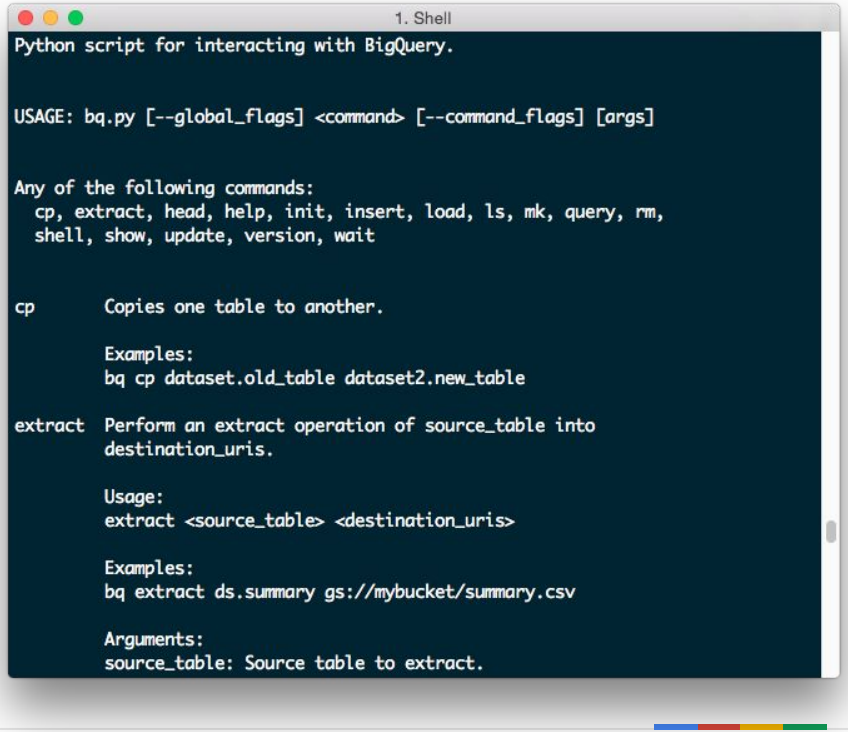
Agenda

1. BigQuery Overview
2. BigQuery Resources
3. Accessing BigQuery
4. Lab

Accessing BigQuery

BigQuery can be accessed using a variety of tools and methods:

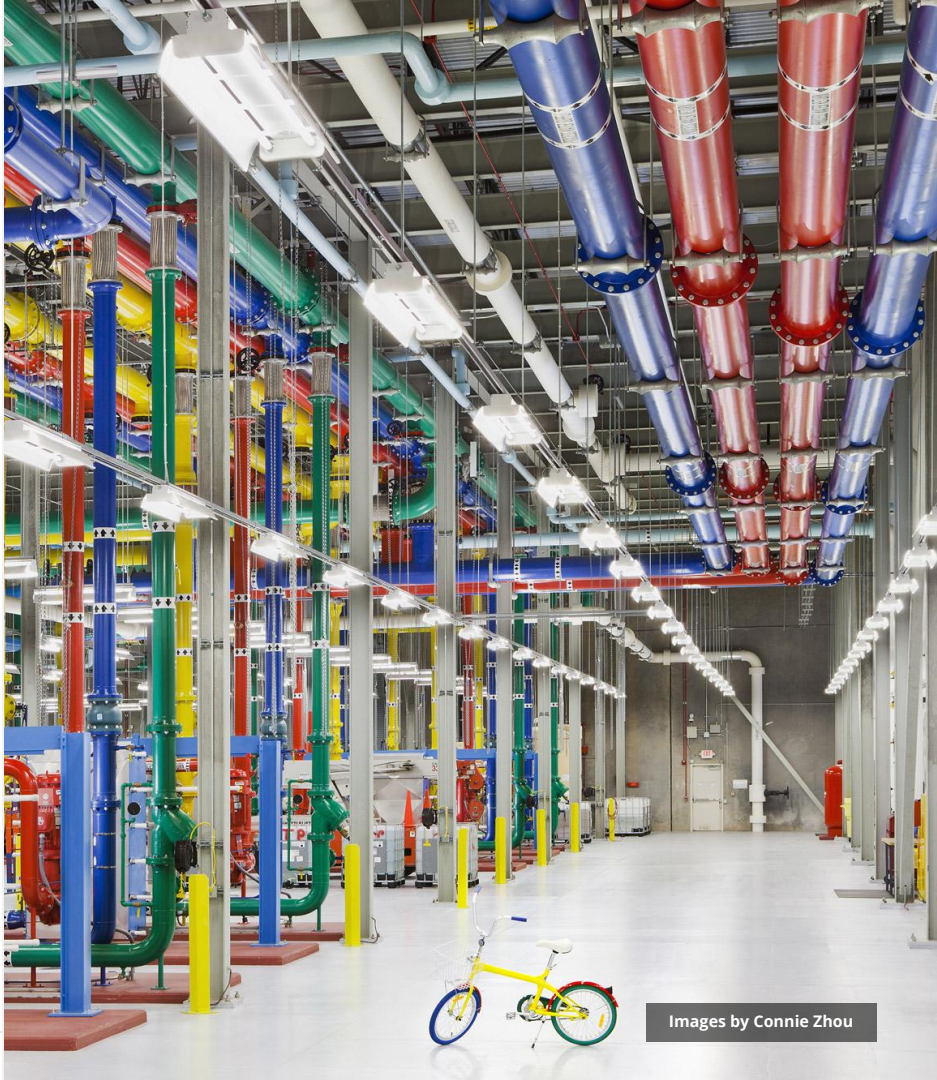
- **Web UI** (introduced in the demo) at https://bigquery.cloud.google.com/
- Cloud SDK using the **bq** utility with
  - The `query` command
  - Interactively via `shell` command
- RESTful JSON **API** via the client libraries
- Using a variety of **third party** tools
- Experimental support in **Apps Script**



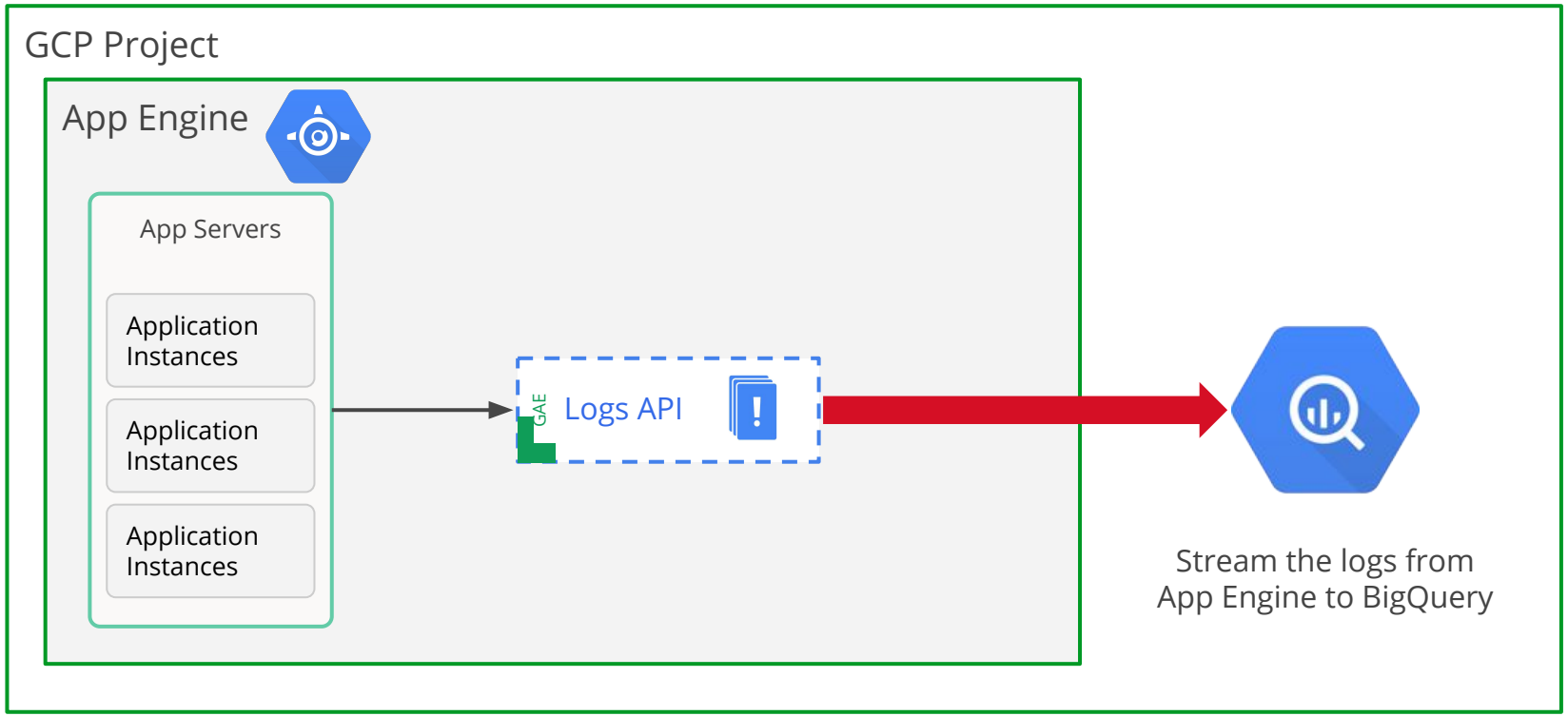
Lab (1 of 2)

Analyse the logs for the Guestbook application.

1. Enable BigQuery log streaming in App Engine
2. Generate some additional web requests in App Engine
3. Explore and query the log data as it streams into BigQuery



Lab (2 of 2)



Resources

- BigQuery: product demo, features, case studies, pricing, and documentation
  https://cloud.google.com/bigquery/
- DevBytes - What is BigQuery?
  https://www.youtube.com/watch?v=aupC-Wj7XDY
- BigQuery Web UI Quickstart
  https://cloud.google.com/bigquery/web-ui-quickstart
- Sample Tables
  https://cloud.google.com/bigquery/docs/sample-tables


cloud.google.com
Images by Connie Zhou