

[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 281 DISCUSSION

Actual exam question from Google's Professional Data Engineer

Question #: 281

Topic #: 1

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

You work for a large ecommerce company. You store your customer's order data in Bigtable. You have a garbage collection policy set to delete the data after 30 days and the number of versions is set to 1. When the data analysts run a query to report total customer spending, the analysts sometimes see customer data that is older than 30 days. You need to ensure that the analysts do not see customer data older than 30 days while minimizing cost and overhead. What should you do?

- A. Set the expiring values of the column families to 29 days and keep the number of versions to 1.
- B. Use a timestamp range filter in the query to fetch the customer's data for a specific range.
- C. Schedule a job daily to scan the data in the table and delete data older than 30 days.
- D. Set the expiring values of the column families to 30 days and set the number of versions to 2.

[Show Suggested Answer](#)

by [scaenrui](#) at Jan. 4, 2024, 6:28 a.m.

### Comments

Type your comment...


[Submit](#)

**Matt\_108** [Highly Voted](#) 9 months, 3 weeks ago

**Selected Answer: B**

Agree with others <https://cloud.google.com/bigtable/docs/garbage-collection>

   upvoted 8 times

  **AllenChen123** 9 months, 2 weeks ago

Agree. <https://cloud.google.com/bigtable/docs/garbage-collection#data-removed>

"Because it can take up to a week for expired data to be deleted, you should never rely solely on garbage collection policies to ensure that read requests return the desired data. Always apply a filter to your read requests that excludes the same values as your garbage collection rules. You can filter by limiting the number of cells per column or by specifying a timestamp range."

   upvoted 10 times

  **cuadradobertolinisebastiancami** **Highly Voted**  8 months, 1 week ago

**Selected Answer: B**

Agree with MAtt\_108 and AllenChen 123.

"Garbage collection is a continuous process in which Bigtable checks the rules for each column family and deletes expired and obsolete data accordingly. In general, it can take up to a week from the time that data matches the criteria in the rules for the data to actually be deleted. You are not able to change the timing of garbage collection."

"Always apply a filter to your read requests that exclude the same values as your garbage collection rules. "


Ref: <https://cloud.google.com/bigtable/docs/garbage-collection#data-removed>

   upvoted 6 times

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

**Selected Answer: B**

Because it can take up to a week for expired data to be deleted, you should never rely solely on garbage collection policies to ensure that read requests return the desired data. Always apply a filter to your read requests that excludes the same values as your garbage collection rules. You can filter by limiting the number of cells per column or by specifying a timestamp range.

   upvoted 1 times

  **m\_a\_p\_s** 4 months, 3 weeks ago

**Selected Answer: B**

"Because it can take up to a week for expired data to be deleted, you should never rely solely on garbage collection policies to ensure that read requests return the desired data. Always apply a filter to your read requests that excludes the same values as your garbage collection rules. You can filter by limiting the number of cells per column or by specifying a timestamp range."

<https://cloud.google.com/bigtable/docs/garbage-collection#data-removed>



   upvoted 1 times

  **Sofia98** 9 months, 4 weeks ago

**Selected Answer: B**

I will go for B too

   upvoted 1 times

  **GCP001** 10 months ago

B. Use a timestamp range filter in the query to fetch the customer's data for a specific range.

Always use query filter as garbage collector runs on it's way - <https://cloud.google.com/bigtable/docs/garbage-collection>

   upvoted 3 times

  **scaenrui** 10 months ago

**Selected Answer: B**

B. Use a timestamp range filter in the query to fetch the customer's data for a specific range.

   upvoted 1 times

> Home

> All Exams

> Examtopics PRO

> Training Courses



© 2024 ExamTopics