

[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 308 DISCUSSION

Actual exam question from Google's Professional Data Engineer

Question #: 308

Topic #: 1

[\[All Professional Data Engineer Questions\]](#)

You are migrating a large number of files from a public HTTPS endpoint to Cloud Storage. The files are protected from unauthorized access using signed URLs. You created a TSV file that contains the list of object URLs and started a transfer job by using Storage Transfer Service. You notice that the job has run for a long time and eventually failed. Checking the logs of the transfer job reveals that the job was running fine until one point, and then it failed due to HTTP 403 errors on the remaining files. You verified that there were no changes to the source system. You need to fix the problem to resume the migration process. What should you do?

- A. Set up Cloud Storage FUSE, and mount the Cloud Storage bucket on a Compute Engine instance. Remove the completed files from the TSV file. Use a shell script to iterate through the TSV file and download the remaining URLs to the FUSE mount point.
- B. Renew the TLS certificate of the HTTPS endpoint. Remove the completed files from the TSV file and rerun the Storage Transfer Service job.
- C. Create a new TSV file for the remaining files by generating signed URLs with a longer validity period. Split the TSV file into multiple smaller files and submit them as separate Storage Transfer Service jobs in parallel.
- D. Update the file checksums in the TSV file from using MD5 to SHA256. Remove the completed files from the TSV file and rerun the Storage Transfer Service job.

[Show Suggested Answer](#)

by [scaenruy](#) at Jan. 4, 2024, 1:05 p.m.

Comments

Type your comment...

Submit

  **raaad** Highly Voted  1 year, 4 months ago

Selected Answer: C

- It addresses the likely issue: that the signed URLs have expired or are otherwise invalid. By creating a new TSV file with freshly generated signed URLs (with a longer validity period), you're ensuring that the Storage Transfer Service has valid authorization to access the files.
- Splitting the TSV file and running parallel jobs might help in managing the workload more efficiently and overcoming any limitations related to the number of files or transfer speed.

   upvoted 7 times

  **srivastavas08** Highly Voted  1 year, 2 months ago

C. Create a new TSV file for the remaining files by generating signed URLs with a longer validity period. Split the TSV file into multiple smaller files and submit them as separate Storage Transfer Service jobs in parallel.

Here's why:

HTTP 403 errors: These errors indicate unauthorized access, but since you verified the source system and signed URLs, the issue likely lies with expired signed URLs. Renewing the URLs with a longer validity period prevents this issue for the remaining files.

Separate jobs: Splitting the file into smaller chunks and submitting them as separate jobs improves parallelism and potentially speeds up the transfer process.

Avoid manual intervention: Options A and D require manual intervention and complex setups, which are less efficient and might introduce risks.

Longer validity: While option B addresses expired URLs, splitting the file offers additional benefits for faster migration.

   upvoted 6 times

  **Pime13** Most Recent  3 months, 4 weeks ago

Selected Answer: C

C. Create a new TSV file for the remaining files by generating signed URLs with a longer validity period. Split the TSV file into multiple smaller files and submit them as separate Storage Transfer Service jobs in parallel.

The HTTP 403 errors likely occurred because the signed URLs expired before the transfer could complete. By generating new signed URLs with a longer validity period, you can ensure that the URLs remain valid for the duration of the transfer. Splitting the TSV file into smaller files and submitting them as separate jobs can help manage the load and reduce the risk of timeouts or other issues

   upvoted 1 times

  **iooj** 9 months ago

Selected Answer: C

got this one on the exam, aug 2024, passed


   upvoted 2 times

  **Matt_108** 1 year, 3 months ago

Selected Answer: C

Option C - agree with Raaad

   upvoted 3 times

  **scaenrui** 1 year, 4 months ago

Selected Answer: C

C. Create a new TSV file for the remaining files by generating signed URLs with a longer validity period. Split the TSV file into multiple smaller files and submit them as separate Storage Transfer Service jobs in parallel.

   upvoted 2 times

Platform

> Home

> Examtopics PRO

> All Exams

> Training Courses

