Google Discussions

EXAM PROFESSIONAL DATA ENGINEER TOPIC 1 QUESTION 189 DISCUSSION

Actual exam question from Google's Professional Data Engineer
Question #: 189
Topic #: 1
[All Professional Data Engineer Questions]

You need to migrate 1 PB of data from an on-premises data center to Google Cloud. Data transfer time during the migration should take only a few hours. You want to follow Google-recommended practices to facilitate the large data transfer over a secure connection. What should you do?

A. Establish a Cloud Interconnect connection between the on-premises data center and Google Cloud, and then use the Storage Transfer Service.

B. Use a Transfer Appliance and have engineers manually encrypt, decrypt, and verify the data.

C. Establish a Cloud VPN connection, start gcloud compute scp jobs in parallel, and run checksums to verify the data.

D. Reduce the data into 3 TB batches, transfer the data using gsutil, and run checksums to verify the data.

Show Suggested Answer

by AWSandeep at *Sept. 2, 2022, 11:13 p.m.*

Comments

Type your comment...

Submit

devaid Highly Voted 2 years, 7 months ago

Selected Answer: A

Well it doesn't mentions anything about not enough bandwidth to meet your project deadline. I guess you can assume they have 200GBps+ of bandwith, otherwise it shouldn't take only a few hours.

upvoted 5 times

iooj Most Recent 9 months ago

One who wanted to use Transfer Appliance to migrate data in a few hours, you should live near Google office and run really fast :D

upvoted 3 times

MaxNRG 1 year, 4 months ago

Selected Answer: A

Cloud Interconnect provides a dedicated private connection between on-prem and Google Cloud for high bandwidth (up to 100 Gbps) and low latency. This facilitates large, fast data transfers.
Storage Transfer Service supports parallel data transfers over Cloud Interconnect. It can transfer petabyte-scale datasets faster by transferring objects in parallel.
Storage Transfer Service uses HTTPS encryption in transit and at rest by default for secure data transfers.
It follows Google-recommended practices for large data migrations vs ad hoc methods like gsutil or scp.
The other options would take too long for a 1 PB transfer (VPN capped at 3 Gbps, manual transfers) or introduce extra steps like batching and checksums. Cloud Interconnect + Storage Transfer is the recommended Google solution.

upvoted 2 times

LanaOjisan 1 year, 6 months ago

It is believed that A.
One reason is that for "secure" and "in a few hours," the communication can be done securely using a direct physical line without going through an ISP. Also, depending on the case, in the case of "Dedicated Interconnect," the maximum transfer can be as high as 200 Gbps, and the fastest data transfer of 1 PB can be completed in 11 hours.
Therefore, A.

upvoted 2 times

arien_chen 1 year, 8 months ago

Selected Answer: A

A
https://cloud.google.com/storage-transfer/docs/transfer-options#:~:text=Transferring%20more%20than%201%20TB%20from%20on%2Dpremises

upvoted 3 times

knith66 1 year, 9 months ago

Selected Answer: A

Dedicated Interconnect provides direct physical connections between your on-premises network and Google's network. Dedicated Interconnect enables you to transfer large amounts of data between networks, which can be more cost-effective than purchasing additional bandwidth over the public internet. https://cloud.google.com/network-connectivity/docs/interconnect/concepts/dedicated-overview

upvoted 1 times

knith66 1 year, 9 months ago

This link has additional clarity
https://cloud.google.com/network-connectivity/docs/interconnect/concepts/terminology

upvoted 1 times

vaga1 1 year, 10 months ago

Selected Answer: B

1 PB and "few hours". It is clearly referring to Transfer Appliance

https://cloud.google.com/architecture/migration-to-google-cloud-transferring-your-large-datasets#time

upvoted 3 times

knith66 1 year, 9 months ago

Transfer Appliance is a slow process. wont be able to do in few hours

upvoted 3 times

Oleksandr0501 2 years ago

Selected Answer: A

gpt: Based on security and speed, if the data is highly sensitive and security is the top priority, then option B (using a Transfer Appliance) may be a better choice. Transfer Appliance uses hardware encryption to transfer data and is designed to securely transfer large amounts of data. However, if speed is the primary concern, then option A (using Cloud Interconnect and Storage Transfer Service) may be a better choice as it allows for faster transfer speeds over a dedicated and secure connection. It ultimately depends on the specific needs and priorities of the organization.

A vague tricky question. Bad author of it

A vague freaky question. Bad author of it...
B is also good. As were said in discuss. by smb, a question asks "safe connection", so - a Cloud Interconnect (A)

upvoted 1 times

midgoo 2 years, 1 month ago

Selected Answer: B

Either this question is very tricky or very poor written. It says 'Data transfer time during the migration should take only a few hours'. We should not add the 20days for overhead time for Appliance into the total time of migration.

If 'a few hours' = 30hours or more, A will be good enough.
If 'a few hours' = 10 or less, B is the only way (with multiple devices to copy at the same time)

upvoted 3 times

spicebits 1 year, 6 months ago

B can't be the answer - You have to wait 25 days to receive the appliance and another 25 days to get the appliance back and data loaded to cloud storage: https://cloud.google.com/transfer-appliance/docs/4.0/overview#transfer-speeds

upvoted 1 times

Nandhu95 2 years, 1 month ago

Selected Answer: A

Expected time via transfer appliance is around 20 days , and achieving the same using Storage transfer service with highest bandwidth of 100GPS is 30 hrs, so hence its been asked for hrs .. its A
Acquiring a Transfer Appliance is straightforward. In the Google Cloud console, you request a Transfer Appliance, indicate how much data you have, and then Google ships one or more appliances to your requested location. You're given a number of days to transfer your data to the appliance ("data capture") and ship it back to Google.

The expected turnaround time for a network appliance to be shipped, loaded with your data, shipped back, and rehydrated on Google Cloud is 20 days. If your online transfer timeframe is calculated to be substantially more than this timeframe, consider Transfer Appliance. The total cost for the 300 TB device process is less than $2,500.

upvoted 1 times

vaga1 1 year, 10 months ago

it says data transfer during the migration. It mean from when the migration is "activated", which means from when the Transfer Appliace device is plugged and ready to be used

upvoted 1 times

wjtb 2 years, 1 month ago

Selected Answer: B

Even with 100gbps bandwith, you will not reach a data transfer time within the range of "hours" for 1PB. Transfer appliance is the way to go. https://cloud.google.com/architecture/migration-to-google-cloud-transferring-your-large-datasets#time

upvoted 2 times

musumusu 2 years, 2 months ago

Answer A,
One time transfer is cheaper and less secure always using Transfer Appliance.
you need to do it in faster way, set up Interconnect speed limit is 50mbps - 10GBps
and Transfer Appliance speed can goes up to 40GBps.
I am choosing A for security concern only.

upvoted 1 times

AzureDP900 2 years, 4 months ago

A is right

upvoted 1 times

AzureDP900 2 years, 4 months ago

A. Establish a Cloud Interconnect connection between the on-premises data center and Google Cloud, and then use the Storage Transfer Service. Most Voted

upvoted 1 times

zellck 2 years, 5 months ago

Selected Answer: A

A is the answer.

https://cloud.google.com/architecture/migration-to-google-cloud-transferring-your-large-datasets#storage-transfer-service-for-large-transfers-of-on-premises-data
Like gsutil, Storage Transfer Service for on-premises data enables transfers from network file system (NFS) storage to Cloud Storage. Although gsutil can support small transfer sizes (up to 1 TB), Storage Transfer Service for on-premises data is designed for large-scale transfers (up to petabytes of data, billions of files).

upvoted 2 times

Atnafu 2 years, 5 months ago

B
https://cloud.google.com/architecture/migration-to-google-cloud-transferring-your-large-datasets#:~:text=Few%20things%20in,not%20be%20obtained.

upvoted 2 times

Atnafu 2 years, 5 months ago

B
It takes 30hrs with 100Gbps bandwidth- more than a day to transfer
https://cloud.google.com/architecture/migration-to-google-cloud-transferring-your-large-datasets#:~:text=addresses%20or%20NATs.-,Online%20versus%20offline%20transfer,A%20certain%20amount%20of%20management%20overhead%20is%20built%20into%20these%20calculations.,-As%20noted%20earlier

upvoted 4 times

pluiedust 2 years, 7 months ago

Selected Answer: A

A is correct.

upvoted 1 times

bigquery1102 2 years, 7 months ago

Selected Answer: A

A is correct
https://cloud.google.com/architecture/migration-to-google-cloud-transferring-your-large-datasets#transfer_appliance_for_larger_transfers

upvoted 1 times

Load full discussion...