Administering Clusters on Google Kubernetes Engine (GKE)

Integrating Cisco Container Platform with Google Kubernetes Engine (GKE), allows you to deploy and run containerized applications on Google Cloud Platform (GCP).

This chapter contains the following topics:

Prerequisites for Configuring Clusters on GKE

The prerequisites for configuring clusters on GKE are as follows:

See also Adding Google Kubernetes Engine Provider Profile.

Creating New Project on GCP

You need to create a new project on Google Cloud Platform (GCP) to use the GKE services.

Step 1   Go to the New Project page on GCP.

Step 2   In the **Project name** field, enter a name for your project, and then click **CREATE**.

For more information, see Creating Your Project.

Creating Service Account

A service account represents a Google Cloud service identity. You need to create a service account to interact with the google APIs.

Administering Clusters on Google Kubernetes Engine (GKE)

Step 1   Open the GCP console:

https://console.cloud.google

Step 2   In the left pane, click **IAM & Admin** > **Service Accounts**.

The **Service Accounts** page appears on the GCP console.

Step 3   Click **SELECT PROJECT**, choose your project, and then click **OPEN**.

Step 4   Click **CREATE SERVICE ACCOUNT**.

The **Create service account** page appears.

Step 5   In the **Service account details** area, enter a service account name, a description, and then click **CREATE**.

Step 6   In the **Service account permissions** page, add the following roles:

- Compute Viewer

- Kubernetes Engine Admin

- Service Account User

- Viewer

Step 7   Click **CONTINUE**.

Step 8   Grant users access to this service account and click **DONE**.

For more information, see Creating and managing service accounts.

Creating User Credentials on GCP

Step 1   Open the GCP console:

https://console.cloud.google.com

Step 2   In the left pane, click **IAM & Admin** > **Service Accounts**.

The **Service accounts** page appears on the GCP console.

Step 3   In the left pane, click **Service accounts**.

Step 4   In the row of the service account for which you want to create a key, click **Actions** button, and then click **Create key**.

Step 5   Select a **JSON** key type and click **CREATE**.

The credentials are saved to your computer in the `credentials.json` file. You need to copy and paste the contents of this file when adding a GKE provider. For more information, see Adding Google Kubernetes Engine Provider Profile.

For more information on creating credentials in GCP, see Creating and managing service account keys.

Enabling GCP APIs

You need to enable the GCP APIs to allow querying and responding to the Google APIs.

Step 1   Open the GCP console:

         https://console.cloud.google

Step 2   In the left pane, click **APIs & Services** > **Dashboard**.
The **APIs & Services** page appears on the GCP console.

Step 3   Click **ENABLE APIS AND SERVICES**.

Step 4   In the search box, enter **Kubernetes Engine API**, and then click **ENABLE**.

Step 5   In the search box, enter **Cloud Resource Manager API**, and then click **ENABLE**.

Configuring Control Plane Proxy for GKE Access

If your Control Plane VMs need proxy configuration to access the internet, specifically GKE API endpoints, you must configure the proxy information on Cisco Container Platform.

Step 1   SSH to the master node of the control plane.

Step 2   Run the following commands to specify the proxy information:

     Note   You need to replace *<Proxy_URL_or_IPAddress:Port>* with the URL/IP address of your proxy server and the no_proxy list with a list of your internal IP addresses.

```
kubectl patch deploy kaas-api
```

```
kubectl patch deploy kaas-ccp-gke-operator --patch
```

Creating Clusters on GKE

Before you begin

Ensure that you have completed the prerequisites for configuring clusters on GKE. For more information, see Prerequisites for Configuring Clusters on GKE, on page 1.

Step 1   In the left pane, click **Clusters**, and then click the **GKE** tab.

Step 2   Click **NEW CLUSTER**.

Step 3   In the **Basic Information** screen, specify the following information:

    a)   From the **INFRASTRUCTURE PROVIDER** drop-down list, choose the provider related to the appropriate GKE account.

    b)   In the **KUBERNETES CLUSTER NAME** field, enter a name for your cluster.

    c)   From the **LOCATION** drop-down list, choose a GKE region.

        Note   For more information on the supported regions, see Cloud Locations.

d) The **KUBERNETES VERSION** drop-down list, choose the Kubernetes version for your cluster.

e) Click **NEXT**.

Step 4 In the **Node Configuration** screen, specify the following information:

a) In the **WORKER NODES** field, enter an appropriate number of worker nodes.

b) Use the **AUTOSCALING** toggle button to enable autoscaling, and then set the minimum and maximum number of worker nodes.

c) In the **NODE POOL NAME** field, enter an alphanumeric name for the primary node pool of your cluster.

d) From the **MACHINE TYPE** drop-down list, choose an appropriate machine type for your VMs.

For more information, see Machine types.

e) From the **IMAGE TYPE** drop-down list, choose an appropriate image type for your VMs.

f) Use the **USE PREEMPTIBLE VMS** toggle button to enable or disable preemptible VMs.

For more information, see Running preemptible VMs.

g) Click **NEXT**.

Step 5 In the **Summary** screen, review the cluster information, and then click **FINISH**.

Cluster creation can take up to 20 minutes. You can monitor cluster creation status on the **Clusters** screen.

Scaling Clusters on GKE

Cisco Container Platform supports both manual scaling and autoscaling of GKE cluster node pools. Autoscaling is enabled or disabled during cluster creation. If autoscaling is disabled, you can manually scale the node pool by modifying the number of worker nodes. Whereas, if autoscaling is enabled, in addition to modifying the number of worker nodes, you can modify the minimum and maximum limits of a node pool range.

Follow these steps to scale the node pool in the cluster:

Before you begin

Ensure that the cluster is in READY state.

Step 1 Click on the GKE cluster that you want to scale.

The **Cluster Details** screen appears displaying the node pools of the cluster that you have chosen.

Step 2 From the drop-down list next to the name of the node pool, click **Edit**.

The **Edit node pool** screen appears.

Step 3 If autoscaling is disabled, increase or decrease the number of worker nodes.

Step 4 If autoscaling is enabled, you can perform these actions:

• Increase or decrease the number of worker nodes.

• Modify the node pool range, by changing the minimum and maximum number of nodes configured for the node pool.

Step 5 Click **UPDATE**.

In the **Cluster Details** screen, the status of the cluster is indicated as **OPERATION_IN_PROGRESS**.

After the scaling operation is complete, the status of the cluster is changed to **READY** and the updated number of worker nodes is displayed in the **Cluster Details** page.

Deleting Clusters on GKE

Before you begin

Ensure that the GKE cluster that you want to delete is not currently in use, as deleting a cluster removes the containers and data associated with it.

Step 1   In the left pane, click **Clusters**, and then click the **GKE** tab.

Step 2   From the drop-down list displayed under the **ACTIONS** column, choose **Delete** for the cluster that you want to delete.

Step 3   Click **DELETE** in the confirmation dialog box.

Upon deleting a GKE cluster, it takes about 15 minutes for the cluster resources to be released.