PROFESSIONAL DATA ENGINEER *- Google Cloud Platform*

TITLE: "Introduction to Google Cloud - Cloud BigTable"

Author:

- *Name: "Vignesh Sekar S"*
- *Designation: "Multi Cloud Architect"*

- *Tags: [Google Cloud, DataEngineer, Python, PySpark, SQL, BigData]*

Cloud Bigtable

1. Cloud Bigtable is Google's NoSQL Big Data database service.
2. It's the same database that powers many core Google services, including Search, Analytics, Maps, and Gmail.
3. Bigtable is designed to handle massive workloads at consistent low latency and high throughput, so it's a great choice for both operational and analytical applications, including IoT, user analytics, and financial data analysis.

Cloud Shell

1. Cloud Shell is a virtual machine that is loaded with development tools. It offers a persistent 5GB home directory and runs on the Google Cloud.
2. Cloud Shell provides command-line access to your Google Cloud resources.

Objective:

- How to use the ***cbt*** command line to connect to a Cloud Bigtable instance, perform basic administrative tasks, and read and write data in a table.

- Enable through gcli - bigtable gcloud services enable bigtable.googleapis.com bigtableadmin.googleapis.com

Task 1. Create a Cloud Bigtable instance

1. In the Cloud Console, go to Navigation menu menu, click on Bigtable in the Databases section, then click Create instance.

2. Fill in the fields for your instance as follows:

Field Value Instance name: bigtable-instance Instance ID: bigtable-instance Storage type: SSD Cluster ID: bigtable-instance-c1 Region: us-east1 Zone: us-east1-c

Task 2. Connect to your instance

1. In Cloud Shell, configure cbt to use your project and instance by modifying the .cbtrc file:

    a. echo project = `gcloud config get-value project` > ~/.cbtrc
    b. echo instance = myfirst-instance-kt-001 >> ~/.cbtrc

2. To get the lastest version of cbt in cloud shell cbt version

3. To get help in cbt version cbt help

4. Print godoc-suitable documentation for cbt cbt doc

5. List instances in a project cbt listinstances

6. List clusters in an instance cbt listclusters

Task 3. Read and write data

1. Cloud Bigtable stores data in tables, which contain rows. Each row is identified by a row key.

2. Data in a row is organized into column families, or groups of columns. A column qualifier identifies a single column within a column family.A cell is the intersection of a row and a column. Each cell can contain multiple versions of a value.

    a. Create a table named my-table:
        1. cbt createtable my-table
    b. List your tables:
        1. cbt ls
    c. Add one column family named cf1:
        1. cbt createfamily my-table cf1
        2. cbt createfamily my-table cf2
    d. List your column families:
        1. cbt ls my-table
    e. Put the value test-value in the row r1, using the column family cf1 and the column qualifier c1:
        1. cbt set my-table r1 cf1:c1=test-value
        2. cbt set my-table r1 cf1:c2=value
        3. cbt set my-table r3 cf1:c1=final-value
        4. cbt set my-table r2 cf1:c2=test
    f. Count rows in a table
        1. cbt count my-table
    g. Use the cbt read command to read the data you added to the table:
        1. cbt read my-table

h. Delete the table my-table:
    1. cbt deletetable my-table
i. Delete all rows
    1. cbt deleteallrows my-table

3. Delete a column family cbt deletefamily my-table cf2

4. Delete a cluster from the configured instance cbt deletecluster demoinstancebigtable1-c1

5. Delete an instance cbt deleteinstance demoinstancebigtable1

resources:

1. https://cloud.google.com/bigtable/docs/cbt-reference (https://cloud.google.com/bigtable/docs/cbt-reference)