

[Google Discussions](#)

Exam Professional Data Engineer All Questions

View all questions & answers for the Professional Data Engineer exam

[Go to Exam](#)

EXAM PROFESSIONAL DATA ENGINEER TOPIC 1 QUESTION 103 DISCUSSION

Actual exam question from Google's Professional Data Engineer

Question #: 103

Topic #: 1

[\[All Professional Data Engineer Questions\]](#)

You have a data stored in BigQuery. The data in the BigQuery dataset must be highly available. You need to define a storage, backup, and recovery strategy of this data that minimizes cost. How should you configure the BigQuery table that have a recovery point objective (RPO) of 30 days?

- A. Set the BigQuery dataset to be regional. In the event of an emergency, use a point-in-time snapshot to recover the data.
- B. Set the BigQuery dataset to be regional. Create a scheduled query to make copies of the data to tables suffixed with the time of the backup. In the event of an emergency, use the backup copy of the table.
- C. Set the BigQuery dataset to be multi-regional. In the event of an emergency, use a point-in-time snapshot to recover the data.
- D. Set the BigQuery dataset to be multi-regional. Create a scheduled query to make copies of the data to tables suffixed with the time of the backup. In the event of an emergency, use the backup copy of the table.

[Show Suggested Answer](#)

by [AWSandeep](#) at *Sept. 3, 2022, 2:04 p.m.*

Comments

Type your comment...

[Submit](#)

  **DeepakVenkatachalam** Highly Voted  1 year, 7 months ago

Answer is B. Timetravel only covers for 7 days and a scheduled query is needed for creating Table snapshots for 30 days. Also table snapshot must remain in the same region as base table (please refer to limitation of table snapshot from below link) <https://cloud.google.com/bigquery/docs/table-snapshots-intro>

   upvoted 10 times

  **desertlotus1211** Highly Voted  2 years, 3 months ago

Answer is C: <https://cloud.google.com/bigquery/docs/table-snapshots-intro>
"Benefits of using table snapshots include the following:



Keep a record for longer than seven days. With BigQuery time travel, you can only access a table's data from seven days ago or more recently. With table snapshots, you can preserve a table's data from a specified point in time for as long as you want.

Minimize storage cost. BigQuery only stores bytes that are different between a snapshot and its base table, so a table snapshot typically uses less storage than a full copy of the table."

But the wording is foolish... It's table snapshot, NOT point in time snapshot!

<https://cloud.google.com/bigquery/docs/time-travel#restore-a-table>
this is point in time using time travel window - max is 7 days...

   upvoted 5 times

  **desertlotus1211** 1 month, 2 weeks ago

Sorry folks - I change my answer to D... C is not correct at it can ONLY go back for 7 days, max!

   upvoted 1 times

  **Parandhaman_Margan** Most Recent  1 month, 3 weeks ago

Selected Answer: B

Minimizing Cost

   upvoted 2 times

  **imarri876** 1 month, 3 weeks ago

Selected Answer: B



BigQuery multi-regional is more expensive than BQ regional.

   upvoted 2 times

  **imarri876** 1 month, 3 weeks ago

After deeper review, with pricing calculator and other resources, the answer is C.

   upvoted 2 times

  **loki82** 3 months, 1 week ago

Selected Answer: A

RPO is NOT the same as a PITR window. Point-in-time recovery (PITR) is a process that allows users to restore data or settings from a previous point in time. A recovery point objective (RPO) is the maximum amount of data loss that an organization can tolerate after a data loss event. So a PITR snapshot easily meets an RPO of 30 days. A regional bucket minimizes cost.

   upvoted 1 times

  **hussain.sain** 4 months, 1 week ago

Selected Answer: C

Answer is C. As question is related to highly available so this rules out A and B.

   upvoted 1 times

  **clouditis** 4 months, 3 weeks ago

Selected Answer: C

Because this option uses Multiregional & BQ Snapshot, others are not right/cumbersome

   upvoted 1 times

  **cloud_rider** 5 months ago

Selected Answer: A



A is the right Answer

   upvoted 1 times

  **Erg_de** 6 months ago

Selected Answer: B

Best choice, minimized cost

   upvoted 2 times

🗄️ 👤 **Gcpteamprep** 6 months, 1 week ago

Selected Answer: B

Minimized Cost with Regional Storage: Regional datasets are less costly than multi-regional datasets in BigQuery. Since there is no requirement here for multi-regional availability, regional storage meets the high availability need while keeping costs lower.

RPO Compliance with Scheduled Backups: A scheduled query that periodically creates copies of the data (e.g., daily or weekly, depending on the requirements) allows for recovery within the 30-day RPO, meeting the requirement for data retention and recovery.

Point-in-Time Recovery Not Native in BigQuery: Although BigQuery provides a limited "table snapshot" feature, it's not a true point-in-time recovery option for the last 30 days. Creating periodic backups through scheduled queries gives you control over retention, enabling you to keep backups for 30 days and reducing dependency on more costly or limited snapshot capabilities.

👍 🔄 🚩 upvoted 4 times

🗄️ 👤 **Vogangster** 6 months, 3 weeks ago

D. Create monthly snapshots of a table by using a service account that runs a scheduled query. Link: <https://cloud.google.com/bigquery/docs/table-snapshots-scheduled>

👍 🔄 🚩 upvoted 1 times

🗄️ 👤 **AlizCert** 11 months ago

Selected Answer: D

HA => multi-region
30-days RPO => manual backups as max time-travel is 7 days

👍 🔄 🚩 upvoted 3 times

🗄️ 👤 **Lestrang** 11 months, 2 weeks ago

This is in one of google's training practice questions and the answer for it is C.

👍 🔄 🚩 upvoted 3 times

🗄️ 👤 **NickNtaken** 11 months, 2 weeks ago

Agreed. Multi-regional datasets offer higher availability by replicating data across multiple regions

👍 🔄 🚩 upvoted 1 times

🗄️ 👤 **rocky48** 1 year, 5 months ago

Selected Answer: A

ou should consider option A.

Setting the BigQuery dataset to be regional and using a point-in-time snapshot to recover the data in the event of an emergency can help you achieve the desired level of availability and minimize cost. This approach can help you avoid the additional cost of creating and maintaining backup copies of the data, which can be expensive.

Setting the BigQuery dataset to be multi-regional (options C and D) can provide additional redundancy and availability. However, this approach can be more expensive than setting the dataset to be regional, especially if you do not require the additional level of redundancy.

👍 🔄 🚩 upvoted 4 times

🗄️ 👤 **Nirca** 1 year, 7 months ago

Selected Answer: A

I'm going for A:

1. Set the BigQuery dataset to be regional. This will reduce the cost of storage compared to a multi-regional dataset.
2. building Snapshot: `bq snapshot --dataset <dataset_id> --table <table_id> <snapshot_id>`

👍 🔄 🚩 upvoted 4 times

🗄️ 👤 **ffggrrre** 1 year, 6 months ago

typically Multi-region cost is equal or less than a region. <https://cloud.google.com/bigquery/pricing#storage>

👍 🔄 🚩 upvoted 1 times

🗄️ 👤 **ckanaar** 1 year, 7 months ago

I think the answer is A:

This option meets the 30-day RPO requirement, assuming that the snapshot is maintained for that long. It offers high availability as data is written synchronously to 2 zones within a region: <https://cloud.google.com/blog/topics/developers-practitioners/backup-disaster-recovery-strategies-bigquery/>. The cost would be lower than maintaining a multi-regional dataset, but you'll incur the cost of the snapshot.

👍 🔄 🚩 upvoted 4 times

🗄️ 👤 **lucaluca1982** 2 years, 1 month ago

Why not B? Setting dataset regional or multi does not affect the backup and recovery strategy.

   upvoted 3 times

[Load full discussion...](#)



Platform

> [Home](#)

> [Examtopics PRO](#)

> [All Exams](#)

> [Training Courses](#)



© 2024 ExamTopics