

[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 122 DISCUSSION

Actual exam question from Google's Professional Data Engineer

Question #: 122

Topic #: 1

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

You decided to use Cloud Datastore to ingest vehicle telemetry data in real time. You want to build a storage system that will account for the long-term data growth, while keeping the costs low. You also want to create snapshots of the data periodically, so that you can make a point-in-time (PIT) recovery, or clone a copy of the data for Cloud Datastore in a different environment. You want to archive these snapshots for a long time. Which two methods can accomplish this?
(Choose two.)

- A. Use managed export, and store the data in a Cloud Storage bucket using Nearline or Coldline class.
- B. Use managed export, and then import to Cloud Datastore in a separate project under a unique namespace reserved for that export.
- C. Use managed export, and then import the data into a BigQuery table created just for that export, and delete temporary export files.
- D. Write an application that uses Cloud Datastore client libraries to read all the entities. Treat each entity as a BigQuery table row via BigQuery streaming insert. Assign an export timestamp for each export, and attach it as an extra column for each row. Make sure that the BigQuery table is partitioned using the export timestamp column.
- E. Write an application that uses Cloud Datastore client libraries to read all the entities. Format the exported data into a JSON file. Apply compression before storing the data in Cloud Source Repositories.








[Show Suggested Answer](#)































by [rickywk](#) at March 17, 2020, 12:49 p.m.











Comments

Type your comment...

Submit

-   **Ganshank** Highly Voted 4 years, 6 months ago
A,B
<https://cloud.google.com/datastore/docs/export-import-entities>
   upvoted 38 times
-   **salsabilsf** 3 years, 6 months ago
"while keeping the costs"

should be A,D
   upvoted 6 times
-   **MrCastro** 3 years, 2 months ago
Big query streaming inserts ARE NOT cheap
   upvoted 9 times
-   **hellofrnds** 3 years, 1 month ago
If you use B , not D , how can we do "point in time" recovery? is it possible?
Point in time recovery needs export along with timestamp, so that we can recover for a particular timestamp.
   upvoted 4 times
-   **atnafu2020** Highly Voted 4 years, 2 months ago
AC
<https://cloud.google.com/datastore/docs/export-import-entities>
C: To import only a subset of entities or to import data into BigQuery, you must specify an entity filter in your export.
B: Not correct since you want to store in a different environment than Datastore. Tho this statment is true: Data exported from one Datastore mode database can be imported into another Datastore mode database, even one in another project.
A is correct
Billing and pricing for managed exports and imports in Datastore
Output files stored in Cloud Storage count towards your Cloud Storage data storage costs.
Steps to Export all the entities
1. Go to the Datastore Entities Export page in the Google Cloud Console.
2. Go to the Datastore Export page
2. Set the Namespace field to All Namespaces, and set the Kind field to All Kinds.
3. Below Destination, enter the name of your "Cloud Storage bucket".
4. Click Export.
   upvoted 23 times
-   **aparna4387** 2 years, 12 months ago
<https://cloud.google.com/datastore/docs/export-import-entities#import-into-bigquery>
Data exported without specifying an entity filter cannot be loaded into BigQuery. This is not mentioned explicitly. Safe to assume there is no filter on the exports. So options are AB
   upvoted 6 times
-   **AzureDP900** 1 year, 10 months ago
A, B is perfect
   upvoted 1 times
-   **tavva_prudhvi** 2 years, 7 months ago
As you've mentioned in B, does the environment meant to be a project or a resource? As, we can clone a copy of the data in a datastore even in another project!?. Then, it's B.

Also, in point C they didn't mention any entity filter hence we eliminate C how can you support your own statement with a different answer?
   upvoted 2 times
-   **Yiouk** 3 years, 3 months ago
C is valid because of table snapshots. Else standard time travel is valid only for 7 days
https://cloud.google.com/bigquery/docs/table-snapshots-intro#table_snapshots
<https://cloud.google.com/bigquery/docs/time-travel#limitation>
   upvoted 1 times
-   **Shashank** 3 years, 1 month ago

 **Cniseajcole** 3 years, 1 month ago

you wanna say invalid?



   upvoted 1 times

  **CGS22** Most Recent 7 months ago

Selected Answer: AB

<https://cloud.google.com/datastore/docs/export-import-entities>

   upvoted 1 times

  **kskssk** 1 year, 2 months ago

AB chatgpt

A. Use managed export, and store the data in a Cloud Storage bucket using Nearline or Coldline class:

Managed export is a feature provided by Cloud Datastore to export your data.



Storing the data in a Cloud Storage bucket, especially using Nearline or Coldline storage classes, helps keep storage costs low while allowing you to retain the snapshots for a long time.

B. Use managed export, and then import to Cloud Datastore in a separate project under a unique namespace reserved for that export:

This method allows you to create snapshots by exporting data from Cloud Datastore (using managed export) and then importing it into a separate project under a unique namespace.

By importing into a separate project, you can keep a copy of the data in a different environment, which is useful for point-in-time recovery or creating clones of the data.

   upvoted 5 times

  **zelck** 1 year, 11 months ago

Selected Answer: AB

AB is the answer.

   upvoted 3 times

  **NicolasN** 1 year, 11 months ago

Selected Answer: AB

A rather complicated question, of a kind I wish I won't face in the exam. My opinion:

[A] A valid and cost-effective solution satisfying the requirement for PIT recovery

[B] A valid solution but far from ideal for archiving. It satisfies the requirement part "you can ... clone a copy of the data for Cloud Datastore in a different environment" (an objection to the word "namespace", I think it should be just "name")

   upvoted 11 times

  **NicolasN** 1 year, 11 months ago

[C] There is the limitation "Data exported without specifying an entity filter cannot be loaded into BigQuery". The entity filter for this case should contain all the kinds of entities but there is another limitation of "100 entity filter combinations". We have no knowledge of the kinds or the namespaces of the entities.

Sources:

<https://cloud.google.com/datastore/docs/export-import-entities#import-into-bigquery>

https://cloud.google.com/datastore/docs/export-import-entities#exporting_specific_kinds_or_namespaces

[D] seems a detailed candidate solution but it violates the limitation "You cannot append Datastore export data to an existing table."

https://cloud.google.com/bigquery/docs/loading-data-cloud-datastore#appending_to_or_overwriting_a_table_with_cloud_datastore_data

[E] Cloud Source Repositories are for source code and not a suitable storage for this case.

   upvoted 11 times

  **John_Pongthorn** 2 years, 1 month ago

Selected Answer: AB

<https://cloud.google.com/datastore/docs/export-import-entities>

   upvoted 1 times

  **John_Pongthorn** 2 years, 1 month ago

Selected Answer: AB

<https://cloud.google.com/datastore/docs/export-import-entities>

   upvoted 1 times

  **John_Pongthorn** 2 years, 1 month ago

Selected Answer: AB

The answer is nothing to do with bigquery , so you can skip what mention to bigquery.

A B is the final answer

   upvoted 1 times

👍 🔄 🚩 upvoted 1 times

🗨️ 👤 **DataEngineer_WideOps** 2 years, 3 months ago

A,B

For those who say using BQ as archival, How can we achieve that while datastore are NO-SQL whereas BQ are SQL , will that work? also BQ are not created for achieving purposes.

👍 🔄 🚩 upvoted 2 times

🗨️ 👤 **AmirN** 2 years, 4 months ago

Option B is 36 times more expensive than C

👍 🔄 🚩 upvoted 1 times

🗨️ 👤 **Nico1310** 2 years, 9 months ago

Selected Answer: AB

AB. for sure streaming to BQ its quite expensive!

👍 🔄 🚩 upvoted 2 times

🗨️ 👤 **MaxNRG** 2 years, 10 months ago

Selected Answer: AD

A - Cloud Storage (long-term data + costs low)

D - BigQuery (timestamp for point-in-time (PIT) recovery)

👍 🔄 🚩 upvoted 3 times

🗨️ 👤 **tavva_prudhvi** 2 years, 6 months ago

D is wrong, BQ Streaming inserts costs are high!

👍 🔄 🚩 upvoted 2 times

🗨️ 👤 **MaxNRG** 10 months, 3 weeks ago

Agreed, AB

<https://cloud.google.com/datastore/docs/export-import-entities>

👍 🔄 🚩 upvoted 1 times

🗨️ 👤 **medeis_jar** 2 years, 10 months ago

Selected Answer: AB

Option A; Cheap storage and it is a supported method <https://cloud.google.com/datastore/docs/export-import-entities>

Option B; Rationale - "Data exported from one Datastore mode database can be imported into another Datastore mode database, even one in another project." <<https://cloud.google.com/datastore/docs/export-import-entities>>

👍 🔄 🚩 upvoted 2 times

🗨️ 👤 **squishy_fishy** 3 years ago

Answer is A, B.

https://cloud.google.com/datastore/docs/export-import-entities#exporting_specific_kinds_or_namespaces

👍 🔄 🚩 upvoted 1 times

🗨️ 👤 **sergio6** 3 years ago

A, D

A: Option for storage system that will account for the long-term data growth

D: Option for snapshots, PIT recovery, copy of the data for Cloud Datastore in a different environment and, above all, archive snapshots for a long time

B: not a good solution for archiving snapshots for a long time

C: to import data into BigQuery, you must specify an entity filter

E: Cloud Source Repositories is for code

One note: E --> would be my second choice if there was Cloud Storage instead of Source Repositories (typo?)

👍 🔄 🚩 upvoted 4 times

🗨️ 👤 **Chelseajcole** 3 years, 1 month ago

Vote A B . What's the purpose load into bigquery?

👍 🔄 🚩 upvoted 1 times

🗨️ 👤 **Chelseajcole** 3 years, 1 month ago

<https://cloud.google.com/datastore/docs/export-import-entities#import-into-bigquery>

Importing into BigQuery

To import data from a managed export into BigQuery, see Loading Datastore export service data.

Data exported without specifying an entity filter cannot be loaded into BigQuery. If you want to import data into BigQuery, your export request must include one or more kind names in the entity filter.

You have to specify an entity filter before you can load from datastore to BQ. It didn't mention that at all. So C is incorrect

👍 🔄 🚩 upvoted 3 times

[Load full discussion...](#)



Platform

> [Home](#)

> [Examtopics PRO](#)

> [All Exams](#)

> [Training Courses](#)



© 2024 ExamTopics