Imagen que contiene Interfaz de usuario gráfica

Descripción generada automáticamente

**Indice**

Contenido

[1](#_Toc170120908)

[**Indice** 2](#_Toc170120909)

[**Autoscaling an Instance Group with Custom Cloud Monitoring Metrics** 2](#_Toc170120910)

[Autoscaling an Instance Group with Custom Cloud Monitoring Metrics | Google Cloud Skills Boost 2](#_Toc170120911)

[**GSP087** 2](#_Toc170120912)

[**Overview** 3](#_Toc170120913)

[Application architecture 3](#_Toc170120914)

[**Objectives** 4](#_Toc170120915)

[**Task 1. Creating the application** 5](#_Toc170120916)

[Uploading the script files to Cloud Storage 5](#_Toc170120917)

[**Task 2. Create a bucket** 6](#_Toc170120918)

[Understanding the code components 9](#_Toc170120919)

[**Task 3. Creating an instance template** 10](#_Toc170120920)

[**Task 4. Creating the instance group** 13](#_Toc170120921)

[**Task 5. Verifying that the instance group has been created** 16](#_Toc170120922)

[**Task 6. Verifying that the Node.js script is running** 17](#_Toc170120923)

[**Task 7. Configure autoscaling for the instance groups** 20](#_Toc170120924)

[**Task 8. Watching the instance group perform autoscaling** 25](#_Toc170120925)

[**Task 9. Autoscaling example** 27](#_Toc170120926)

**Autoscaling an Instance Group with Custom Cloud Monitoring Metrics**

[Autoscaling an Instance Group with Custom Cloud Monitoring Metrics | Google Cloud Skills Boost](https://www.cloudskillsboost.google/focuses/611?parent=catalog)

**GSP087**



**Overview**

In this lab you will create a [Compute Engine](https://cloud.google.com/compute/docs/) managed instance group that autoscales based on the value of a custom [Cloud Monitoring](https://cloud.google.com/monitoring/docs/) metric.

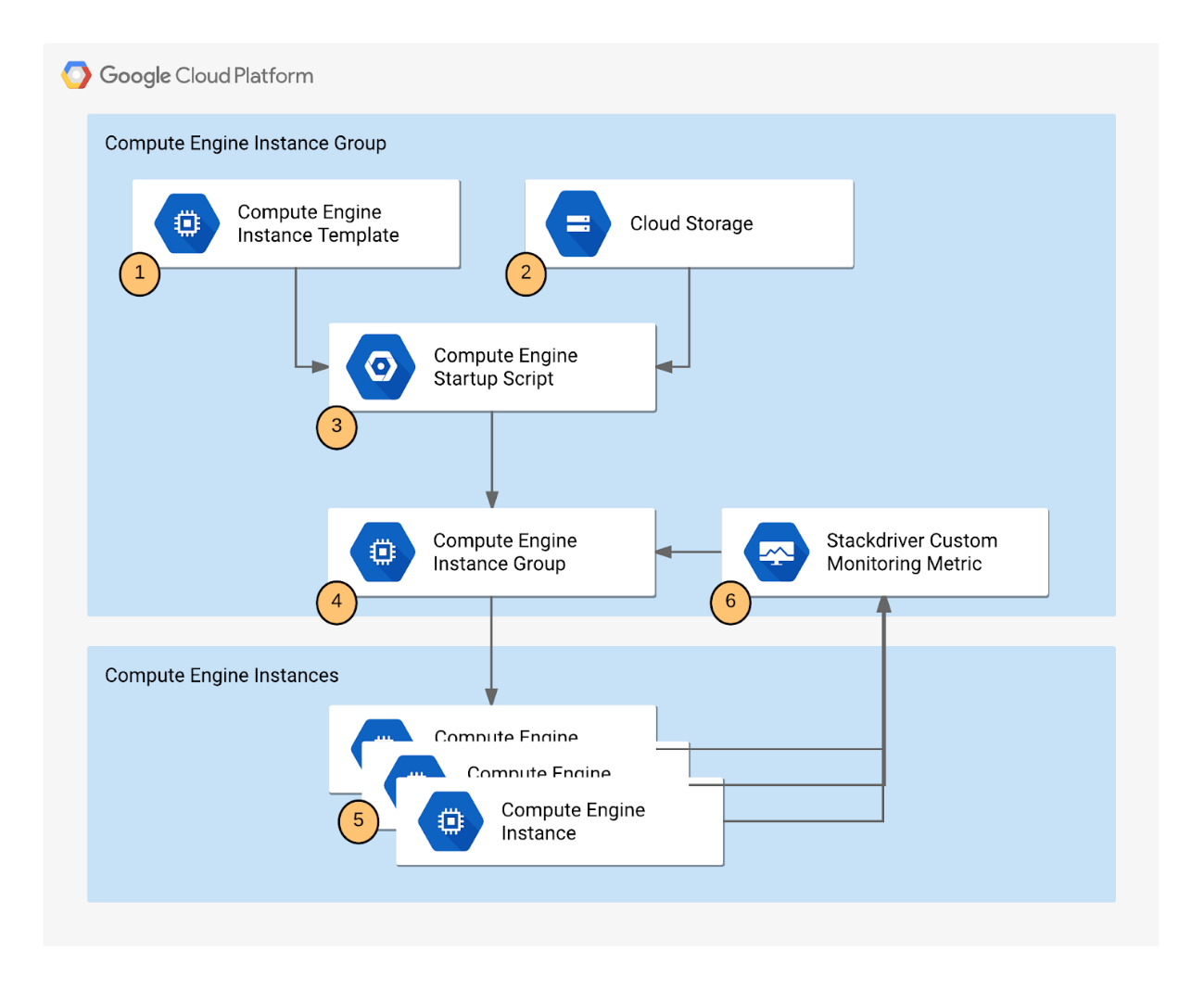
Application architecture

The autoscaling application uses a Node.js script installed on Compute Engine instances. The script reports a numeric value to a Cloud monitoring metric. You do not need to know Node.js or JavaScript for this lab. In response to the value of the metric, the application autoscales the Compute Engine instance group up or down as needed.

The Node.js script is used to seed a custom metric with values that the instance group can respond to. In a production environment, you would base autoscaling on a metric that is relevant to your use case.

The application includes the following components:

1. **Compute Engine instance template** - A template used to create each instance in the instance group.
2. **Cloud Storage** - A bucket used to host the startup script and other script files.
3. **Compute Engine startup script** - A startup script that installs the necessary code components on each instance. The startup script is installed and started automatically when an instance starts. When the startup script runs, it in turn installs and starts code on the instance that writes values to the Cloud monitoring custom metric.
4. **Compute Engine instance group** - An instance group that autoscales based on the Cloud monitoring metric values.
5. **Compute Engine instances** - A variable number of Compute Engine instances.
6. **Custom Cloud Monitoring metric** - A custom monitoring metric used as the input value for Compute Engine instance group autoscaling.



**Objectives**

In this lab, you will learn how to perform the following tasks:

* Deploy an autoscaling Compute Engine instance group.
* Create a custom metric used to scale the instance group.
* Use the [Cloud Console](https://cloud.google.com/storage/docs/cloud-console) to visualize the custom metric and instance group size.

**Task 1. Creating the application**

Creating the autoscaling application requires downloading the necessary code components, creating a managed instance group, and configuring autoscaling for the managed instance group.

Uploading the script files to Cloud Storage

During autoscaling, the instance group will need to create new Compute Engine instances. When it does, it creates the instances based on an instance template. Each instance needs a startup script. Therefore, the template needs a way to reference the startup script. Compute Engine supports using Cloud Storage buckets as a source for your startup script. In this section, you will make a copy of the startup script and application files for a sample application used by this lab that pushes a pattern of data into a custom Cloud logging metric that you can then use to configure as the metric that controls the autoscaling behavior for an autoscaling group.

**Note:** There is a pre-existing instance template and group that has been created automatically by the lab that is already running. Autoscaling requires at least 30 minutes to demonstrate both scale-up and scale-down behavior, and you will examine this group later to see how scaling is controlled by the variations in the custom metric values generated by the custom metric scripts.

**Task 2. Create a bucket**

1. In the Cloud Console, from the **Navigation menu** select **Cloud Storage** > **Buckets**, then click **Create**.

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

1. Give your bucket a unique name, but don't use a name you might want to use in another project. For details about how to name a bucket, see the [bucket naming guidelines](https://cloud.google.com/storage/docs/naming). You can use your Project ID for the bucket. This bucket will be referenced as YOUR\_BUCKET throughout the lab.

Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

Descripción generada automáticamente

1. Accept the default values then click **Create**.

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

Click **Confirm** for Public access will be prevented pop-up if prompted.

Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

Descripción generada automáticamente

When the bucket is created, the **Bucket details** page opens.

**Test completed task**

Click **Check my progress** to verify your performed task. If you have successfully created a Cloud Storage bucket, you will see an assessment score.

Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente

|  |
| --- |
| Create a Cloud Storage bucket  Check my progress |

Escala de tiempo

Descripción generada automáticamente con confianza baja

1. Next, run the following command in Cloud Shell to copy the startup script files from the lab default Cloud Storage bucket to your Cloud Storage bucket. Remember to replace <YOUR BUCKET> with the name of the bucket you just made:

gsutil cp -r gs://spls/gsp087/\* gs://<YOUR BUCKET>

Texto, Chat o mensaje de texto

Descripción generada automáticamente

1. After you upload the scripts, click **Refresh** on the **Bucket details** page. Your bucket should list the added files.

Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente

Understanding the code components

* Startup.sh - A shell script that installs the necessary components to each Compute Engine instance as the instance is added to the managed instance group.
* writeToCustomMetric.js - A Node.js snippet that creates a custom monitoring metric whose value triggers scaling. To emulate real-world metric values, this script varies the value over time. In a production deployment, you replace this script with custom code that reports the monitoring metric that you're interested in, such as a processing queue value.
* Config.json - A Node.js config file that specifies the values for the custom monitoring metric and used in writeToCustomMetric.js.
* Package.json - A Node.js package file that specifies standard installation and dependencies for writeToCustomMetric.js.
* writeToCustomMetric.sh - A shell script that continuously runs the writeToCustomMetric.js program on each Compute Engine instance.

**Task 3. Creating an instance template**

Now create a template for the instances that are created in the instance group that will use autoscaling. As part of the template, you specify the location (in Cloud Storage) of the startup script that should run when the instance starts.

1. In the Cloud Console, click **Navigation menu > Compute Engine > Instance templates**.
2. Click **Create Instance Template** at the top of the page.

Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

Descripción generada automáticamente

1. Name the instance template autoscaling-instance01.

Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

Descripción generada automáticamente

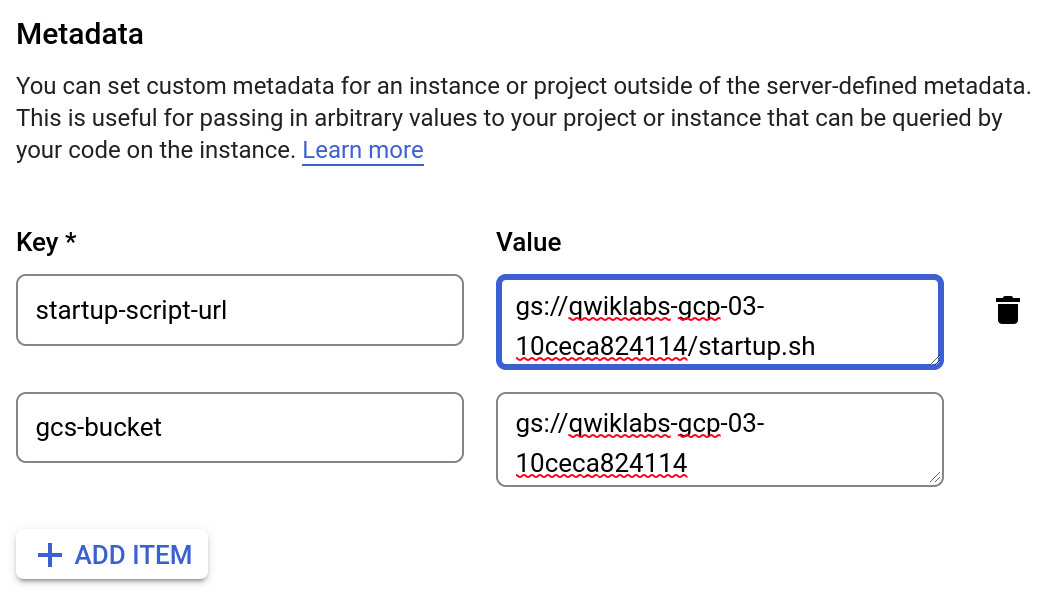
1. Scroll down, click **Advanced options**.

Interfaz de usuario gráfica, Texto

Descripción generada automáticamente

1. In the **Metadata** section of the **Management** tab, enter these metadata keys and values, clicking the **+ Add item** button to add each one. Remember to substitute your bucket name for the [YOUR\_BUCKET\_NAME] placeholder:

|  |  |
| --- | --- |
| **Key** | **Value** |
| startup-script-url | gs://[YOUR\_BUCKET\_NAME]/startup.sh |
| gcs-bucket | gs://[YOUR\_BUCKET\_NAME] |



Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

1. Click **Create**.

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

**Test completed task**

Click **Check my progress** to verify your performed task. If you have successfully created an instance template, you will see an assessment score.

|  |
| --- |
| Create an instance template  Check my progress |

Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

Descripción generada automáticamente

**El Check my progress en este punto no funciona**

**Task 4. Creating the instance group**

1. In the left pane, click **Instance groups**.

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

1. Click **Create instance group** at the top of the page.

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

1. **Name:** autoscaling-instance-group-1.

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

1. For **Instance template**, select the instance template you just created.

Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

Descripción generada automáticamente

1. For **Location**, select **Single Zone** and use region and zone for the region and zone, respectively.

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

1. Set **Autoscaling mode** to **Off: do not autoscale**.

Texto

Descripción generada automáticamente

You'll edit the autoscaling setting after the instance group has been created. Leave the other settings at their default values.

1. Click **Create**.

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

**Note:** You can ignore the Autoscaling is turned off. The number of instances in the group won't change automatically. The autoscaling configuration is preserved. warning next to your instance group.

**Test completed task**

Click **Check my progress** to verify your performed task. If you have successfully created an instance group, you will see an assessment score.

|  |
| --- |
| Create an instance group  Check my progress |

Escala de tiempo

Descripción generada automáticamente con confianza media

**Task 5. Verifying that the instance group has been created**

Wait to see the green check mark next to the new instance group you just created. It might take the startup script several minutes to complete installation and begin reporting values. Click **Refresh** if it seems to be taking more than a few minutes.

**Note:** If you see a red icon next to the other instance group that was pre-created by the lab, you can ignore this warning. The instance group reports a warning for up to 10-15 minutes as it is initializing. This is expected behavior.

Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente

**Task 6. Verifying that the Node.js script is running**

The custom metric custom.googleapis.com/appdemo\_queue\_depth\_01 isn't created until the first instance in the group is created and that instance begins reporting custom metric values.

You can verify that the writeToCustomMetric.js script is running on the first instance in the instance group by checking whether the instance is logging custom metric values.

1. Still in the Compute Engine Instance groups window, click the name of the autoscaling-instance-group-1 to display the instances that are running in the group.

Interfaz de usuario gráfica, Texto, Aplicación, Chat o mensaje de texto

Descripción generada automáticamente

1. Scroll down and click the instance name. Because autoscaling has not started additional instances, there is just a single instance running.

Imagen que contiene Tabla

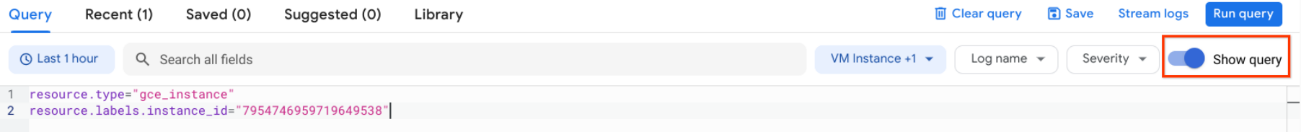
Descripción generada automáticamente

1. In the **Details** tab, in the **Logs** section, click the **Logging** link to view the logs for the VM instance.

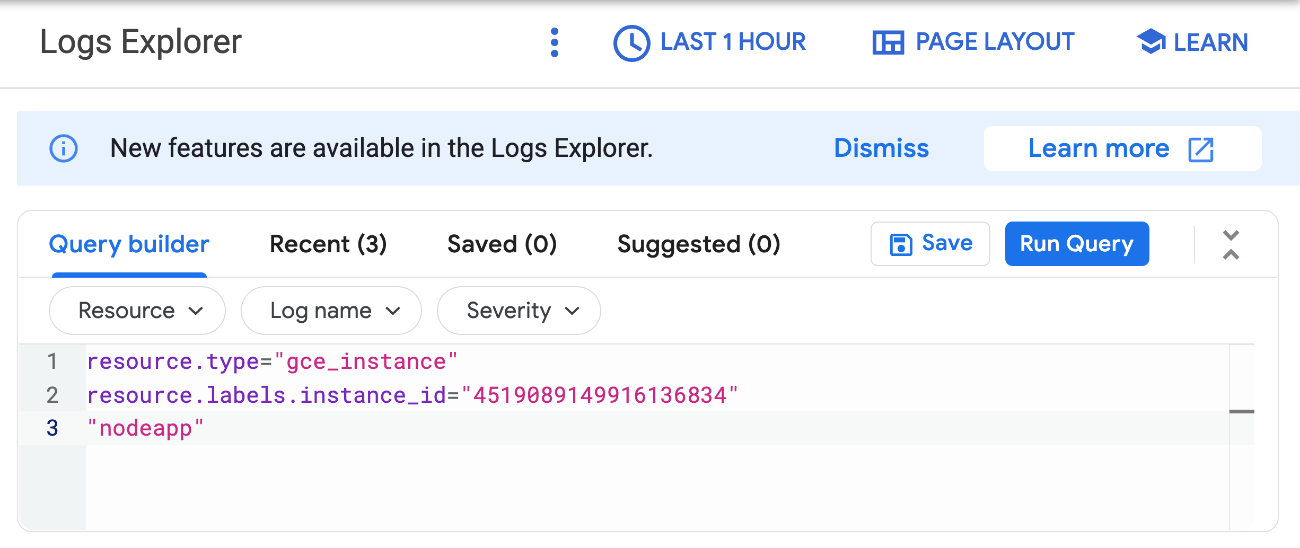
Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente

1. Wait a minute or 2 to let some data accumulate. Enable the **Show query** toggle, you will see resource.type and resource.labels.instance\_id in the **Query** preview box.



1. Add "nodeapp" as line 3, so the code looks similar to this:



1. Click **Run query**.

Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente

If the Node.js script is being executed on the Compute Engine instance, a request is sent to the API, and log entries that say nodeapp: available is displayed.

**Note:**If you don't see this log entry, the Node.js script isn't reporting the custom metric values. Check that the metadata was entered correctly. If the metadata is incorrect, it might be easiest to restart the lab. It may take around 10 minutes for the app to start up.

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

**Task 7. Configure autoscaling for the instance groups**

After you've verified that the custom metric is successfully reporting data from the first instance, the instance group can be configured to autoscale based on the value of the custom metric.

1. In the Cloud Console, go to **Compute Engine > Instance groups**.
2. Click the autoscaling-instance-group-1 group.

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

1. Click **Edit**.

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

1. Under **Autoscaling** set **Autoscaling mode** to **On: add and remove instances to the group**.
2. Set **Minimum number of instances**: 1 and **Maximum number of instances**: 3

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

1. Under **Autoscaling signals** click **ADD SIGNAL** to edit metric. Set the following fields, leave all others at the default value.

Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

Descripción generada automáticamente

* + **Signal type**: Cloud Monitoring metric new. Click **Configure**.

Interfaz de usuario gráfica, Texto, Aplicación, Chat o mensaje de texto

Descripción generada automáticamente

* + Under **Resource and metric** click **SELECT A METRIC** and navigate to **VM Instance > Custom metrics >**

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

* + **Custom/appdemo\_queue\_depth\_01**.

Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente

* + Click **Apply**.

Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente

* + **Utilization target**: 150

When custom monitoring metric values are higher or lower than the **Target** value, the autoscaler scales the managed instance group, increasing or decreasing the number of instances. The target value can be any [double](https://cloud.google.com/monitoring/api/ref_v3/rest/v3/projects.metricDescriptors#valuetype) value, but for this lab, the value 150 was chosen because it matches the values being reported by the custom monitoring metric.

* + **Utilization target type**: Gauge. Click **Select**.

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

The **Gauge** setting specifies that the autoscaler should compute the average value of the data collected over the last few minutes and compare it to the target value. (By contrast, setting **Target mode** to **DELTA\_PER\_MINUTE** or **DELTA\_PER\_SECOND** autoscales based on the *observed* rate of change rather than an *average* value.)

1. Click **Save**.

Interfaz de usuario gráfica, Texto, Aplicación, Chat o mensaje de texto

Descripción generada automáticamente

**Test completed task**

Click **Check my progress** to verify your performed task. If you have successfully configured autoscaling for the instance group, you will see an assessment score.

|  |
| --- |
| Configure autoscaling for the instance group  Check my progress |

Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

Descripción generada automáticamente

Interfaz de usuario gráfica

Descripción generada automáticamente

**Task 8. Watching the instance group perform autoscaling**

The Node.js script varies the custom metric values it reports from each instance over time. As the value of the metric goes up, the instance group scales up by adding Compute Engine instances. If the value goes down, the instance group detects this and scales down by removing instances. As noted earlier, the script emulates a real-world metric whose value might similarly fluctuate up and down.

Next, you will see how the instance group is scaling in response to the metric by clicking the **Monitoring** tab to view the **Autoscaled size** graph.

1. In the left pane, click **Instance groups**.
2. Click the builtin-igm instance group in the list.

Interfaz de usuario gráfica, Texto, Aplicación

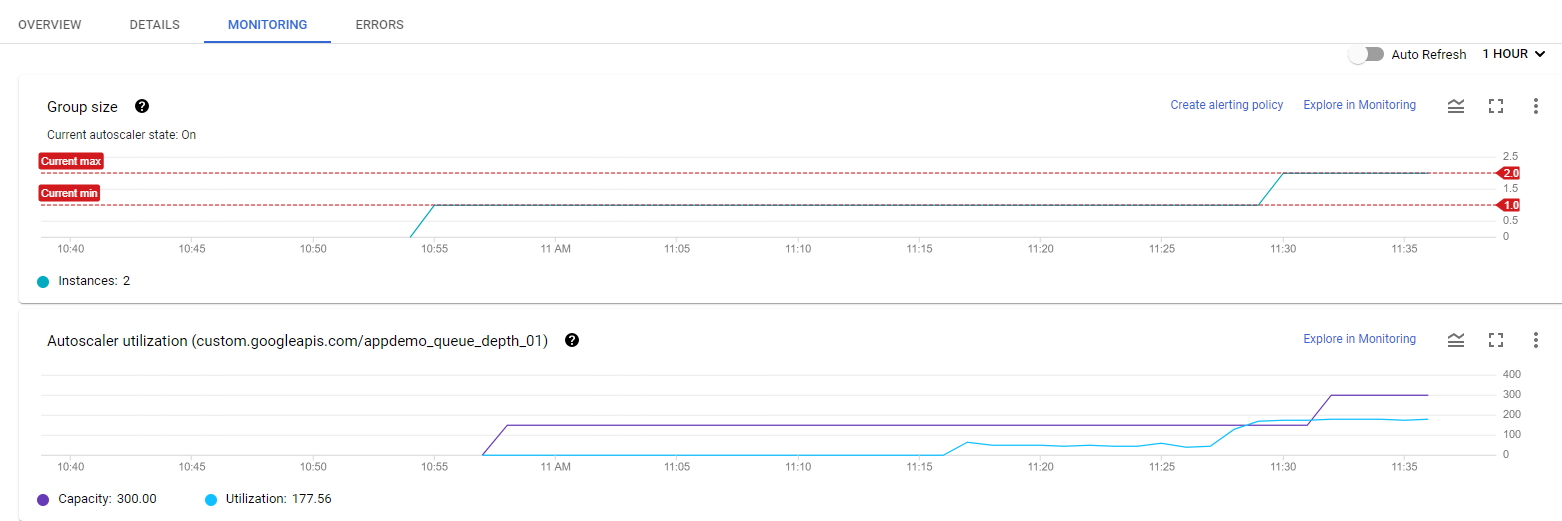
Descripción generada automáticamente

1. Click the **Monitoring** tab.
2. Enable **Auto Refresh**.

Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente

Since this group had a head start, you can see the autoscaling details about the instance group in the autoscaling graph. The autoscaler will take about five minutes to correctly recognize the custom metric and it can take up to 10-15 minutes for the script to generate sufficient data to trigger the autoscaling behavior.



Hover your mouse over the graphs to see more details.

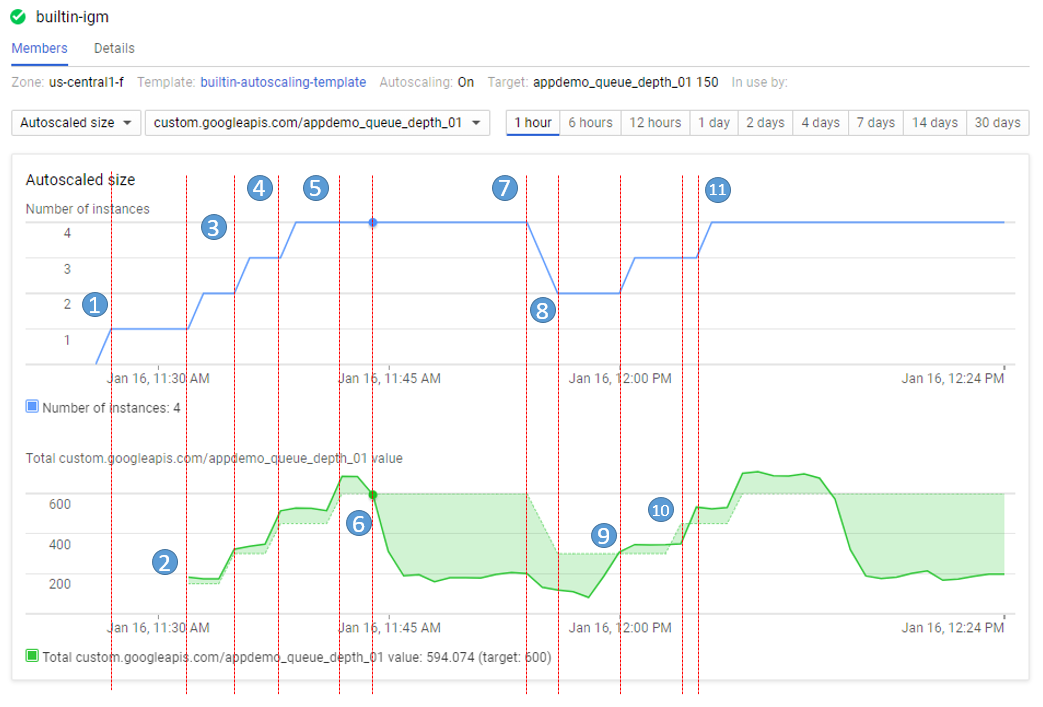
You can switch back to the instance group that you created to see how it's doing (there may not be enough time left in the lab to see any autoscaling on your instance group).

For the remainder of the time in your lab, you can watch the autoscaling graph move up and down as instances are added and removed.

**Task 9. Autoscaling example**

Read through this autoscaling example to see how capacity and number of autoscaled instances can work in a larger environment.

The number of instances depicted in the top graph changes as a result of the varying aggregate level of the custom metric property values reported in the lower graph. There is a slight delay of up to five minutes after each instance starts up before that instance begins to report its custom metric values. While your autoscaling starts up, read through this graph to understand what will be happening:



The script starts by generating high values for approximately 15 minutes in order to trigger scale-up behavior.

* **11:27** Autoscaling Group starts with a single instance. The aggregate custom metric target is 150.
* **11:31** Initial metric data acquired. As the metric is greater than the target of 150 the autoscaling group starts a second instance.
* **11:33** Custom metric data from the second instance starts to be acquired. The aggregate target is now 300. As the metric value is above 300 the autoscaling group starts the third instance.
* **11:37** Custom metric data from the third instance starts to be acquired. The aggregate target is now 450. As the cumulative metric value is above 450 the autoscaling group starts the fourth instance.
* **11:42** Custom metric data from the fourth instance starts to be acquired. The aggregate target is now 600. The cumulative metric value is now above the new target level of 600 but since the autoscaling group size limit has been reached no additional scale-up actions occur.
* **11:44** The application script has moved into a low metric 15 minute period. Even though the cumulative metric value is below the target of 600 scale-down must wait for a ten minute built-in scale-down delay to pass before making any changes.
* **11:54** Custom metric data has now been below the aggregate target level of 600 for a four node cluster for over 10 minutes. Scale-down down now removes two instances in quick succession.
* **11:56** Custom metric data from the removed nodes is eliminated from the autoscaling calculation and the aggregate target is reduced to 300.
* **12:00** The application script has moved back into a high metric 15 minute period. The cumulative custom metric value has risen above the aggregate target level of 300 again so the autoscaling group starts a third instance.
* **12:03** Custom metric data from the new instance have been acquired but the cumulative values reported remain below the target of 450 so autoscaling makes no changes.
* **12:04** Cumulative custom metric values rise above the target of 450 so autoscaling starts the fourth instance.

**Comprobación más abajo**

Interfaz de usuario gráfica

Descripción generada automáticamente