Google Discussions

EXAM ASSOCIATE CLOUD ENGINEER TOPIC 1 QUESTION 85 DISCUSSION

Actual exam question from Google's Associate Cloud Engineer
Question #: 85
Topic #: 1
[All Associate Cloud Engineer Questions]

You have a large 5-TB AVRO file stored in a Cloud Storage bucket. Your analysts are proficient only in SQL and need access to the data stored in this file. You want to find a cost-effective way to complete their request as soon as possible. What should you do?

A. Load data in Cloud Datastore and run a SQL query against it.

B. Create a BigQuery table and load data in BigQuery. Run a SQL query on this table and drop this table after you complete your request.

C. Create external tables in BigQuery that point to Cloud Storage buckets and run a SQL query on these external tables to complete your request.

D. Create a Hadoop cluster and copy the AVRO file to NDFS by compressing it. Load the file in a hive table and provide access to your analysts so that they can run SQL queries.

Show Suggested Answer

by [deleted] at *June 11, 2020, 6:42 p.m.*

Comments

Type your comment...

Submit

mohdafiuddin Highly Voted 3 years, 9 months ago

Breaking down the question into key points -

1. 5-TB AVRO file stored in a Cloud Storage bucket.
2. Analysts are proficient only in SQL
3. cost-effective way to complete their request as soon as possible

A. ....Load data in Cloud Datastore... (Not Correct because Cloud Datastore is not a good option to run SQL Queries)

B. ...Load data in BigQuery.... (Not Cost Effective because loading the data which is already present in the bucket into BigQuery again is expensive)

C. Create external tables in BigQuery that point to Cloud Storage buckets and run a SQL query on these external tables to complete your request.
(This is the right answer as it meets all the requirements from the question)

D. Create a Hadoop cluster and copy the AVRO file to NDFS by compressing it. Load the file in a hive table and provide access to your analysts so that they can run SQL queries.
(Too roundabout and indirect. Not the right option)

upvoted 154 times

pondai 3 years, 7 months ago

listem this guy

upvoted 17 times

Ciumela Highly Voted 4 years, 4 months ago

C is correct: https://cloud.google.com/bigquery/external-data-sources

upvoted 23 times

dck4893 Most Recent 6 months, 1 week ago

D - The ONLY answer that provides data for the analysts is D. The question states that the analysts need access to the data and that they only know SQL (not that *you* only know SQL). The other 3 answers don't provide the data to analysts. You might have to fill in the blanks that you will pass it to them in a spreadsheet format, but that very likely won't satisfy their needs to query the data using SQL and at 5TB size, that isn't ideal. Therefore D is the only answer that satisfies the requirement.

upvoted 1 times

JB28 9 months, 2 weeks ago

The most cost-effective and efficient option would be Option C: Create external tables in BigQuery that point to Cloud Storage buckets and run a SQL query on these external tables to complete your request.

This approach allows you to query data directly from the AVRO file stored in the Cloud Storage bucket without having to load the data into BigQuery first. This saves both time and money as you are not charged for the storage of data within BigQuery. Plus, BigQuery is designed to be able to handle large datasets, making it a suitable choice for a 5-TB AVRO file. Your analysts, who are proficient in SQL, can easily work with BigQuery as it uses a SQL interface.

upvoted 1 times

BAofBK 11 months, 2 weeks ago

The correct answer is C

upvoted 1 times

Captain1212 1 year, 1 month ago

Selected Answer: C

C is the right answer , as

upvoted 1 times

Neha_Pallavi 1 year, 1 month ago

C. Create external tables in BigQuery that point to Cloud Storage buckets and run a SQL query on these external tables to complete your request.

upvoted 2 times

Neha_Pallavi 1 year, 1 month ago

C. Create external tables in BigQuery that point to Cloud Storage buckets and run a SQL query on these external tables to complete your request.

upvoted 1 times

Buruguduystunstugudunstuy 1 year, 8 months ago

Selected Answer: C

Answer C is the most cost-effective and efficient way to provide analysts access to the data stored in the 5-TB AVRO file in Cloud Storage.

Here's why:

You can create external tables in BigQuery that point to the 5-TB AVRO file stored in Cloud Storage. External tables allow you to query data stored in Cloud Storage without the need to load the data into BigQuery. This is a cost-effective way to provide your analysts' access to the data they need, and it is also an efficient solution since you can run SQL queries on the data directly in BigQuery.

upvoted 7 times

Emmanski08 1 year, 9 months ago

External tables in BigQuery

upvoted 1 times

cslince 1 year, 10 months ago

Selected Answer: C

C is correct: https://cloud.google.com/bigquery/external-data-sources

upvoted 1 times

leogor 1 year, 11 months ago

Selected Answer: C

external tables in BigQuery

upvoted 1 times

Untamables 2 years ago

Selected Answer: C

Similar to Athena

upvoted 1 times

Charumathi 2 years ago

C. is correct,
An external data source is a data source that you can query directly from BigQuery, even though the data is not stored in BigQuery storage.

BigQuery supports the following external data sources:

Amazon S3
Azure Storage
Cloud Bigtable
Cloud Spanner
Cloud SQL
Cloud Storage
Drive

upvoted 2 times

fifi1907 2 years, 1 month ago

answer is c

upvoted 1 times

AzureDP900 2 years, 4 months ago

mohdafiuddin explanation is very detailed .. C is right

upvoted 1 times

haroldbenites 2 years, 4 months ago

Go for C

upvoted 2 times

Load full discussion...