☰ 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 117 DISCUSSION**

Actual exam question from Google's Professional Data Engineer

Question #: 117

Topic #: 1

[All Professional Data Engineer Questions]

You are designing a data processing pipeline. The pipeline must be able to scale automatically as load increases. Messages must be processed at least once and must be ordered within windows of 1 hour. How should you design the solution?

    A. Use Apache Kafka for message ingestion and use Cloud Dataproc for streaming analysis.

    B. Use Apache Kafka for message ingestion and use Cloud Dataflow for streaming analysis.

    C. Use Cloud Pub/Sub for message ingestion and Cloud Dataproc for streaming analysis.

    D. Use Cloud Pub/Sub for message ingestion and Cloud Dataflow for streaming analysis.

Show Suggested Answer

by 👤 madhu1171 at *March 15, 2020, 4:23 a.m.*

Comments

Type your comment…

Submit

⊟ 👤 **madhu1171** `Highly Voted 👍` 4 years, 1 month ago
Answer should be D
👍 ↩ 🚩 upvoted 27 times

⊟ 👤 [Removed] `Highly Voted 👍` 4 years, 1 month ago

☐ 📦 [Removed] `Highly Voted` 4 years, 1 month ago

Answer - D

👍 ↩ ⚑ upvoted 14 times

☐ 👤 **grshankar9** `Most Recent ⊘` 3 months, 2 weeks ago

`Selected Answer: D`

Kafka is recommended over Pub/Sub only when there is requirement of high throughput, complex streaming, more flexibility with customization and fine-tuning configurations or when application spans multiple cloud providers and requires more flexibility in deployment across different platforms.

👍 ↩ ⚑ upvoted 1 times

☐ 👤 **NeoNitin** 9 months ago

Data proc is serverbased
Dataflow is serverless which is used to run pipelines which uses apache framework in the background. Just need to mention the number of workers needed.

so question saying we need scale automatically . so dataproc eliminate ho gaya
now Dataflow is correct , pub/sub is recommended for this scenario. D

👍 ↩ ⚑ upvoted 1 times

☐ 👤 **dconesoko** 1 year, 4 months ago

`Selected Answer: D`

google's preferred choice

👍 ↩ ⚑ upvoted 2 times

☐ 👤 **zellck** 1 year, 5 months ago

D is the answer.

👍 ↩ ⚑ upvoted 1 times

☐ 👤 **Pime13** 1 year, 10 months ago

`Selected Answer: D`

Answer should be D

👍 ↩ ⚑ upvoted 1 times

☐ 👤 **VictorBa** 2 years ago

`Selected Answer: D`

It cannot be C because Dataproc is more suitable for Hadoop jobs.

👍 ↩ ⚑ upvoted 1 times

☐ 👤 **medeis_jar** 2 years, 3 months ago

`Selected Answer: D`

Pub/Sub + Dataflow

👍 ↩ ⚑ upvoted 1 times

☐ 👤 **MaxNRG** 2 years, 4 months ago

`Selected Answer: D`

D: Pub/Sub + Dataflow
https://cloud.google.com/solutions/stream-analytics/
https://cloud.google.com/blog/products/data-analytics/streaming-analytics-now-simpler-more-cost-effective-cloud-dataflow

👍 ↩ ⚑ upvoted 4 times

☐ 👤 **hendrixlives** 2 years, 4 months ago

`Selected Answer: D`

D: "at least once and must be ordered within windows" means Pub/Sub (at least once) with Dataflow (windows).

👍 ↩ ⚑ upvoted 3 times

☐ 👤 **JG123** 2 years, 5 months ago

Correct: D

👍 ↩ ⚑ upvoted 3 times

☐ 👤 **Chelseajcole** 2 years, 7 months ago

rule of thumb: If you see Kafka and Pub/Sub, always go with Pub/Sub in Google exam

👍 ↩ ⚑ upvoted 8 times

 ☐ 👤 **hendrixlives** 2 years, 4 months ago

 Careful doing that: I got a question where you had to choose between Kafka and Pub/Sub... and the solution required to be able to replay all messages without time limit. So no Pub/Sub there.
 This being a Google cert does not mean that they always force Google solutions.

 👍 ↩ ⚑ upvoted 7 times

sandipk91 2 years, 8 months ago

Answer is D

👍 ↰ ⚑ upvoted 2 times

awssp12345 2 years, 10 months ago

https://cloud.google.com/architecture/migrating-from-kafka-to-pubsub#comparing_features

👍 ↰ ⚑ upvoted 2 times

sumanshu 2 years, 10 months ago

Vote for D

Scaling - Dataflow.
Delivery of confimed atleast 1 message - Pub/Sub

👍 ↰ ⚑ upvoted 3 times

Sush12 3 years, 2 months ago
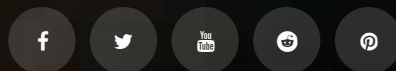
Answer is D

👍 ↰ ⚑ upvoted 2 times

Load full discussion…