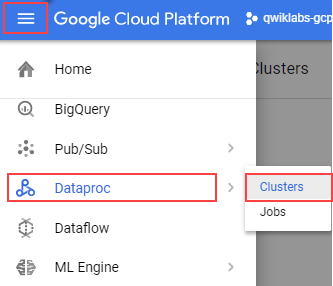
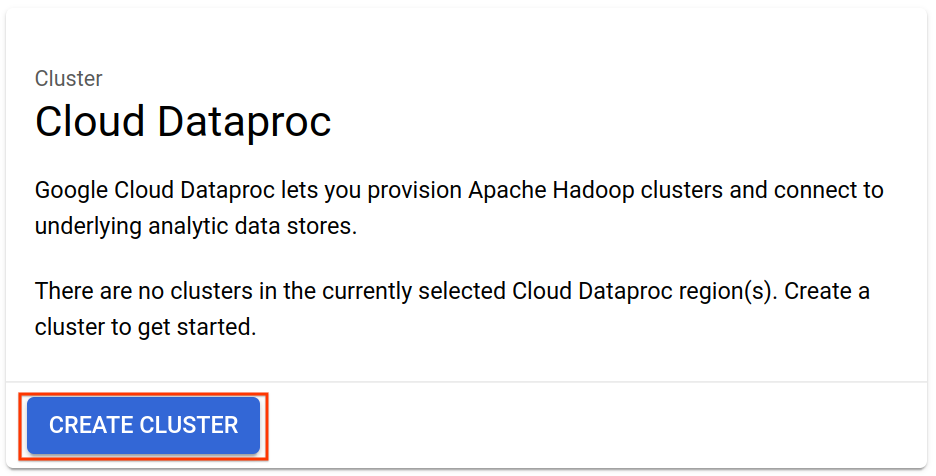
# **Introduction to Cloud DataProc: Hadoop and Spark on Google Cloud**

# **Create a Cloud Dataproc cluster**

In the console, click **Navigation menu** > **Dataproc** > **Clusters** on the top left of the screen:



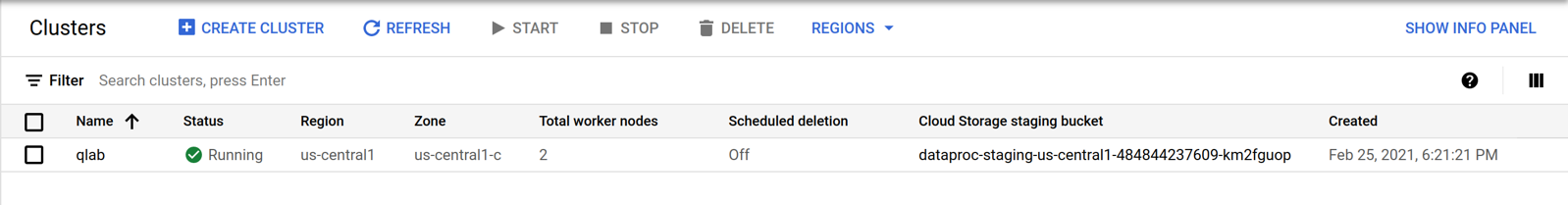
To create a ne.w cluster, click **Create cluster** and for Cluster on Compute Engine click **Create**



There are many parameters you can configure when creating a new cluster. Set values for the parameters listed below, leave the default settings for the other parameters.

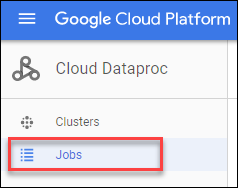
|  |  |
| --- | --- |
| **Parameter** | **Value** |
| Name | qlab |
| Region | us-east1 |
| Zone | us-east1-c |
| Click Configure nodes, for Master node - Machine type | 4 vCPUs (n1-standard-4) |
| Worker node - Machine type | 2 vCPUs (n1-standard-2) |

Click on **Create** to create the new cluster. You will see the Status go from Provisioning to Running—move on to the next step once you're output resembles the following:

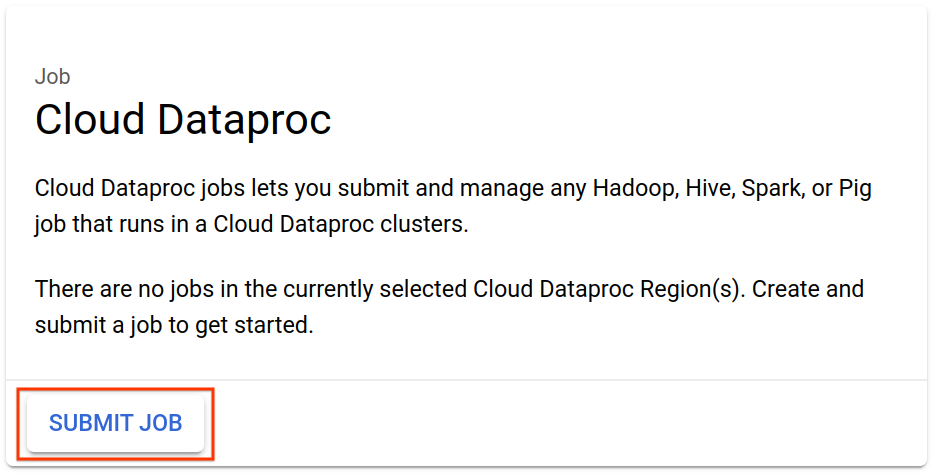


**Submit a Spark job to your cluster**

Select **Jobs** to switch to Dataproc's jobs view.

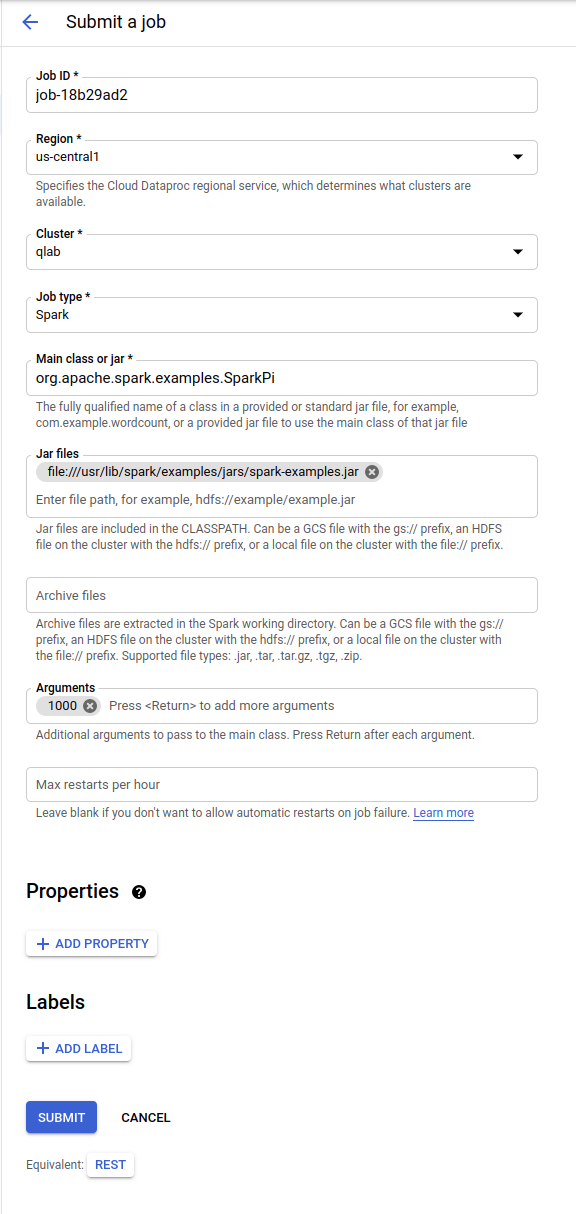


Click **Submit job**.



Set values for the parameters listed below, leave the default settings for the other parameters.

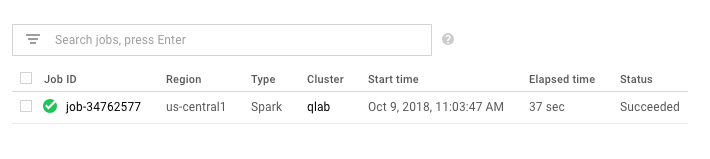
|  |  |
| --- | --- |
| **Parameter** | **Value** |
| Region | us-east1 |
| Cluster | qlab |
| Job type | Spark |
| Main class or jar | org.apache.spark.examples.SparkPi |
| Jar files | file:///usr/lib/spark/examples/jars/spark-examples.jar |
| Arguments | 1000 (This sets the number of tasks) |



Click **Submit**.

Your job should appear in the Jobs list, which shows all your project's jobs with their cluster, type, and current status. The new job displays as "Running"—move on once you see "Succeeded" as the Status.

To see your completed job's output, click the job ID in the **Jobs** list.



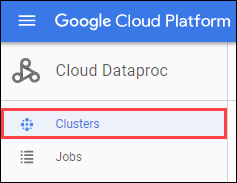
To avoid scrolling, select **Line Wrap** to **ON**.



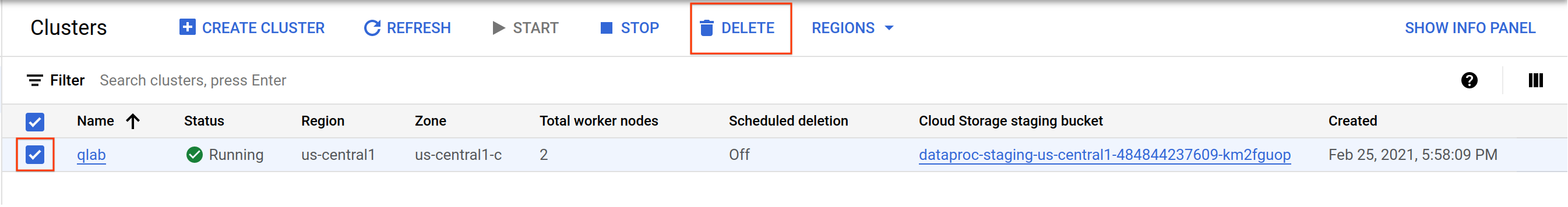
You should see that your job has successfully calculated a rough value for pi!

**Shut down your cluster**

You can shut down a cluster on the Clusters page.



Select the checkbox next to the **qlab** cluster and click **Delete**.



Click **CONFIRM** to confirm deletion.