

EXAMTOPICS

- Expert Verified, Online, Free.

≡ MENU



← Google Discussions



Exam Professional Machine Learning Engineer All Questions

View all questions & answers for the Professional Machine Learning Engineer exam

Go to Exam

EXAM PROFESSIONAL MACHINE LEARNING ENGINEER TOPIC 1 QUESTION 17 DISCUSSIO..

Actual exam question from Google's Professional Machine Learning Engineer

Question #: 17

Topic #: 1

[\[All Professional Machine Learning Engineer Questions\]](#)

You are building a real-time prediction engine that streams files which may contain Personally Identifiable Information (PII) to Google Cloud. You want to use the Cloud Data Loss Prevention (DLP) API to scan the files. How should you ensure that the PII is not accessible by unauthorized individuals?

- A. Stream all files to Google Cloud, and then write the data to BigQuery. Periodically conduct a bulk scan of the table using the DLP API.
- B. Stream all files to Google Cloud, and write batches of the data to BigQuery. While the data is being written to BigQuery, conduct a bulk scan of the data using the DLP API.
- C. Create two buckets of data: Sensitive and Non-sensitive. Write all data to the Non-sensitive bucket. Periodically conduct a bulk scan of that bucket using the DLP API, and move the sensitive data to the Sensitive bucket.
- D. Create three buckets of data: Quarantine, Sensitive, and Non-sensitive. Write all data to the Quarantine bucket. Periodically conduct a bulk scan of that bucket using the DLP API, and move the data to either the Sensitive or Non-Sensitive bucket.




Show Suggested Answer

by  [chohan](#) at June 16, 2021, 1:58 p.m.

Comments

Type your comment...

Submit

  **chohan** Highly Voted  3 years, 4 months ago

Should be D

https://cloud.google.com/architecture/automating-classification-of-data-uploaded-to-cloud-storage#building_the_quarantine_and_classification_pipeline

   upvoted 25 times

  **Swagluke** 3 years, 2 months ago

All PII should be Sensitive data, that's why I think the answer is A.

   upvoted 1 times




  **u_phoria** 2 years, 3 months ago

Option D, as documented in that link (a fully automated process, using Cloud Functions - rather than a "periodic" scan as worded in the question), would be my choice.

It's easier than B, which would work for a real-time scenario - but would require loads more custom work to implement (things like batching, segmentation, triggering).



A and C are 'reactive' / periodic, and so not appropriate for the given scenario.

   upvoted 1 times

  **maartenalexander** Highly Voted  3 years, 4 months ago

D; others pose risks

   upvoted 5 times

  **Choisus** Most Recent  1 week, 5 days ago

Selected Answer: B

why not b, it requires real-time right?

   upvoted 1 times

  **PhilipKoku** 5 months ago

Selected Answer: D

D) The best choice for this scenario would be D. Create three buckets of data: Quarantine, Sensitive, and Non-sensitive.

Write all data to the Quarantine bucket. Periodically conduct a bulk scan of that bucket using the DLP API, and move the data to either the Sensitive or Non-Sensitive bucket.

   upvoted 1 times

  **fragkris** 11 months, 1 week ago

Selected Answer: D

D - Quarantine bucket is the google recommended approach

   upvoted 2 times

  **tavva_prudhvi** 1 year ago

Selected Answer: D

Option B does not provide a clear separation between sensitive and non-sensitive data before it is written to BigQuery, which means that PII might be exposed during the process.

But, in D offers a better level of security by writing all the data to a Quarantine bucket first. This way, the DLP API can scan and categorize the data into Sensitive or Non-sensitive buckets before it is further processed or stored. This ensures that PII is not accessible by unauthorized individuals, as the sensitive data is identified and separated from the non-sensitive data before any further actions are taken.

   upvoted 1 times



  **harithacML** 1 year, 4 months ago

Selected Answer: D

real-time prediction engine, that streams files to Google Cloud. PII is not accessible by unauthorized individuals.

D

   upvoted 1 times

  **Liting** 1 year, 4 months ago

Selected Answer: D

D should be the correct answer

   upvoted 1 times

🗨️ **M25** 1 year, 6 months ago

Selected Answer: D

Went with D

👍 ↩️ 🚩 upvoted 2 times

🗨️ **lucaluca1982** 1 year, 6 months ago

Selected Answer: B

B is real time

👍 ↩️ 🚩 upvoted 1 times

🗨️ **dfdrin** 1 year, 7 months ago

Selected Answer: D

It's D

👍 ↩️ 🚩 upvoted 1 times

🗨️ **enghabeth** 1 year, 9 months ago

Selected Answer: B

A, D, C they do not apply to a realtime case, all three say that the scan is applied periodically
Then it's B

👍 ↩️ 🚩 upvoted 3 times

🗨️ **tavva_prudhvi** 1 year, 8 months ago

Never mentioned periodically in the question, if I'm not wrong?

👍 ↩️ 🚩 upvoted 1 times

🗨️ **guilhermebutzke** 1 year, 9 months ago

Selected Answer: B

I think that is the correct because of the "real time" application.

👍 ↩️ 🚩 upvoted 1 times

🗨️ **EFIGO** 1 year, 11 months ago

Selected Answer: D

D is the right answer: you can temporarily store the sensitive data in a Quarantine bucket with restricted access, then move the data to the relative buckets once the PII have been protected.

👍 ↩️ 🚩 upvoted 1 times

🗨️ **GCP72** 2 years, 2 months ago

Selected Answer: D

Correct answer is "D"

👍 ↩️ 🚩 upvoted 1 times

🗨️ **dasouna** 2 years, 5 months ago

Answer is D : Question says that there MAY be sensitive data, so not all data is sensitive. This is why we need 3 buckets : Quarantine as a landing bucket, sensitive for sensitive data after DLP scan, non-sensitive for non-sensitive after DLP scan.
<https://cloud.google.com/architecture/automating-classification-of-data-uploaded-to-cloud-storage>

👍 ↩️ 🚩 upvoted 1 times

🗨️ **atuls287** 2 years, 7 months ago

Selected Answer: B

Reason being "Real Time" DLP scanning. Option A would scan all the data again and again. For others - Buckets etc is overkill and offline process.

👍 ↩️ 🚩 upvoted 2 times

[Load full discussion...](#)

Start Learning for free



Social Media

[Facebook](#) , [Twitter](#)

[YouTube](#) , [Reddit](#)

[Pinterest](#)



We are the biggest and most updated IT certification exam material website.

Using our own resources, we strive to strengthen the IT professionals community for free.



© 2024 ExamTopics

ExamTopics doesn't offer Real Microsoft Exam Questions. ExamTopics doesn't offer Real Amazon Exam Questions. ExamTopics Materials do not contain actual questions and answers from Cisco's Certification Exams.

CFA Institute does not endorse, promote or warrant the accuracy or quality of ExamTopics. CFA® and Chartered Financial Analyst® are registered trademarks owned by CFA Institute.