≡ MENU

🔍

↩ **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 23 DISCUSSION**

Actual exam question from Google's Professional Data Engineer

Question #: 23

Topic #: 1

[All Professional Data Engineer Questions]

You are deploying 10,000 new Internet of Things devices to collect temperature data in your warehouses globally. You need to process, store and analyze these very large datasets in real time. What should you do?

A. Send the data to Google Cloud Datastore and then export to BigQuery.

B. Send the data to Google Cloud Pub/Sub, stream Cloud Pub/Sub to Google Cloud Dataflow, and store the data in Google BigQuery.

C. Send the data to Cloud Storage and then spin up an Apache Hadoop cluster as needed in Google Cloud Dataproc whenever analysis is required.

D. Export logs in batch to Google Cloud Storage and then spin up a Google Cloud SQL instance, import the data from Cloud Storage, and run an analysis as needed.

Show Suggested Answer

by [deleted] at *March 18, 2020, 4:38 p.m.*

Comments

Type your comment...

Submit

👤 **[Removed]** `Highly Voted 👍` 4 years, 7 months ago

Answer: B

👍 ↩ 🚩 upvoted 30 times

👤 **[Removed]** 4 years, 7 months ago

https://cloud.google.com/blog/products/iot-devices/quick-and-easy-way-set-end-end-iot-solution-google-cloud-platform

👍 ↩ 🚩 upvoted 9 times

👤 **[Removed]** `Highly Voted 👍` 4 years, 7 months ago

Answer: B
Description: Pubsub for realtime, Dataflow for pipeline, Bigquery for analytics

👍 ↩ 🚩 upvoted 26 times

👤 **Abizi** `Most Recent ⊘` 2 months ago

`Selected Answer: B`

B is the correct answer

👍 ↩ 🚩 upvoted 1 times

👤 **regal_2010** 6 months, 3 weeks ago

`Selected Answer: B`

Answer is B

👍 ↩ 🚩 upvoted 1 times

👤 **axantroff** 11 months, 2 weeks ago

`Selected Answer: B`

In short, B is less complex and more recommended other than D

👍 ↩ 🚩 upvoted 1 times

👤 **rtcpost** 1 year ago

`Selected Answer: B`

B. Send the data to Google Cloud Pub/Sub, stream Cloud Pub/Sub to Google Cloud Dataflow, and store the data in Google BigQuery.

Here's why this approach is preferred:

Google Cloud Pub/Sub allows for efficient ingestion and real-time data streaming.
Google Cloud Dataflow can process and transform the streaming data in real-time.
Google BigQuery is a fully managed, highly scalable data warehouse that is well-suited for real-time analysis and querying of large datasets.

👍 ↩ 🚩 upvoted 1 times

👤 **GCP_PDE_AG** 1 year, 2 months ago

Obviously B.

👍 ↩ 🚩 upvoted 1 times

👤 **Maurilio_Cardoso** 1 year, 5 months ago

`Selected Answer: B`

PubSub for queue in real time, Dataflow for processing (pipeline) and Bigquery for analyses.

👍 ↩ 🚩 upvoted 2 times

👤 **bha11111** 1 year, 7 months ago

`Selected Answer: B`

B is correct

👍 ↩ 🚩 upvoted 1 times

👤 **DGames** 1 year, 10 months ago

`Selected Answer: B`

GCP recommend best practice for streaming data pipeline as option B - pub/sub, dataflow & Bigquery

👍 ↩ 🚩 upvoted 1 times

👤 **Nirca** 1 year, 10 months ago

`Selected Answer: B`

B. Send the data to Google Cloud Pub/Sub, stream Cloud Pub/Sub to Google Cloud Dataflow, and store the data in Google BigQuery.

👍 ↩ 🚩 upvoted 1 times

👤 **gitaexams** 1 year, 11 months ago

B aris tqve yleebo

👍 ↩ 🚩 upvoted 1 times

⊟ 👤 **devaid** 2 years ago

Selected Answer: B

B of course

👍 ↩ 🚩 upvoted 1 times

⊟ 👤 **Dip1994** 2 years, 3 months ago

B is the correct answer

👍 ↩ 🚩 upvoted 1 times

⊟ 👤 **noob_master** 2 years, 4 months ago

Selected Answer: B

Answer: B

Deafult ETL streaming process: Pub/Sub + Dataflow + Bigquery.

👍 ↩ 🚩 upvoted 1 times

⊟ 👤 **nexus1_** 2 years, 4 months ago

Definitely B

👍 ↩ 🚩 upvoted 1 times

⊟ 👤 **vw13** 2 years, 6 months ago

Selected Answer: B

B is the only option for real time process & analysis

👍 ↩ 🚩 upvoted 1 times

Load full discussion...