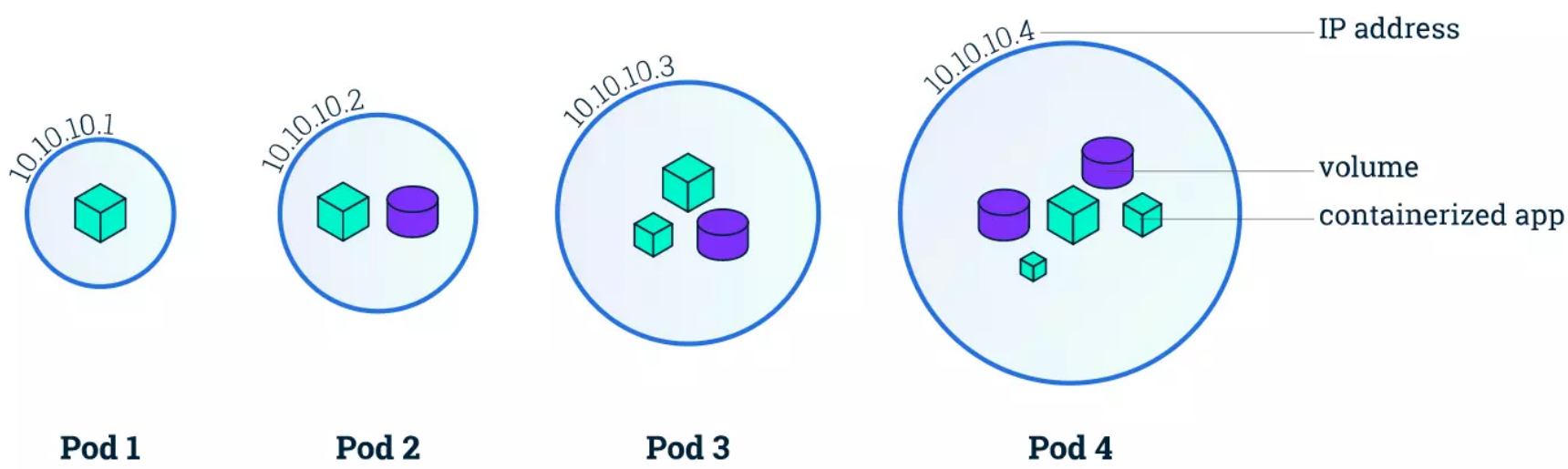
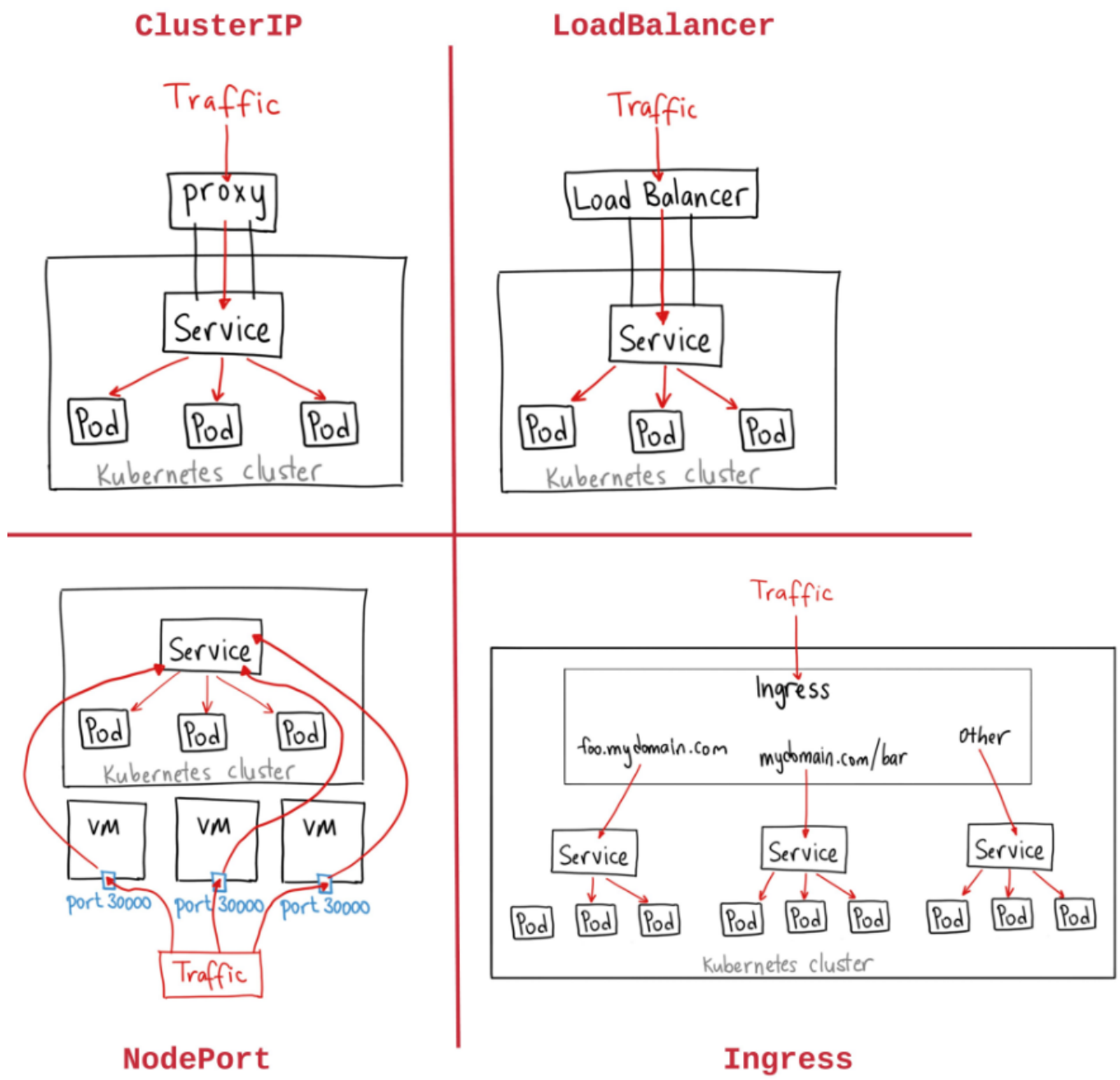
Kubernetes Pod?

A Kubernetes pod is the smallest deployable units in kubernetes. Containers within a pod share the same network namespace, enabling them to communicate easily.
Here's a simple guide to understand how it works:



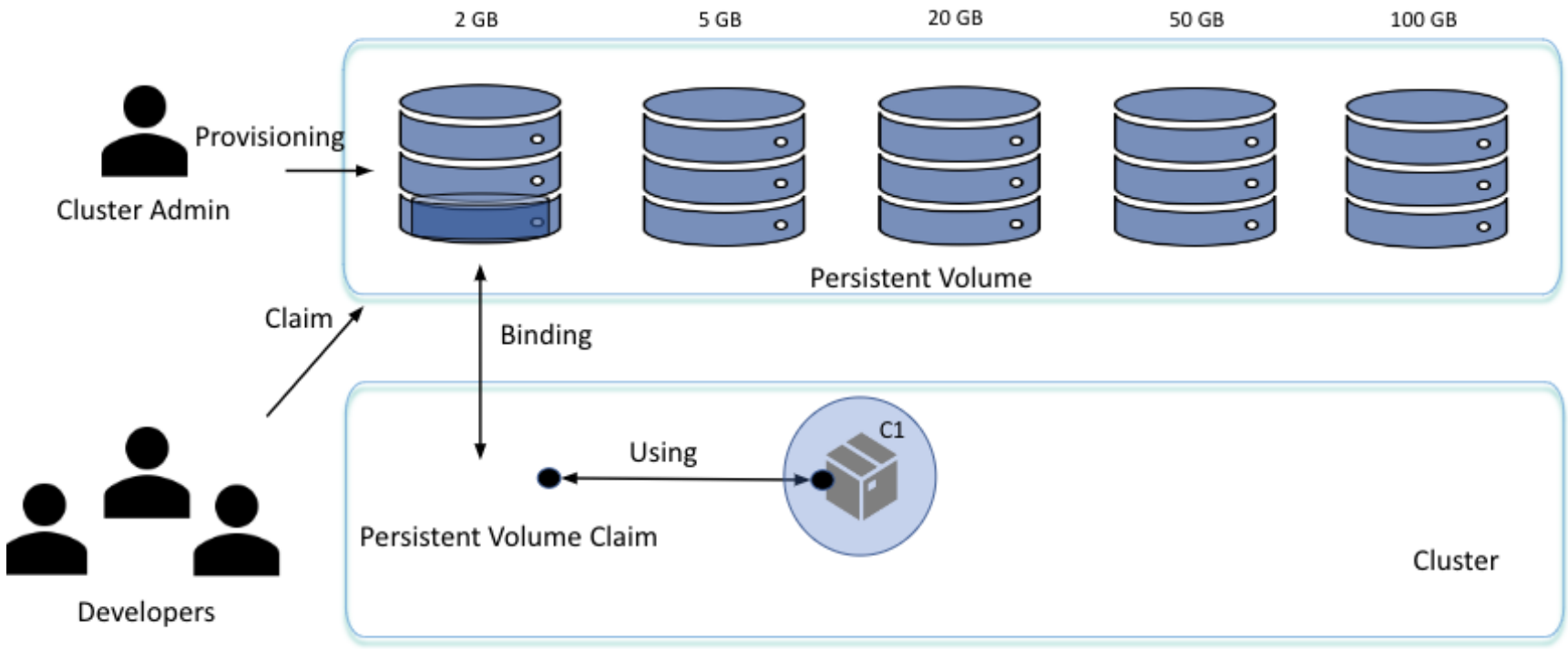
Kubernetes Services?

Coupling of a set of pods to a policy by which to access them. Services are used to expose containerized applications to origins from outside the cluster.



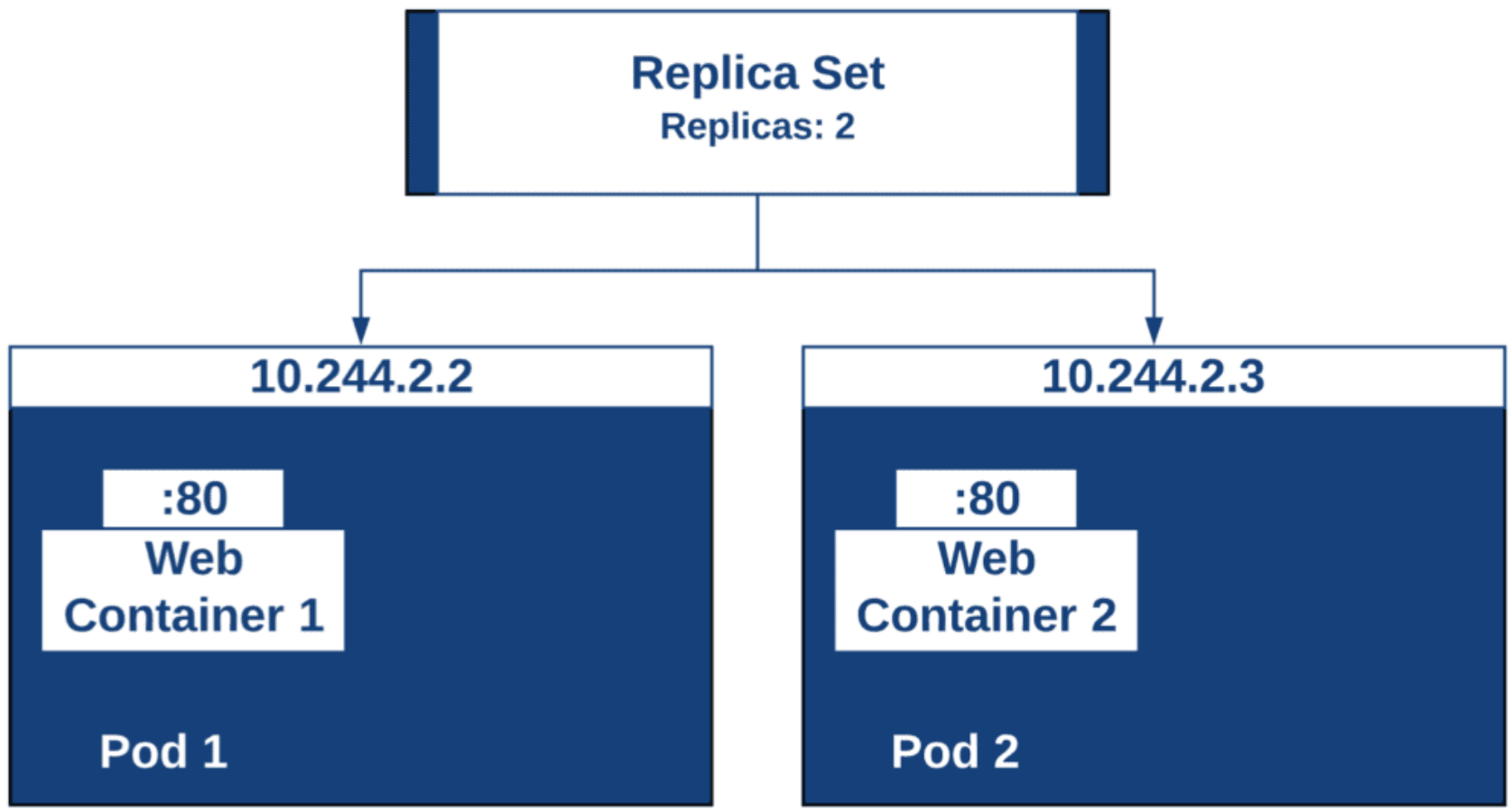
Persistent Volume?

Persistent Volume is a piece of storage within the cluster that has been provisioned by an administrator or dynamically provisioned using Storage Classes.



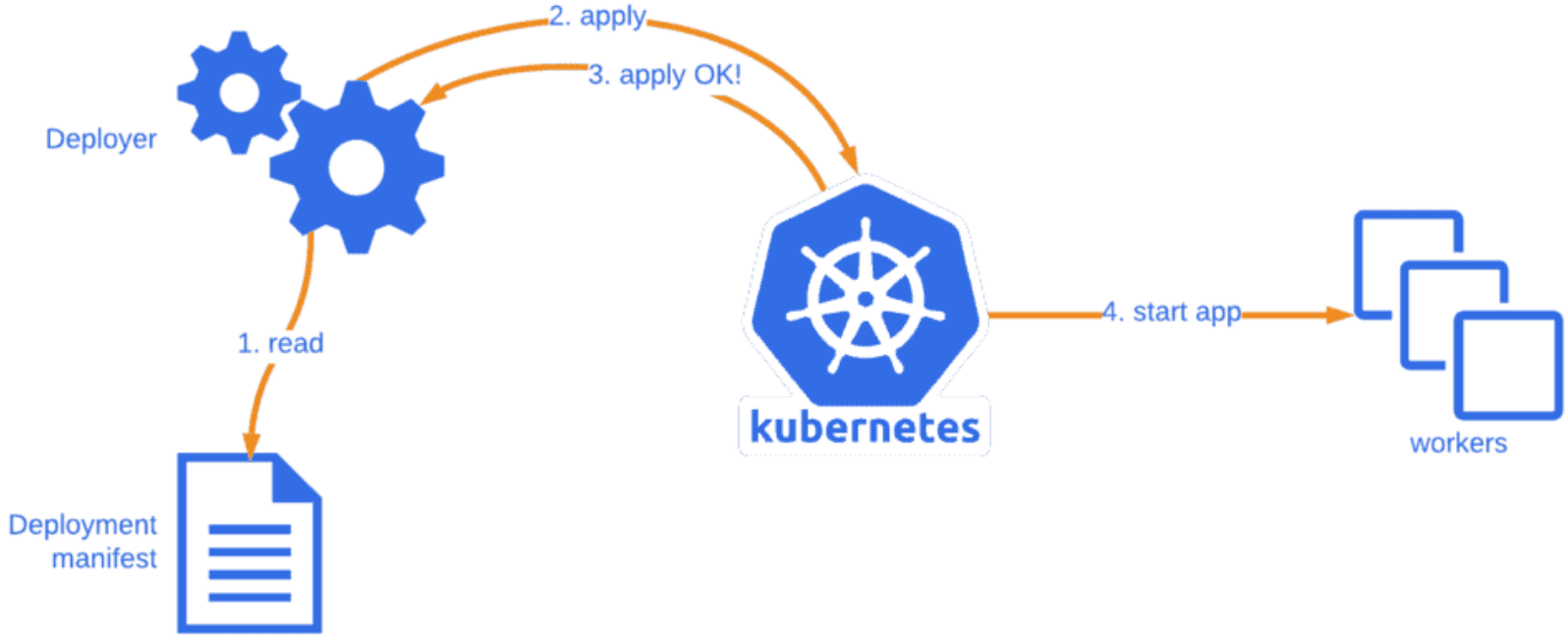
ReplicaSet (RS)?

A Kubernetes replica set ensures that the specified number of pods in a replica set is running at all times.



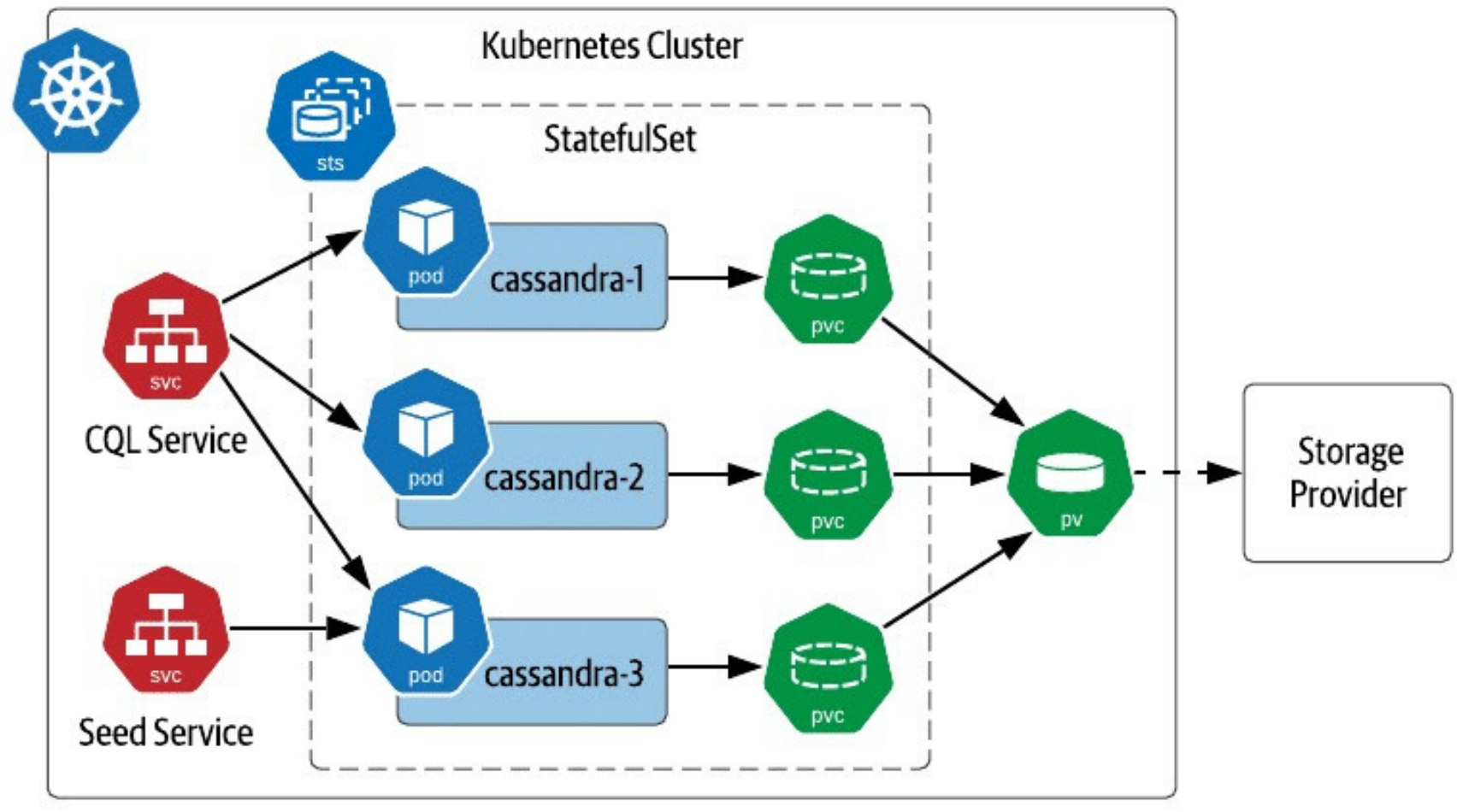
Deployment?

A way to define the desired state of pods or a replica set. Deployments are used to define HA policies to your containers by defining policies around how many of each container must be running at any one time.



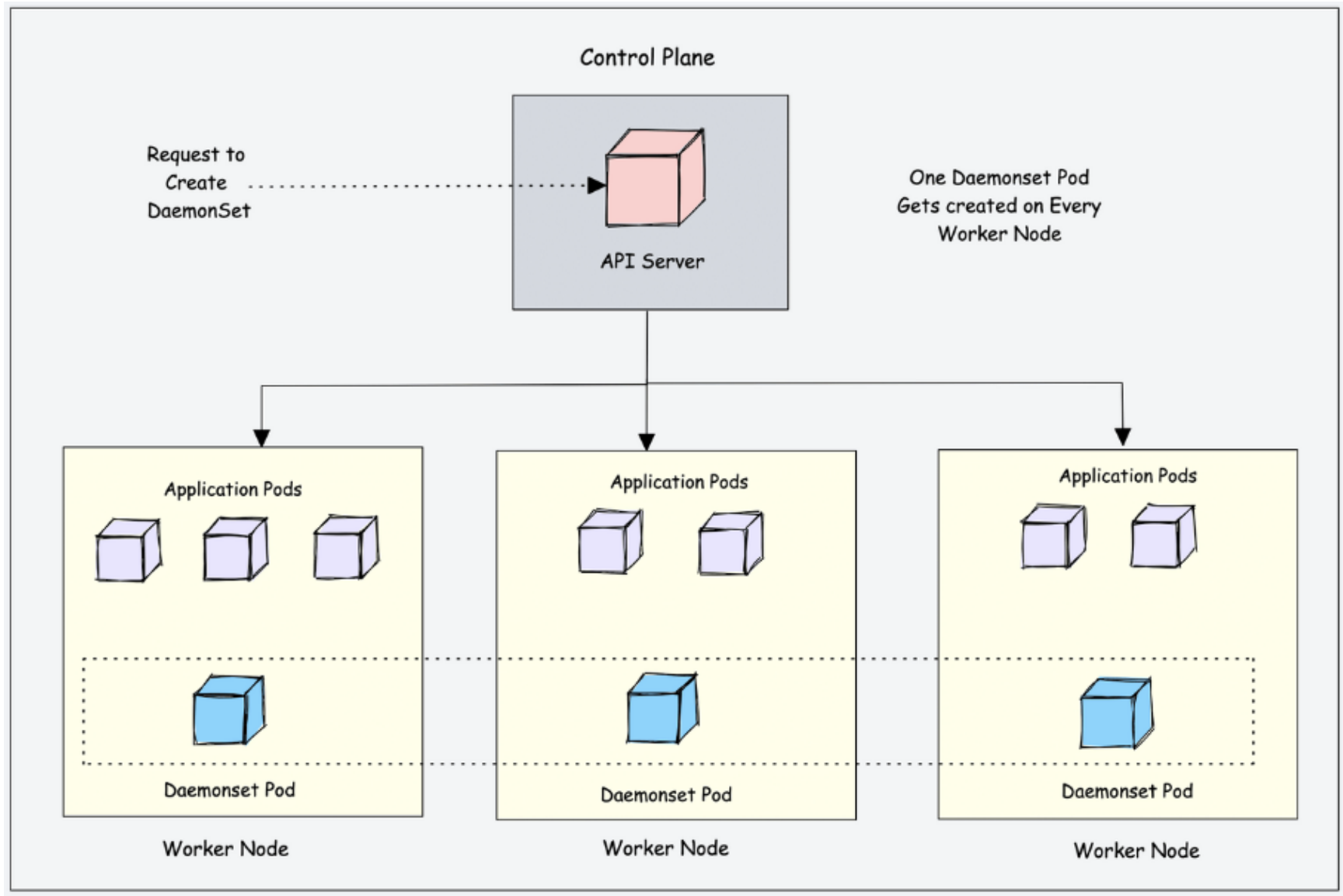
StatefulSet?

A workload API object that manages stateful applications, such as databases.



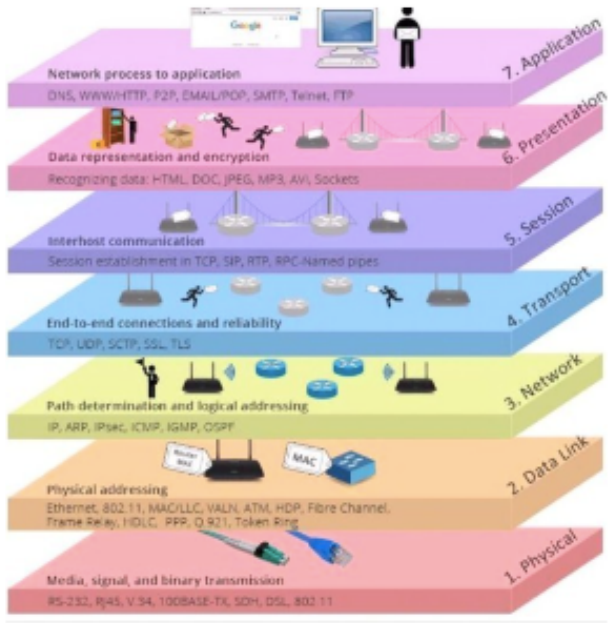
DaemonSet?

Assume there is some functionality for kubernetes platform administrators to run platform-specific applications on all the nodes. For example, running a logging agent on all the Kubernetes nodes. Here is where Daemonset comes into the picture. As Name suggests, Daemonset is designed to run system daemons.



Ingress Controller?

Ingress Controller is an intelligent Load Balancer. Ingress is a high-level abstraction responsible for allowing simple host or URL based HTTP routing. It is always implemented using a third-party proxy.



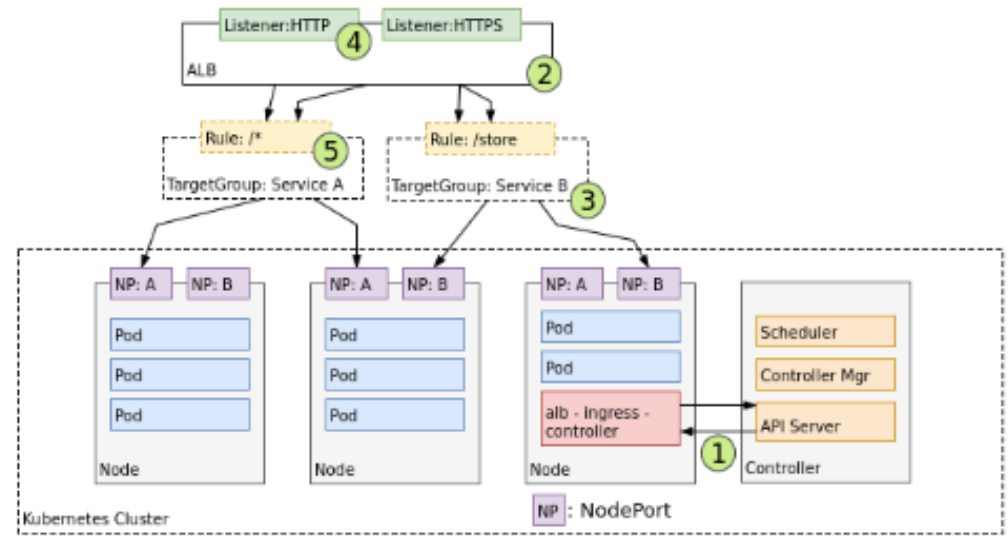
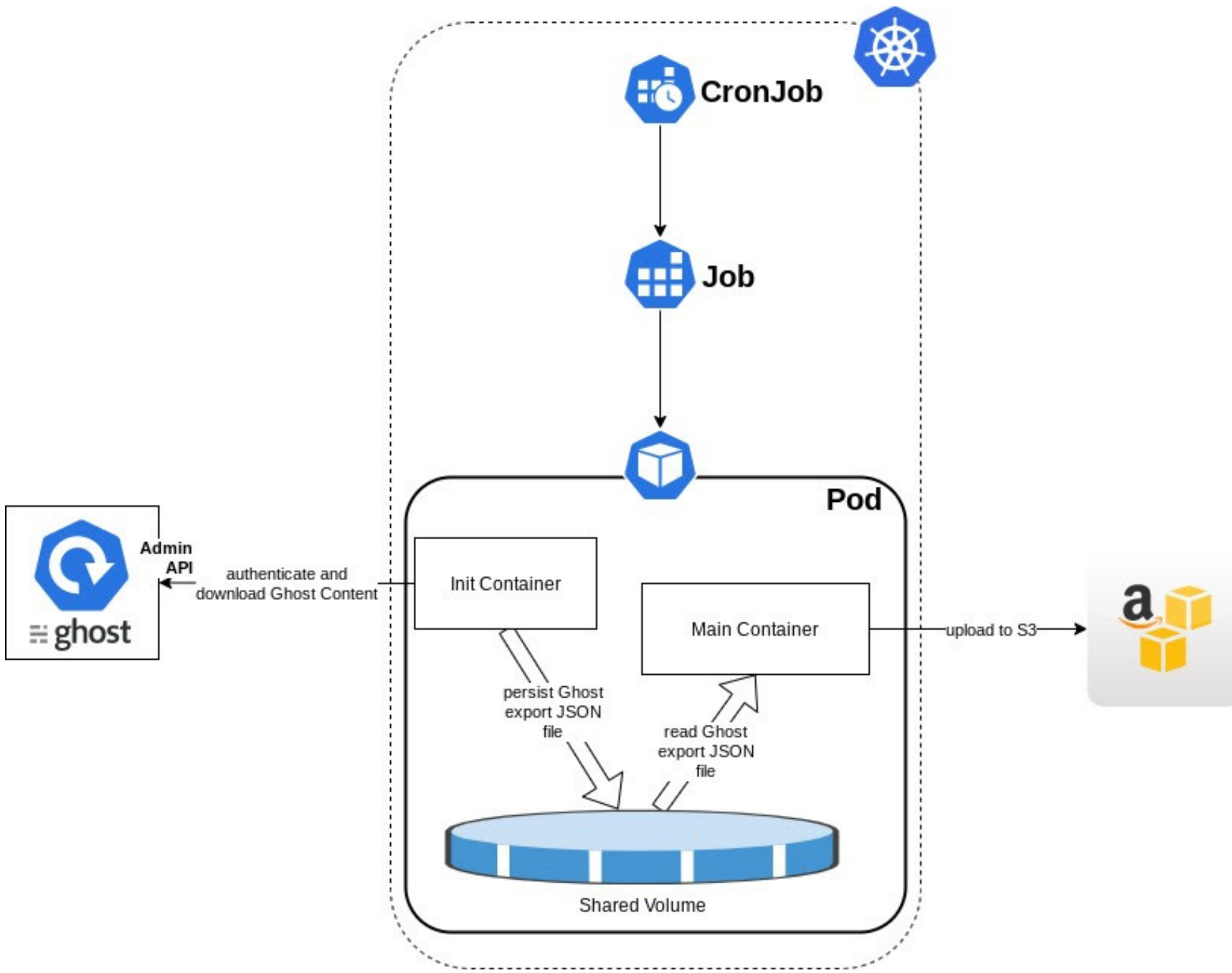


Image Courtesy: Kevin Sookocheff

Kubernetes Jobs?

Kubernetes Jobs provide a way to run short-lived, parallel, or sequential batch tasks within the cluster. Jobs ensure that a specified number of pods successfully complete their tasks before considering the job as finished.



High Availability?

High availability focuses on minimizing downtime by ensuring that applications and services continue to function despite failures at various levels, including nodes, containers, and networking components.