

🔗 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 120 DISCUSSION

Actual exam question from Google's Professional Data Engineer

Question #: 120

Topic #: 1

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

You are operating a Cloud Dataflow streaming pipeline. The pipeline aggregates events from a Cloud Pub/Sub subscription source, within a window, and sinks the resulting aggregation to a Cloud Storage bucket. The source has consistent throughput. You want to monitor an alert on behavior of the pipeline with Cloud Stackdriver to ensure that it is processing data. Which Stackdriver alerts should you create?

- A. An alert based on a decrease of subscription/num\_undelivered\_messages for the source and a rate of change increase of instance/storage/ used\_bytes for the destination
- B. An alert based on an increase of subscription/num\_undelivered\_messages for the source and a rate of change decrease of instance/storage/ used\_bytes for the destination
- C. An alert based on a decrease of instance/storage/used\_bytes for the source and a rate of change increase of subscription/ num\_undelivered\_messages for the destination
- D. An alert based on an increase of instance/storage/used\_bytes for the source and a rate of change decrease of subscription/ num\_undelivered\_messages for the destination

[Show Suggested Answer](#)

by [deleted] at March 22, 2020, 12:46 p.m.

## Comments

Type your comment...

Submit

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

You would want to get alerted only if Pipeline fails & not if it is running fine. I think Option [B] is correct, because in event of Pipeline failure :

- 1) subscription/ num\_undelivered\_messages would pile up at a constant rate as the source has consistent throughput
- 2) instance/storage/ used\_bytes will get closer to zero. Hence need to monitor it's rate of change

   upvoted 28 times

  **Barniyah** 3 years, 4 months ago

Yes, you are right, it should be B:  
Thank you

   upvoted 5 times

  **marioferrulli** 1 year, 10 months ago

Why would the instance/storage/used\_bytes get closer to zero? If there's an error at a certain point, wouldn't we just see that the used\_bytes remain constant while the num\_undelivered\_messages increases? I don't get why the destination's used bytes should decrease.



   upvoted 1 times

  **baubaumiaomiao** 1 year, 10 months ago

"If there's an error at a certain point, wouldn't we just see that the used\_bytes remain constant while the num\_undelivered\_messages increases?"

It's the rate of change, not the absolute value

   upvoted 2 times



  **szefco** 1 year, 10 months ago

"rate of change decrease of instance/storage/ used\_bytes" - if rate of instance/storage/ used\_bytes decreases that means less data is written - so something is wrong with the pipeline.

It's not used bytes that decreases - it's rate of change decreases.

Example: if everything works fine your pipeline writes 5MB/s to the sink. If it decreases to 0.1MB/s it means something is wrong

   upvoted 6 times

  **[Removed]** Highly Voted  3 years, 7 months ago

Correct - B

   upvoted 21 times

  **desertlotus1211** Most Recent  1 month, 2 weeks ago

Selected Answer: A

It should be Answer A. You want to see in being processed versus looking for a bottleneck.

   upvoted 1 times

  **midgoo** 7 months, 3 weeks ago

Selected Answer: B

For those who may get confuse at the start by the term 'subscription/num\_undelivered\_messages', it is not a division. It is the full path of the metric. So we should just read it as 'num\_undelivered\_messages'. The same for 'used\_bytes'.



So if we see the source have more backlog (more num\_undelivered\_messages), or the destination utilization going down, that is the indicator of something going wrong

   upvoted 15 times

  **desertlotus1211** 1 month, 2 weeks ago

the Answer is A. You want to see it working.

   upvoted 1 times

  **desertlotus1211** 1 month, 2 weeks ago

you're looking for evidence that it's working:  
'that it is processing data...'

   upvoted 1 times

  **kryzo** 4 months, 3 weeks ago

great explanation thanks !

   upvoted 2 times

  **musumusu** 8 months, 2 weeks ago

Answer B:

Trick: In steeldriver always put Alert for Subscriber CPU

THICK. In Stackdriver always put Alert for Subscriber + CPU  
Subscriber - num of undelivered message INCREASE alert  
CPU - Instance or storage DECREASE alert.  
Make sense right !

   upvoted 3 times

  **atlan** 10 months ago


Nobody seems to pay attention to instance/storage/used\_bytes. I only find this metric for Spanner.  
[https://cloud.google.com/monitoring/api/metrics\\_gcp#gcp-spanner](https://cloud.google.com/monitoring/api/metrics_gcp#gcp-spanner)

While Dataflow processes and stores everything in Cloud Storage, Spanner could only be the source.  
<https://cloud.google.com/spanner/docs/change-streams>

Also, if it is either A or B, the instance/storage/used\_bytes metric does not make sense for the destination, which is Cloud Storage.

Can anyone help me understand?

   upvoted 1 times

  **desertlotus1211** 9 months, 2 weeks ago

look here: [https://cloud.google.com/monitoring/api/metrics\\_gcp](https://cloud.google.com/monitoring/api/metrics_gcp)

instance/storage/used\_bytes GA  
Storage used.

   upvoted 2 times

  **AzureDP900** 10 months, 1 week ago


B. An alert based on an increase of subscription/num\_undelivered\_messages for the source and a rate of change decrease of instance/storage/ used\_bytes for the destination

   upvoted 1 times

  **AzureDP900** 10 months, 1 week ago

B is right

   upvoted 1 times

  **Catweazle1983** 10 months, 1 week ago

**Selected Answer: A**

An alert based on a decrease of subscription/num\_undelivered\_messages for the source and a rate of change increase of instance/storage/ used\_bytes for the destination

10 subscriptions / 1 undelivered messages = 10

10 subscriptions / 5 undelivered messages = 2

You clearly want to be alerted when the number of undelivered messages increases. The ratio then decreases. In my example from 10 to 2.

   upvoted 1 times

  **squishy\_fishy** 7 months, 2 weeks ago

subscription/num\_undelivered\_messages is a path, not a division.

   upvoted 1 times

  **zellck** 11 months ago

**Selected Answer: B**

B is the answer.



[https://cloud.google.com/pubsub/docs/monitoring#monitoring\\_the\\_backlog](https://cloud.google.com/pubsub/docs/monitoring#monitoring_the_backlog)

Monitor message backlog

To ensure that your subscribers are keeping up with the flow of messages, create a dashboard. The dashboard can show the following backlog metrics, aggregated by resource, for all your subscriptions:

- Unacknowledged messages (subscription/num\_undelivered\_messages) to see the number of unacknowledged messages.

   upvoted 3 times

  **A1000** 1 year, 2 months ago

**Selected Answer: B**

Increase subscription/num delivered message




decrease instance/storage/used bytes

   upvoted 1 times

  **Pime13** 1 year, 4 months ago

**Selected Answer: B**

Correct - B

👍 ↩ 🚩 upvoted 1 times

🗨️ 👤 **JG123** 1 year, 11 months ago

Correct: B

👍 ↩ 🚩 upvoted 2 times

🗨️ 👤 **Abhi16820** 1 year, 11 months ago

isn't B and C are same.

👍 ↩ 🚩 upvoted 2 times

🗨️ 👤 **JayZeeLee** 1 year, 12 months ago

B.

It's useful to monitor the source that keeps sending data while the destination that doesn't take anything in.

👍 ↩ 🚩 upvoted 3 times

🗨️ 👤 **squishy\_fishy** 2 years ago

The answer is B.

subscription/num\_undelivered\_messages: the number of messages that subscribers haven't processed

[https://cloud.google.com/pubsub/docs/monitoring#monitoring\\_forwarded\\_undeliverable\\_messages](https://cloud.google.com/pubsub/docs/monitoring#monitoring_forwarded_undeliverable_messages)

👍 ↩ 🚩 upvoted 2 times

🗨️ 👤 **squishy\_fishy** 2 years, 1 month ago

Silly question: what is subscription/ num\_undelivered\_messages, it is divided by? or per subscription per num\_undelivered\_messages?

👍 ↩ 🚩 upvoted 2 times

🗨️ 👤 **910** 1 year, 7 months ago

yes is misleading:

the metric "subscription/num\_undelivered\_messages" is just the path of the API URL

[actions.googleapis.com/...subscription/num\\_undelivered\\_messages](https://actions.googleapis.com/...subscription/num_undelivered_messages)

ref: [https://cloud.google.com/monitoring/api/metrics\\_gcp#pubsub/subscription/num\\_undelivered\\_messages](https://cloud.google.com/monitoring/api/metrics_gcp#pubsub/subscription/num_undelivered_messages)

👍 ↩ 🚩 upvoted 3 times

[Load full discussion...](#)



## Platform

> Home

> Examtopics PRO

> All Exams

> Training Courses

