Google Discussions

EXAM PROFESSIONAL DATA ENGINEER TOPIC 1 QUESTION 79 DISCUSSION

Actual exam question from Google's Professional Data Engineer
Question #: 79
Topic #: 1
[All Professional Data Engineer Questions]

Your company maintains a hybrid deployment with GCP, where analytics are performed on your anonymized customer data. The data are imported to Cloud
Storage from your data center through parallel uploads to a data transfer server running on GCP. Management informs you that the daily transfers take too long and have asked you to fix the problem. You want to maximize transfer speeds. Which action should you take?

A. Increase the CPU size on your server.

B. Increase the size of the Google Persistent Disk on your server.

C. Increase your network bandwidth from your datacenter to GCP.

D. Increase your network bandwidth from Compute Engine to Cloud Storage.

Show Suggested Answer

by [deleted] at *March 21, 2020, 6:17 p.m.*

Comments

Type your comment...

Submit

[Removed] Highly Voted 4 years, 7 months ago

correct: C

upvoted 16 times

[Removed] Highly Voted 4 years, 7 months ago

Answer: C
Description : Speed of data transfer depends on Bandwidth
Few things in computing highlight the hardware limitations of networks as transferring large amounts of data. Typically you can transfer 1 GB in eight seconds over a 1 Gbps network. If you scale that up to a huge dataset (for example, 100 TB), the transfer time is 12 days. Transferring huge datasets can test the limits of your infrastructure and potentially cause problems for your business.

upvoted 15 times

Mathew106 Most Recent 1 year, 3 months ago

Selected Answer: C

We are talking about transfer speed. Network transfer speed does not increase with CPU, but with bandwidth. Since there is no other extra information about what the issue, we have to assume that they imply network transfer speed.

upvoted 1 times

Kiroo 1 year, 5 months ago

Selected Answer: C

To be honest this question is incomplete, I would go increasing the bandwidth, but first I would analyze why it's taking long time maybe I'm uploading many files so I could compress and agregate then and upload just one, maybe the target cpu is overloaded at the time of the upload, maybe the target disk reaching the max iops,

upvoted 3 times

Jarek7 1 year, 6 months ago

Selected Answer: C

Even if transfer server is deployed on the slowest machine available in GCP there is no way it is bottleneck for simple data transfer without any data processing.

upvoted 1 times

Oleksandr0501 1 year, 6 months ago

Selected Answer: A

GPT: Option A, increasing the CPU size on the data transfer server, could potentially increase the transfer speeds if the bottleneck in the data transfer process is the processing power of the server. By increasing the CPU size, the server may be able to process data more quickly, leading to faster transfers.
Option C, increasing the network bandwidth from the datacenter to GCP, could potentially improve the transfer speeds, but it may not be feasible or cost-effective depending on the current infrastructure and network limitations.

upvoted 1 times

Jarek7 1 year, 6 months ago

Please stop using GPT as knowledge source. v3.5 is usually wrong even in simple cases. v4 is much better, but it is not designed to be knowledge source. Looking at the answer you must have used v3.5. The question says nothing about cost-effectivness. The issue is data transfer. No any data processing is done on the data while it is transferred. Simple transfer doesn't need much processing power - the real bottleneck even on slowest machines available on GCP must be data transfer - it is obvious.
BTW for me GPT3.5 said it is C.

upvoted 4 times

Oleksandr0501 1 year, 6 months ago

it should be C, for real, bcz nothing said about cost restrictions in the question. And the user "snamburi3 " found docs.

upvoted 1 times

Oleksandr0501 1 year, 6 months ago

yea, i know it can make mistakes. Thank you.
That`s why i always mark "GPT" at the start of my answer.

upvoted 1 times

izekc 1 year, 6 months ago

Selected Answer: A

it's refer to data transfer server slow here. not transfer data to cloud slow.
100% A

upvoted 1 times

jonathanthezombieboy 1 year, 8 months ago

Selected Answer: C

Answer is C

upvoted 1 times

jin0 1 year, 8 months ago

This question makes people confused only. there is no refer to network or size of data or something could be referred. the answer could be A or C

upvoted 1 times

[Removed] 1 year, 8 months ago

Answer is C

upvoted 1 times

mahdiaqim 1 year, 8 months ago

Selected Answer: A

Very confusing question. I selected A because I assume increasing the CPU size on the cloud server is easier to change, as a data engineer, than the bandwidth.

upvoted 1 times

samdhimal 1 year, 9 months ago

C. Increase your network bandwidth from your datacenter to GCP.
This will likely have the most impact on transfer speeds as it addresses the bottleneck in the transfer between your data center and GCP. Increasing the CPU size or the size of the Google Persistent Disk on the server may help with processing the data once it has been transferred, but will not address the bottleneck in the transfer itself. Increasing the network bandwidth from Compute Engine to Cloud Storage would also help with processing the data once it has been transferred but will not address the bottleneck in the transfer itself as well.

upvoted 3 times

zellck 1 year, 11 months ago

Selected Answer: C

C is the answer.

upvoted 1 times

Nirca 2 years ago

A bit unprofessional question, having performance issues should be addressed by analyzing and looking for saturation in the system and understanding "wait-events". Only than adding more resources.

upvoted 1 times

rr4444 2 years, 4 months ago

"The data are imported to Cloud Storage from your data center through parallel uploads to a data transfer server running on GCP. "

This makes zero sense. Is it to GCS or GCE? Question had to make up its mind. Nonsense, literally.

upvoted 3 times

Thierry_1 2 years, 11 months ago

Vote for C. Mostly because other options seems useless

upvoted 1 times

sumanshu 3 years, 4 months ago

Vote for 'C'

You want to maximize transfer speeds.

upvoted 3 times

Load full discussion...