EXAM PROFESSIONAL DATA ENGINEER TOPIC 1 QUESTION 59 DISCUSSION

Actual exam question from Google's Professional Data Engineer
Question #: 59
Topic #: 1
[All Professional Data Engineer Questions]

An online retailer has built their current application on Google App Engine. A new initiative at the company mandates that they extend their application to allow their customers to transact directly via the application. They need to manage their shopping transactions and analyze combined data from multiple datasets using a business intelligence (BI) tool. They want to use only a single database for this purpose. Which Google Cloud database should they choose?

A. BigQuery

B. Cloud SQL

C. Cloud BigTable

D. Cloud Datastore

Show Suggested Answer

by Remi2021 at *Sept. 7, 2022, 5:35 p.m.*

Comments

Type your comment...

Submit

PolyMoe Highly Voted 2 years, 3 months ago

Selected Answer: B

B. Cloud SQL would be the most appropriate choice for the online retailer in this scenario. Cloud SQL is a fully-managed relational database service that allows for easy management and analysis of data using SQL. It is well-suited for applications built on Google App Engine and can handle the transactional workload of an e-commerce application, as well as the analytical workload of a BI tool.

upvoted 11 times

Aaronn14 Highly Voted 2 years, 1 month ago

A. "They want to use only a single database for this purpose" is a key requirement. You can use BigQuery for transactions, though it is not efficient. You can not use CloudSQL for analytics. So it is probably BQ.

upvoted 5 times

baimus 7 months, 2 weeks ago

Yeah my thinking was the same, but actually cloud SQL is fine to connect BI tools to, which is specified in this question.

upvoted 1 times

Mathew106 Most Recent 1 year, 9 months ago

Selected Answer: B

Cloud SQL seems to fit the best. It supports transactions and can be used to run queries and do analytics.

BigQuery is good for the analysis part but it's not good for managing transactions. If the question needed a database just to store the data for analysis it would be ok. But if we want to update single transactions or add them row by row, then it's not good. BigQuery is not made to support an application. It's a DW.

BigTable is can not carry transactions over multiple rows and is better for large scale analytics jobs. Also we should pick it for use-cases with high throughput/low latency requirements. Seems redundant.

upvoted 3 times

Siddhesh05 2 years ago

Selected Answer: A

Big Query because of analysis

upvoted 5 times

izekc 2 years ago

Selected Answer: C

Should be bigtable

upvoted 1 times

baimus 7 months, 2 weeks ago

I can't really see that. Bigtable is only ever the right choice for noSql at vast scale.

upvoted 2 times

juliobs 2 years, 1 month ago

Selected Answer: A

I think BigQuery makes sense here. It works for transactions too.

upvoted 4 times

juliobs 2 years, 1 month ago

I just did a session with an official trainer from Google that said BigTable is better.

upvoted 2 times

Fotofilico 1 year, 6 months ago

I'm an official trainer from Google and I can say that my best two options for this scenario would be Cloud SQL and BigQuery in that order.
Also we can consider datastore since we're using it with a web app, but it's another topic.

upvoted 2 times

certs4pk 5 months ago

but, how to analyze 'combined data from multiple datasets' in cloud sql?

upvoted 1 times

ninjatech 2 years, 2 months ago

Transactional Data need to written first by application before it could be analysed so cloudsql.

upvoted 2 times

samdhimal 2 years, 3 months ago

Both BigQuery and Cloud Bigtable are valid options for this use case, but BigQuery is better suited for this specific scenario where the retailer needs to manage and analyze large amounts of data from multiple datasets using a BI tool.

BigQuery is a fully-managed, cloud-native data warehouse that enables super-fast SQL queries using the processing power of Google's infrastructure. It can handle large, complex datasets and is well-suited for both transactional and analytical workloads. It can also handle data from multiple datasets and can be integrated with other Google Cloud services, such as Dataflow, Dataproc and Looker for BI analysis.

While Cloud Bigtable is also a good option for this use case as it is a highly scalable and performant NoSQL database that is well-suited for handling large amounts of data and high-write loads. It is not as good as BigQuery for analytical workloads and it may not be as well-suited for this specific scenario where the retailer needs to manage and analyze large amounts of data from multiple datasets using a BI tool.

upvoted 2 times

jin0 2 years, 2 months ago

Bigquery is a OLAP. So it could be not a answer I think.

upvoted 1 times

samdhimal 2 years, 3 months ago

Cloud SQL and Cloud Datastore are also good options for certain use cases, but they may not be as well-suited for this specific scenario where the retailer needs to manage and analyze large amounts of data from multiple datasets using a BI tool.

upvoted 2 times

desertlotus1211 2 years, 3 months ago

The Community is choosing Answer B - Cloud SQL, as per the question.
However when they explain - they're speaking about BQ[????]

So is it BigQuery or Cloud SQL?

upvoted 3 times

DipT 2 years, 4 months ago

Selected Answer: B

https://cloud.google.com/bigquery/docs/partitioned-tables

upvoted 1 times

DipT 2 years, 4 months ago

Selected Answer: B

It needs support for transaction so cloud sql is the choice of database and with Bigquery we can still analyze cloud sql data via federated queries https://cloud.google.com/bigquery/docs/reference/legacy-sql

upvoted 4 times

DGames 2 years, 4 months ago

Selected Answer: B

Most important part of question is transaction (RDBMS) strong ACID property database. Second part analysis of data, yes possible using any BI tool its possible with RDBMS db.

upvoted 1 times

zellck 2 years, 4 months ago

Selected Answer: B

B is the answer.

https://cloud.google.com/bigquery/docs/cloud-sql-federated-queries
BigQuery Cloud SQL federation enables BigQuery to query data residing in Cloud SQL in real time, without copying or moving data. Query federation supports both MySQL (2nd generation) and PostgreSQL instances in Cloud SQL.

upvoted 4 times

sjtesla 2 years ago

Agreed. Two catches here: transactional, and BI tool. Although BigQuery nowadays can handle everything, if we specifically deal with questions highlighting transactional data, I believe to differenticate services, we should choose what they primarily mean to be .

upvoted 2 times

odacir 2 years, 4 months ago

Selected Answer: B

C and D are not able to work with BI directly, so discard.
A: It's the best option for BI for awful for transactions
B: it's the best option for transaction, and works for BI, so this must be the answer

upvoted 2 times

Leeeeee 2 years, 5 months ago

Selected Answer: B

BigQuery for Analytics and BI

upvoted 1 times

Leelas 2 years, 6 months ago

Selected Answer: B

Cloud Sql is Used to store Transactional Data and supports Sql Transactions. Where as Big Query is used for Analytics.

upvoted 2 times

Zion0722 2 years, 6 months ago

Cloud SQL supports transactions as well as analysis through a BI tool. Firestore/Datastore does not support SQL syntax typically needed to do analysis done by a BI tool. BigQuery is not suitable for transactional use case. BigTable does not support SQL.
It's A.

upvoted 2 times

Load full discussion...