

# The Complete HorizontalPodAutoscaler Flow of Events

In this lesson, we will describe the complete flow of HPA events.

## WE'LL COVER THE FOLLOWING ^

- Complete **HPA** flow of events
- **HPA** manipulation

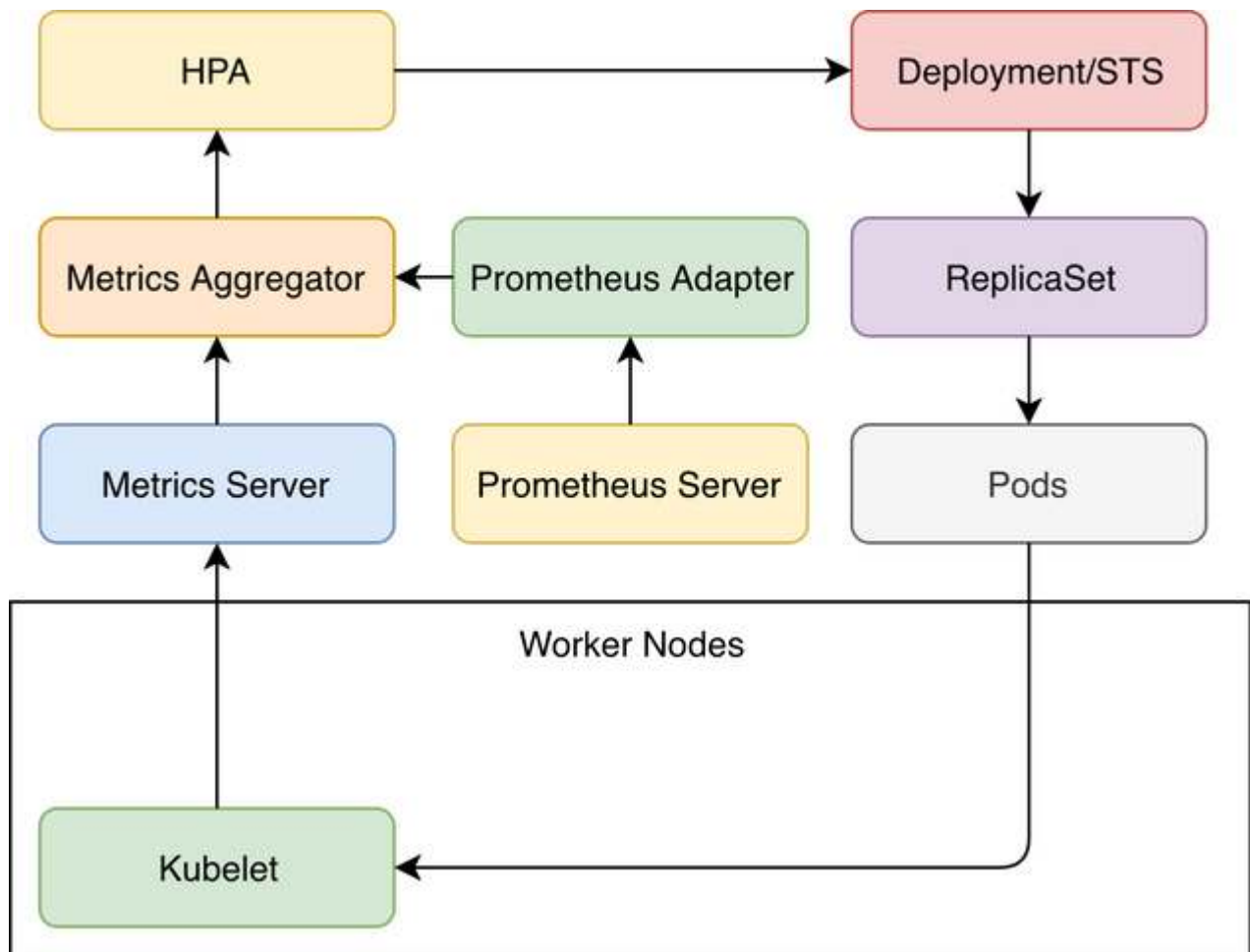
## Complete **HPA** flow of events #

**Metrics Server** is fetching memory and CPU data from Kubelets running on the worker nodes. In parallel, **Prometheus Adapter** is fetching data from **Prometheus Server** which, as you already know, pulls data from different sources. Data from both **Metrics Server** and **Prometheus Adapter** is combined in *Metrics Aggregator*.

🔍 **HPA** is periodically evaluating metrics defined as scaling criteria. It's fetching data from **Metrics Aggregator**, and it does not really care whether they're coming from **Metrics Server**, **Prometheus Adapter**, or any other tool we could have used.

## **HPA** manipulation #

Once scaling criteria are met, **HPA** manipulates Deployments and StatefulSets by changing their number of replicas. As a result, rolling updates are performed by creating and updating ReplicaSets which, in turn, create or remove Pods.



HPA using a combination of metrics from Metrics Server and those provided by Prometheus Adapter (arrows show the flow of data)



Data from both **Metrics Server** and **Prometheus Adapter** is combined in *Metrics Aggregator*.

COMPLETED 0%

1 of 1



In the next lesson, we will discuss reaching Nirvana.

