Google Discussions

EXAM PROFESSIONAL DATA ENGINEER TOPIC 1 QUESTION 121 DISCUSSION

Actual exam question from Google's Professional Data Engineer
Question #: 121
Topic #: 1
[All Professional Data Engineer Questions]

You currently have a single on-premises Kafka cluster in a data center in the us-east region that is responsible for ingesting messages from IoT devices globally.
Because large parts of globe have poor internet connectivity, messages sometimes batch at the edge, come in all at once, and cause a spike in load on your
Kafka cluster. This is becoming difficult to manage and prohibitively expensive. What is the Google-recommended cloud native architecture for this scenario?

A. Edge TPUs as sensor devices for storing and transmitting the messages.

B. Cloud Dataflow connected to the Kafka cluster to scale the processing of incoming messages.

C. An IoT gateway connected to Cloud Pub/Sub, with Cloud Dataflow to read and process the messages from Cloud Pub/Sub.

D. A Kafka cluster virtualized on Compute Engine in us-east with Cloud Load Balancing to connect to the devices around the world.

Show Suggested Answer

by [deleted] at *March 22, 2020, 11:22 a.m.*

Comments

Type your comment...

Submit

[Removed] Highly Voted 3 years, 7 months ago

Should be C

upvoted 21 times

[Removed] Highly Voted 3 years, 7 months ago

Answer: C
Description: Pubsub is global and dataflow can scale workers

upvoted 19 times

ga8our Most Recent 5 months, 1 week ago

Can anyone pls explain what's wrong with D, the load balancing solution?

upvoted 1 times

musumusu 8 months, 2 weeks ago

Answer C:
What is wrong with D, nothing, Cloud load balancing can shift traffic for high volume and low internet in one region. It cost avg. 0.01-0.25 $ per GB, or if volume is too high. 0.05 $ per Hour http request. This might be the answer if your exam for network engineer.

upvoted 2 times

musumusu 8 months, 3 weeks ago

Answer C, but it will not solve bad internet connection, make sure 100mbps speed of internet is at sensor side.

upvoted 1 times

zellck 11 months ago

Selected Answer: C

C is the answer.

https://cloud.google.com/architecture/iot-overview#cloud-pubsub
Pub/Sub can act like a shock absorber and rate leveller for both incoming data streams and application architecture changes. Many devices have limited ability to store and retry sending telemetry data. Pub/Sub scales to handle data spikes that can occur when swarms of devices respond to events in the physical world, and buffers these spikes to help isolate them from applications monitoring the data.

upvoted 10 times

AzureDP900 10 months, 1 week ago

Agree with your explanation

upvoted 1 times

MisuLava 1 year ago

"single on-premises Kafka cluster in a data center in the us-east region"
is it on-prem or in a datacenter in us-east ?

upvoted 1 times

JamesKarianis 1 year, 3 months ago

Selected Answer: C

Answer is C

upvoted 1 times

Prasanna_kumar 1 year, 8 months ago

Answer is option C

upvoted 1 times

ivanhsiav 2 years, 3 months ago

Answer c
kafka cluster in on-premise for streaming msgs
pub/sub for streaming msgs in cloud

upvoted 4 times

sumanshu 2 years, 4 months ago

Vote for C

upvoted 4 times

Allan222 2 years, 8 months ago

Should be C

upvoted 4 times

daghayeghi 2 years, 8 months ago

C is correct:
the main trick come from A, and response is that TPU only use when we have a deployed machine learning model that we don't have now.

upvoted 5 times

ArunSingh1028 2 years, 8 months ago

Answer - C

upvoted 4 times

Alasmindas 2 years, 12 months ago

Easy Question : ANswer is Option C.
Alterative to Kafka in google cloud native service is Pub/Sub and Dataflow punched with Pub/Sub is the google recommended option

upvoted 5 times

atnafu2020 3 years, 2 months ago

C
the issue is with a single Kafka cluster is the need to scale automatically with Dataflow

upvoted 4 times

haroldbenites 3 years, 2 months ago

C is correct

upvoted 4 times

Load full discussion...