Kubernetes

Overview of Kubernetes

- Kubernetes also called "K8s" is an open-source system for automating deployment, management, and scaling of containerized applications
- It was designed by Google and donated to the Cloud Native Computing Foundation
- Kubernetes provides a platform for automating deployment, scaling, and operations of application containers
- It supports a range of container tools, including Docker
- Was founded by Joe Beda, Brendan Burns, and Craig McLuckie and was quickly joined by other Google engineers including Brian Grant and Tim Hockin and was first announced by Google in mid-2014

Kubernetes Design

• Kubernetes has a set of building blocks that come together to provide mechanisms for deploying, maintaining, and scaling applications

• Kubernetes components are designed to be loosely coupled and extensible to meet a wide variety of workloads

• Kubernetes APIs which provide extensibility, extensions, and containers running on Kubernetes are:

- Pods
- Labels and Selectors
- Controllers
- Services

Kubernetes Design - Pods

- Basic scheduling unit in Kubernetes is called a "pod"
- Pod consists of one or more containers that are to be co -located on the host machine and share resources
- Each pod in Kubernetes has a unique IP address that allows applications to use ports without conflict
- Pod can define a volume, like a local disk directory or a network disk and expose it to the containers in the pod
- Pods are manually managed by Kubernetes API
- Pods management can be delegated to a controller

Kubernetes Design—Labels and Selectors:

- Kubernetes allows clients to attach key-value pairs called "labels" to any API objects like pods and nodes
- Labels and selectors are the primary grouping mechanism in Kubernetes
- Labels and selectors are used to determine the components to which an operation applies

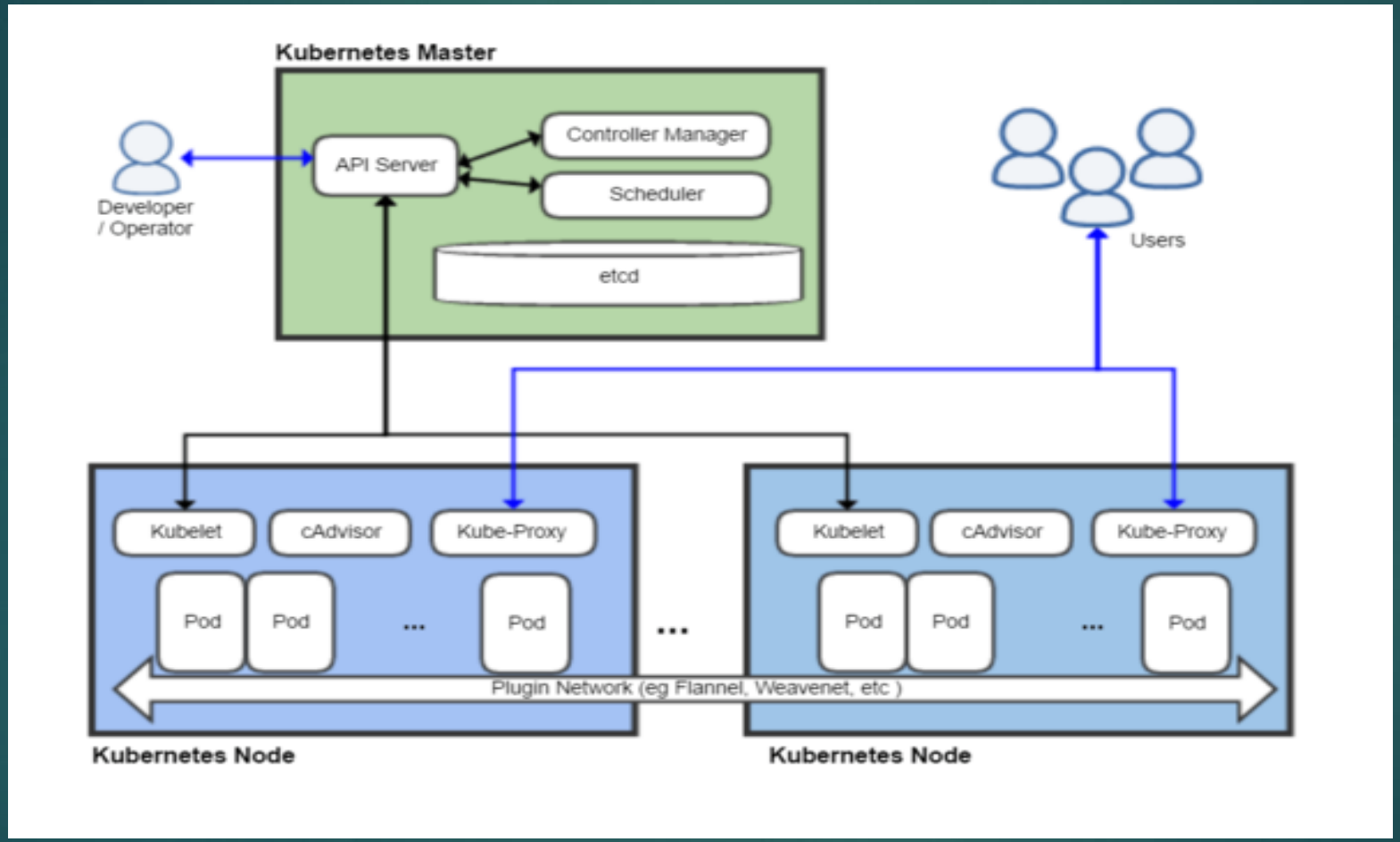
Kubernetes Design—Controllers:

- Controller is a resolution loop that brings actual cluster state toward the desired cluster state by managing a set of pods
- Replication Controller handles replication and scaling by running a number of copies of a pod
- DaemonSet Controller runs one pod on every machine
- Job Controller runs pods as part of a batch job
- Label selector defines the set of pods that a controller manages and are part of controller's definition

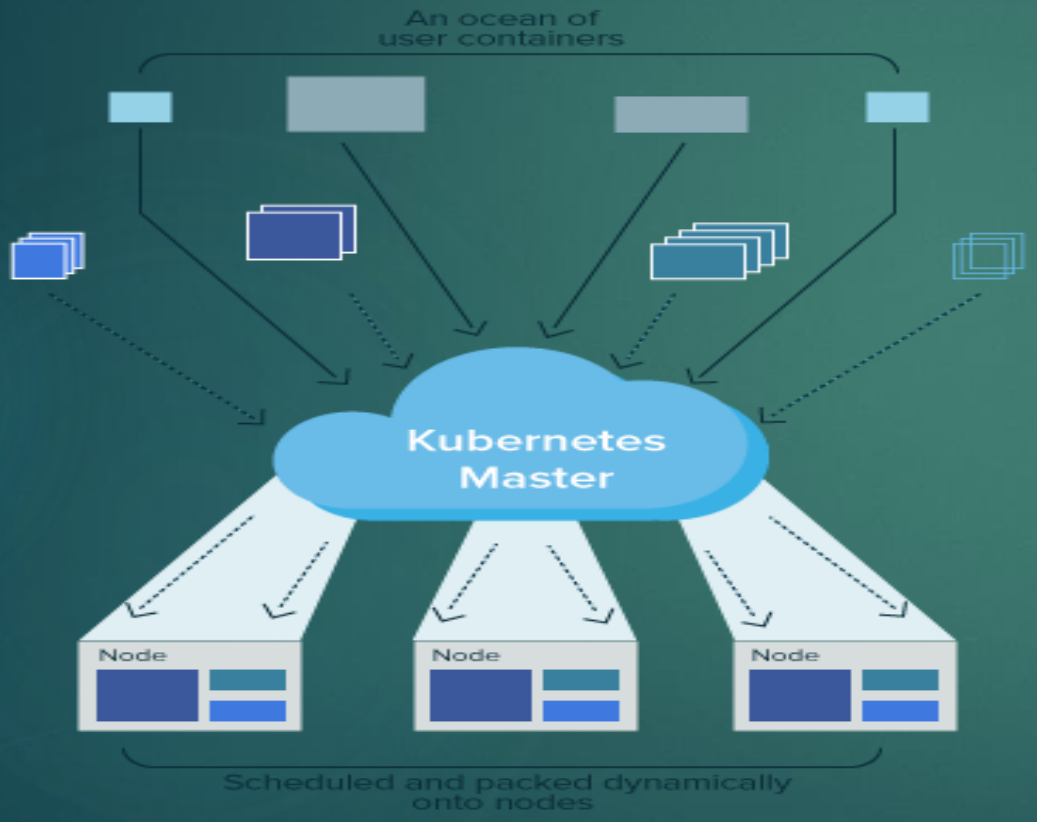
Kubernetes Design—Services:

- Kubernetes service is a set of pods that work together like a single tier of a multi-tier software application
- Label selector defines set of pods that constitute a service
- Kubernetes assigns a stable IP address and DNS name to the service, thus enabling service discovery and request routing
- Kubernetes load balances traffic in a round-robin manner to network connections of that IP address among the pods
- By default, a service gets exposed inside a cluster, and it is possible to expose service outside a cluster

Kubernetes Architecture:



Kubernetes …



- Kubernetes has a master component called as ‘Kubernetes Master’.
- Lot of application containers that await implementation.
- So the users send a request to the Kubernetes Master.
- The Kubernetes Master then allocates these containers to nodes or Operating Systems attached to Kubernetes where these applications can run.