Today's agenda

01 Introduction

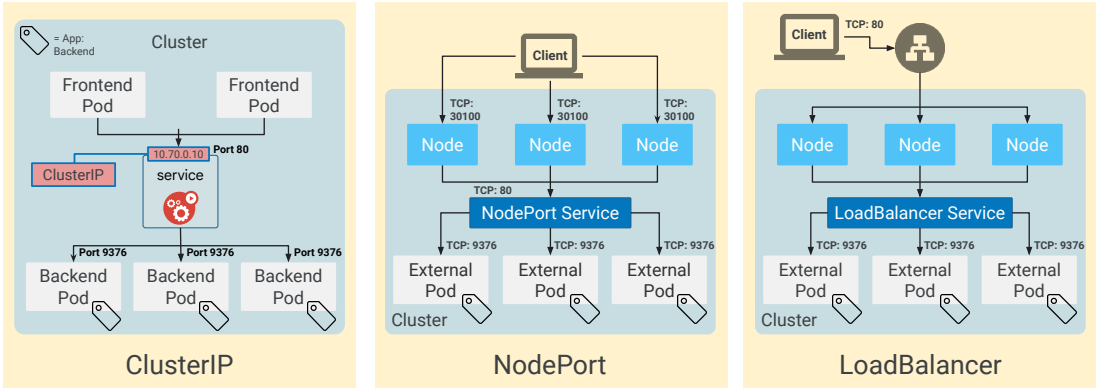
02 Clusters and services

03 Lab: Creating Services and Ingress Resources

04 Pod DNS

05 Quiz

Review: service types

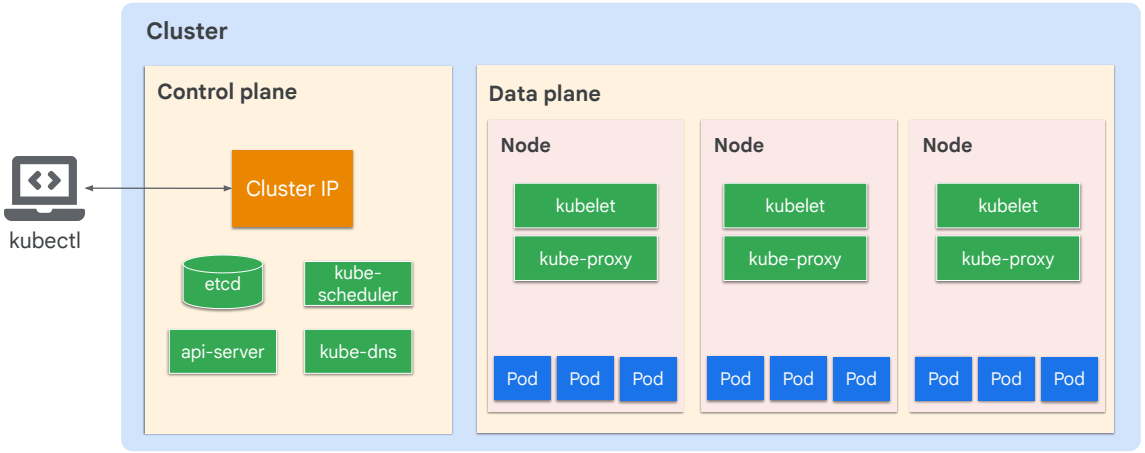


Google Cloud

In prerequisite materials to this course, you learned about ClusterIP, NodePort, and LoadBalancer services. These service types are also covered in detail in the *Architecting with Google Kubernetes Engine* course.

This module covers the Ingress and Gateway service types.

Review: GKE cluster components



Google Cloud

The information on this slide is presented in more detail in *Architecting with Google Kubernetes Engine*, which is a prerequisite for this course.

Control plane: The control plane is responsible for managing the cluster, including scheduling pods (kube-scheduler), managing the cluster state (etcd), and

Data plane: The data plane is responsible for running the pods which make up a containerized workload. kubelet runs on each node in the data plane; it helps create and remove pods. Kube-proxy is a network proxy that maintains network rules on the nodes.

Nodes: Nodes are the worker VMs in a GKE cluster. They run the Kubernetes software and host the pods in the cluster. Each node is managed by a kubelet, an agent that receives commands from the control plane. kube-proxy maintains network rules on the node.

Pods: Pods are the basic unit of deployment in GKE. They are a group of one or more containers that are scheduled together on a node. (The containers themselves are not shown here.)

Developers can interact with cluster at the command line using kubectl.

Making microservices available

- Sasha, a cloud architect, must make services in Cymbal's core banking application available to the applications run by its subsidiaries.
- This application is deployed on Google Kubernetes Engine (GKE).
- Cymbal plans to release some additional applications on GKE for its global customers.
- How can they make these services accessible internally and externally in a cost-effective manner?