Read the Linux Academy Oath (https://go.acloudguru.com/faq-la-customer-oath) from Sam Kroonenburg, our CEO.

See What the LA Product Team Has Been Up To! (/updates)

Google Certified Professional Data Engineer - NoSQL Data with Cloud Bigtable Quiz

30 minutes | 10 Questions | 3 Minutes per Question | Advanced (/search?type=Practice Exam Challenge&difficulty=Advanced&categories=Cloud)

Go Back | Start Challenge

90%

Question List Show All Passed

← Go Back

Congratulations!

You passed this challenge on this attempt.

1 2 3 4 5 6 7 8 9 10

Expectations Report Card

Google Certified Professional Data Engineer - NoSQL Data with Cloud Bigtable Quiz 90%

Exam Breakdown

Google Certified Professional Data Engineer - NoSQL Data with Cloud Bigtable Quiz

1. If you have 2 replicating clusters in your Bigtable instance, how can you ensure that your application will be guaranteed strong consistency for its transactions?

A Use an application profile that specifies single-cluster routing.

B Refactor your application so that strong consistency is no longer required.

C Use an application profile that specifies multi-cluster routing, but place both clusters within the same region.

D Refactor your application so that it expects to wait 2-3 seconds after every mutation for data to become eventually consistent.

Correct Answer: A

Why is this correct?

Strong consistency can only be achieved using single-cluster routing. Eventual consistency is normally quick but can take several minutes depending on the distance between clusters. If your application requires strong consistency, refactoring is unlikely to be an option without a complete redesign.

---

2. What kind of database is Bigtable?

A NoSQL wide-column tabular database

B NoSQL document store

C Managed SQL-like relational database

D Unmanaged SQL-like relational database

Correct Answer: A

Why is this correct?

Cloud Bigtable is a petabyte-scale, fully managed NoSQL database service for large analytical and operational workloads that typically uses wide-column tables.

INCORRECT

3. What is the maximum number of clusters per Bigtable instance?

A 5

B 3

C 4

D There is no limit, providing you are within the Compute Engine quotas of your GCP project.

Your Answer: D

Why is this incorrect?

Correct Answer: C

Why is this correct?

A Bigtable instance can contain up to 4 clusters. Compute Engine is a red herring, it has nothing to do with Cloud Bigtable.

---

4. What action should you take if you observe that Bigtable performance is poor on a subset of queries?

A Add an additional cluster to the instance to increase the read and write throughput.

B Use the Key Visualizer tool to identify hot spots and consider changing how the row key distributes rows.

C Add more nodes to the cluster to increase the overall storage capacity.

D Add debugging steps to your application to identify the problematic queries.

Correct Answer: B

Why is this correct?

If a subset of queries are performing poorly, there is likely a hot spot being caused by the row key design. Key Visualizer provides a heatmap of the load across your entire table in a way that is difficult to understand using standard debugging techniques. Increasing the size of a cluster is not a long term solution, and adding additional clusters will not affect write throughput.

---

5. How does Bigtable manage rows, tables, and tablets?

A Tables contain unsorted rows and are stored in a single tablet.

B Tables are sorted into contiguous rows and are sharded into tablets.

C Tables are sorted into contiguous rows and are stored in a single tablet.

D Tables contain unsorted rows, which are shared into tablets.

Correct Answer: B

Why is this correct?

A Cloud Bigtable table is sharded into blocks of contiguous rows, called tablets, to help balance the workload of queries.

---

6. How does Bigtable manage the storage of tablets?

A Tablets are stored on cluster nodes, which must be sized accordingly for their storage needs.

B Tablets are stored in Google Colossus, so storage capacity is not an issue when defining nodes.

C Tablets are stored in Google Colossus, but a cluster node has a limit on how much storage it can process.

D Tablets are stored on cluster nodes, but storage can dynamically grow as part of the managed service.

Correct Answer: C

Why is this correct?

Data is never stored in Cloud Bigtable nodes themselves; each node has pointers to a set of tablets that are stored on Colossus. However, CPU resources are required for a node to manage all of its associated tablets.

---

7. Bigtable is compatible with which open source project's client library for Java?

A Apache Cassandra

B MongoDB

C Apache HBase

D Apache Kafka

Correct Answer: C

Why is this correct?

Cloud Bigtable is exposed to applications through multiple client libraries, including a supported extension to the Apache HBase library for Java. As a result, it integrates with the existing Apache ecosystem of open-source Big Data software.

---

8. Which of these identifiers would make a good component of a row key?

A 64-character UUID

B Reverse domain name

C Domain name

D String identifier

E Hashed string identifier

Correct Answer: B

Why is this correct?

A reverse domain name would make a good row key, especially if each row's data overlaps with adjacent rows. This allows Cloud Bigtable to compress your data more efficiently.
See: https://cloud.google.com/bigtable/docs/schema-design (https://cloud.google.com/bigtable/docs/schema-design)

Correct Answer: D

Why is this correct?

Good row key design is essential. Cloud Bigtable queries use the row key, a row key prefix, or a row range to retrieve the data. Other types of queries trigger a full table scan, which is much less efficient. Using human-readable values makes it much easier to troubleshoot issues with Cloud Bigtable. See: https://cloud.google.com/bigtable/docs/schema-design

(https://cloud.google.com/bigtable/docs/schema-design)

9. Which of these identifiers would make a bad component of a row key?

A String relating to a logical namespace.

B String relating to a logical group of devices.

C Timestamp at the end of a multi-component row key.

D Sequential numeric ID.

E Timestamp at the start of a multi-component row key.

Correct Answer: D

Why is this correct?

Good row key design is essential. Cloud Bigtable queries use the row key, a row key prefix, or a row range to retrieve the data. Other types of queries trigger a full table scan, which is much less efficient.

Correct Answer: E

Why is this correct?

Good row key design is essential. Cloud Bigtable queries use the row key, a row key prefix, or a row range to retrieve the data. Other types of queries trigger a full table scan, which is much less efficient.

---

10. For which types of data workload would Bigtable ***not*** be a good fit?

A Time-series data

B Small datasets below 300GB in size

C Financial data

D Internet of things data

E Relational data requiring SQL-type queries

Correct Answer: B

Why is this correct?

Cloud Bigtable does not support SQL queries or the multiple indexes required for true relational data. If your dataset is too small, Bigtable won't be able to balance tables in a way that optimizes performance.

Correct Answer: E

Why is this correct?

Cloud Bigtable does not support SQL queries or the multiple indexes required for true relational data. If your dataset is too small, Bigtable won't be able to balance tables in a way that optimizes performance.