

[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 93 DISCUSSION

Actual exam question from Google's Professional Data Engineer

Question #: 93

Topic #: 1

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

You're using Bigtable for a real-time application, and you have a heavy load that is a mix of read and writes. You've recently identified an additional use case and need to perform hourly an analytical job to calculate certain statistics across the whole database. You need to ensure both the reliability of your production application as well as the analytical workload. What should you do?

- A. Export Bigtable dump to GCS and run your analytical job on top of the exported files.
- B. Add a second cluster to an existing instance with a multi-cluster routing, use live-traffic app profile for your regular workload and batch-analytics profile for the analytics workload.
- C. Add a second cluster to an existing instance with a single-cluster routing, use live-traffic app profile for your regular workload and batch-analytics profile for the analytics workload.
- D. Increase the size of your existing cluster twice and execute your analytics workload on your new resized cluster.

[Show Suggested Answer](#)

by [AWSandeep](#) at *Sept. 3, 2022, 2:01 p.m.*

Comments

Type your comment...

[Submit](#)

🗄️ 👤 **[Removed]** Highly Voted 2 years, 8 months ago

Answer is C

When you use a single cluster to run a batch analytics job that performs numerous large reads alongside an application that performs a mix of reads and writes, the large batch job can slow things down for the application's users. With replication, you can use app profiles with single-cluster routing to route batch analytics jobs and application traffic to different clusters, so that batch jobs don't affect your applications' users.

<https://cloud.google.com/bigtable/docs/replication-overview#use-cases>

👍 🔄 🚩 upvoted 22 times

🗄️ 👤 **nadavw** 8 months, 2 weeks ago

full example here: <https://cloud.google.com/bigtable/docs/replication-settings#batch-vs-serve>

👍 🔄 🚩 upvoted 1 times

🗄️ 👤 **HarshKothari21** 2 years, 7 months ago

Agreed :)

👍 🔄 🚩 upvoted 2 times

🗄️ 👤 **[Removed]** 2 years, 5 months ago

"When you use a single cluster", here we are creating a 2nd cluster, so we'll be using 2 different clusters. We want to redirect analysis jobs to the 2nd cluster, and the other job to the 1st cluster. Thus, I think that D is more adequate

👍 🔄 🚩 upvoted 1 times

🗄️ 👤 **somilaseeja** 2 years, 4 months ago

Option D didnt say to create a new cluster, rather it said to increase the size of the cluster. There is a difference. Hence c is the correct answer to run the batch processing in a single cluster mode

👍 🔄 🚩 upvoted 2 times

🗄️ 👤 **aewis** Highly Voted 1 year, 9 months ago

Selected Answer: C

It was actually illustrated here

<https://cloud.google.com/bigtable/docs/replication-settings#batch-vs-serve>

👍 🔄 🚩 upvoted 5 times

🗄️ 👤 **oussama7** Most Recent 1 month, 2 weeks ago

Selected Answer: B

Option C offers single-cluster routing, meaning that each query is directed to a single specific cluster. This does not protect the transactional workload from heavy analytical loads, which can lead to performance degradation. With multi-cluster routing (option B), Bigtable can automatically distribute the load and avoid congestion on a single cluster.

👍 🔄 🚩 upvoted 1 times

🗄️ 👤 **desertlotus1211** 1 month, 2 weeks ago

Selected Answer: B

Multi-cluster satisfy reliability aspect of the question. single cluster may and will cause contention for resources...

👍 🔄 🚩 upvoted 1 times

🗄️ 👤 **baimus** 7 months, 1 week ago

Selected Answer: C

To answer some confusion - "single cluster routing" is routing to one cluster per profile, rather than having failover options per profile. So we have two clusters, but it's not multicluster, because we have two profiles, so it's one cluster per profile, so "single cluster routing". We COULD use multicluster, but none of the answers give the steps required to do so, so the assumption as to be that we're using single.

👍 🔄 🚩 upvoted 1 times

🗄️ 👤 **baimus** 7 months, 1 week ago

(an example of multicluster in this case would be 4 clusters, 2 for the transactional load and 2 for the analytical load)

👍 🔄 🚩 upvoted 1 times

🗄️ 👤 **47767f9** 10 months, 1 week ago

Selected Answer: B

B better than C. Multi-cluster routing to handle failovers automatically. Reference:

<https://cloud.google.com/bigtable/docs/replication-settings#regional-failover>

👍 🔄 🚩 upvoted 1 times

🗄️ 👤 **opt_sub** 1 year, 1 month ago

Selected Answer: B

B is correct.

Two different job profiles to redirect traffic to two different cluster. C is incorrect because there is no point in creating app profile for two different workloads in the same cluster. One cluster handles writes and another handle reads.

👍 ↩ 🚩 upvoted 2 times

🗨️ 👤 **carbino** 1 year, 4 months ago

Selected Answer: C

It is C:

"Workload isolation:

Using separate app profiles lets you use different routing policies for different purposes. For example, consider a situation when you want to prevent a batch read job (workload A) from increasing CPU usage on clusters that handle an application's steady reads and writes (workload B). You can create an app profile for workload B that routes to a cluster group that excludes one cluster. Then you create an app profile for workload A that specifies single-cluster routing to the cluster that workload B doesn't send requests to.

You can change the settings for one application or function without affecting other applications that connect to the same data."

<https://cloud.google.com/bigtable/docs/app-profiles>

👍 ↩ 🚩 upvoted 3 times

🗨️ 👤 **DevShah** 2 years ago

Selected Answer: C

<https://cloud.google.com/bigtable/docs/replication-settings#batch-vs-serve>

👍 ↩ 🚩 upvoted 3 times

🗨️ 👤 **A4M** 2 years ago

I see what you say on C but the question states high availability how do you handle that with option C when you have a single region cluster hence answer needs to be with multi-region cluster - To configure your instance for a high availability (HA) use case, create a new app profile that uses multi-cluster routing, or update the default app profile to use multi-cluster routing.

👍 ↩ 🚩 upvoted 1 times

🗨️ 👤 **zevexWM** 1 year ago

It actually addresses the issue of High availability in that same link if you scroll down a bit more.

<https://cloud.google.com/bigtable/docs/replication-settings#high-availability>

👍 ↩ 🚩 upvoted 1 times

🗨️ 👤 **A4M** 2 years ago

i meant single-cluster routing

👍 ↩ 🚩 upvoted 1 times

🗨️ 👤 **juliobs** 2 years, 1 month ago

Selected Answer: C

C. This is exactly the example in the documentation.

<https://cloud.google.com/bigtable/docs/replication-settings#batch-vs-serve>

👍 ↩ 🚩 upvoted 3 times

🗨️ 👤 **DevShah** 2 years ago

Correct

2 jobs >> 2 cluster

3 jobs >> 3 cluster

app profiles with single-cluster routing used to route to specific cluster

Job1 >> Cluster 1

Job2 >> Cluster 2

👍 ↩ 🚩 upvoted 1 times

🗨️ 👤 **musumusu** 2 years, 2 months ago

Answer B:

reason 1: If you don't have any cost constraint use multi-cluster routing,

reason 2: Single cluster is less scalable as we need high scalability i would go with B

👍 ↩ 🚩 upvoted 1 times

🗨️ 👤 **samdhimal** 2 years, 2 months ago

Selected Answer: C

I am going for C?

👍 ↩ 🚩 upvoted 1 times

🗨️ 👤 **slade_wilson** 2 years, 4 months ago



Selected Answer: C

When you use a single cluster to run a batch analytics job that performs numerous large reads alongside an application that performs a mix of reads and writes, the large batch job can slow things down for the application's users. With replication, you can use app profiles with single-cluster routing to route batch analytics jobs and application traffic to different clusters, so that batch jobs don't affect your applications' users.

Single cluster routing - You can use single-cluster routing for this use case if you don't want your Bigtable cluster to automatically fail over if a zone or region becomes unavailable.

Multi-cluster routing - If you want Bigtable to automatically fail over to one region if your application cannot reach the other region, use multi-cluster routing.

   upvoted 2 times

  **zelck** 2 years, 5 months ago

Selected Answer: C

C is the answer.

<https://cloud.google.com/bigtable/docs/replication-settings#batch-vs-serve>

When you use a single cluster to run a batch analytics job that performs numerous large reads alongside an application that performs a mix of reads and writes, the large batch job can slow things down for the application's users. With replication, you can use app profiles with single-cluster routing to route batch analytics jobs and application traffic to different clusters, so that batch jobs don't affect your applications' users.



   upvoted 2 times

  **Siant_137** 2 years, 5 months ago

Answer is C

"When you use a single cluster to run a batch analytics job that performs numerous large reads alongside an application that performs a mix of reads and writes, the large batch job can slow things down for the application's users. With replication, you can use app profiles with single-cluster routing to route batch analytics jobs and application traffic to different clusters, so that batch jobs don't affect your applications' users."

<https://cloud.google.com/bigtable/docs/replication-overview#batch-vs-serve>

   upvoted 2 times

  **sfsdeniso** 2 years, 5 months ago

Answer is C

   upvoted 1 times

  **dish11dish** 2 years, 5 months ago

Selected Answer: B

Option B is correct

An app profile specifies the routing policy that Bigtable should use for each request.

Single-cluster routing routes all requests to 1 cluster in your instance. If that cluster becomes unavailable, you must manually fail over to another cluster.

Multi-cluster routing automatically routes requests to the nearest cluster in an instance. If the cluster becomes unavailable, traffic automatically fails over to the nearest cluster that is available. Bigtable considers clusters in a single region to be equidistant, even though they are in different zones. You can configure an app profile to route to any cluster in an instance, or you can specify a cluster group that tells the app profile to route to only some of the clusters in the instance.

Cluster group routing sends requests to the nearest available cluster within a cluster group that you specify in the app profile settings.

Reference:-<https://cloud.google.com/bigtable/docs/app-profiles#routing>

   upvoted 3 times

[Load full discussion...](#)

> Home

> All Exams

> Examtopics PRO

> Training Courses



© 2024 ExamTopics