Cluster

A Kubernetes Cluster is a group of connected machines (nodes) that work together to run container-based applications.
It is not a single server. It is a complete environment managed by Kubernetes.

A Kubernetes cluster has two main types of nodes:

1. Master Node (Control Plane)
2. Worker Node

1.Master Node (Control Plane)

The Master Node is the brain of the cluster.
It manages the cluster and takes decisions like where pods should run and how many should run.

API Server

API Server is the main entry point of Kubernetes.
All commands and requests go through the API Server.

- It accepts requests from kubectl or other tools
- It checks permissions and validates requests
- It communicates with etcd to store and read cluster data

etcd

etcd is the database of Kubernetes.
It stores the full cluster state and configuration.

- Stores pod, node, deployment, and service details
- Stores ConfigMaps and Secrets
- It is the main source of truth for the cluster

Controller Manager

Controller Manager ensures the cluster always stays in the desired state.

- If a pod fails, it creates a new one
- It maintains the required number of pods
- This is why Kubernetes is self-healing

Scheduler

Scheduler selects the best worker node for a new pod.

- It checks CPU, RAM, and node availability
- It assigns the pod to the best worker node
- Scheduler only selects the node, it does not run the pod

2.Worker Node

Worker Nodes are the machines where the application actually runs. Pods and containers run only on worker nodes.

Kubelet

Kubelet is the main agent inside every worker node.

- It receives pod instructions from the control plane
- It ensures containers are running properly
- It restarts containers if they fail

Kube-proxy

Kube-proxy manages networking on the worker node.

- It enables pod-to-pod communication
- It routes service traffic to the correct pods

Pod

A Pod is the smallest unit in Kubernetes.
Kubernetes deploys pods, not containers directly.

- A pod usually contains one container
- Containers in a pod share the same IP and network

Container Runtime

Container Runtime is the software that runs containers on the worker node.

- Kubelet uses the runtime to start and stop containers
- Common runtimes are containerd and CRI-O (Docker was used earlier)

CLUSTER WORKFLOW: