

🔗 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 286 DISCUSSION

Actual exam question from Google's Professional Data Engineer

Question #: 286

Topic #: 1

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

You have thousands of Apache Spark jobs running in your on-premises Apache Hadoop cluster. You want to migrate the jobs to Google Cloud. You want to use managed services to run your jobs instead of maintaining a long-lived Hadoop cluster yourself. You have a tight timeline and want to keep code changes to a minimum. What should you do?

- A. Move your data to BigQuery. Convert your Spark scripts to a SQL-based processing approach.
- B. Rewrite your jobs in Apache Beam. Run your jobs in Dataflow.
- C. Copy your data to Compute Engine disks. Manage and run your jobs directly on those instances.
- D. Move your data to Cloud Storage. Run your jobs on Dataproc.

[Show Suggested Answer](#)

by [scaenry](#) at Jan. 4, 2024, 10:41 a.m.

Comments

Type your comment...

[Submit](#)

🗨️ [hussain.sain](#) 4 months, 1 week ago

Selected Answer: D

D is correct.

Dataproc is the most suitable choice for migrating your existing Apache Spark jobs to Google Cloud because it is a fully managed service that supports Apache Spark and Hadoop workloads with minimal changes to your existing code. Moving your data to Cloud Storage and running jobs on Dataproc offers a fast, efficient, and scalable solution for your needs.

👍 ↩ 🚩 upvoted 1 times

🗄️ 👤 **meh_33** 8 months, 4 weeks ago

Selected Answer: D

option D, minimum code changes

👍 ↩ 🚩 upvoted 1 times

🗄️ 👤 **hanoverquay** 1 year, 1 month ago

Selected Answer: D

option D, minimum code changes

👍 ↩ 🚩 upvoted 2 times

🗄️ 👤 **JyoGCP** 1 year, 2 months ago

Selected Answer: D

Option D

👍 ↩ 🚩 upvoted 2 times

🗄️ 👤 **ML6** 1 year, 2 months ago

Selected Answer: D

D) That is what Dataproc is made for. It is a fully managed and highly scalable service for running Apache Hadoop, Apache Spark, etc.

👍 ↩ 🚩 upvoted 3 times

🗄️ 👤 **Matt_108** 1 year, 3 months ago

Selected Answer: D

Clearly D

👍 ↩ 🚩 upvoted 2 times

🗄️ 👤 **Sofiia98** 1 year, 3 months ago

Selected Answer: D

of course D

👍 ↩ 🚩 upvoted 2 times

🗄️ 👤 **GCP001** 1 year, 3 months ago

D. Move your data to Cloud Storage. Run your jobs on Dataproc. Dataproc is managed service and not needed much code changes.

👍 ↩ 🚩 upvoted 3 times

🗄️ 👤 **scaenrui** 1 year, 4 months ago

Selected Answer: D

D. Move your data to Cloud Storage. Run your jobs on Dataproc.

👍 ↩ 🚩 upvoted 3 times



Platform

> Home

> Examtopics PRO

> All Exams

> Training Courses

