



PROFESSIONAL-DATA-ENGINEER^{Q&As}

Professional Data Engineer on Google Cloud Platform

**Pass Google PROFESSIONAL-DATA-ENGINEER
Exam with 100% Guarantee**

Free Download Real Questions & Answers **PDF** and **VCE** file from:

<https://www.pass4itsure.com/professional-data-engineer.html>

100% Passing Guarantee
100% Money Back Assurance

Following Questions and Answers are all new published by Google
Official Exam Center



VCE & PDF

Pass4itSure.com

<https://www.pass4itsure.com/professional-data-engineer.html>

2024 Latest pass4itsure PROFESSIONAL-DATA-ENGINEER PDF and VCE
dumps Download

- ⚙️ **Instant Download** After Purchase
- ⚙️ **100% Money Back** Guarantee
- ⚙️ **365 Days** Free Update
- ⚙️ **800,000+** Satisfied Customers



**QUESTION 1**

Your company produces 20,000 files every hour. Each data file is formatted as a comma separated values (CSV) file that is less than 4 KB. All files must be ingested on Google Cloud Platform before they can be processed. Your company site has a 200 ms latency to Google Cloud, and your Internet connection bandwidth is limited as 50 Mbps. You currently deploy a secure FTP (SFTP) server on a virtual machine in Google Compute Engine as the data ingestion point. A local SFTP client runs on a dedicated machine to transmit the CSV files as is. The goal is to make reports with data from the previous day available to the executives by 10:00 a.m. each day. This design is barely able to keep up with the current volume, even though the bandwidth utilization is rather low. You are told that due to seasonality, your company expects the number of files to double for the next three months. Which two actions should you take? (choose two.)

- A. Introduce data compression for each file to increase the rate file of file transfer.
- B. Contact your internet service provider (ISP) to increase your maximum bandwidth to at least 100 Mbps.
- C. Redesign the data ingestion process to use gsutil tool to send the CSV files to a storage bucket in parallel.
- D. Assemble 1,000 files into a tape archive (TAR) file. Transmit the TAR files instead, and disassemble the CSV files in the cloud upon receiving them.
- E. Create an S3-compatible storage endpoint in your network, and use Google Cloud Storage Transfer Service to transfer on-premises data to the designated storage bucket.

Correct Answer: CE

QUESTION 2

You are designing a Dataflow pipeline for a batch processing job. You want to mitigate multiple zonal failures at job submission time. What should you do?

- A. Specify a worker region by using the --region flag.
- B. Set the pipeline staging location as a regional Cloud Storage bucket.
- C. Submit duplicate pipelines in two different zones by using the --zone flag.
- D. Create an Eventarc trigger to resubmit the job in case of zonal failure when submitting the job.

Correct Answer: A

By specifying a worker region, you can run your Dataflow pipeline in a multi- zone or multi-region configuration, which provides higher availability and resilience in case of zonal failures¹. The --region flag allows you to specify the regional endpoint for your pipeline, which determines the location of the Dataflow service and the default location of the Compute Engine resources¹. If you do not specify a zone by using the --zone flag, Dataflow automatically selects a zone within the region for your job workers¹. This option is recommended over submitting duplicate pipelines in two different zones, which would incur additional costs and complexity. Setting the pipeline staging location as a regional Cloud Storage bucket does not affect the availability of your pipeline, as the staging location only stores the pipeline code and dependencies². Creating an Eventarc trigger to resubmit the job in case of zonal failure is not a reliable solution, as it depends on the availability of the Eventarc service and the zonal resources at the time of resubmission. References:

1: Pipeline troubleshooting and debugging | Cloud Dataflow | Google Cloud



3: Regional endpoints | Cloud Dataflow | Google Cloud

QUESTION 3

You are updating the code for a subscriber to a Pub/Sub feed. You are concerned that upon deployment the subscriber may erroneously acknowledge messages, leading to message loss. Your subscriber is not set up to retain acknowledged messages. What should you do to ensure that you can recover from errors after deployment?

- A. Use Cloud Build for your deployment if an error occurs after deployment, use a Seek operation to locate a timestamp logged by Cloud Build at the start of the deployment
- B. Create a Pub/Sub snapshot before deploying new subscriber code. Use a Seek operation to re-deliver messages that became available after the snapshot was created
- C. Set up the Pub/Sub emulator on your local machine. Validate the behavior of your new subscriber code before deploying it to production
- D. Enable dead-lettering on the Pub/Sub topic to capture messages that aren't successfully acknowledged if an error occurs after deployment, re-deliver any messages captured by the dead-letter queue

Correct Answer: B

QUESTION 4

You have a BigQuery table that ingests data directly from a Pub/Sub subscription. The ingested data is encrypted with a Google-managed encryption key. You need to meet a new organization policy that requires you to use keys from a centralized Cloud Key Management Service (Cloud KMS) project to encrypt data at rest. What should you do?

- A. Create a new BigQuery table by using customer-managed encryption keys (CMEK), and migrate the data from the old BigQuery table.
- B. Create a new BigQuery table and Pub/Sub topic by using customer-managed encryption keys (CMEK), and migrate the data from the old BigQuery table.
- C. Create a new Pub/Sub topic with CMEK and use the existing BigQuery table by using Google-managed encryption key.
- D. Use Cloud KMS encryption key with Dataflow to ingest the existing Pub/Sub subscription to the existing BigQuery table.

Correct Answer: A

To use CMEK for BigQuery, you need to create a key ring and a key in Cloud KMS, and then specify the key resource name when creating or updating a BigQuery table. You cannot change the encryption type of an existing table, so you need to create a new table with CMEK and copy the data from the old table with Google-managed encryption key.

References:

Customer-managed Cloud KMS keys | BigQuery | Google Cloud Creating and managing encryption keys | Cloud KMS Documentation | Google Cloud



QUESTION 5

Your financial services company is moving to cloud technology and wants to store 50 TB of financial timeseries data in the cloud. This data is updated frequently and new data will be streaming in all the time. Your company also wants to move their existing Apache Hadoop jobs to the cloud to get insights into this data.

Which product should they use to store the data?

- A. Cloud Bigtable
- B. Google BigQuery
- C. Google Cloud Storage
- D. Google Cloud Datastore

Correct Answer: A

Reference: <https://cloud.google.com/bigtable/docs/schema-design-time-series>

[PROFESSIONAL-DATA-ENGINEER VCE Dumps](#)

[PROFESSIONAL-DATA-ENGINEER Practice Test](#)

[PROFESSIONAL-DATA-ENGINEER Exam Questions](#)