Google Discussions

EXAM PROFESSIONAL DATA ENGINEER TOPIC 1 QUESTION 106 DISCUSSION

Actual exam question from Google's Professional Data Engineer
Question #: 106
Topic #: 1
[All Professional Data Engineer Questions]

You are managing a Cloud Dataproc cluster. You need to make a job run faster while minimizing costs, without losing work in progress on your clusters. What should you do?

A. Increase the cluster size with more non-preemptible workers.

B. Increase the cluster size with preemptible worker nodes, and configure them to forcefully decommission.

C. Increase the cluster size with preemptible worker nodes, and use Cloud Stackdriver to trigger a script to preserve work.

D. Increase the cluster size with preemptible worker nodes, and configure them to use graceful decommissioning.

Show Suggested Answer

by AWSandeep at *Sept. 3, 2022, 2:06 p.m.*

Comments

Type your comment...

Submit

AzureDP900 Highly Voted 1 year, 4 months ago
D is right
https://cloud.google.com/dataproc/docs/concepts/configuring-clusters/scaling-clusters#using_graceful_decommissioning
upvoted 6 times

rocky48 Most Recent 4 months, 4 weeks ago

Selected Answer: D

D is right
https://cloud.google.com/dataproc/docs/concepts/configuring-clusters/scaling-clusters#using_graceful_decommissioning

upvoted 1 times

Prakzz 1 year, 4 months ago

Selected Answer: A

Should be A. You can configure the preemptible worker to gracefull decommission, its for non preemptible worker nodes.

upvoted 1 times

wan2three 1 year, 4 months ago

nope, they are not only for non-preeemtible workers

upvoted 1 times

yafsong 1 year, 4 months ago

graceful decommissioning: to finish work in progress on a worker before it is removed from the Cloud Dataproc cluster.
https://cloud.google.com/dataproc/docs/concepts/configuring-clusters/scaling-clusters

upvoted 2 times

odacir 1 year, 4 months ago

Selected Answer: D

All your workers need to be the same kind. Use Graceful Decommissioning for don't lose any data and add more(increase the cluster) preemptible workers because there are more cost-effective .

upvoted 1 times

skp57 1 year, 5 months ago

A. "graceful decommissioning" is not a configuration value but a parameter passed with scale down action - to decrease the number of workers to save money (see Graceful Decommissioning as an option to use when downsizing a cluster to avoid losing work in progress)

upvoted 2 times

John_Pongthorn 1 year, 7 months ago

Selected Answer: D

https://cloud.google.com/dataproc/docs/concepts/configuring-clusters/scaling-clusters
Why scale a Dataproc cluster?
to increase the number of workers to make a job run faster
to decrease the number of workers to save money (see Graceful Decommissioning as an option to use when downsizing a cluster to avoid losing work in progress).
to increase the number of nodes to expand available Hadoop Distributed Filesystem (HDFS) storage

upvoted 3 times

hauhau 1 year, 6 months ago

This weird.
The question mentions that increase cluster, but Graceful Decommissioning use in downscale the cluster

upvoted 2 times

odacir 1 year, 4 months ago

All your workers need to be the same kind. Use Graceful Decommissioning for don't lose any data and add more preemptible workers because there are more cost-effective

upvoted 1 times

AWSandeep 1 year, 8 months ago

Selected Answer: D

D. Increase the cluster size with preemptible worker nodes, and configure them to use graceful decommissioning.

upvoted 1 times