◉ **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 212 DISCUSSION**

Actual exam question from Google's Professional Data Engineer

Question #: 212

Topic #: 1

[All Professional Data Engineer Questions]

You are troubleshooting your Dataflow pipeline that processes data from Cloud Storage to BigQuery. You have discovered that the Dataflow worker nodes cannot communicate with one another. Your networking team relies on Google Cloud network tags to define firewall rules. You need to identify the issue while following Google-recommended networking security practices. What should you do?

A. Determine whether your Dataflow pipeline has a custom network tag set.

B. Determine whether there is a firewall rule set to allow traffic on TCP ports 12345 and 12346 for the Dataflow network tag.

C. Determine whether there is a firewall rule set to allow traffic on TCP ports 12345 and 12346 on the subnet used by Dataflow workers.

D. Determine whether your Dataflow pipeline is deployed with the external IP address option enabled.

Show Suggested Answer

by 👤 **e70ea9e** at *Dec. 30, 2023, 9:36 a.m.*

Comments

Type your comment...

Submit

MaxNRG `Highly Voted` 10 months ago

`Selected Answer: B`

The best approach would be to check if there is a firewall rule allowing traffic on TCP ports 12345 and 12346 for the Dataflow network tag.
Dataflow uses TCP ports 12345 and 12346 for communication between worker nodes. Using network tags and associated firewall rules is a Google-recommended security practice for controlling access between Compute Engine instances like Dataflow workers.

So the key things to check would be:

1. Ensure your Dataflow pipeline is using the Dataflow network tag on the worker nodes. This tag is applied by default unless overridden.
2. Check if there is a firewall rule allowing TCP 12345 and 12346 ingress and egress traffic for instances with the Dataflow network tag. If not, add the rule.

Options A, C and D relate to other networking aspects but do not directly address the Google recommended practice of using network tags and firewall rules.

👍 ↩ 🚩 upvoted 10 times

JyoGCP `Most Recent ⊙` 8 months, 3 weeks ago

`Selected Answer: B`

B. Determine whether there is a firewall rule set to allow traffic on TCP ports 12345 and 12346 for the Dataflow network tag.

👍 ↩ 🚩 upvoted 1 times

Matt_108 9 months, 3 weeks ago

`Selected Answer: B`

B, check if there is a firewall rule allowing traffic on TCP ports 12345 and 12346 for the Dataflow network tag.

👍 ↩ 🚩 upvoted 1 times

Smakyel79 10 months ago

`Selected Answer: B`

Because network tags are used and Dataflow uses TCP ports 12345 and 12346 as stated on
https://cloud.google.com/dataflow/docs/guides/routes-firewall

👍 ↩ 🚩 upvoted 3 times

raaad 10 months, 1 week ago

`Selected Answer: B`

This option focuses directly on ensuring that the firewall rules are set up correctly for the network tags used by Dataflow worker nodes. It specifically addresses the potential issue of worker nodes not being able to communicate due to restrictive firewall rules blocking the necessary ports.

👍 ↩ 🚩 upvoted 3 times

e70ea9e 10 months, 1 week ago

`Selected Answer: B`

Focus on Network Tags:

Adheres to the recommended practice of using network tags for firewall configuration, enhancing security and flexibility.
Avoids targeting specific subnets, which can be less secure and harder to manage.

👍 ↩ 🚩 upvoted 2 times