

Kubernetes services are replicated. At least one of the Kubernetes services has `replicas:` defined with a value greater than 1 in its `deployment.yml` file.

Screenshot of Kubernetes cluster of command `kubectl describe hpa` has autoscaling configured with CPU metrics.



Kubernetes services are replicated. At least one of the Kubernetes services has `replicas:` defined with a value greater than 1 in its `deployment.yml` file.

Screenshot of Kubernetes cluster of command `kubectl describe hpa` has autoscaling configured with CPU metrics.

✖ Attention!!! [Requires Changes]:

- Kubernetes service has replicas ✔ Please note that replicas will accrue the most cost in AWS so we don't need to enforce every single service to have multiple replicas.
- `kubectl describe hpa` has autoscaling configured with CPU metrics ✖ Good job while including the screenshot of `kubectl describe hpa` command however, the HPA does not work correctly currently. Pay attention to the reason for the error described in your screenshot which is **Failed to get resource metric**. Please note that we are expecting some numeric value under metric instead of unknown (which means currently it's unable to fetch the metric count).

Check out the red highlights in your screenshots, 📌

```
Lenovo@LAPTOP-IG6VBS7F MINGW64 /d/HDC-TAP/Udacity/Project-Submit/Project-3/AWS-Udacity-Project-3 (main)
$ kubectl describe hpa
Name:                               api-feed
Namespace:                          default
Labels:                             <none>
Annotations:                        <none>
CreationTimestamp:                  Sat, 28 Oct 2023 01:48:00 +0700
Reference:                          Deployment/api-feed
Metrics:                            ( current / target )
  resource cpu on pods  (as a percentage of request): <unknown> / 70%
Min replicas:                       3
Max replicas:                       5
Deployment pods:                     0 current / 0 desired
Events:                             <none>

Name:                               api-user
Namespace:                          default
Labels:                             <none>
Annotations:                        <none>
CreationTimestamp:                  Sat, 28 Oct 2023 01:46:30 +0700
Reference:                          Deployment/api-user
Metrics:                            ( current / target )
  resource cpu on pods  (as a percentage of request): <unknown> / 70%
Min replicas:                       3
Max replicas:                       5
Deployment pods:                     3 current / 3 desired
Conditions:
  Type            Status  Reason
  ----            -
  AbleToScale     True    SucceededGetScale
  ScalingActive   False   FailedGetResourceMetric the HPA controller was able to get the target's current scale
                                the HPA was unable to compute the replica count: failed to get cpu utilization: unable to get metrics for resource cpu: unable to fetch metrics from resource metrics API: the server could not find the requested resource (get pods.metrics.k8s.io)
Events:
  Type            Reason              Age           From              Message
  ----            -
  Normal          SuccessfulRescale    86s          horizontal-pod-autoscaler    New size: 3; reason: Current number of replicas below Spec.MinReplicas
```

=> I have fixed it according to your request. Refer to the image below.

```
PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS CODEWHISPERER REFERENCE LOG
Lenovo@LAPTOP-IG6VBSTF MINGW64 /d/HOC-TAP/Udacity/Project-Submit/Project-3/AWS-Udacity-Project-3 (main)
$ kubectl describe hpa
Name: api-feed
Namespace: default
Labels: <none>
Annotations: <none>
CreationTimestamp: Sat, 28 Oct 2023 01:48:00 +0700
Reference: Deployment/api-feed
Metrics: ( current / target )
  resource cpu on pods (as a percentage of request): 0% (0) / 70%
Min replicas: 3
Max replicas: 5
Deployment pods: 3 current / 3 desired
Conditions:
  Type           Status Reason          Message
  ----           -
  AbleToScale    True  ReadyForNewScale recommended size matches current size
  ScalingActive  True  ValidMetricFound the HPA was able to successfully calculate a replica count from cpu resource utilization (percentage of request)
  ScalingLimited True  TooFewReplicas   the desired replica count is less than the minimum replica count
Events:
  Type           Reason          Age           From           Message
  ----           -
  Warning        FailedGetResourceMetric 5m20s (x2099 over 8h) horizontal-pod-autoscaler failed to get cpu utilization: unable to get metrics for resource cpu: unable to fetch metrics from resource metrics API: the server could not find the requested resource (get pods.metrics.k8s.io)

Name: api-user
Namespace: default
Labels: <none>
Annotations: <none>
CreationTimestamp: Sat, 28 Oct 2023 01:46:30 +0700
Reference: Deployment/api-user
Metrics: ( current / target )
  resource cpu on pods (as a percentage of request): 0% (0) / 70%
Min replicas: 3
Max replicas: 5
Deployment pods: 3 current / 3 desired
Conditions:
  Type           Status Reason          Message
  ----           -
  AbleToScale    True  ReadyForNewScale recommended size matches current size
  ScalingActive  True  ValidMetricFound the HPA was able to successfully calculate a replica count from cpu resource utilization (percentage of request)
  ScalingLimited True  TooFewReplicas   the desired replica count is less than the minimum replica count
Events:
  Type           Reason          Age           From           Message
  ----           -
  Warning        FailedGetResourceMetric 6m50s (x2099 over 8h) horizontal-pod-autoscaler failed to get cpu utilization: unable to get metrics for resource cpu: unable to fetch metrics from resource metrics API: the server could not find the requested resource (get pods.metrics.k8s.io)

Lenovo@LAPTOP-IG6VBSTF MINGW64 /d/HOC-TAP/Udacity/Project-Submit/Project-3/AWS-Udacity-Project-3 (main)
$
```