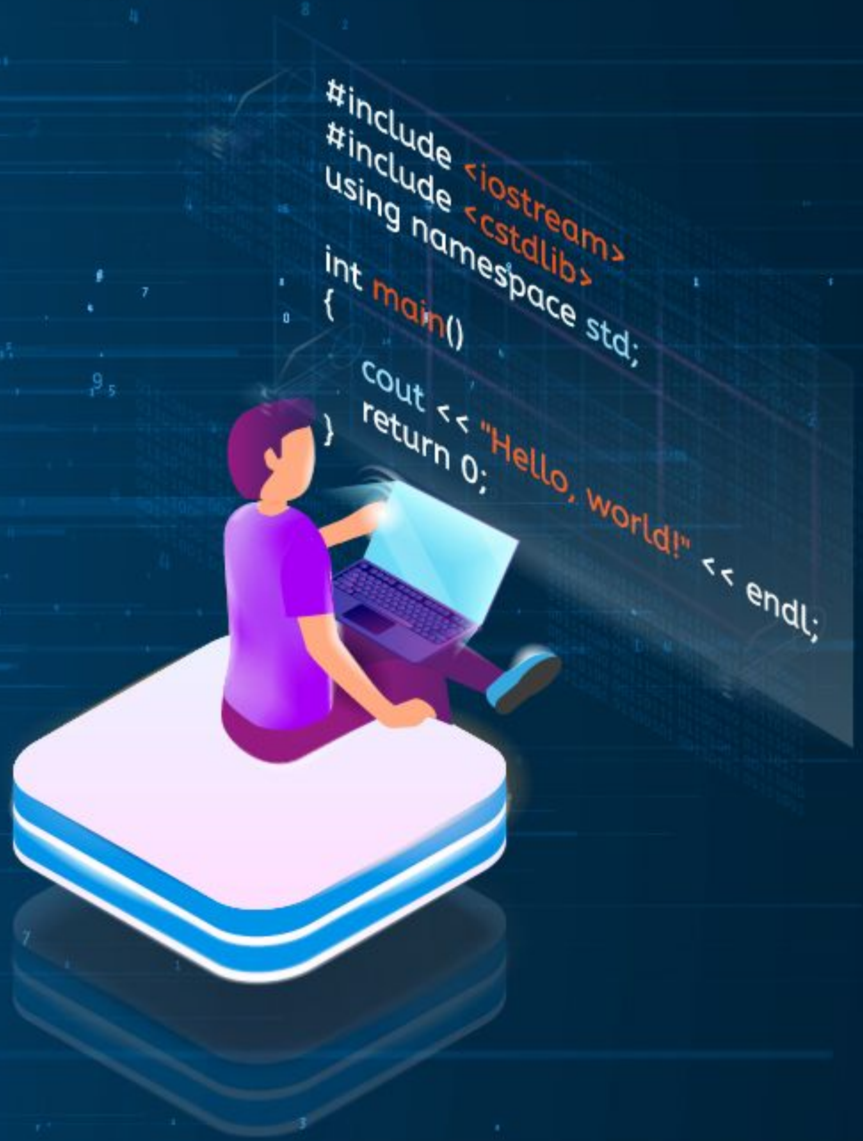
Certified Kubernetes Administrator

TECHNOLOGY

Kubernetes Cluster

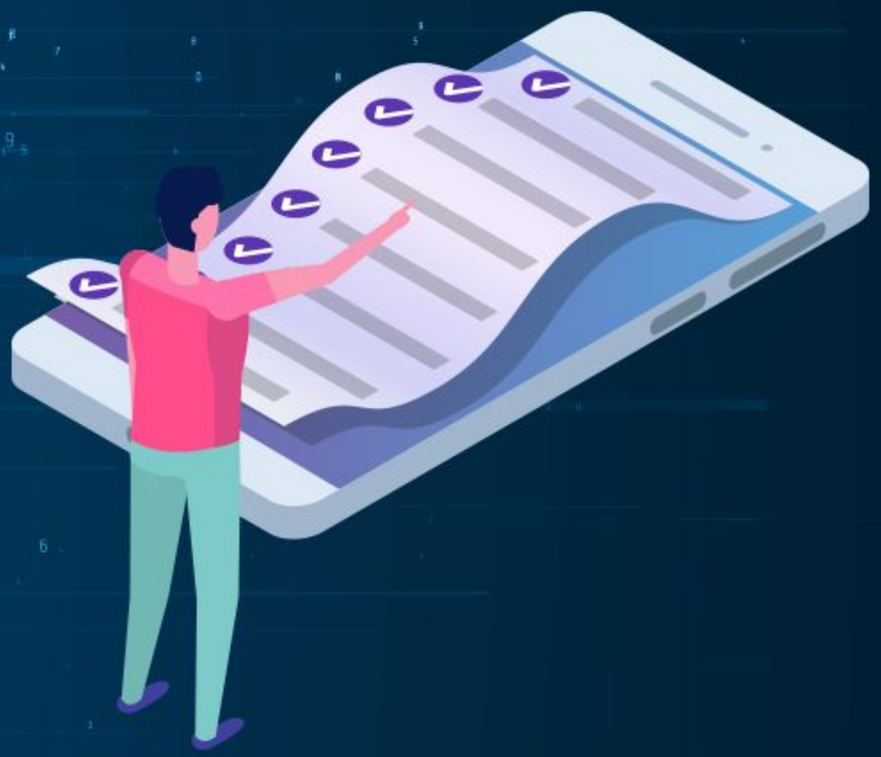


```
#include <iostream>
#include <cstdlib>
using namespace std;
int main()
{
    cout << "Hello, world!" << endl;
    return 0;
}
```

Learning Objectives

By the end of this lesson, you will be able to:

- Discuss Cluster Architecture
- Present an overview of Nodes
- Discuss Controller
- Discuss the working of kubeadm
- Define API server

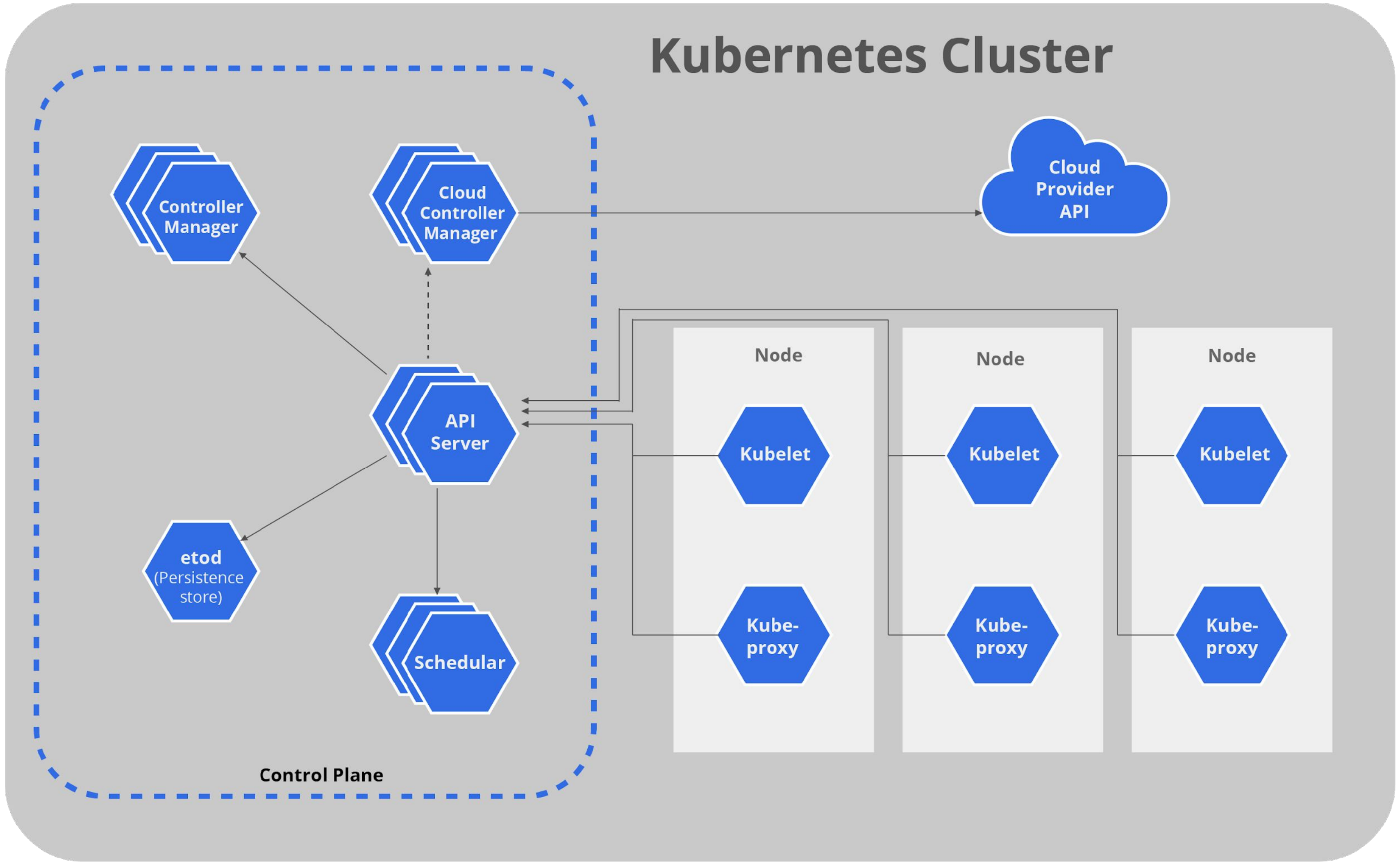


TECHNOLOGY

Overview

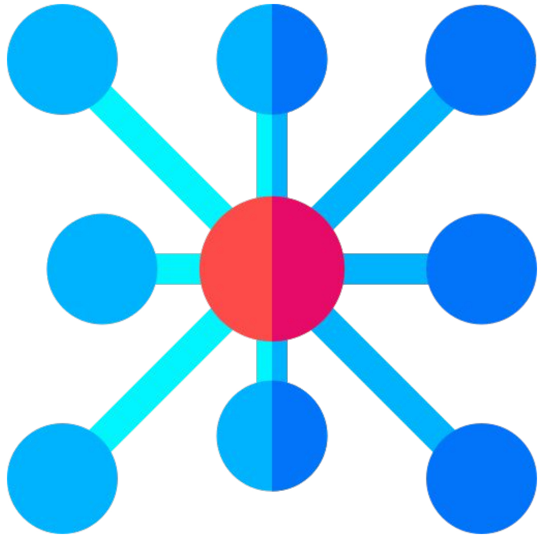
Introduction

Kubernetes Cluster is a set of Nodes for running containerized applications. It contains a Control Plane and one or more Nodes.



Control Plane

A Control Plane is responsible for maintaining the desired state of the cluster. It manages all the applications running in the cluster and the container images associated with the applications.



Nodes are the components that run the applications and workloads associated with it.

Control Plane is the Master Node in Kubernetes.

Components of Control Plane

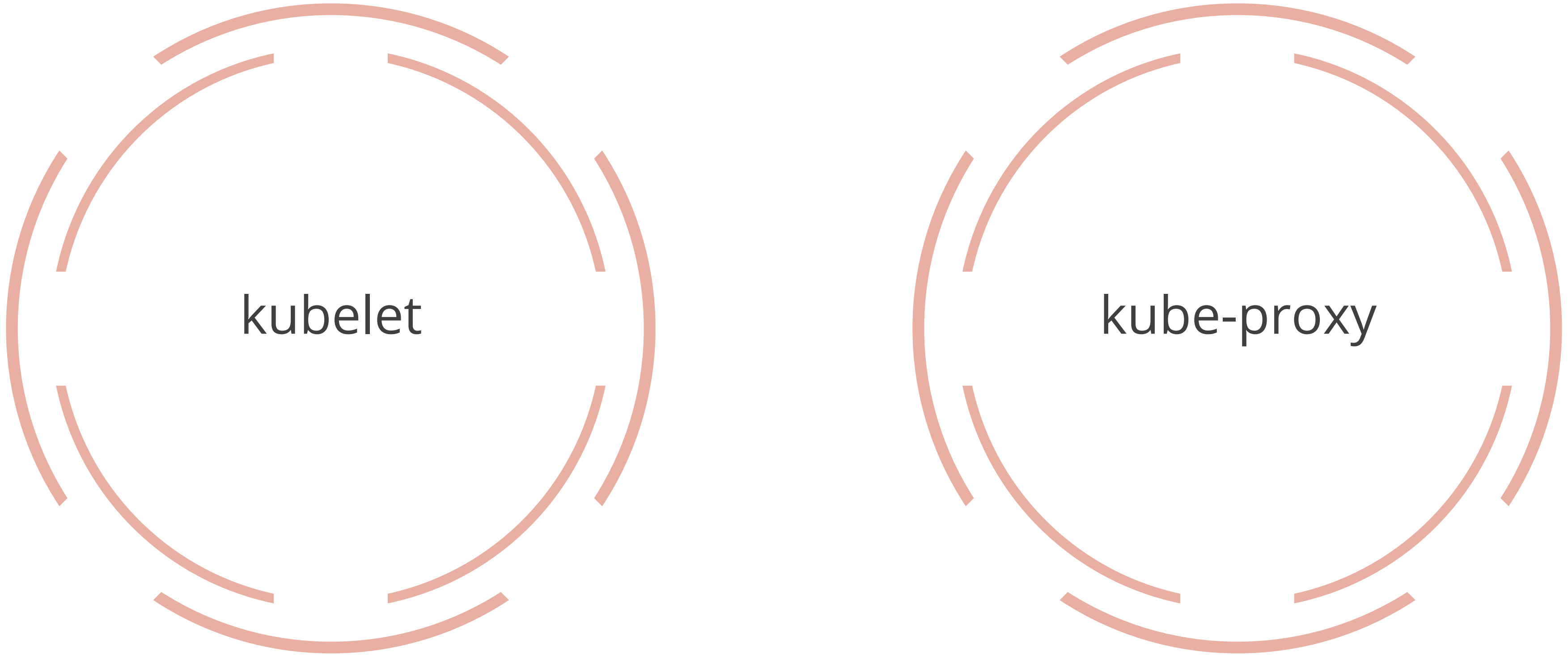
The Control Plane has five important components.

1. kube-apiserver
2. etcd
3. kube-scheduler
4. kube-controller-manager
5. Cloud-controller-manager

Nodes

To run the workload, Kubernetes places Containers in Pods. These Pods run on Nodes. Nodes are physical or virtual machines that are a part of the cluster.

A Node has two components.



kubelet

kube-proxy

Container Runtime

Container Runtime is the software responsible for running Containers. Kubernetes supports many Container Runtimes including docker, containerd, and CRI-O.

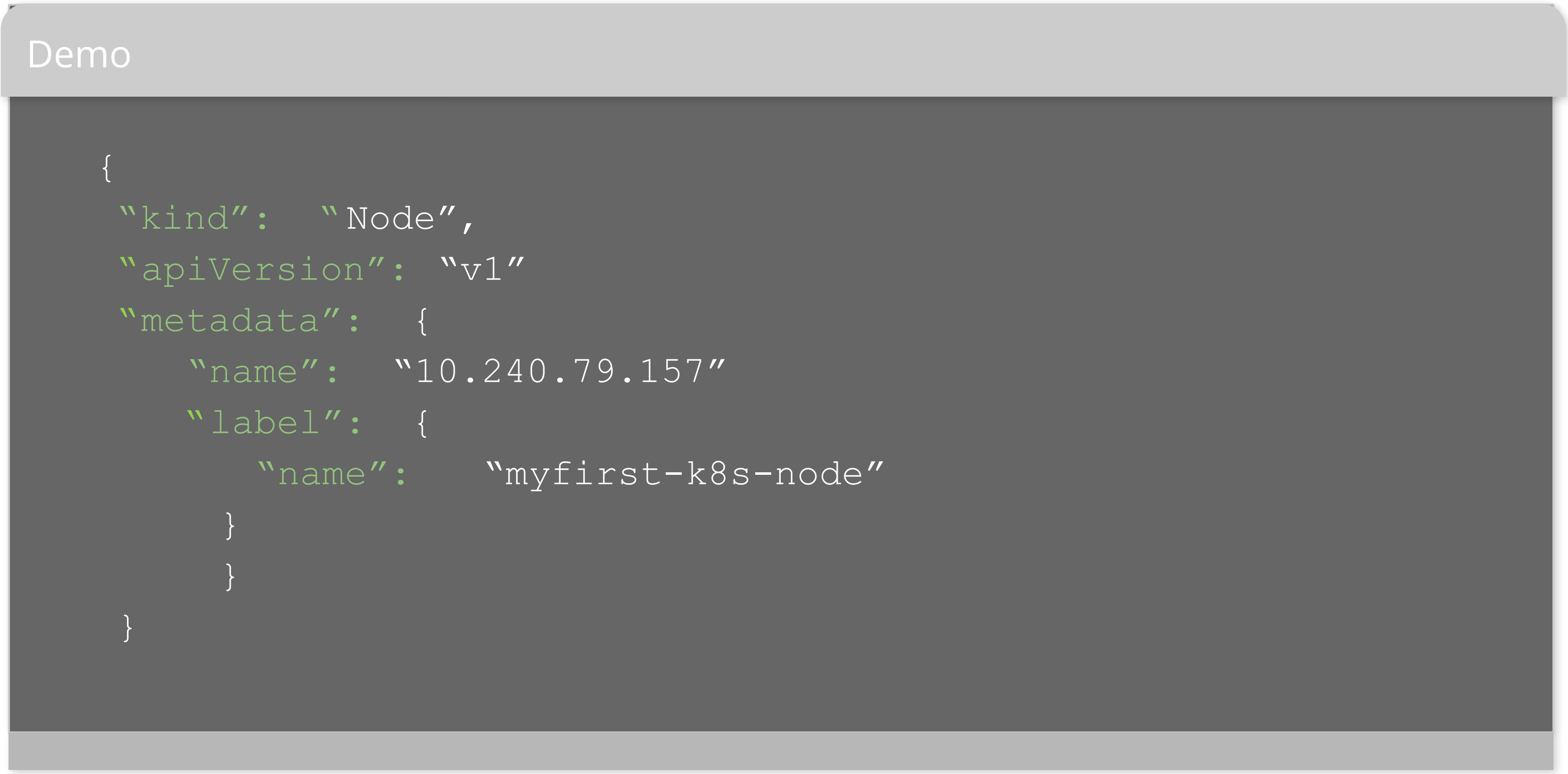
The two most common ways to add Nodes to an API server are:

Kubelet registers by itself to the Control Plane

User manually adds a Node object

Container Runtime

Following the Node registration, the Control Plane checks the validity of the new Node object. A sample JSON manifest from which a Node is being created is shown below:



```
{
 “kind”:  “Node”,
 “apiVersion”: “v1”
 “metadata”:  {
    “name”:  “10.240.79.157”
    “label”:  {
       “name”:   “myfirst-k8s-node”
        }
        }
 }
```

An unhealthy Node is ignored for any cluster activity.

Control Plane-Node Communication

Kubernetes has a **hub and spoke** API Pattern. All the API usage from Nodes (or Pods) terminates at the apiserver.

There are two communication paths from Control Plane to Node.

1. Apiserver to kubelet
2. Apiserver to Nodes, Pods, and services

TECHNOLOGY

Configuring a Cluster

General Configuration Tips

The latest stable API version must be specified while defining the configurations.

Store the configuration files in version control

Use YAML and not JSON to write configuration files

Always group related objects in a single file

Call kubectl commands on a directory

Creating Pods

Pods can be created using:

Deployment

Is the most preferred way to create Pods; it creates a ReplicaSet and specifies a strategy to replace Pods.

Jobs

Are used to create Pods that do a specific job and exit

Naked Pods

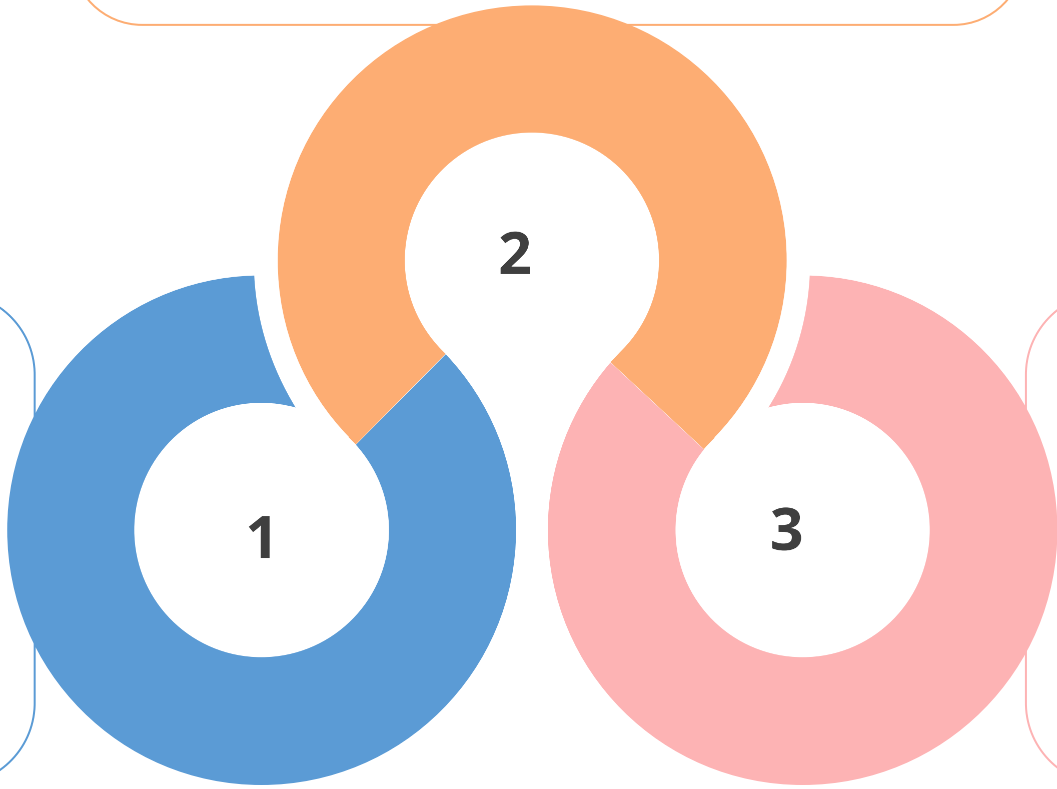
Are Pods created directly using a definition file; they are not bound to a ReplicaSet or Deployment.

Guidelines for Creating a Service

Follow these guidelines to create a service:

1. Avoid using **hostNetwork** for the same reasons as **hostPort**
2. Use **headless** services for service discovery when you don't need kube-proxy load balancing
3. Create a service before its corresponding backend workloads
4. Add DNS server as a cluster add-on to manage new services and create DNS records
5. Specify a hostPort for a Pod only when it is necessary

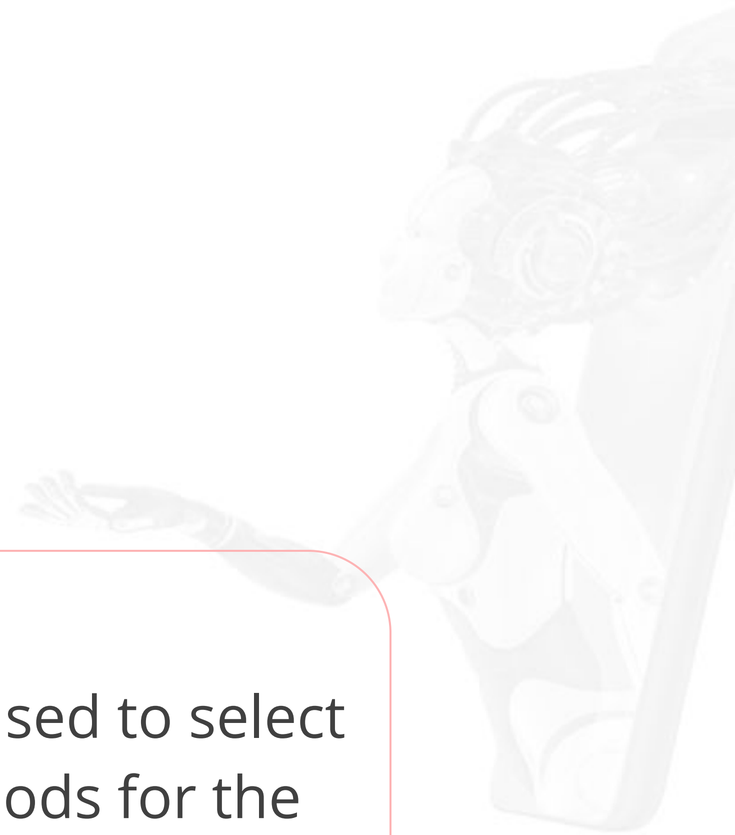
Using Labels



Define and use labels that identify semantic attributes { app: myapp, tier: frontend, phase: test, deployment: v3 }

Omit release-specific labels from their selector to make a service span multiple Deployments

Labels can be used to select appropriate Pods for the resources.



Container Images

When the kubelet attempts to pull a specified image, the **imagePullPolicy** and the tag of the image are affected. Here are some examples:

imagePullPolicy: **IfNotPresent:** The image gets pulled if it is not already present locally.

imagePullPolicy: **Always**: kubelet queries the container image registry to resolve the name to an image digest.

imagePullPolicy:latest or omitted: Omitted.

imagePullPolicy: **Never**: The image is never assumed to exist locally.

Using kubectl

The **kubectl apply** is used to maintain the changes applied to a live object. kubectl apply follows declarative management.

1. To create Single-Container Deployments and services, use **kubectl create deployment** and **kubectl expose.**

2. For **get** and **delete** operations, use label selectors and not specific object names.

Configuring a Kubernetes Cluster

Problem Statement:

Learn to set up and configure a Kubernetes cluster.

ASSISTED PRACTICE

Assisted Practice: Guidelines

Steps to demonstrate configuring a cluster in Kubernetes:

1. Set up the master node and define the cluster
2. Join the worker node to the cluster
3. Verify the cluster

Understanding the Best Practices of Kubernetes Cluster

Problem Statement:

Understand the best practices of Kubernetes cluster configuration.

ASSISTED PRACTICE

Assisted Practice: Guidelines

Steps to demonstrate the best practices to be followed while configuring a cluster in Kubernetes:

1. Having all the configurations of related objects in a single YML file
2. Placing all the related configuration files in the same folder
3. Accessing clusters using Kubernetes API

Understanding the Components of Kubernetes Cluster

Problem Statement:

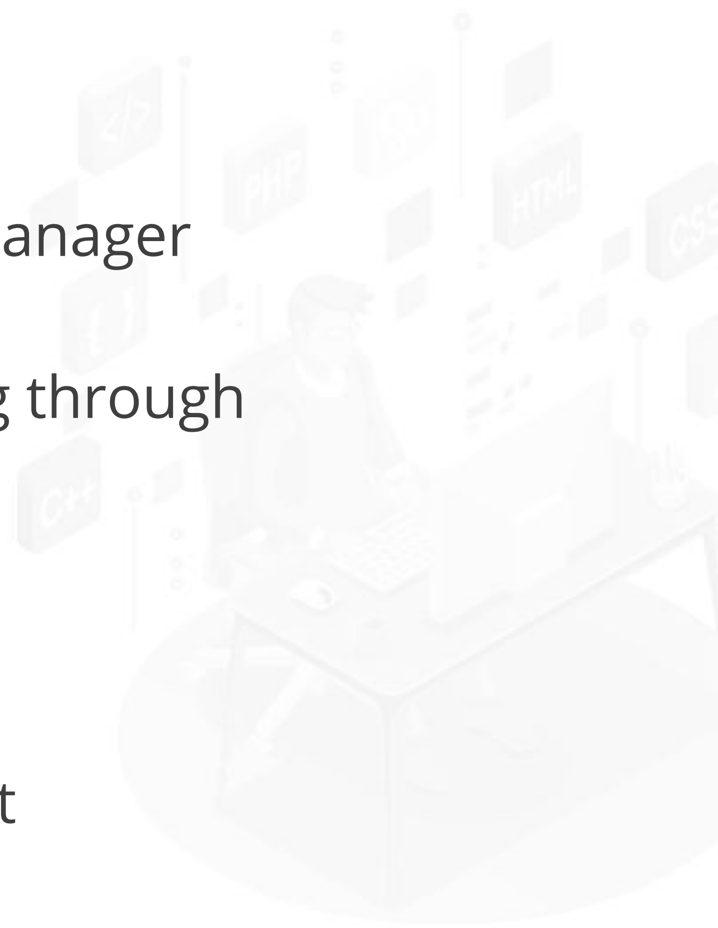
Understand the components of Kubernetes cluster.

ASSISTED PRACTICE

Assisted Practice: Guidelines

Steps to demonstrate the working of cluster components in Kubernetes:

1. Install etcd server and execute basic etcd commands
2. Install kube-controller-manager and execute other options of kube-controller-manager
3. Install kube-scheduler on the master node, learn about kube-scheduler by going through help, and try out some commands
4. Check if the kubelet is installed and try some other kubelet options
5. Install kube-proxy, check the version of kube-proxy, and try out different kubelet commands



Assisted Practice: Guidelines

6. Learn about Pods by going through the help option

7. Learn about ReplicaSets by going through help option

8. Fetch Deployment

9. Fetch a service

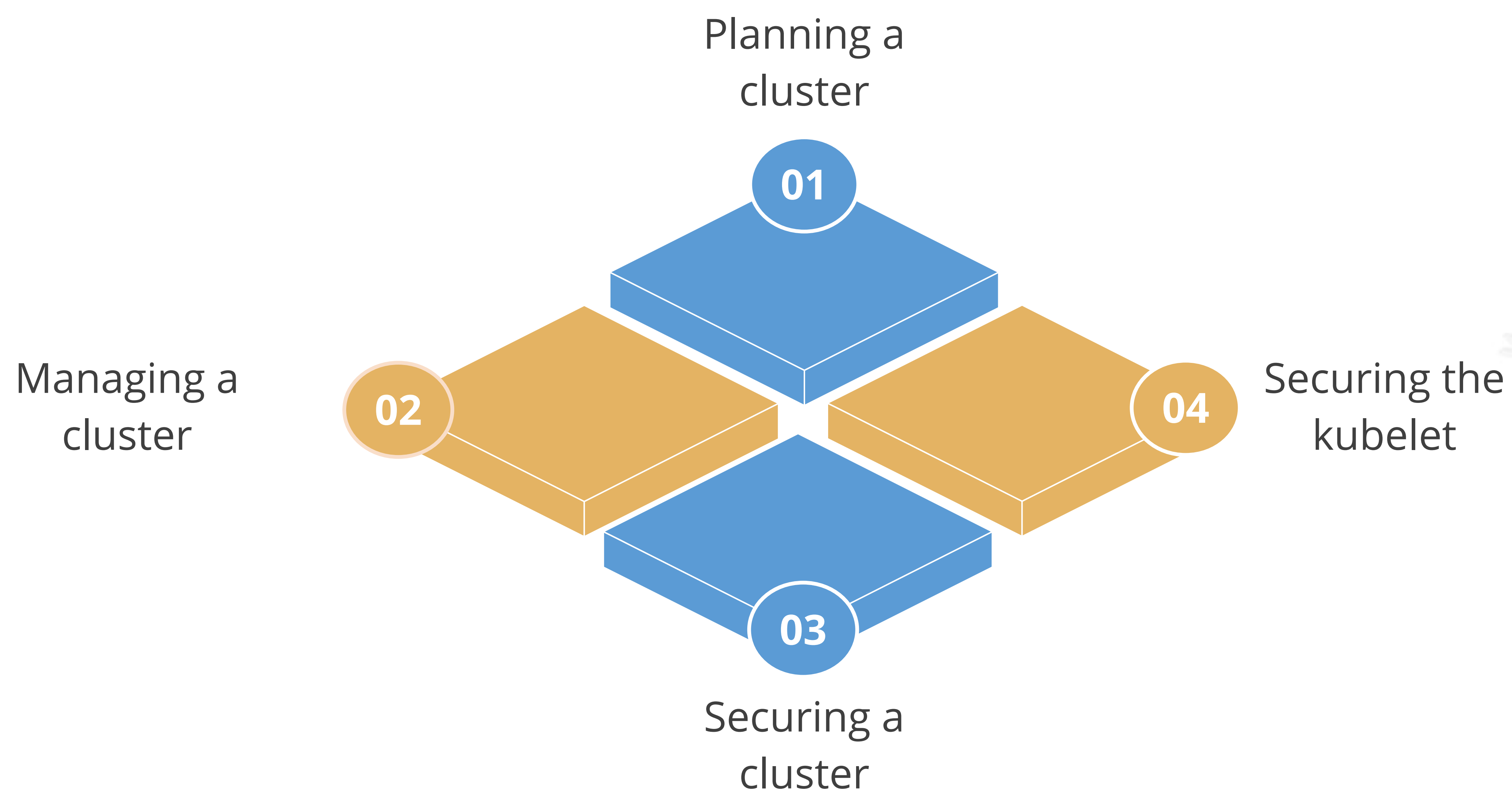
10. Fetch container list from docker

TECHNOLOGY

Managing and Administering Clusters

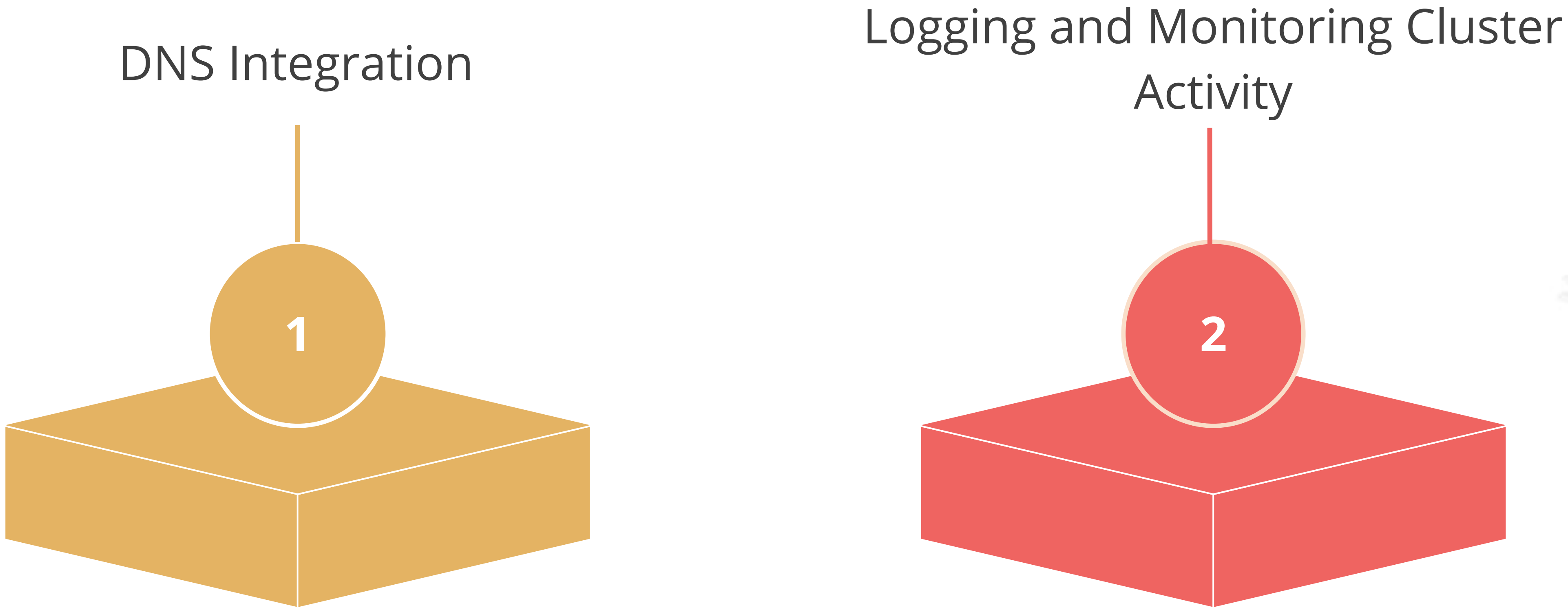
Overview

Cluster administration and management involves four steps.



Optional Cluster Services

The following are optional for Kubernetes Clusters:



1. DNS Integration

2. Logging and Monitoring Cluster Activity

Managing and Administrating a Kubernetes Cluster

Problem Statement:

Manage and administer a cluster to understand the housekeeping concepts like verifying security certificate, creating namespaces, creating Pods, and accessing clusters using Kubernetes API.

ASSISTED PRACTICE

Assisted Practice: Guidelines

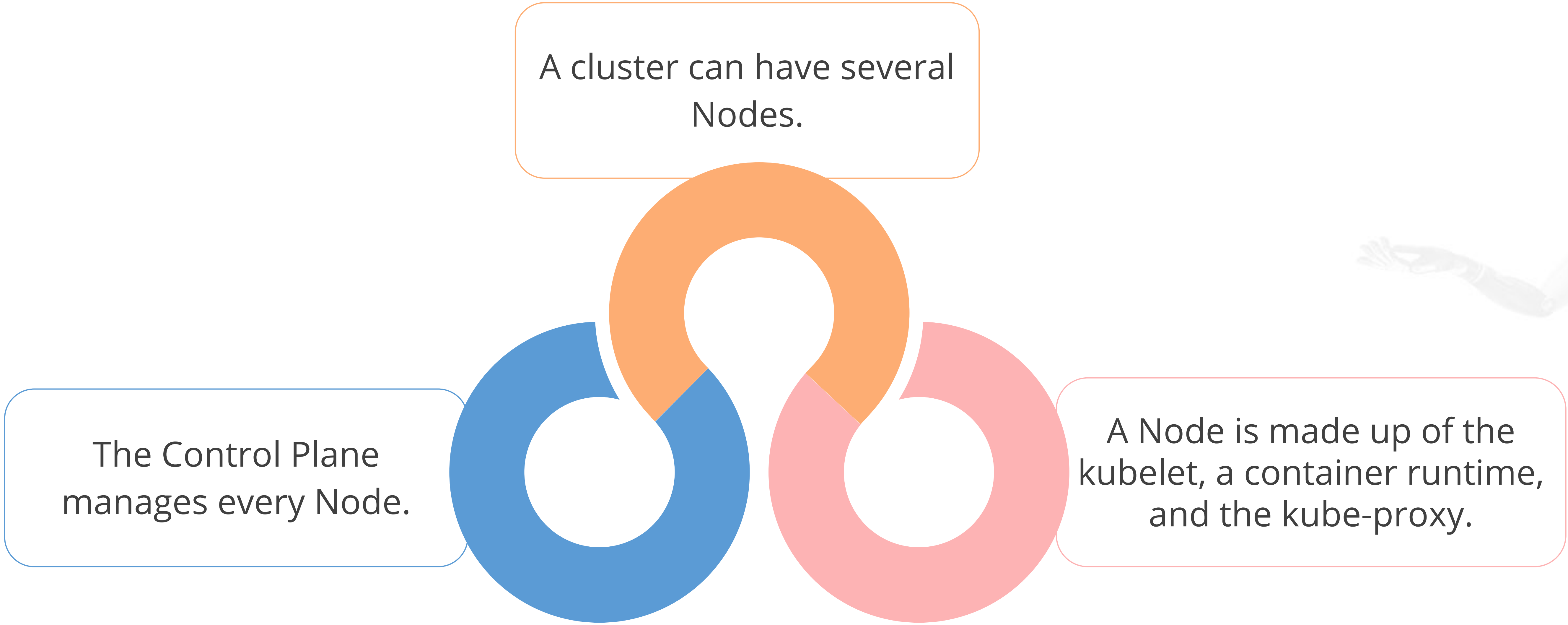
Steps to demonstrate managing and administering clusters in Kubernetes:

1. Verify the certificate of the cluster
2. Viewing the cluster information
3. Creating a namespace
4. Creating a Pod
5. Accessing clusters using Kubernetes API

Node

Introduction

To run the workload, Kubernetes places containers in Pods and runs them on Nodes. A Node can be a physical or a virtual machine.



A cluster can have several Nodes.

The Control Plane manages every Node.

A Node is made up of the kubelet, a container runtime, and the kube-proxy.

Management

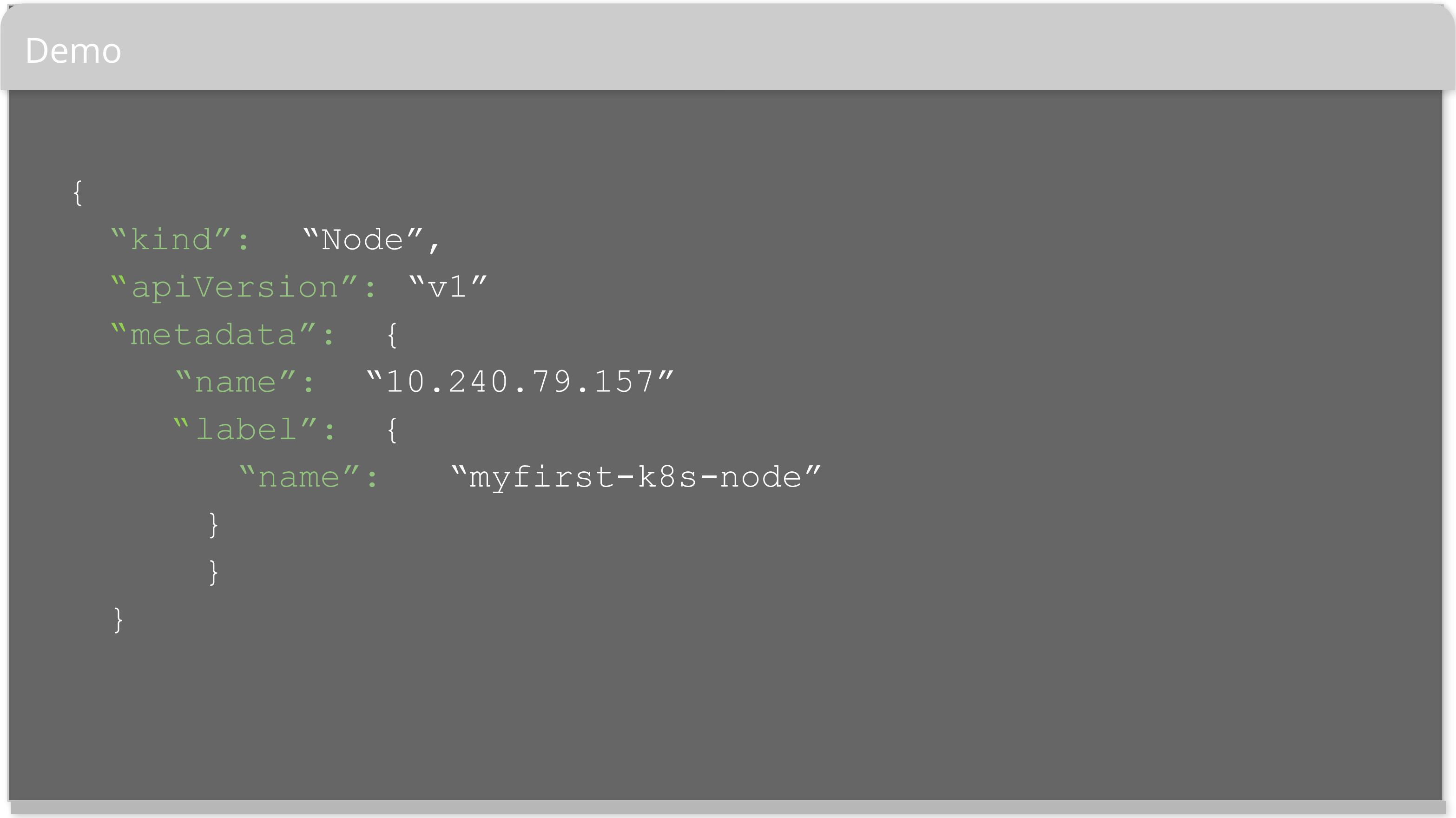
There are two main ways to add Nodes to the API server.

1 Self register

2 Manually add

Management

To create a Node from a JSON manifest, use:

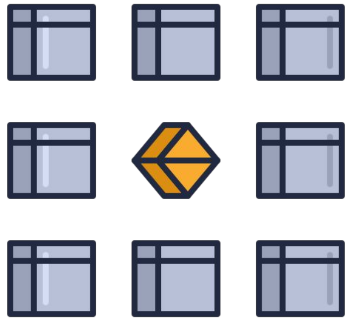


Demo

```
{
  “kind”:  “Node”,
  “apiVersion”: “v1”
  “metadata”:  {
     “name”:  “10.240.79.157”
     “label”:  {
        “name”:   “myfirst-k8s-node”
       }
       }
  }
```

Node Name Uniqueness

The Node name identifies a Node and must, therefore, be unique. Two Nodes cannot have the same name at the same time.



1 Kubernetes assumes that a resource with the same name is the same object.

2 An instance using the same name is assumed to have the same state.

Self-Registration of Nodes

When the kubelet flag **--register-node** is true (the default), the kubelet will attempt to register itself with the API server. To accommodate self-registration, use the following options to start the kubelet:

--kubeconfig

Set a path to credentials to authenticate itself to the API server

--register-with-taints

Register the Node with the given list of Taints

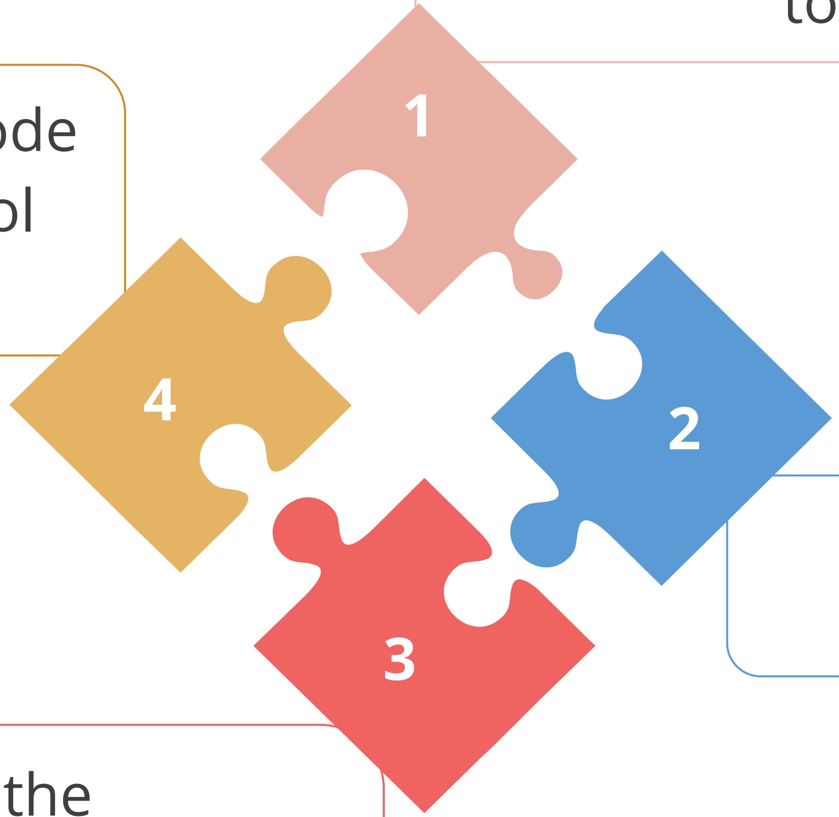
--cloud-provider

Talk to a cloud provider to read metadata about itself

--register-node

Register the Node automatically with the API server

Manual Node Administration



Set kubelet **flag --register-node** to **false**

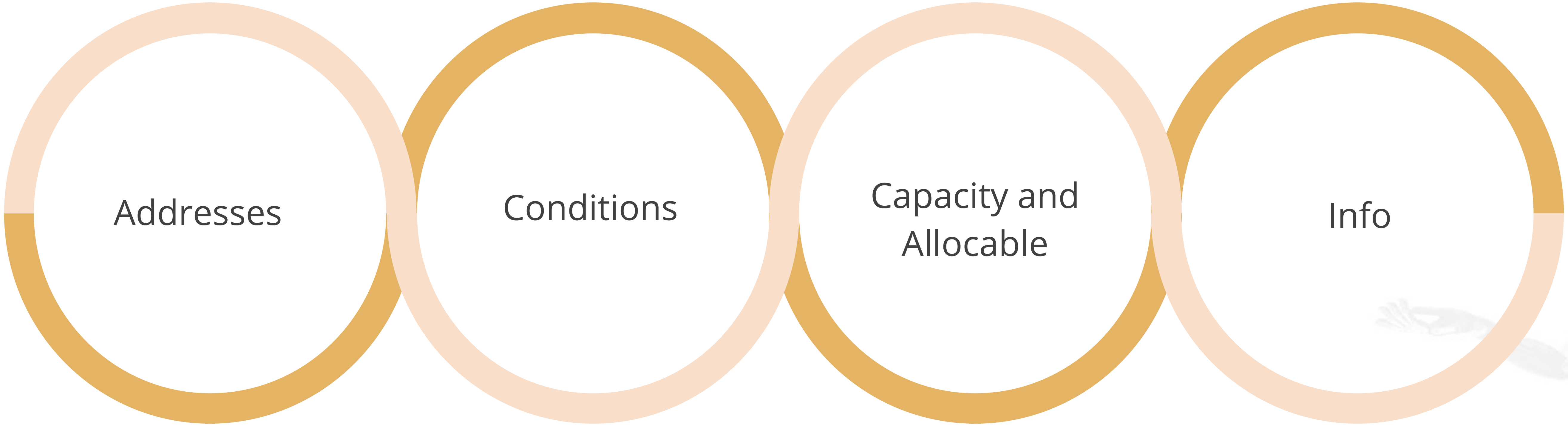
Use Kubectl to create and modify Node objects

Note that the **--register-node** setting has no effect on modification of Node objects

Use labels on Nodes with Node Selectors on Pods to control scheduling

Node Status

A Node's status contains the following information:



Addresses

Conditions

Capacity and Allocable

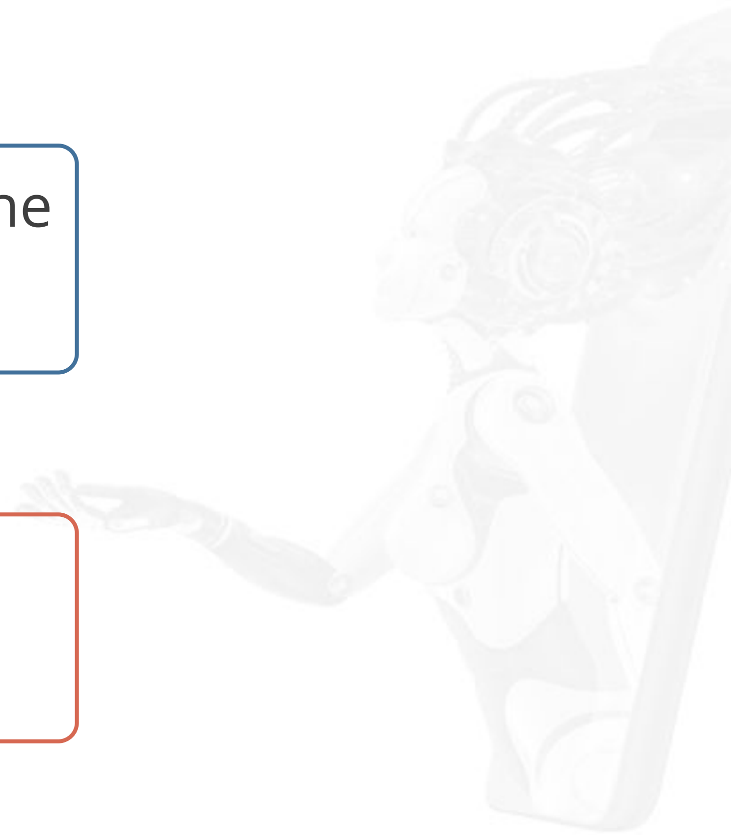
Info

To view the status of a Node, use:

```
kubectl describe node <insert-node-name-here>
```

Addresses

The output of **kubectl describe node** can have three fields. Based on the cloud provider or the bare metal configuration, the usage of these fields will differ.



HostName

The hostname, as reported by the Node's kernel, can be overridden via the kubelet --hostname-override parameter

ExternalIP

The IP address of the Node that is externally routable

InternalIP

The IP address of the Node that is routable only within the cluster

Conditions

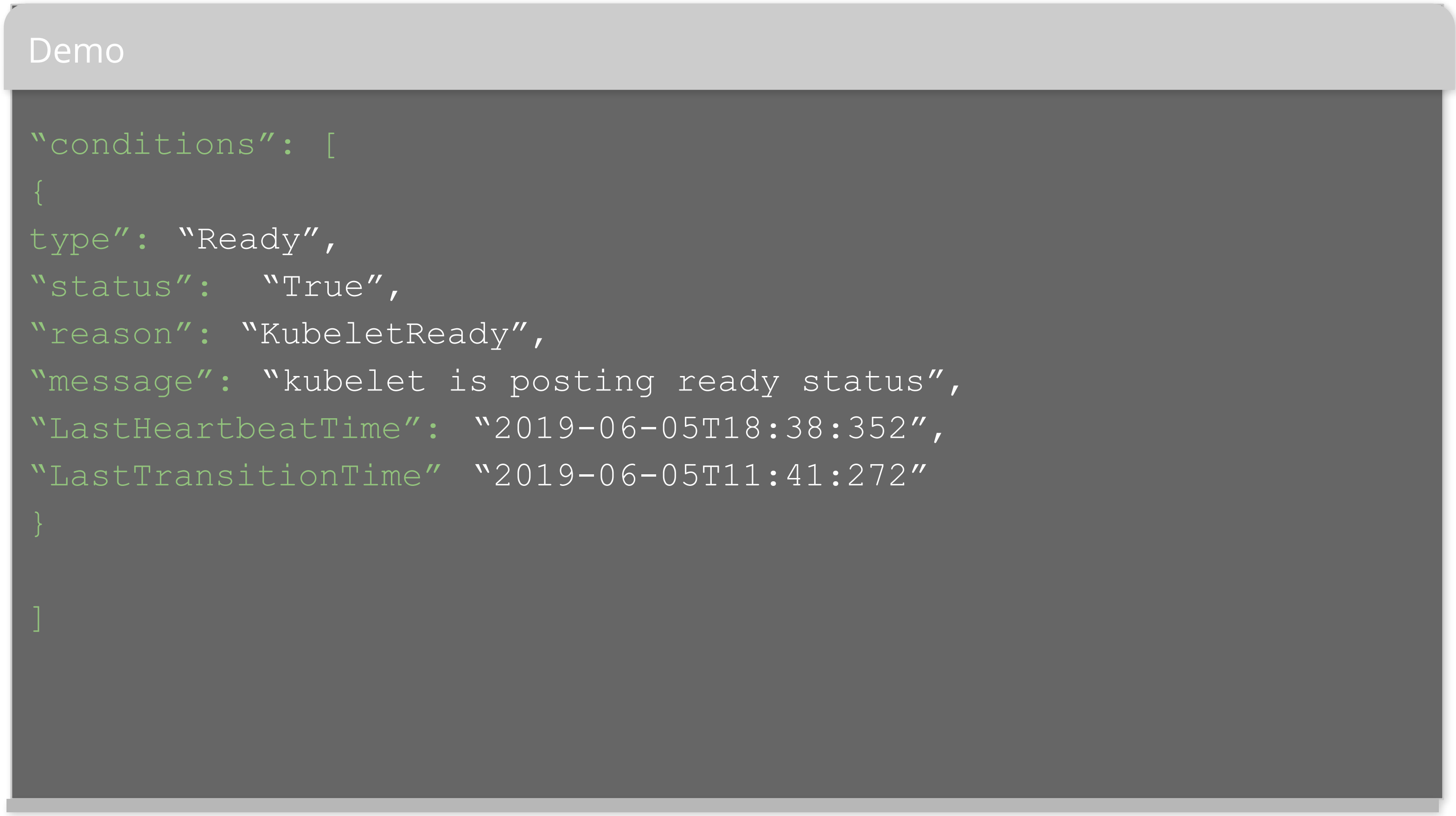
The Conditions field describes the status of all Running Nodes.

Some examples of Conditions are given below:

Node Condition	Description
Ready	True if the Node is healthy, ready to accept Pods; False if the Node is not healthy, not accepting Pods; Unknown if there is no response during the last node-monitor-grace-period
DiskPressure	True if pressure exists on the disk size, that is, if the disk capacity is low; otherwise False
MemoryPressure	True if pressure exists on the Node memory, that is, if the Node memory is low; otherwise False
PIDPressure	True if pressure exists on the processes, that is, if there are too many processes on the Node, otherwise False
NetworkUnavailable	True if the network for the Node is not correctly configured, otherwise False

Conditions

Node Condition is represented as a JSON object. The structure below describes a healthy Node.



Demo

```
"conditions": [
{
type": "Ready",
"status":  "True",
"reason": "KubeletReady",
"message": "kubelet is posting ready status",
"LastHeartbeatTime": "2019-06-05T18:38:352",
"LastTransitionTime" "2019-06-05T11:41:272"
}

]
```

Capacity and Allocatable Blocks

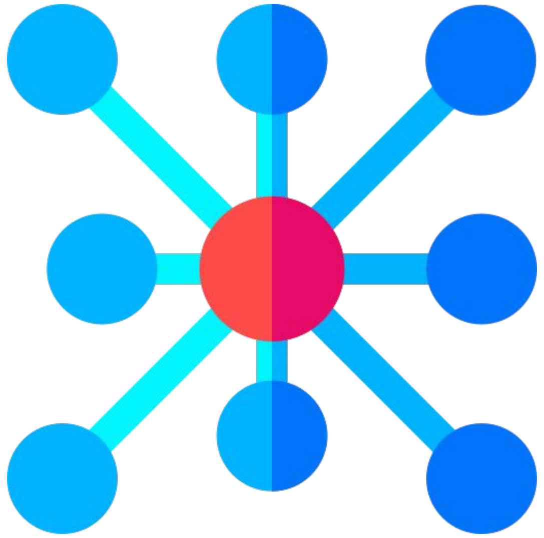
Capacity and Allocatable Blocks describe the resources available on the Node. The resources include CPU, memory, and the maximum number of Pods that can be scheduled on the Node.

Capacity Block specifies the total number of resources present in the Node.

Allocatable Block specifies the number of available resources on a Node.

Info

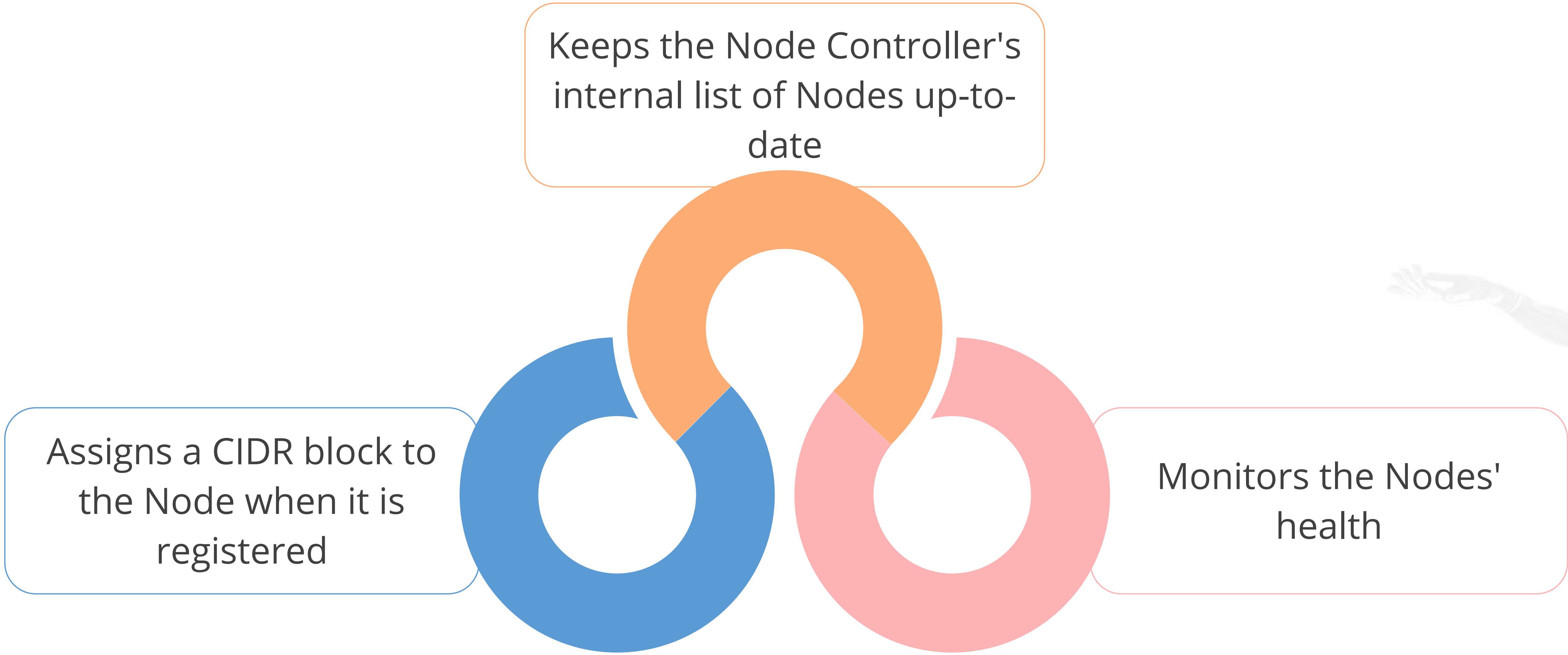
Info provides general information about the Node, such as kernel version, Kubernetes version, Docker version, and OS name.



kubelet is used to gather information about a Node.

Node Controller

Node Controller is a Kubernetes Control Plane component that manages Nodes.



Keeps the Node Controller's internal list of Nodes up-to-date

Assigns a CIDR block to the Node when it is registered

Monitors the Nodes' health

Heartbeats

Heartbeats are used to determine the availability of a Node. Heartbeats can be the updates of nodestatus or the Lease object.

The NodeStatus and the Lease object are updated by the kubelet.

1 The **NodeStatus** gets updated when there is a change in status or if there has been no update for a configured interval.

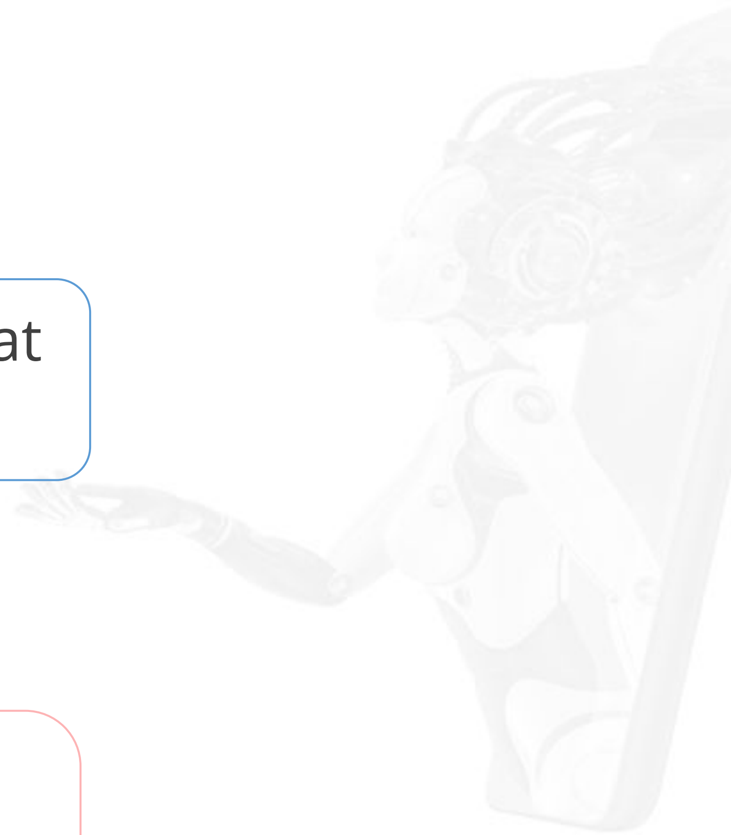
2 Lease objects are created and updated every 10 seconds .

Reliability

The Node eviction behavior changes when a Node becomes unhealthy. Listed below are a few more scenarios that trigger a change in eviction behavior.

01 Eviction rate is reduced if the fraction of unhealthy Nodes is at the least **--unhealthy-zone-threshold.**

02 Eviction rate is stopped if the cluster is small (≤ **--large-cluster-size-threshold** Nodes). Otherwise, the eviction rate is reduced to **--secondary-node-eviction-rate.**

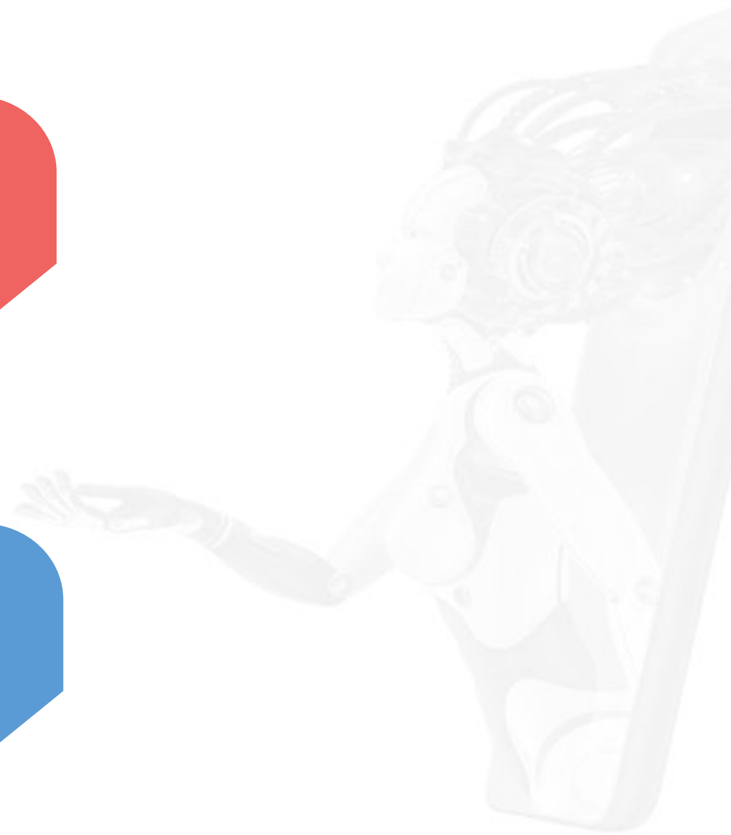


Node Capacity

The resource capacity of a Node is tracked by the Node objects.

1. **Self-registering Nodes:** The Node capacity is reported when the Node is registered.

2. **Manually added Nodes:** The Node capacity must be set when the Node is added.



Graceful Node Shutdown

Kubernetes supports Graceful Node feature. This feature detects a Node system shutdown and terminates Pods running on the Node.

The Graceful Node shutdown feature **depends** on system.

Graceful Node shutdown is **controlled** using the Gracefulnodeshutdown feature gate.

Graceful Node Shutdown

Kubelet carries out the termination process in the two phases.

1 Terminates regular Pods running on the Node

2 Terminates critical Pods running on the Node

Understanding the Working of Nodes

Duration: 5 mins

Problem Statement:

Use kubectl and kubelet utilities to comprehend the working of nodes.

Assisted Practice: Guidelines

Steps to demonstrate Node in Kubernetes:

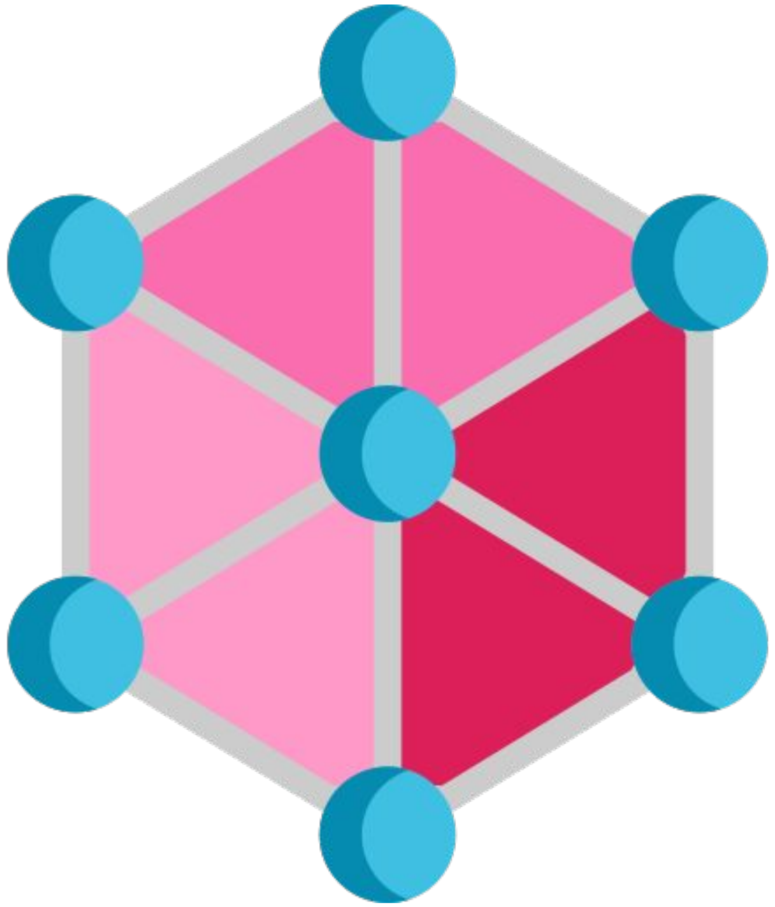
1. Verify the status of the node using kubectl
2. Create a node using configuration file

TECHNOLOGY

Control Plane–Node Communication

Overview

The Control Plane catalogs the communication paths between the Control Plane and the Kubernetes Cluster. This helps users customize installation and harden network configuration.



Communication from Node to Control Plane

Kubernetes has a hub and spoke pattern, which ensures that all API usage from Nodes terminate at the apiserver.

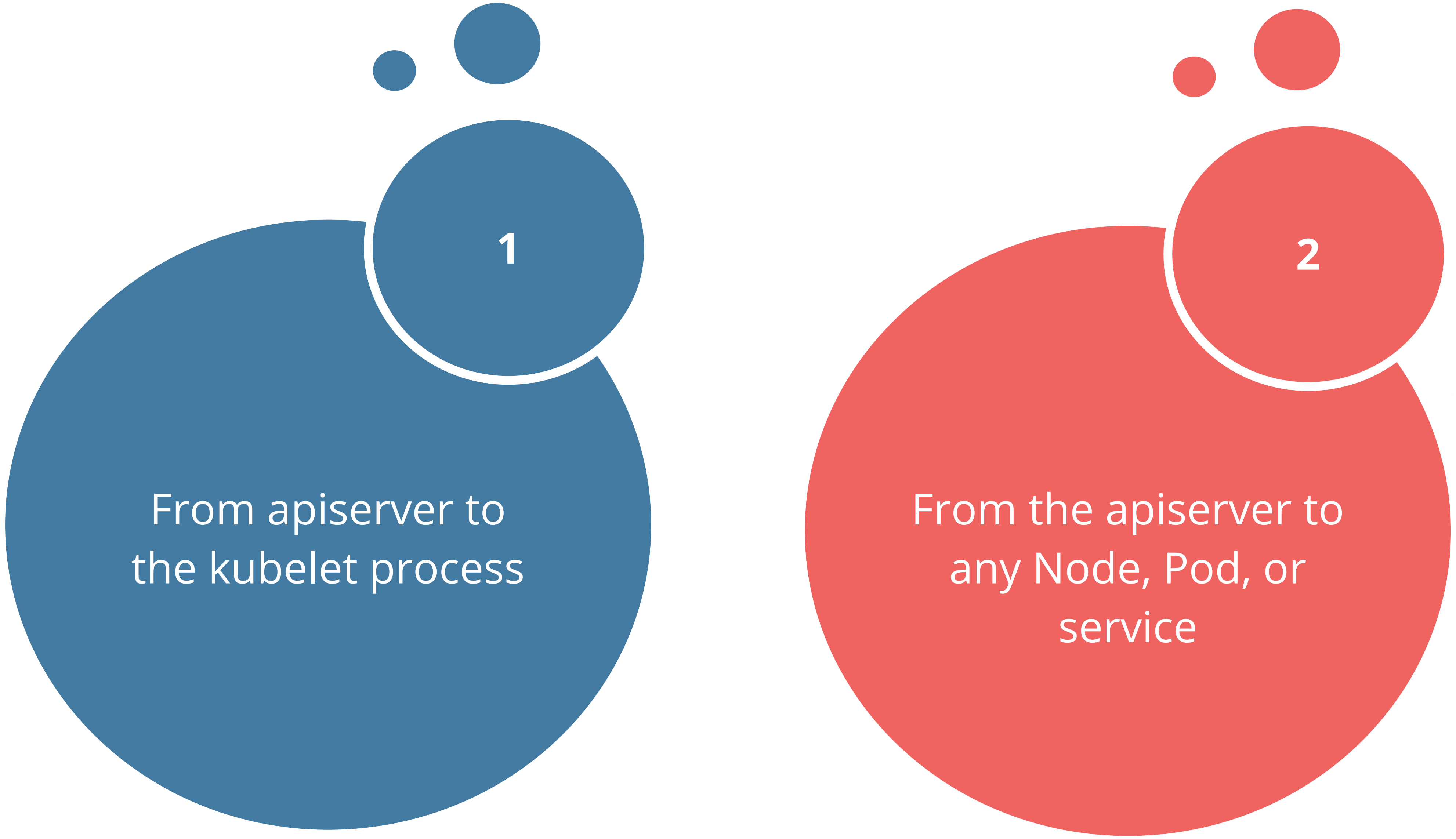
1 The apiserver is configured to listen for remote connections on a secure HTTPS port with one or more forms of client authentication enabled.

2 Nodes should be provisioned with the public root certificate for the cluster in such a way that they can connect securely to the apiserver along with valid client credentials.

3 The components of the Control Plane communicate with the cluster apiserver over the secure port.

Control Plane to Node

Communication from Control Plane to Nodes can follow two paths.



1

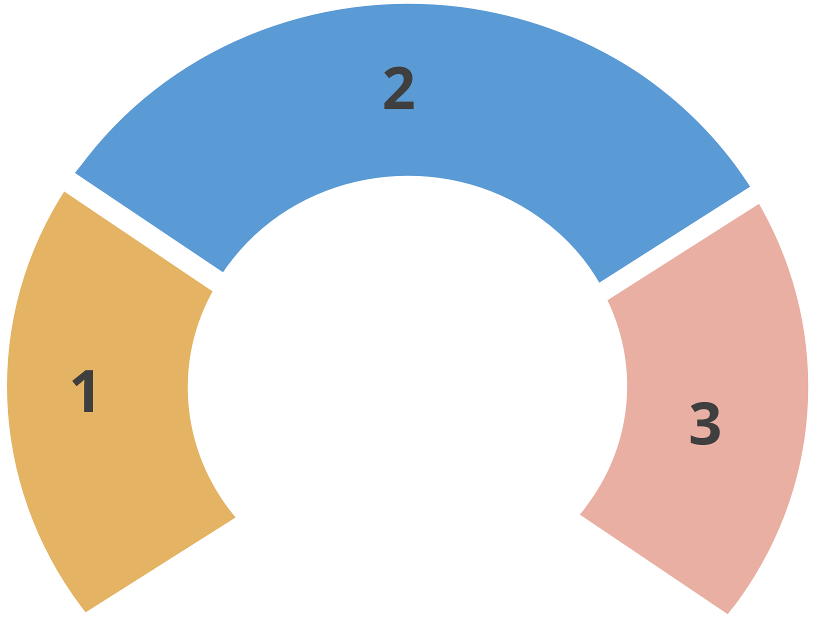
From apiserver to the kubelet process

2

From the apiserver to any Node, Pod, or service

Connection from apiserver to kubelet

The connections from the apiserver to the kubelet serve three purposes.



Attaching (through kubectl) to running Pods

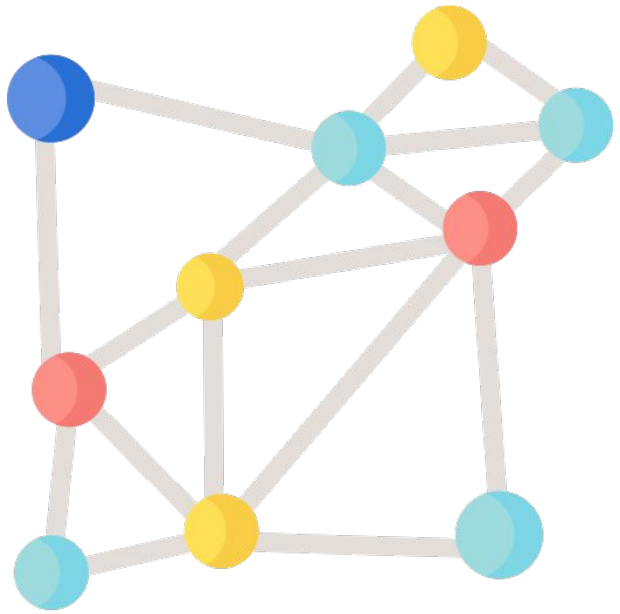
Fetching logs for Pods

Providing kubelet's
port-forwarding functionality

Connection from apiserver to Nodes, Pods, and Services

The apiserver can connect to a Node, Pod or service. These connections that default to a plain HTTP connection are neither encrypted nor authenticated.

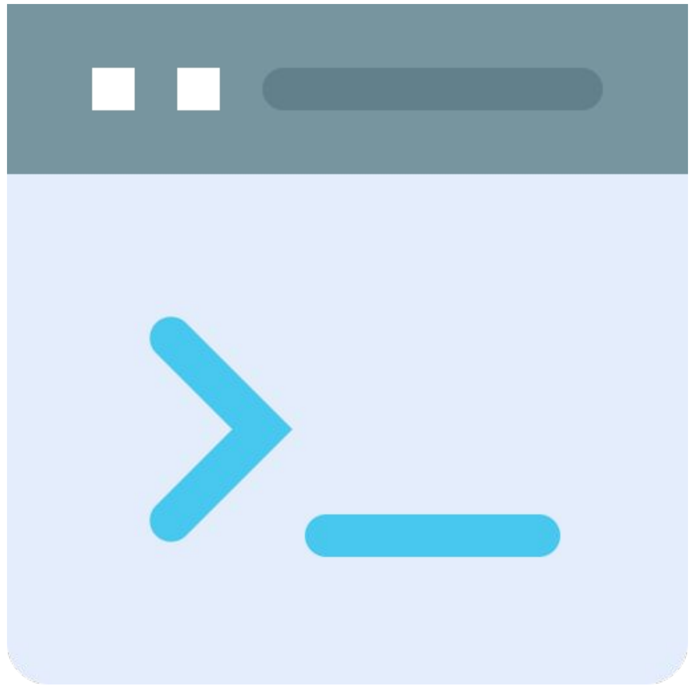
The connections can run over a secure HTTPS. This is done by prefixing **https:** to the Node, Pod, or service name in the API URL.



Connections do not validate the certificate provided by the HTTPS Endpoint. They do not provide client credentials.

SSH Tunnels

SSH Tunnels are used by Kubernetes to protect the Control Plane to Nodes communication paths. They are currently deprecated and, hence, the usage should be minimum.

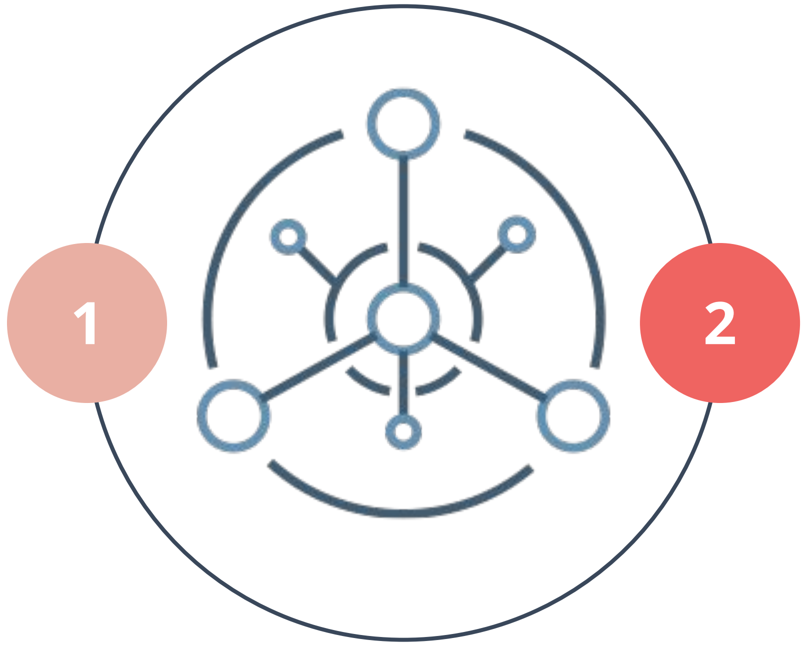


Konnectivity service is used in place of SSH tunnels.

Konnectivity Service

Konnectivity service provides TCP level proxy for the Control Plane to cluster communication.

It consists of two parts.



Konnectivity server

Konnectivity agents

Understanding Controllers

Controller Pattern

The Controller regulates the state of a system.

Tracks at least one Kubernetes resource type

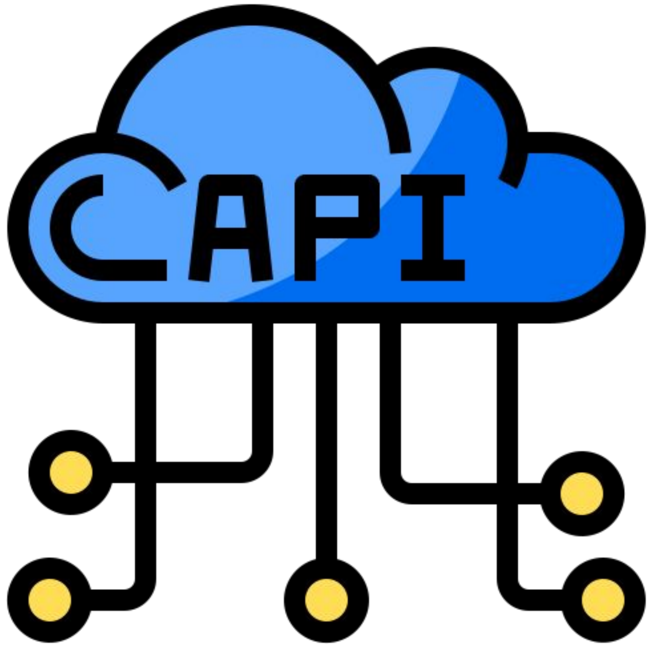
Has a spec field that represents the desired state

Moves the resource’s current state closer to the desired state

Control via API Server

Built-in Controllers interact with the cluster API server and manage the system state.

Job Controller is an example of a Kubernetes built-in Controller.



Controllers can update the objects that are used to configure them.

Direct Control

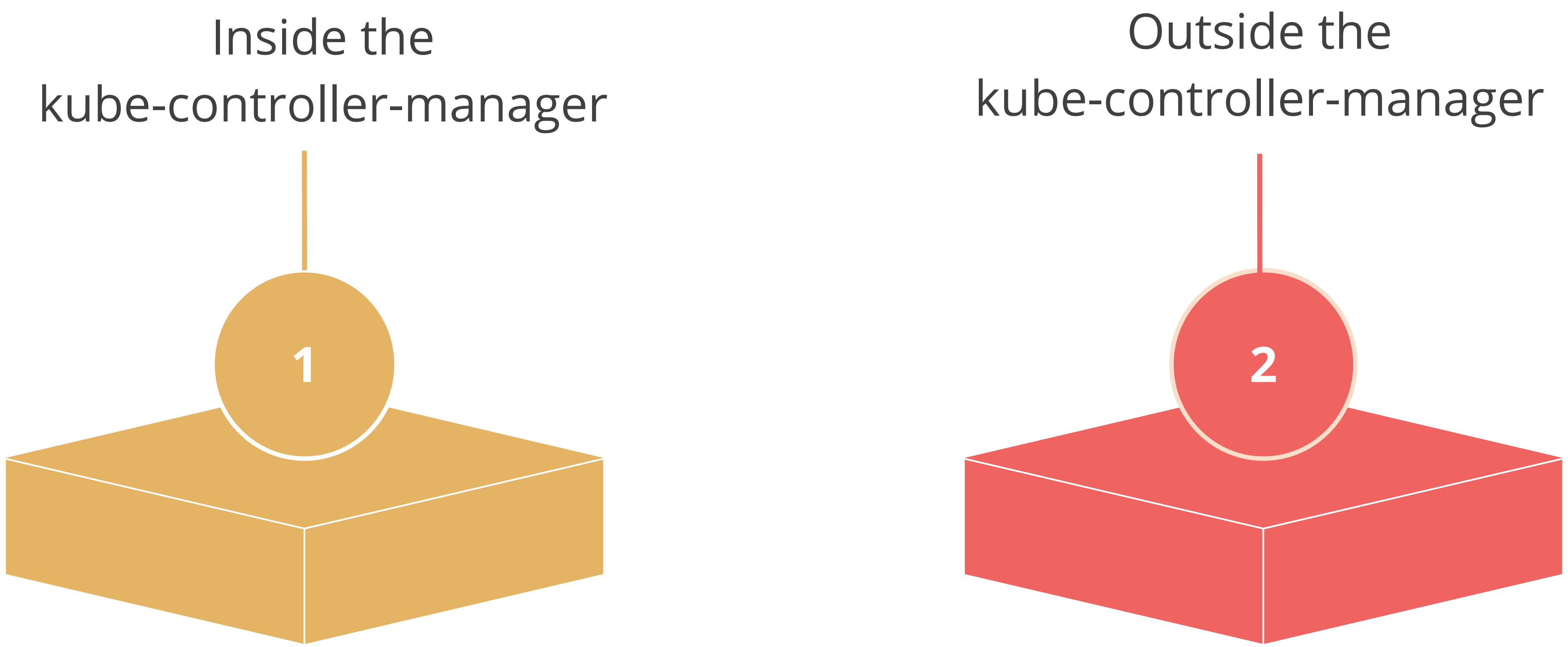
Direct control is employed when Controllers have to make changes outside the cluster.

Controllers find the desired state from API server.

Controllers communicate directly to bring the current state closer in line.

Running Controllers

There are two ways of running Controllers.



Inside the kube-controller-manager

Outside the kube-controller-manager

TECHNOLOGY

Importance of Cloud Controller Manager

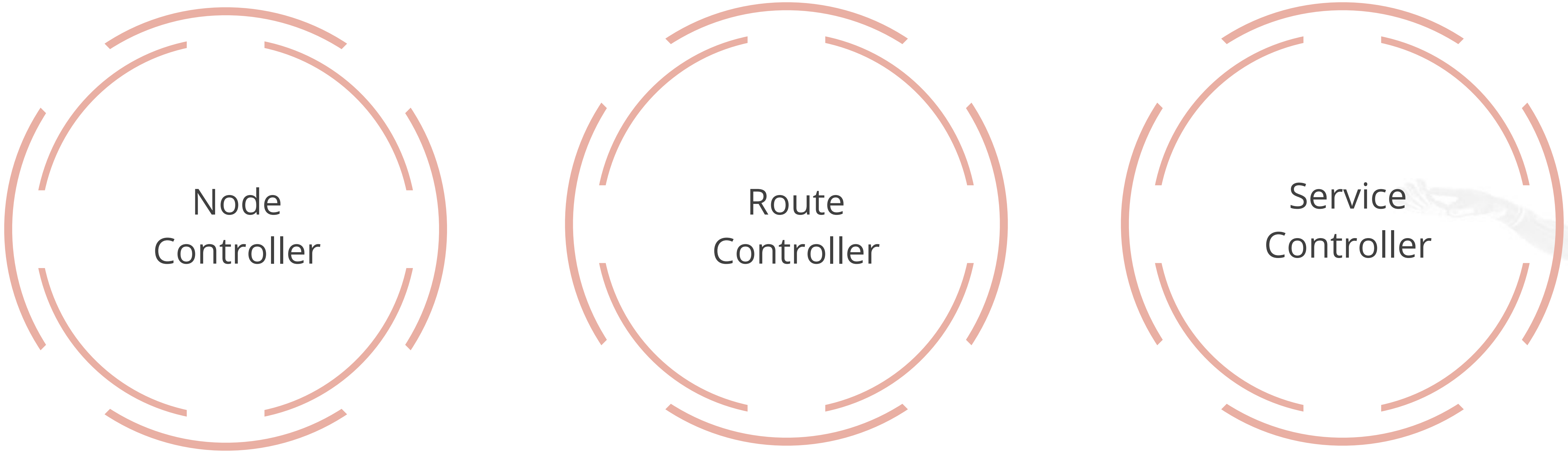
Introduction

Cloud infrastructure technologies:

- ✓ Help run Kubernetes on public, private, and hybrid clouds
- ✓ Embed cloud-specific control logic to run on different cloud platforms
- ✓ Are structured using a plugin mechanism that allows different cloud providers to integrate their platforms with Kubernetes

Cloud Controller Manager Types

There are many types of Controllers inside the Cloud Controller Manager, each of which has a role to play.



TECHNOLOGY

Working with kubeadm

Overview

The Kubeadm toolkit is used to bootstrap a best-practice Kubernetes Cluster. Kubeadm provides **kubeadm init** and **kubeadm join** to achieve this.

kubeadm acts as an installer and a building block that helps to get a minimum viable cluster up and running.

kubeadm Init Command Options

The **kubeadm init [flags]** command is used to initialize and set up a Kubernetes Control Plane Node. The various options available for this command are:

1. --apiserver-advertise-address
2. --apiserver-bind-port
3. --apiserver-cert-extra-sans
4. --cert-dir
5. --certificate-key
6. --config
7. --control-plane-endpoint
8. --cri-socket

kubeadm Init Command Options

9 --dry-run

10 --experimental-patches

11 --feature-gates

12 --h, --help

13 --ignore-preflight-errors

14 --image-repository

15 --node-name

16 --pod-network-cidr

kubeadm Init Command Options

17. --service-cidr
18. --service-dns-domain
19. --skip-certificate-key-print
20. --skip-phases
21. --skip-token-print
22. --token
23. --token-ttl duration
24. --upload-certs

Init Workflow

To bootstrap a Kubernetes Control Plane, kubeadm Init follows these steps:

Step 1

Run a series of pre-flight checks to validate the system state before making changes

Step 2

Set up identities for each cluster component by generating a self-signed CA

Step 3

Write kubeconfig files in /etc/kubernetes/ for the kubelet and controller-manager

Step 4

Generate static Pod manifests for the API server, controller-manager, and scheduler

Init Workflow

Step 5

Apply labels and taints to the Control Plane Node

Step 6

Generate the token that additional Nodes can use to register themselves with a Control Plane

Step 7

Make necessary configurations for allowing Node joining with the Bootstrap Tokens and TLS

Step 8

Install a DNS server and the kube-proxy addon components through the API server

Using Init Phases with kubeadm

To create a Control Plane in phases, use **kubeadm init phase** command.

Certain Control Plane phases have unique flags. To get a list of available options, add **–help.** An example is given below:

```
sudo kubeadm init phase control-plane controller-manager --help
```

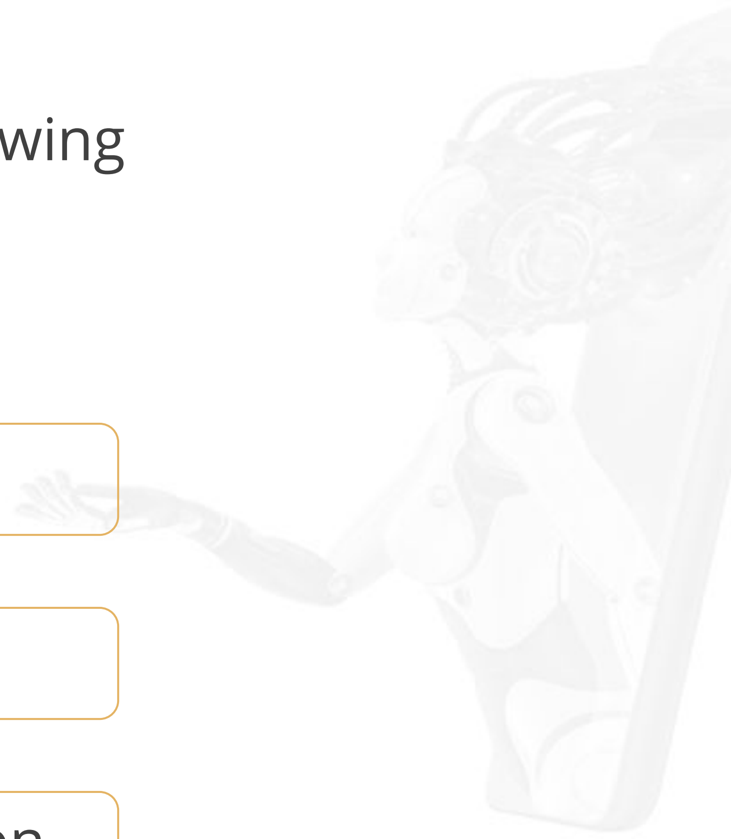
Use **--help** to see the list of sub-phases for a certain parent phase.

```
sudo kubeadm init phase control-plane --help
```

Using Custom Images

By default, **kubeadm** pulls images from **k8s.gcr.io**. **gcr.io/kubernetes-ci-images**. If the requested Kubernetes version is a CI label, **gcr.io/kubernetes-ci-images** is used.

A configuration file helps override this behavior. Kubernetes allows the following customization:



- Provide an alternative **imageRepository** to **k8s.gcr.io**
- Set **useHyperKubeImage** to true to use the HyperKube image
- Provide a specific **imageRepository** and **imageTag** for etcd or DNS add-on

Setting the Node Name

Based on the machine's host address, **kubeadm** assigns a Node name. **--node-name** flag may be used to override this setting.

kubeadm

--node-name flag passes the appropriate **--hostname-override** value to the kubelet.

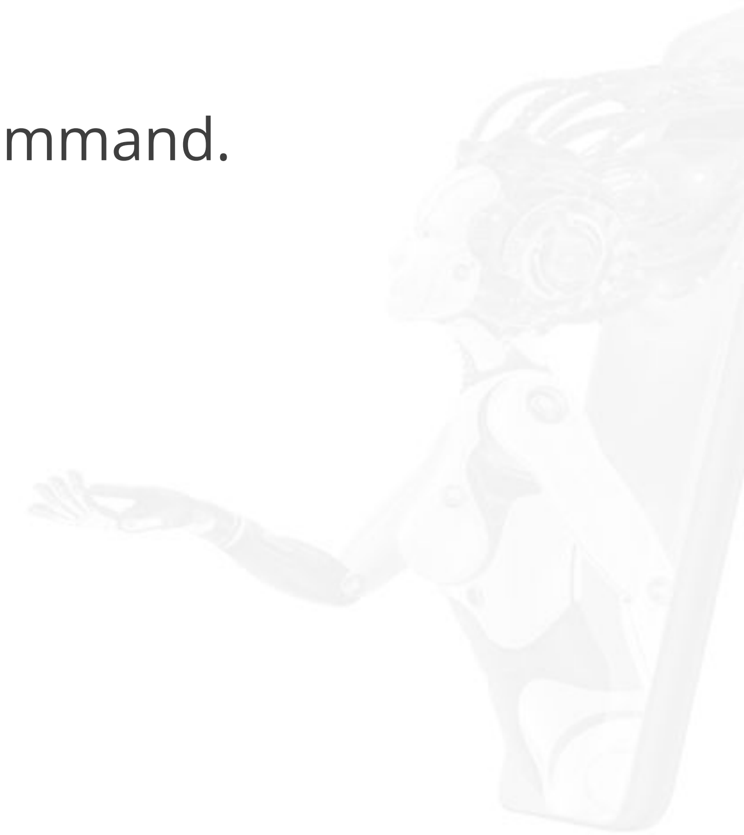
Running kubeadm without Internet Connection

To run kubeadm without an internet connection, the required Control Plane images must be pre-pulled.

The images can be listed and pulled using the **kubeadm config images** sub-command.

Demo

```
kubeadm config images list
kubeadm config images pull
```



Kubeadm Join

The **kubeadm join** command is used to initialize and join a Kubernetes worker to the cluster. The **kubeadm join [api-server-endpoint] [flags]** command supports the following options:

1. --apiserver-advertise-address
2. --apiserver-bind-port
3. --certificate-key
4. --config
5. --control-plane
6. --cri-socket
7. --discovery-file
8. --discovery-token

Kubeadm Join

9 --discovery-token-ca-cert-hash

10 --experimental-patches

11 --discovery-token-unsafe-skip-ca-verification

12 --h, --help

13 --ignore-preflight-errors

14 --node-name

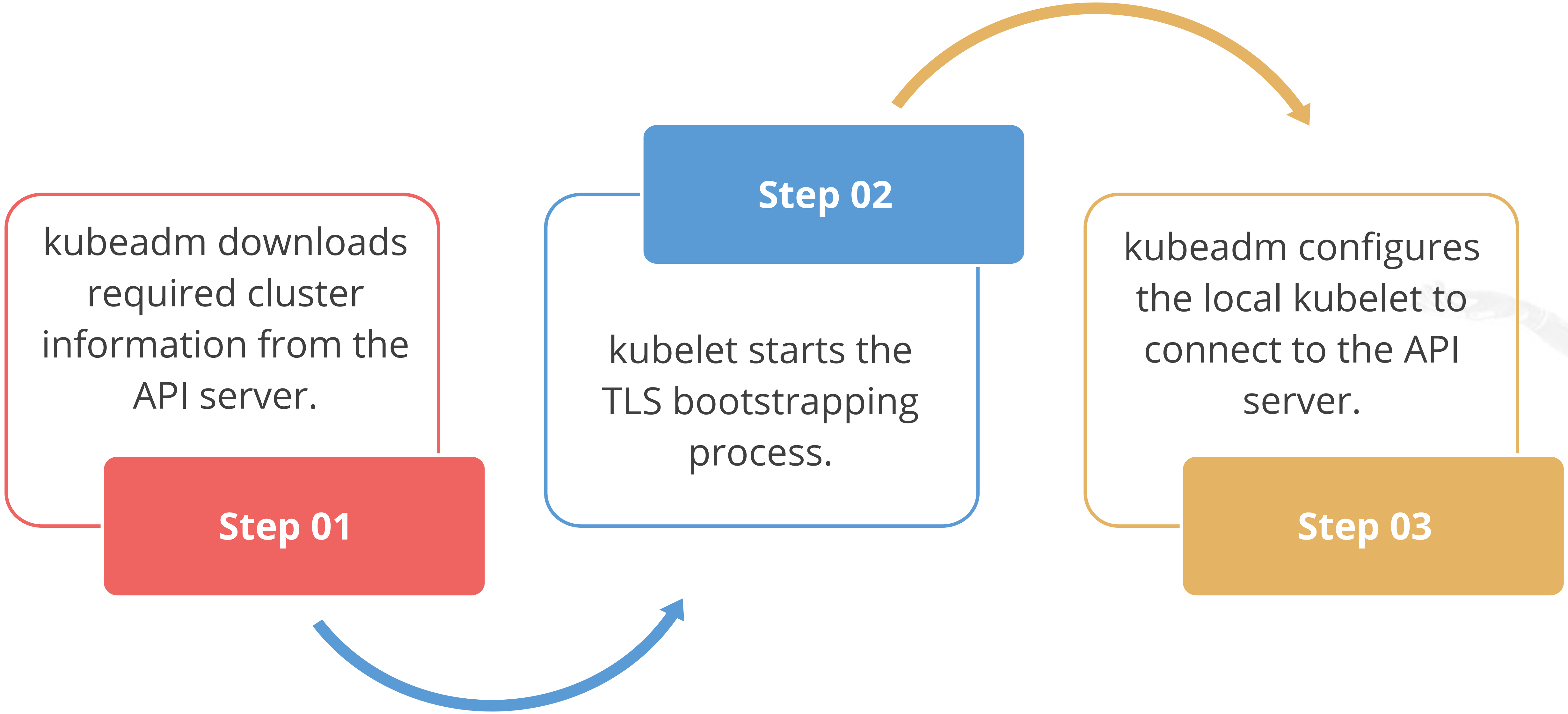
15 --skip-phases

16 --tls-bootstrap-token

17 --token

The Join Workflow

kubeadm join bootstraps a Kubernetes worker Node or a Control Plane Node and adds it to the cluster. This action completed in three steps for worker Nodes.



Step 01

kubeadm downloads required cluster information from the API server.

Step 02

kubelet starts the TLS bootstrapping process.

Step 03

kubeadm configures the local kubelet to connect to the API server.

Using Join Phases with kubeadm

To create a Control Plane in phases, use **kubeadm join phase** command.

Certain Control Plane phases have unique flags. To get a list of available options add **–help.**

An example is given below:

Demo

```
kubeadm join phase kubelet-start --help
```

Working with Kubeadm

Problem Statement:

Generate tokens and manage certificates using kubeadm.

Assisted Practice: Guidelines

Steps to demonstrate managing a cluster using Kubelet in Kubernetes:

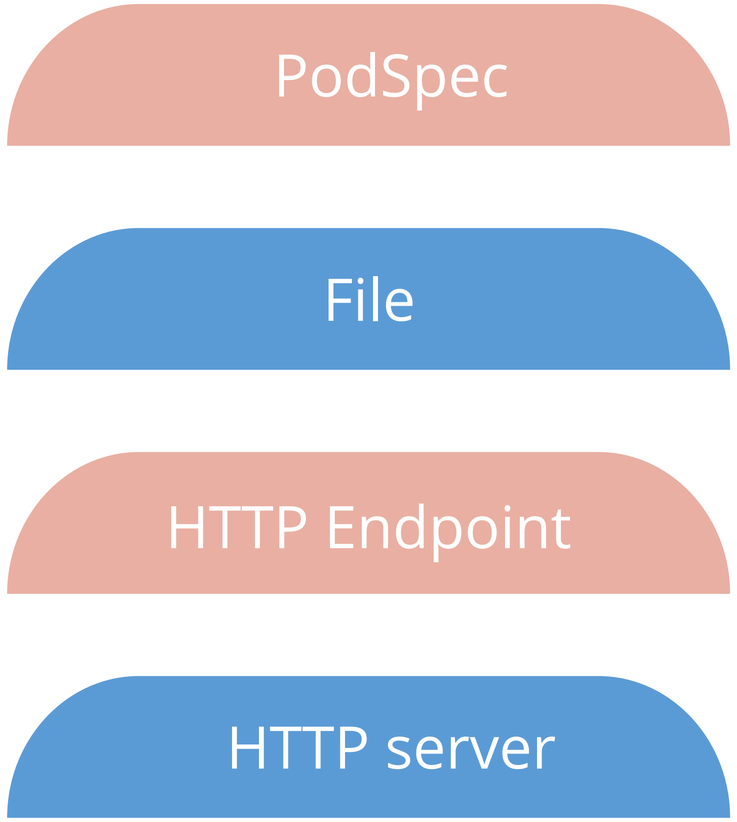
1. Generating tokens for kubeadm
2. Managing kubernetes certificates
3. Viewing the configuration details

Managing a Cluster Using kubelet

Overview

The kubelet is the primary Node agent that runs on each Node. It can register the Node with the apiserver using a hostname, a flag to override the hostname, or specific logic for a cloud provider.

A Container manifest may be provided to the Kubelet using the following:

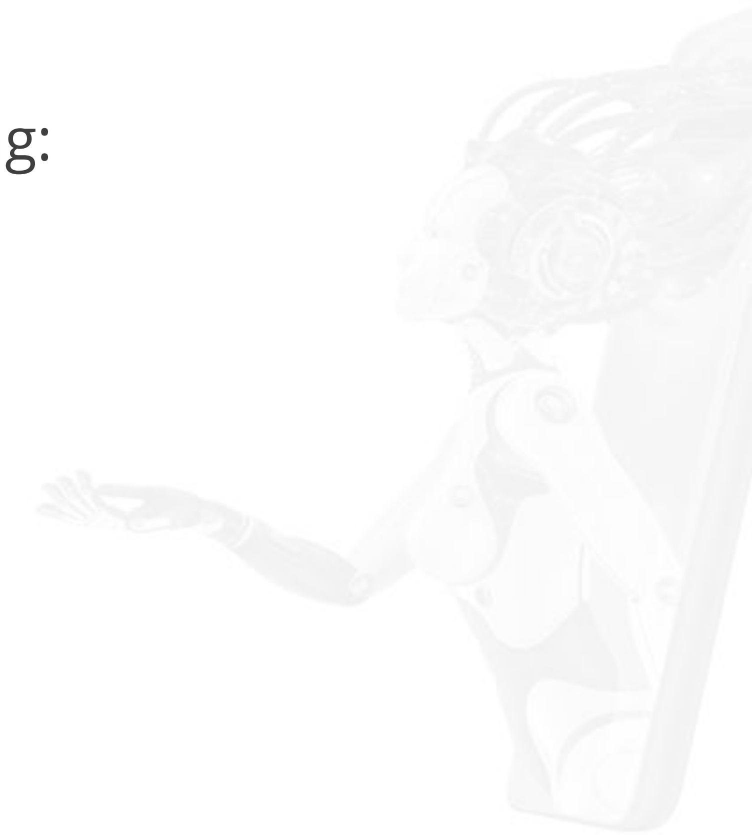


PodSpec

File

HTTP Endpoint

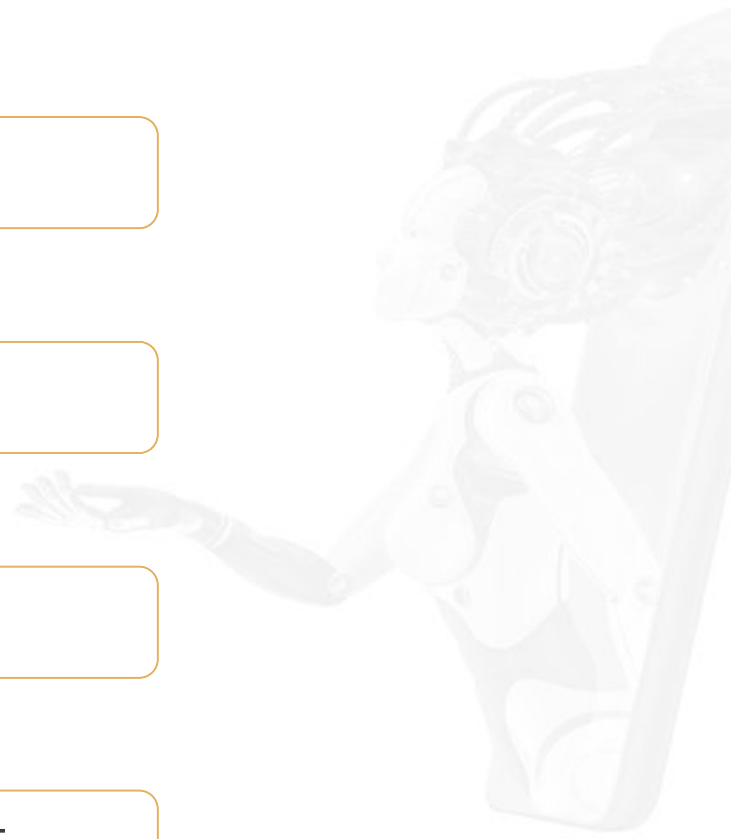
HTTP server



Options

There are non-deprecated and non-alpha version options available with the current version of Kubernetes.

- --add-dir-header
- --alsologtostderr
- --azure-container-registry-config
- --bootstrap-kubeconfig
- --docker-endpoint
- --cert-dir string
- --config
- --container-runtime
- --container-runtime-endpoint
- --dynamic-config-dir



Options

--enable-controller-attach-detach

--exit-on-lock-contention

-h, --help

--hostname-override

--housekeeping-interval

--image-credential-provider-bin-dir

--image-pull-progress-deadline

--image-service-endpoint

-- kubeconfig

-- log-backtrace-at traceLocation

--log-dir

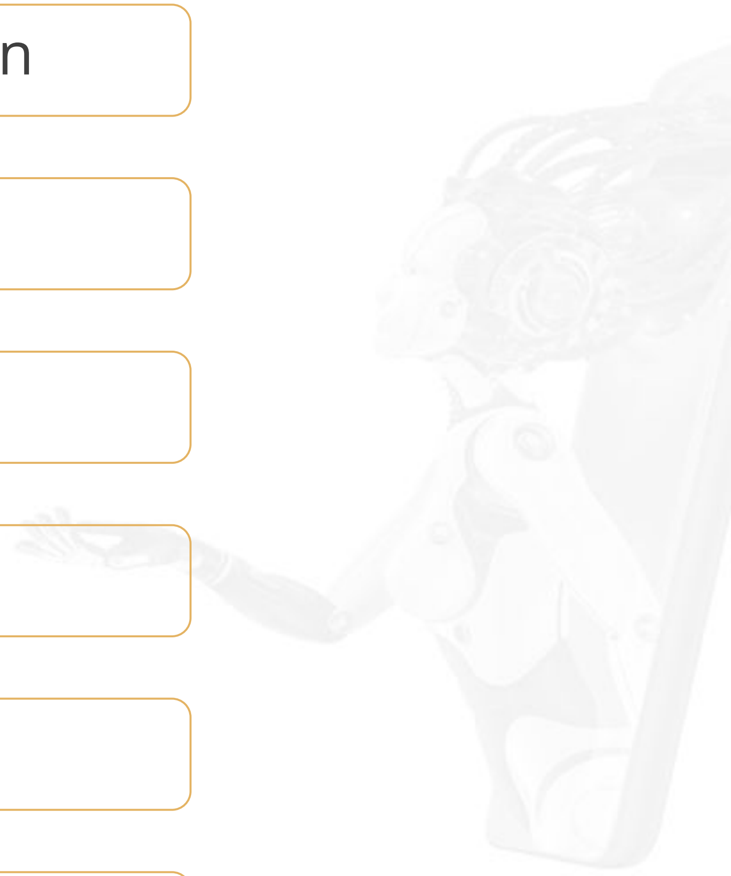
--log-file

--log-file-max-size

--log-flush-frequency

-- logtostderr

-- node-ip



Options

-- one-output

--pod-infra-container-image

-register-node

--register-with-taints

-- root-dir

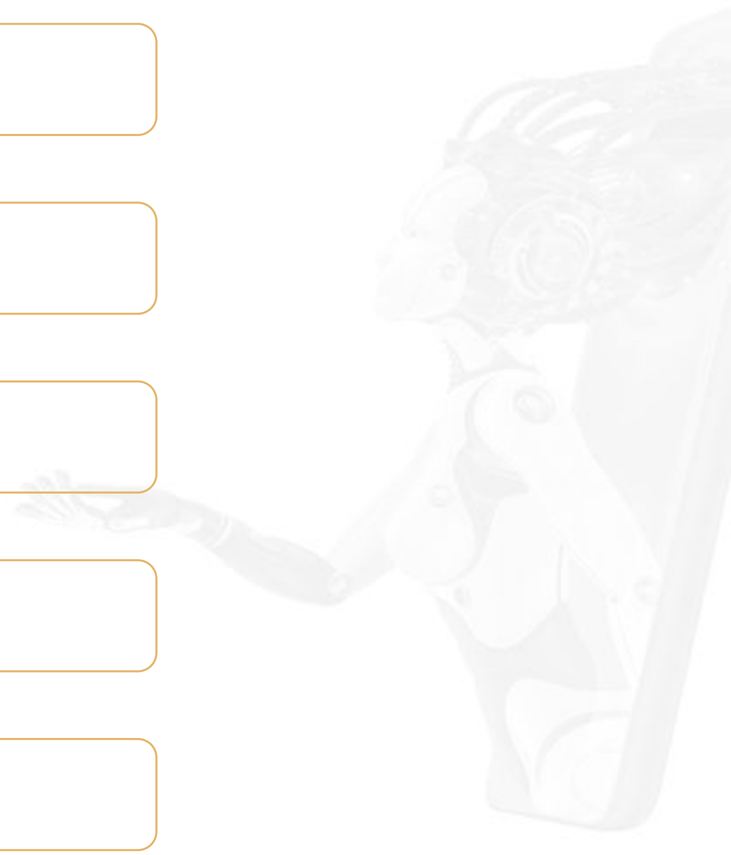
-- runtime-cgroups

-- skip-headers

--skip-log-headers

-- version

-- vmodule



Kubelet Authentication

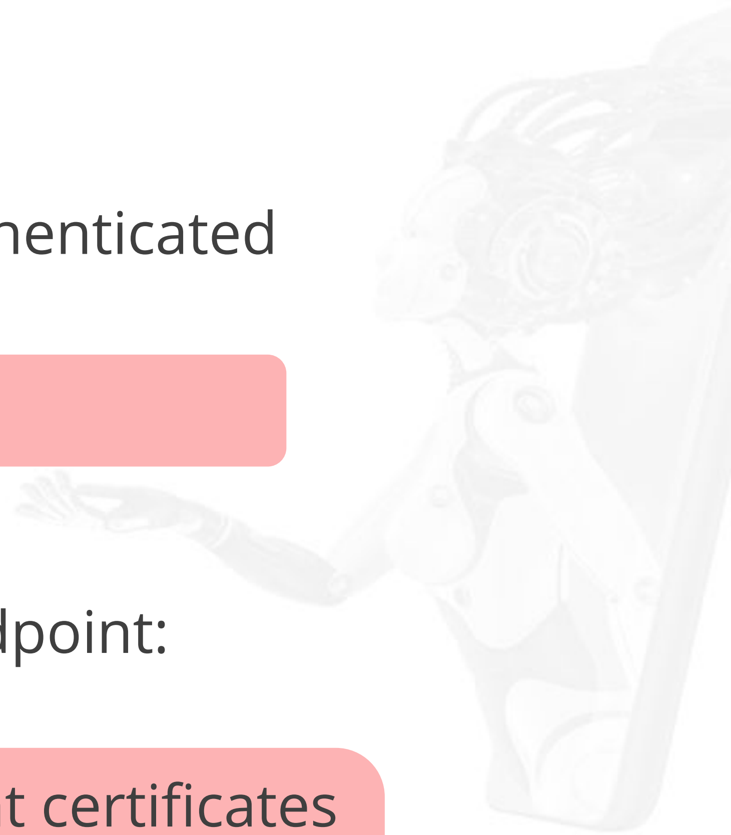
Requests to the kubelet's HTTPS Endpoint that are not rejected by other configured authentication methods are treated as anonymous requests. They will be given a username **system:anonymous** and a group **system:unauthenticated**.

To **disable anonymous access** and send 401 Unauthorized responses to unauthenticated requests:

Start the kubelet with the **--anonymous-auth=false** flag

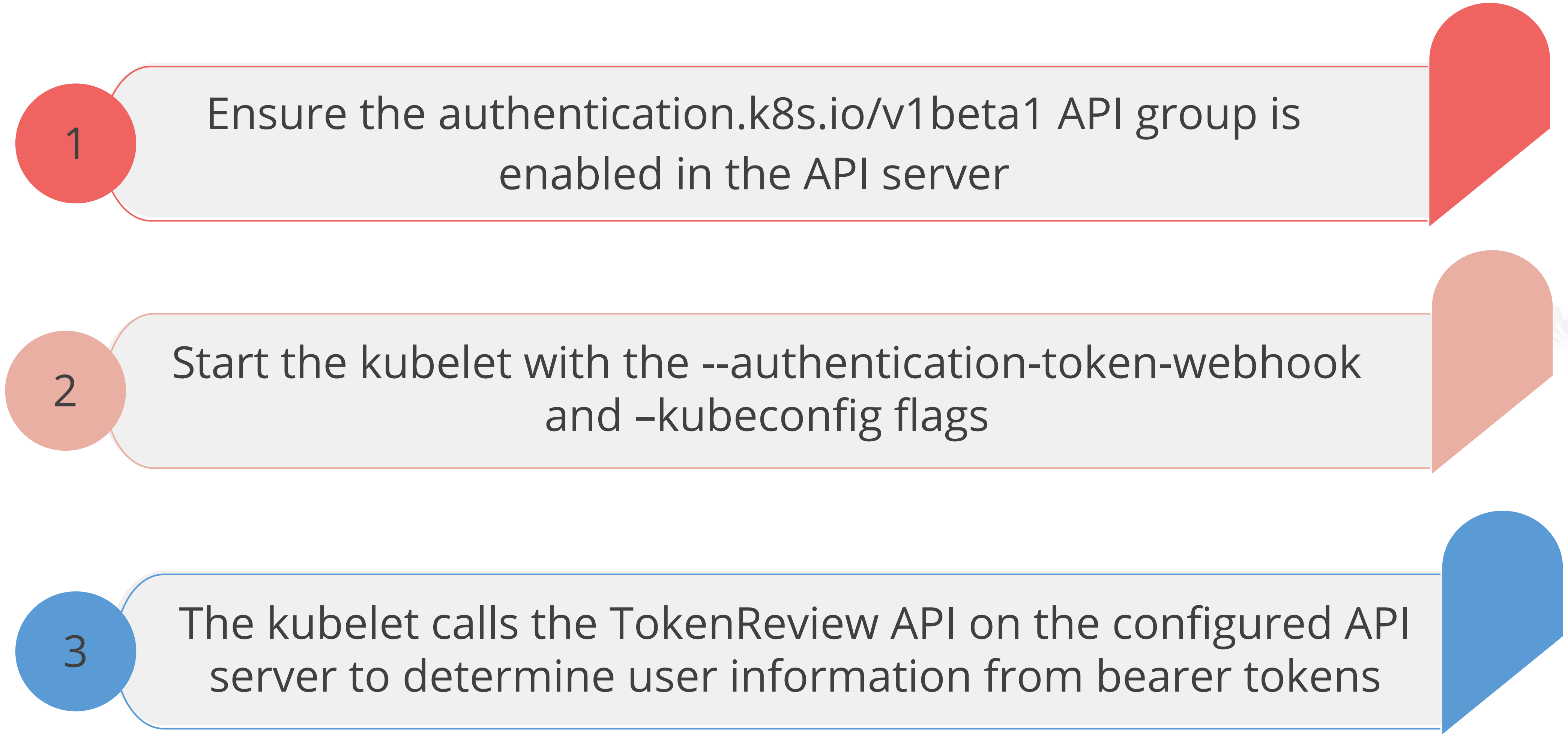
To **enable X509 client certificate authentication** to the kubelet's HTTPS Endpoint:

- Start the kubelet with the **--client-ca-file** flag, providing a CA bundle to verify client certificates with
- Start the apiserver with **--kubelet-client-certificate** and **--kubelet-client-key** flags



Kubelet Authentication

Here are a few steps to **enable API bearer tokens** to be used to authenticate to the kubelet's HTTPS Endpoint.



1 Ensure the authentication.k8s.io/v1beta1 API group is enabled in the API server

2 Start the kubelet with the --authentication-token-webhook and –kubeconfig flags

3 The kubelet calls the TokenReview API on the configured API server to determine user information from bearer tokens

Kubelet Authorization

A successfully authenticated request will be authorized, the default mode being **AlwaysAllow**.

Access to the kubelet API can be subdivided.
Subdivisions occur under the following conditions:

- Anonymous auth is enabled but limit the ability of anonymous users to call the kubelet API.
- Bearer token auth is enabled but limit the ability of arbitrary API users to call the kubelet API.
- Client certificate auth is enabled, but only some of the client certificates must be allowed to use the kubelet API.

Kubelet Authorization

To subdivide access to the kubelet API, delegate authorization to the API server.

The kubelet calls the **SubjectAccessReview** API on the configured API server to determine whether each request is authorized

Start the kubelet with the **--authorization-mode=Webhook** and the **--kubeconfig** flags.

Ensure the **authorization.k8s.io/v1beta1** API group is enabled in the API server.

Kubelet Authorization

To authorize API requests using Request Attributes, kubelet adopts an approach similar to the API server. The verb is determined from the incoming request’s HTTP web.

HTTP verb	Request verb
POST	create
GET.HEAD	get
PUT	update
PATCH	patch
DELETE	delete

Managing Clusters with Kubelet

Problem Statement:

Learn cluster event monitoring using kubelet.

Assisted Practice: Guidelines

Steps to demonstrate managing a cluster using Kubelet in Kubernetes:

1. Define the container runtime
2. Modify configuration files

TECHNOLOGY

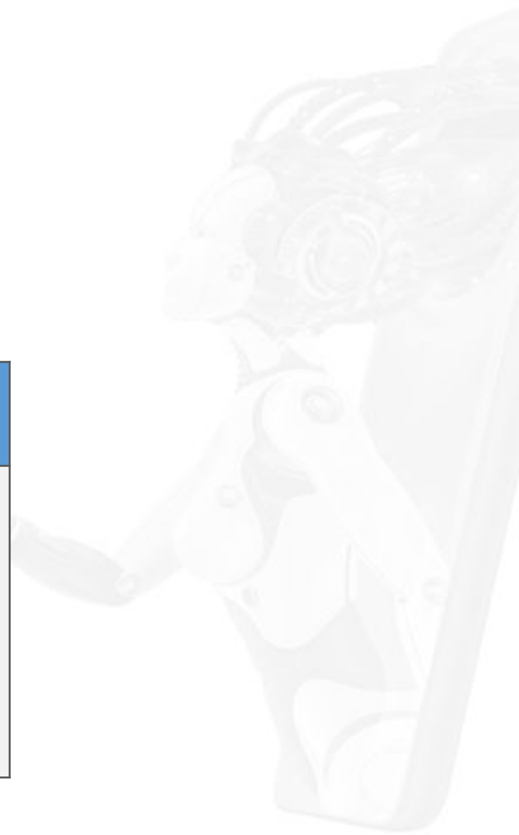
Role-Based Access Controller

RBAC

Role-based Access Controller (RBAC) is used to configure fine-grained and specific sets of permissions within a Kubernetes cluster.

Extensions or declarations can be used to define roles and permissions.

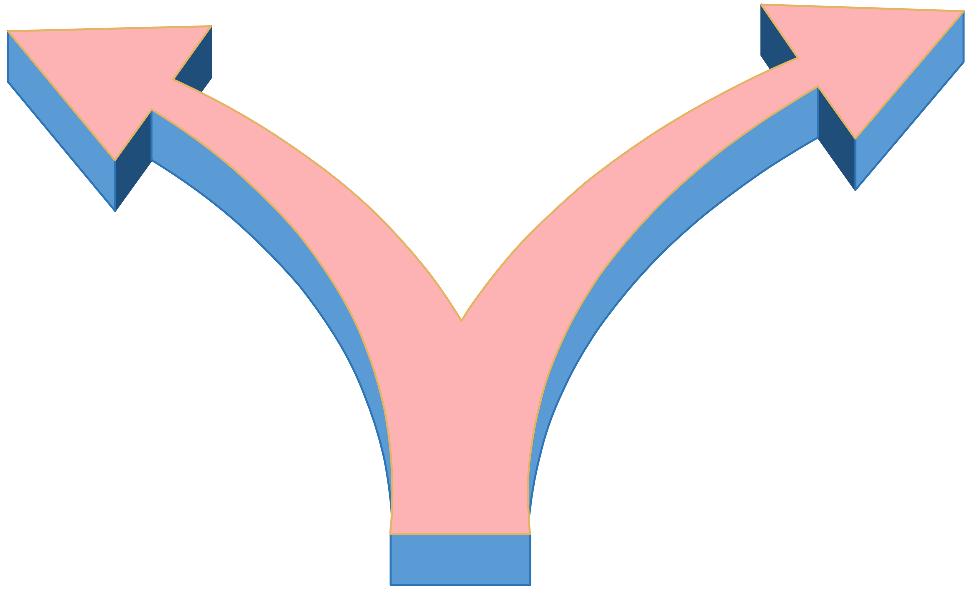
Subject	Operations	Resources
User	List, Get, Create, Update, Delete, watch, patch, post, put	Pods, Nodes, configmaps, secrets, deployment
Group		
Service account		



Operations and Subjects

Access control also connects operations and subjects.

HTTP verbs sent to the API represent the operations on resources.



Subjects are the actors in the Kubernetes API and RBAC.

Resource vs. Non-Resource Requests

Additional attributes can be added by the Kubernetes API server as a resource or a non-resource request. Some of the attributes that can be added include:

Request Type	Attribute	Description
Authentication	User	The user string
	Group	The list of group names
	Extra	A map of arbitrary keys
API	Non-resources or API resources flag	

Resource vs. Non-Resource Requests

Request Type	Attribute	Description
API resource request	API request verb	Lowercase resource verb
	Namespace	The namespace
	API group	The API group being accessed
	Resource	The resources ID
	Subresource	The subresources
Non-resource request	HTTP request verb	Lowercased HTTP method
	Request path	Non-resources request path
Verbs	Common API resource request	Get, list, watch, create, update, patch, delete, delete collection
	Special API resource request	Use, bind, escalate, impersonate, userextras

Authentication Methods for Kubernetes

There are several authentication mechanisms available in Kubernetes. Some of these include:

X509 Client Certs

Bearer token

HTTP Basic auth

Authentication proxy

Impersonate

Implementing Role-Based Access Controller

Problem Statement:

Learn how to implement Role-Based Access Controller.

ASSISTED PRACTICE

Assisted Practice: Guidelines

Steps to demonstrate Role-Based Access Controller in Kubernetes:

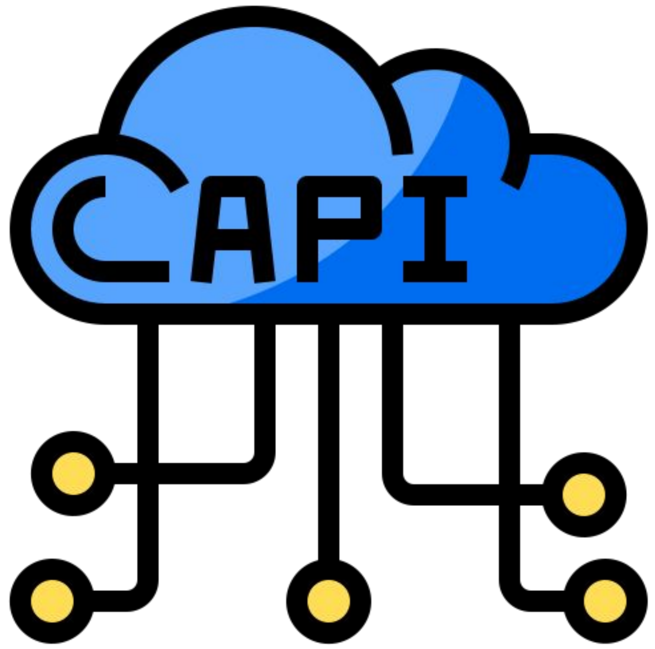
1. Create Client Certificate
2. Add user credentials to the kubeconfig file
3. Create Role
4. Create Role Binding

TECHNOLOGY

API Server

Overview

API server is the core of Kubernetes Control Plane. It exposes an HTTP API that lets end users, clusters, and external components communicate with one another.



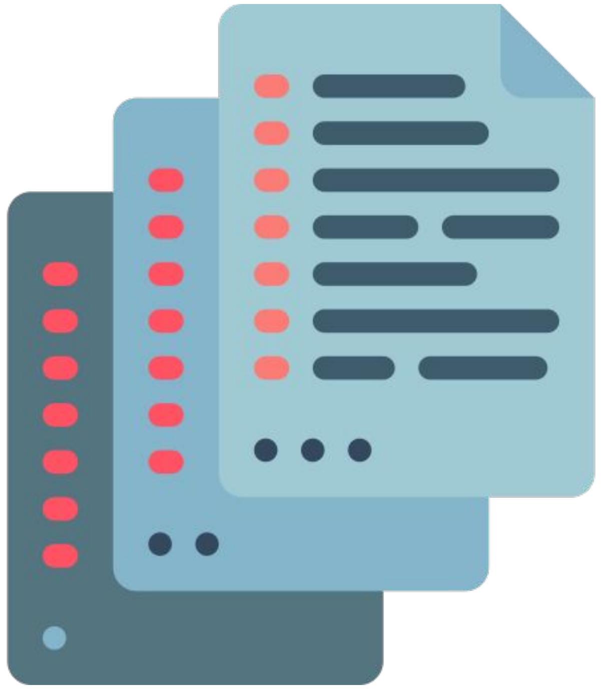
OpenAPI Specification

The Kubernetes API server uses **/openapi/v2 endpoint** to provide OpenAPI spec. The response format can be requested using request headers as shown below:

Header	Possible values	Notes
Accept-Encoding	gzip	not supplying this header is also acceptable
Accept	application/com.github.proto-openapi.spec.v2@v1.0+protobuf	mainly for intra-cluster use
	application/json	default
	*	serves application/json

API Groups and Versioning

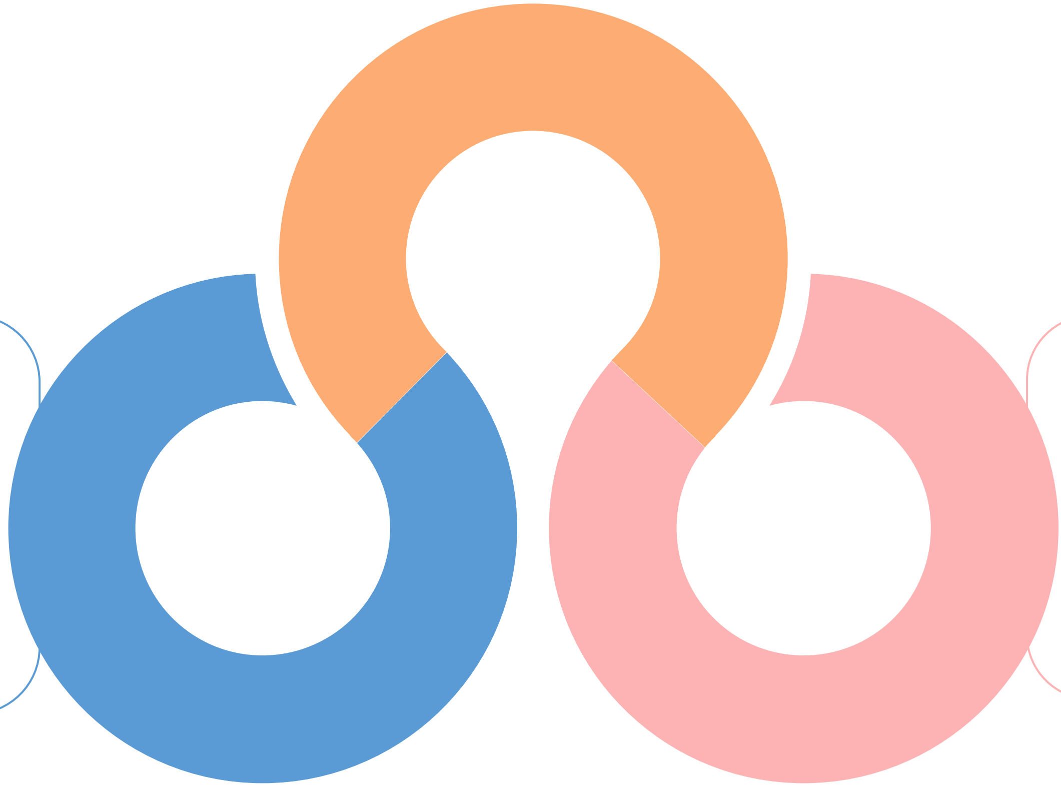
Multiple API versions are supported by Kubernetes. This helps restructure resource representations and eliminate fields.



API resources are distinguished by their API group, resource type, namespace, and name.

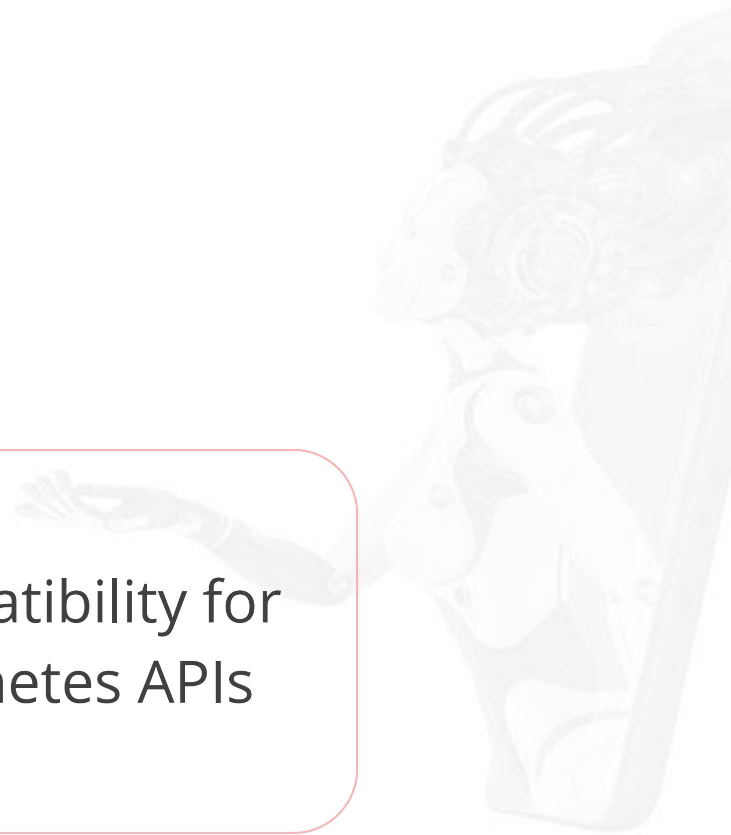
API Changes

Kubernetes is designed to adapt to changes.



Supports addition of new API resources and resource fields

Provides compatibility for official Kubernetes APIs



API Extension

The Kubernetes API can be extended in one of two ways.

1 Using Custom resources

2 Implementing an aggregation layer

Understanding the Working of API Servers

Problem Statement:

Learn how to work with API servers.

Assisted Practice: Guidelines

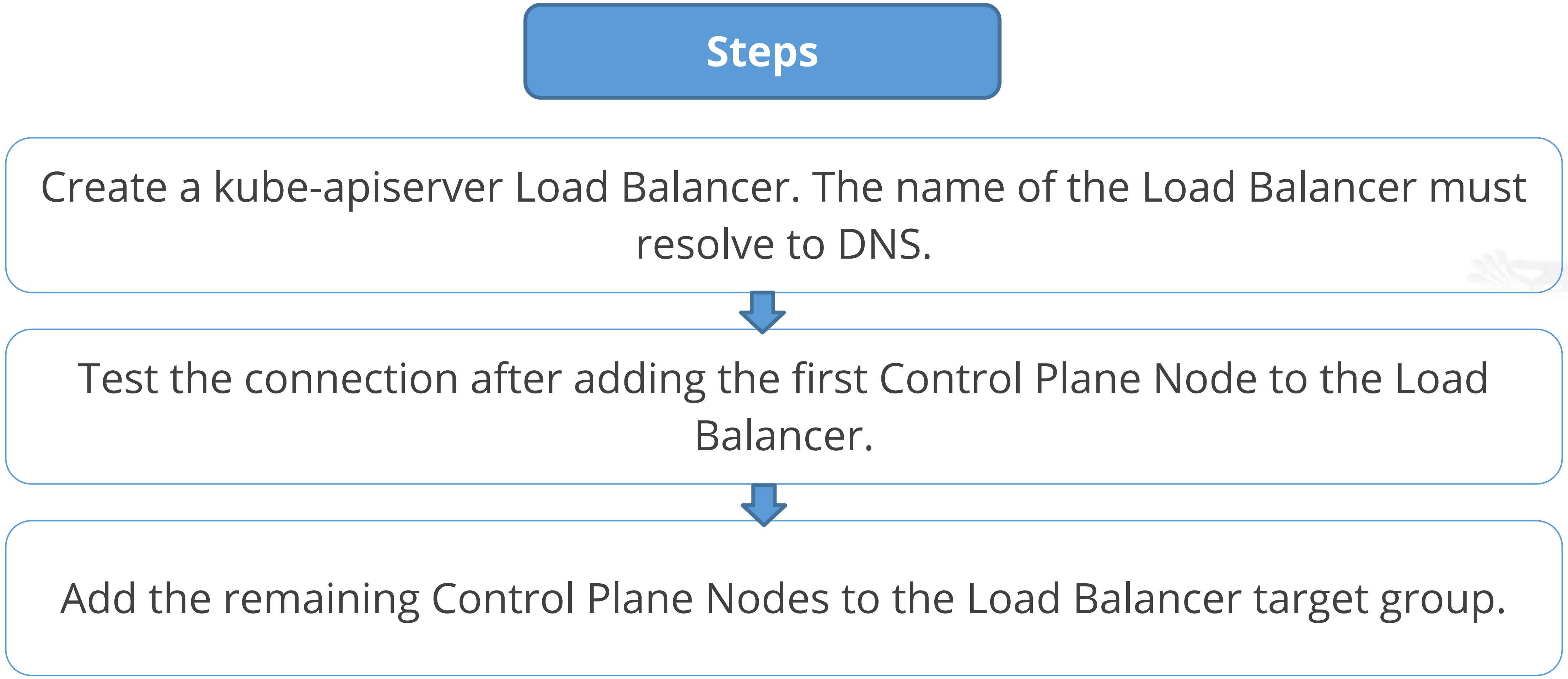
Steps to demonstrate API Server in Kubernetes:

1. Define Custom API Resources with CustomResourcesDefinition
2. Create a custom yaml file
3. Type ***kubectl apply -f customresourcedef.yaml***
4. Check the newly created REST Endpoint by typing ***kubectl api-resources***

Achieving High Availability

Setup

A highly available Kubernetes cluster can be set up with kubeadm by applying two different approaches, namely, using stacked Control Plane Nodes and an external etcd cluster.



Steps

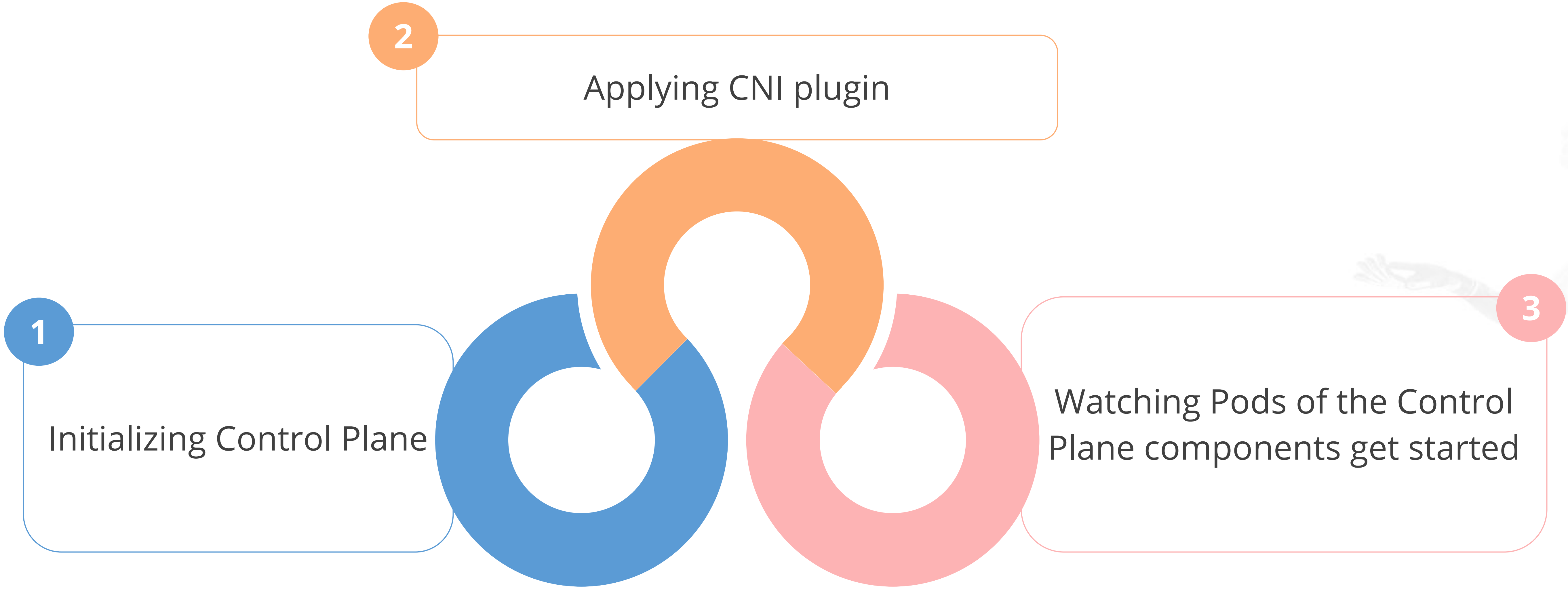
Create a kube-apiserver Load Balancer. The name of the Load Balancer must resolve to DNS.

Test the connection after adding the first Control Plane Node to the Load Balancer.

Add the remaining Control Plane Nodes to the Load Balancer target group.

Stacked Control Plane and etcd Nodes

Steps for the first Control Plane Node:



1 Initializing Control Plane

2 Applying CNI plugin

3 Watching Pods of the Control Plane components get started

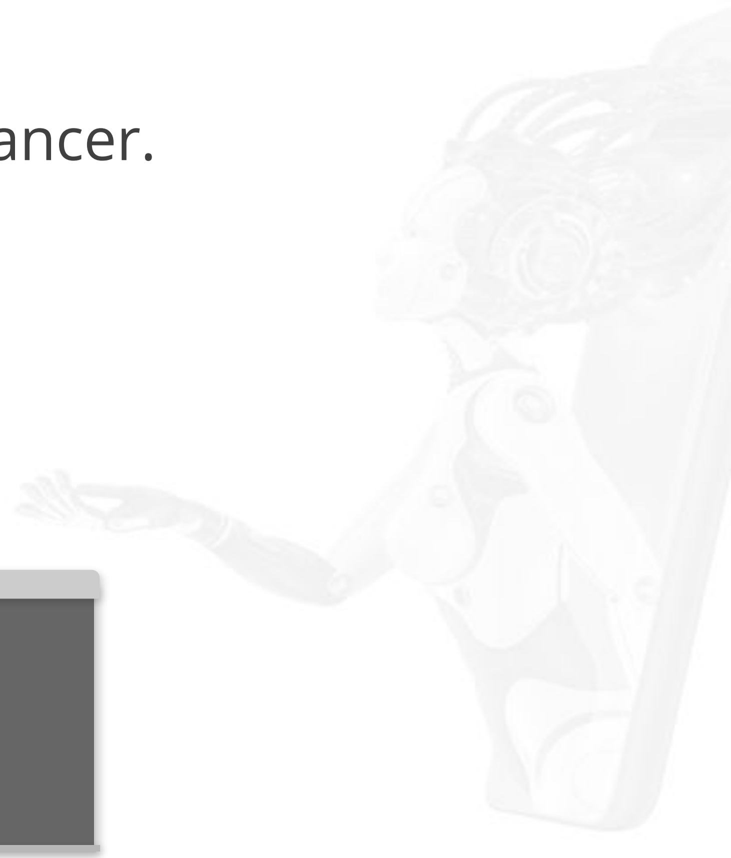
Stacked Control Plane and etcd Nodes

To set the Kubernetes version, use **--kubernetes-version** flag. As per Kubernetes recommendations, **kubeadm**, **kubelet** and **kubectl** versions should match.

Set the **--control-plane-endpoint** flag address/DNS and port of the Load Balancer.

Command to initialize the Control Plane is as given below:

```
sudo kubeadm init --control-plane-endpoint "LOAD_BALANCER_DNS:LOAD_BALANCER_PORT"
--upload-certs
```



Stacked Control Plane and etcd Nodes

To upload the certificates to be shared across all the Control Plane instances, **--upload-certs** flag is used.

Use the following command to re-upload the certificates and generate a new decryption key on a Control Plane Node that is joined to the cluster:

```
sudo kubeadm init phase upload-certs --upload-certs
```

Stacked Control Plane and etcd Nodes

To apply CNI, CNI provider configuration must correspond to the Pod CIDR specified in the kubeadm configuration file. For instance, to apply Weave Net CNI, use the following command:

```
kubectl apply -f "https://cloud.weave.works/k8s/net?k8s-version=$(kubectl version | base64 |
tr -d '\n’)”
```

To watch the Pods of the Control Plane components, use:

```
kubectl get pod -n kube-system -w
```

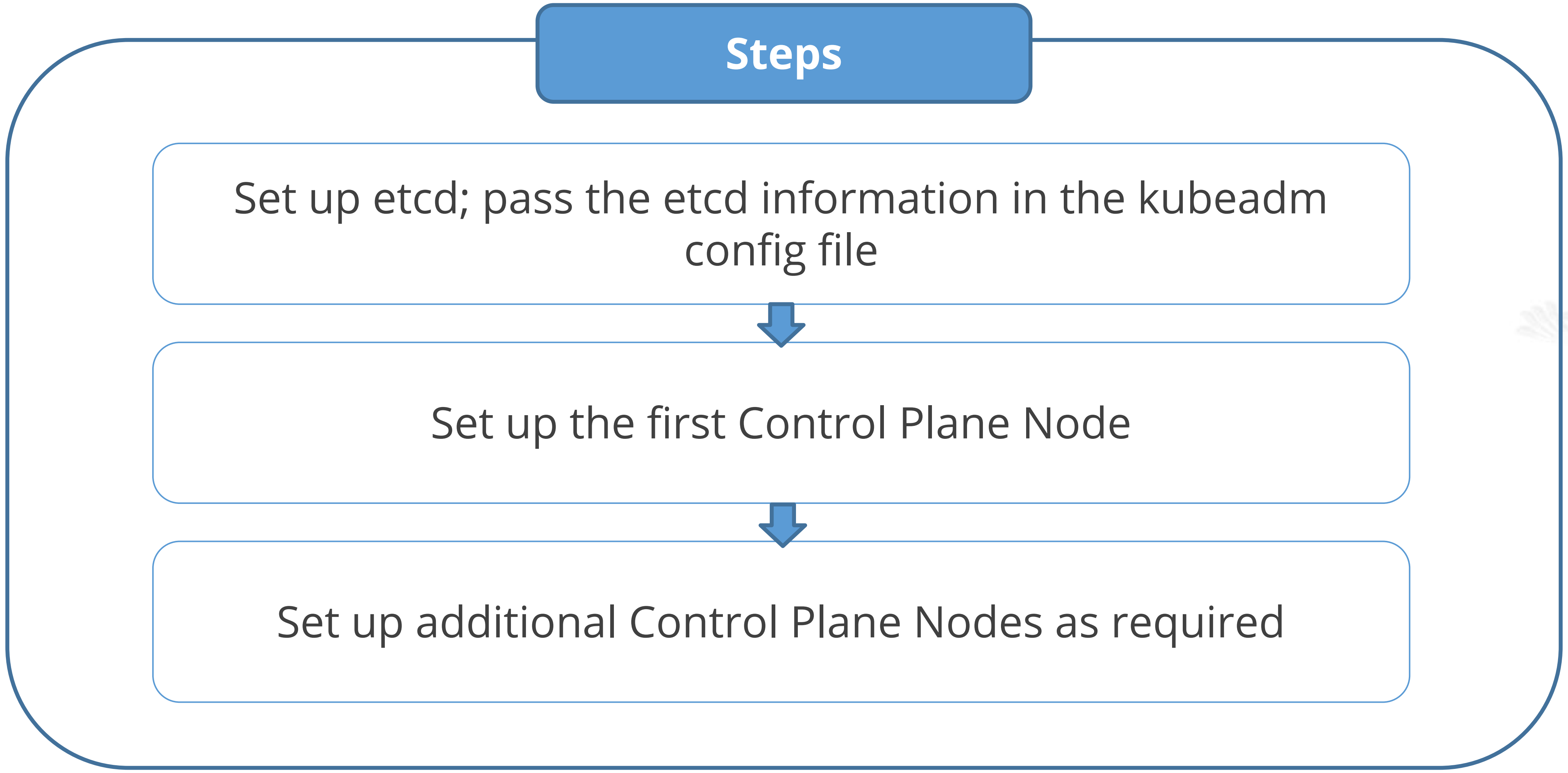
Setting Up Additional Control Plane Nodes

Execute the join command that is generated by kubeadm when the init command on the first Node is run. The command is shown below:

```
sudo kubeadm join 192.168.0.200:6443 --token 9vr73a.a8uxyaju799qwdjv
--discovery-token-ca-cert-hash
sha256:7c2e69131a36ae2a042a339b33381c6d0d43887e2de83720eff5359e26aec866 --control-plane
--certificate-key f8902e114ef118304e561c3ecd4d0b543adc226b7a07f675f56564185ffe0c07
```

External etcd Nodes

Setting up a cluster with external etcd Nodes differs from setting up with stacked etcd in one aspect.



Steps

Set up etcd; pass the etcd information in the kubeadm config file

Set up the first Control Plane Node

Set up additional Control Plane Nodes as required

Setting Up etcd Cluster

To set up the etcd cluster, set up SSH and copy the files given below from any etcd Node in the cluster to the first Control Plane Node.

Demo

```
export CONTROL_PLANE=“ubuntu@10.0.0.7”

scp /etc/kubernetes/pki/etcd/ca.crt “${CONTROL_PLANE}”;

scp /etc/kubernetes/pki/apiserver-etcd.client.crt “${CONTROL_PLANE}”;

scp /etc/kubernetes/pki/apiserver-etcd.client.key “${CONTROL_PLANE}”;
```

Then the value of **CONTROL_PLANE** must be replaced with **user@host** of the first Control Plane Node.

Setting Up First Control Plane Node

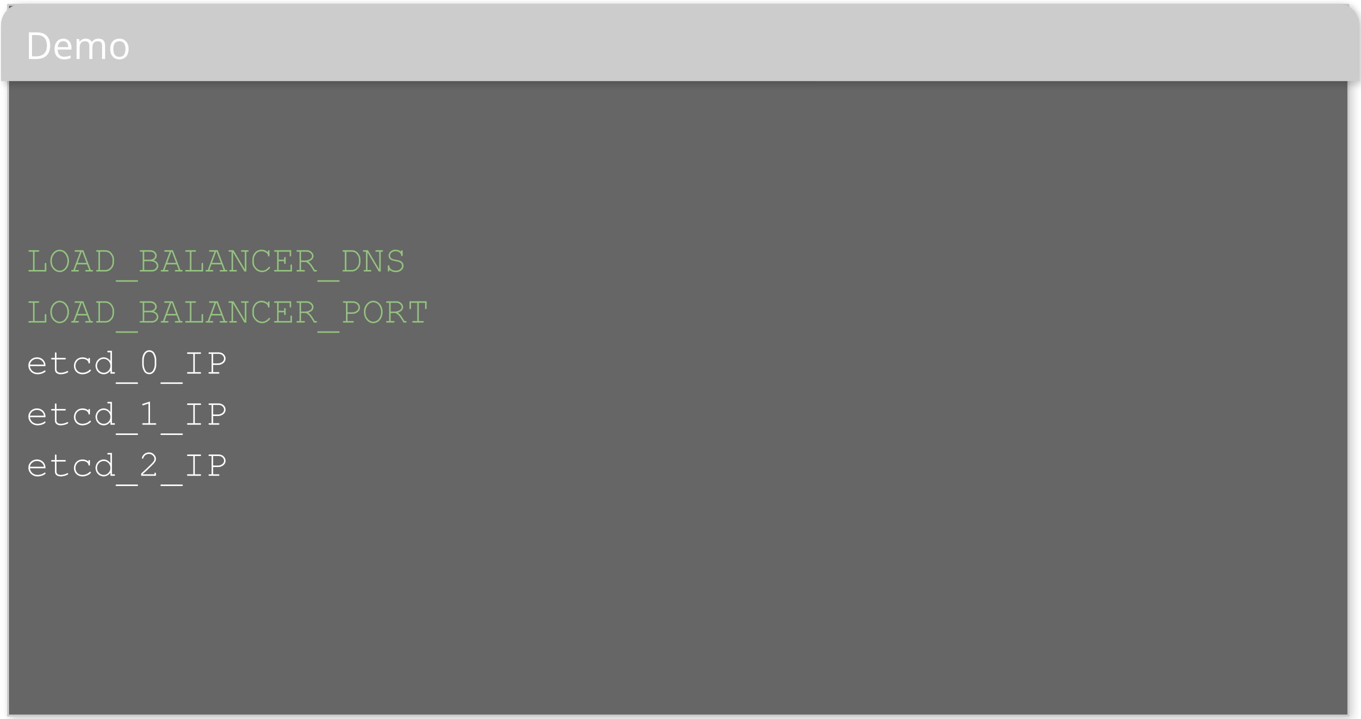
kubeadm-config.yaml with the content shown below must be created.

Demo

```
Apiversion: kubeadm.k8s.io/v1beta2
Kind: clusterConfiguration
kuberenetesVersion: stable
controlPlaneEndpoint: “LOAD_BALANCER_DNS:LOAD_BALANCER_PORT ”
etcd:
    external:
    endpoint:
    https://etcd_0_IP:2379
     https://etcd_1_IP:2379
     https://etcd_2_IP:2379
    caFile: /etc/kubernetes/pki/etcd/ca.crt
    certfilr: /etc/kubernetes/pki/apiserver-etcd-client.crt
    keyfile: /etc/kubernetes/pki/apiserver-etcd-client.key
```

Setting Up First Control Plane Node

The following variables in the config template must be replaced with the appropriate values for the cluster:

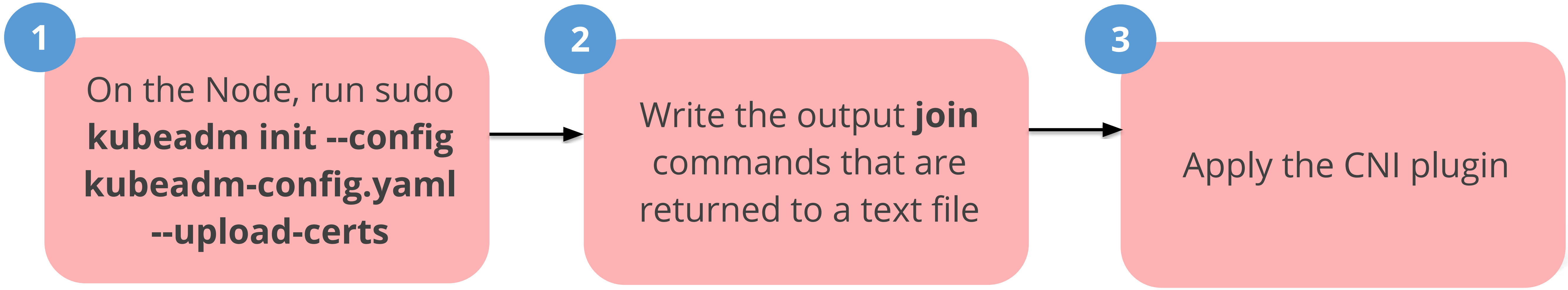


Demo

```
LOAD_BALANCER_DNS
LOAD_BALANCER_PORT
etcd_0_IP
etcd_1_IP
etcd_2_IP
```

Setting Up First Control Plane Node

Once the .yaml file is created, there are three steps to be followed:



1 On the Node, run sudo **kubeadm init --config kubeadm-config.yaml --upload-certs**

2 Write the output **join** commands that are returned to a text file

3 Apply the CNI plugin

Common Tasks After Bootstrapping Control Plane

Installing workers

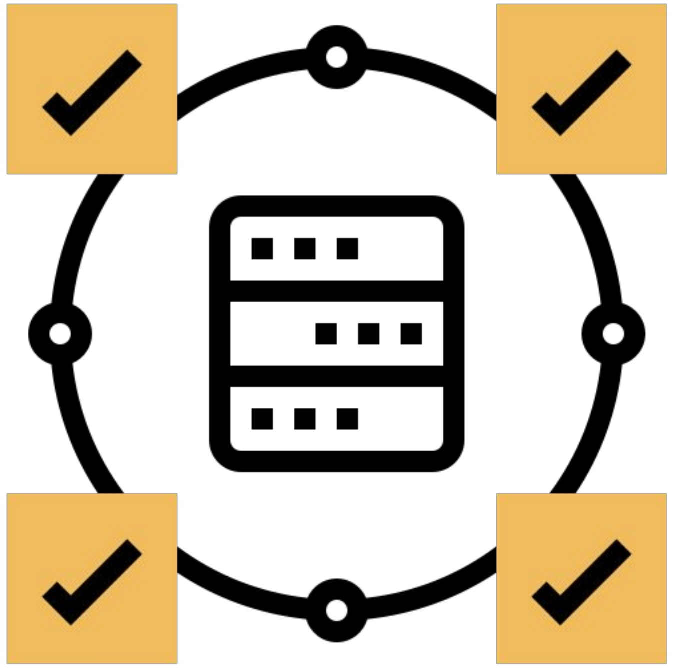
The command generated as the output of kubeadm init command must be used to install worker Nodes to the cluster. This can be done as shown below:

```
sudo kubeadm join 192.168.0.200:6443 --token 9vr73a.a8uxyaju799qwdjv
--discovery-token-ca-cert-hash
sha256:7c2e69131a36ae2a042a339b33381c6d0d43887e2de83720eff5359e26aec866
```

Common Tasks After Bootstrapping Control Plane

Manual Certificate Distribution

Certificates must be manually copied from the primary Control Plane to the joining plane Nodes if –upload-certs flag is not used during initialization.



Manual Certificate Distribution

To distribute certificates using **ssh** and **scp:**

1 Enable ssh-agent on the main device that has access to all other Nodes in the system

3 Check the connection between Nodes

2 Add ssh identity to the session

4 Run the script to copy the certificates from the first Control Plane to the others

5 Run the script on each Control Plane Node to move the copied certificates from home directory to /etc/kubernetes/api directory

Setting Up a Highly Available Cluster

Problem Statement:

Learn how to set up a highly available cluster.

ASSISTED PRACTICE

Assisted Practice: Guidelines

Steps to demonstrate highly available clusters in Kubernetes.

1. Set up a highly available cluster using ***--control-plane-endpoint*** flag
2. Join the nodes to the control plane
3. Add a CNI plug

Backup and Restoration of etcd Cluster Data

Starting a Single-Node etcd Cluster

To start a Single-Node cluster, run the following command:

```
etcd --listen-client-urls=http://$PRIVATE_IP:2379  \
   --advertise-client-urls=http://$PRIVATE_IP:2379
```

Then, start the Kubernetes API server with the flag **--etcd-servers=$PRIVATE_IP:2379**.

Note

Set **PRIVATE_IP** to the etcd client IP

Starting a Multi-Node etcd Cluster

The durability and availability of etcd significantly increases when it is run as a multi-Node cluster.

To start a multi-Node etcd cluster, run the following command:

```
etcd
--listen-client-urls=http://$IP1:2379,http://$IP2:2379,http://$IP3:2379,http://$IP4:2379,http
://$IP5:2379
--advertise-client-urls=http://$IP1:2379,http://$IP2:2379,http://$IP3:2379,http://$IP4:2379,h
ttp://$IP5:2379
```

Then, start the Kubernetes API server with the flag **--etcd-servers=$IP1:2379,$IP2:2379,$IP3:2379,$IP4:2379,$IP5:2379**.

Note

IP<n> variables must be set to the client IP addresses.

Multi-Node etcd Cluster with Load Balancer

To run a Load Balancing etcd cluster, follow three steps.

An etcd cluster must be set up.

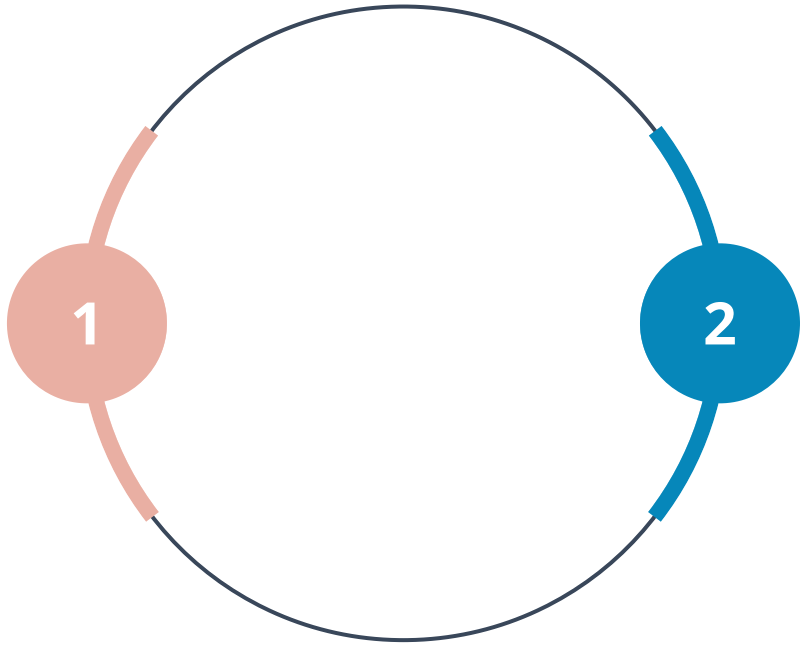
A Load Balancer must be configured.

The Kubernetes API servers must be started with the flag **--etcd-servers** set to **$LB:2379**.

Securing etcd Clusters

To secure etcd clusters, firewall may be set up or etcd-provided security features may be used.

Securing etcd clusters involves:



Securing communication

Limiting access of etcd clusters

Securing Communication

Configuring etcd with secure peer communication

Set **flags --peer-key-file** and **--peer-cert-file** to **=peer.key** and **peer.cert** respectively.

Configuring etcd with secure client communication

Set **flags --key-file** and **--cert-file** to **k8sclient.key** and **k8sclient.cert** respectively.

Limiting Access of etcd Clusters

Access to etcd cluster must be restricted to the Kubernetes API servers through TLS authentication.

Example

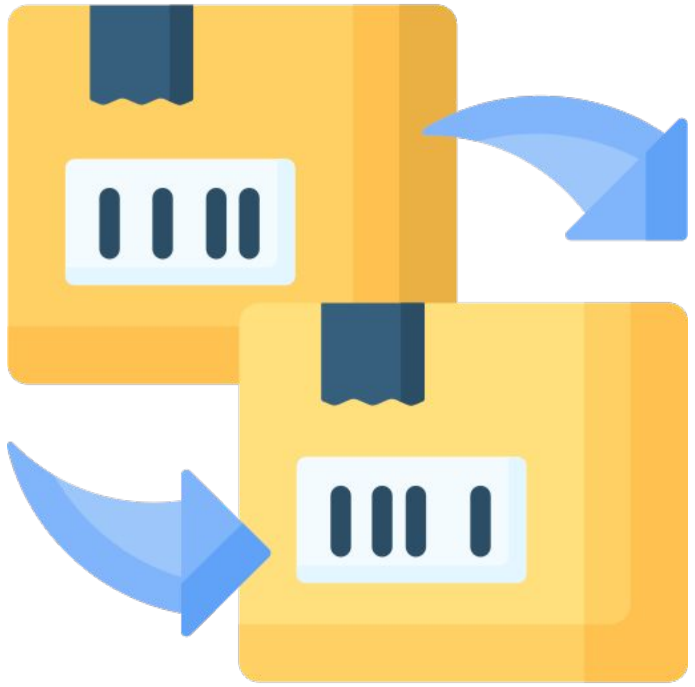
- Consider a key pair - **k8sclient.key** and **k8sclient.cert** trusted by the **CA etcd.ca**.
- etcd, configured with **--client-cert-auth** and TLS, uses system CAs or the CA passed by --trusted-ca-file flag to verify the client certificates.
- Setting **--client-cert-auth** and **--trusted-ca-file** to **true** and **etcd.ca** respectively restricts the access to clients k8sclient.cert.

To provide access to Kubernetes API servers, set **flags --etcd-certfile,--etcd-keyfile,** and **--etcd-cafile** to **k8sclient.cert, k8sclient.key,** and **ca.cert** respectively.

Replacement of a Failed etcd Member

Failed members of a cluster must be replaced to improve the overall health of the cluster.

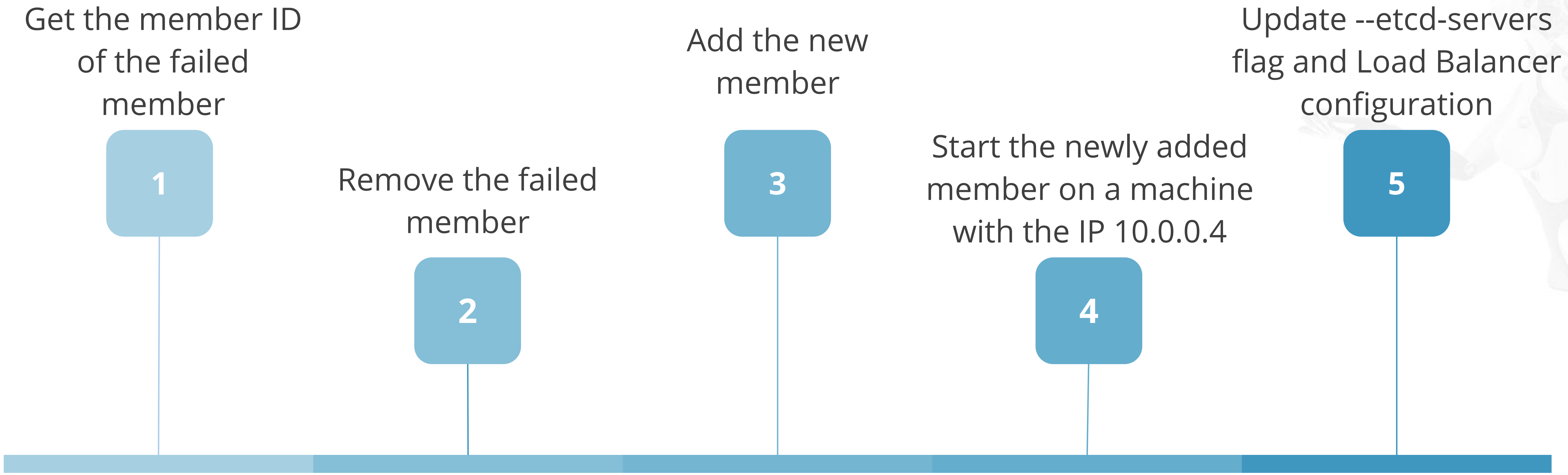
Replacement of a failed member comprises two steps, namely, removing the failed member and adding the new member.



Replacing a Failed etcd Member

Example: Replacing a failed member in a three-member etcd cluster member1=http://10.0.0.1, member2=http://10.0.0.2, and member3=http://10.0.0.3

Scenario: member1 fails and member4=http://10.0.0.4 gets added as a replacement.



1 Get the member ID of the failed member

2 Remove the failed member

3 Add the new member

4 Start the newly added member on a machine with the IP 10.0.0.4

5 Update --etcd-servers flag and Load Balancer configuration

Backing Up an etcd Cluster

An etcd cluster can be backed up either by using a built-in snapshot or a volume snapshot.

etcd built-in snapshot

Snapshot supported by etcd can be taken from a live member or by copying the member/snap/db file from the etcd data directory

Volume snapshot

Snapshots provided by storage volumes on which etcd is running

Scaling Up etcd Clusters

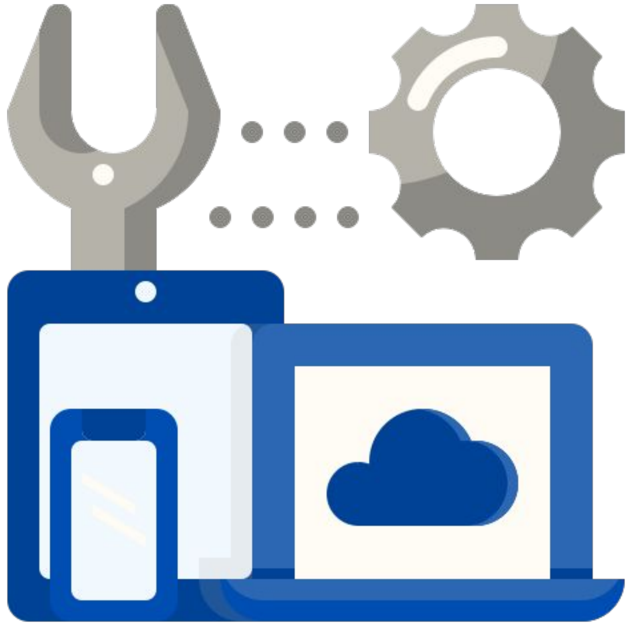
Scaling up etcd clusters increases cluster availability.

Scaling does not increase cluster performance or cluster capability.

Restoring an etcd Cluster

Snapshots taken from an etcd process (major.minor version) can be used to restore an etcd cluster. Kubernetes supports restoring a version from a different patch version of etcd.

To recover data from a failed cluster, a restore operation must be employed.



Backing Up and Restoring Etcd Cluster Data

Problem Statement:

Learn to back up and restore etcd cluster.

ASSISTED PRACTICE

Assisted Practice: Guidelines

Steps to demonstrate backup and restoration of etcd cluster data in Kubernetes:

1. Retrieve etcd cluster name
2. Restore cluster data

Version Upgrade on Kubernetes Cluster

Overview

At a high level, the upgrade workflow of a Kubernetes cluster comprises the following:

Primary Control Plane Node upgrade

Additional Control Plane Nodes upgrade

Worker Nodes upgrade

Determine Version to Upgrade

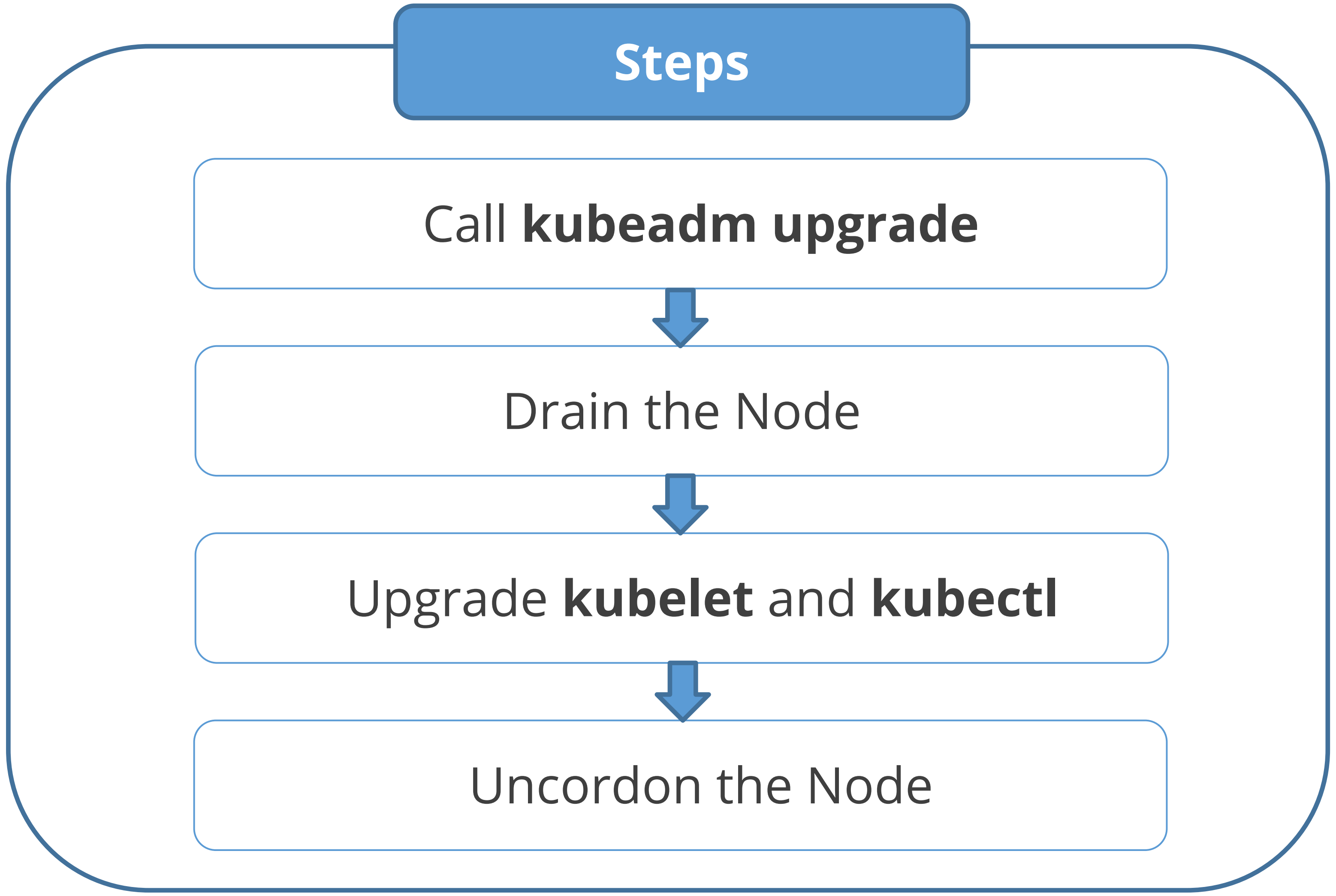
Enter the commands shown below to find the latest stable 1.21 version using the OS package manager.

Demo

```
apt update
apt-cache madison kubeadm
# find the latest 1.21 version in the list
# it should look like 1.21.x-00, where x is the latest patch
```

Upgrading Control Plane Nodes

Control Plane Nodes must be updated one at a time. The Node selected for upgrade must have **/etc/kubernetes/admin.conf** file.



Steps

1. Call **kubeadm upgrade**
2. Drain the Node
3. Upgrade **kubelet** and **kubectl**
4. Uncordon the Node

Upgrading Control Plane Nodes

Upgrade kubeadm for the first Control Plane Node as shown below:

Demo

```
# replace x in 1.21.x-00 with the latest patch version
apt-mark unhold kubeadm && \
apt-get update && apt-get install -y kubeadm=1.21.x-00 && \
apt-mark hold kubeadm
-
# since apt-get version 1.1 the following method can also be used.
apt-get update && \
apt-get install -y --allow-change-held-packages kubeadm=1.21.x-00
```

Upgrading Control Plane Nodes

Verify the working of kubeadm and check its version using the commands shown below:

Demo

```
#Verify that the download works and has the expected version:
kubeadm version

#Verify the upgrade plan:
kubeadm upgrade plan
```

Upgrading Control Plane Nodes

Use this command for other Control Plane Nodes:

Demo

```
#For other control plane nodes, use:

sudo kubeadm upgrade node

#instead of:

sudo kubeadm upgrade apply
```

Upgrading Control Plane Nodes

To prepare the Node for maintenance, mark the unschedulable and evict workloads.

Demo

```
# replace <node-to-drain> with the name of your node you are draining
kubectl drain <node-to-drain> --ignore-daemonsets
```

To upgrade kubelet and kubectl:

Demo

```
   # replace x in 1.21.x-00 with the latest patch version
    apt-mark unhold kubelet kubectl && \
    apt-get update && apt-get install -y kubelet=1.21.x-00 kubectl=1.21.x-00
&& \
    apt-mark hold kubelet kubectl
    -
    # since apt-get version 1.1 the following method can also be used
    apt-get update && \
    apt-get install -y --allow-change-held-packages kubelet=1.21.x-00
kubectl=1.21.x-00
```

Upgrading Control Plane Nodes

To restart the kubelet, use the command:

Demo

```
Restart the kubelet:

sudo systemctl daemon-reload
sudo systemctl restart kubelet
```

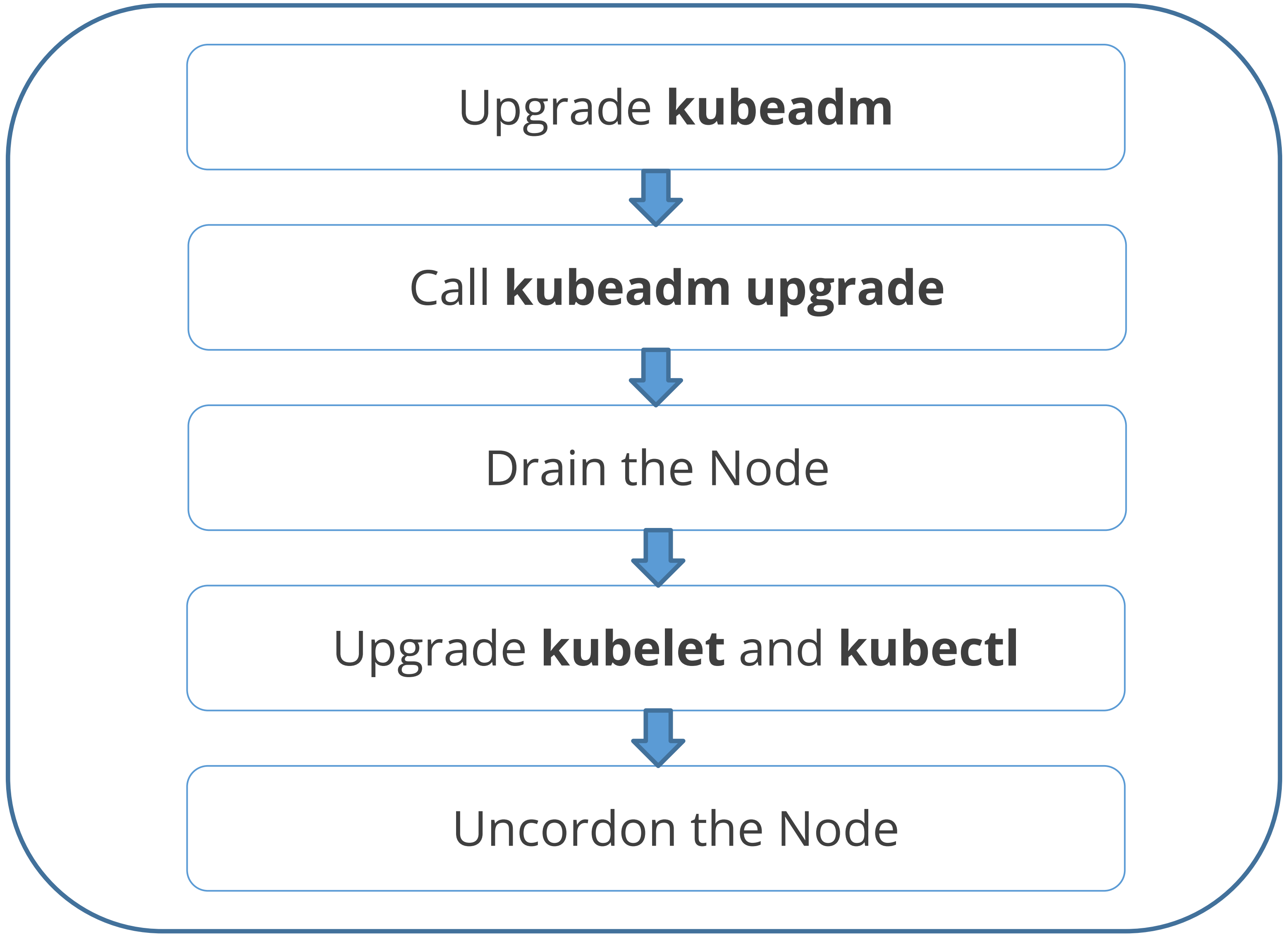
To uncordon the Node, enter:

Demo

```
# replace <node-to-drain> with the name of your node
kubectl uncordon <node-to-drain>
```

Upgrade Worker Nodes

Upgrading worker Nodes involves the following steps:



Upgrade **kubeadm**

Call **kubeadm upgrade**

Drain the Node

Upgrade **kubelet** and **kubectl**

Uncordon the Node

Upgrade Worker Nodes

To upgrade kubeadm:

Demo

```
# replace x in 1.21.x-00 with the latest patch version
apt-mark unhold kubeadm && \
apt-get update && apt-get install -y kubeadm=1.21.x-00 && \
apt-mark hold kubeadm
-
# since apt-get version 1.1 you can also use the following method
apt-get update && \
apt-get install -y --allow-change-held-packages kubeadm=1.21.x-00
```

Upgrade Worker Nodes

Worker Nodes can be upgraded and drained as follows:

Demo

```
#For worker nodes this upgrades the local kubelet configuration:
sudo kubeadm upgrade node


#Drain the node
# replace <node-to-drain> with the name of your node you are
draining
kubectl drain <node-to-drain> --ignore-daemonsets
```

Upgrade Worker Nodes

Upgrading kubelet and kubectl can be done as follows:

Demo

```
# replace x in 1.21.x-00 with the latest patch version
apt-mark unhold kubelet kubectl && \
apt-get update && apt-get install -y kubelet=1.21.x-00
kubectl=1.21.x-00 && \
apt-mark hold kubelet kubectl
-
# since apt-get version 1.1 you can also use the following method
apt-get update && \
apt-get install -y --allow-change-held-packages kubelet=1.21.x-00
kubectl=1.21.x-00
```

Upgrade Worker Nodes

Here is how to restart the kubelet and uncordon the Node:

Demo

```
#Restart the kubelet:
sudo systemctl daemon-reload
sudo systemctl restart kubelet

#Uncordon the node
#Bring the node back online by marking it schedulable:

# replace <node-to-drain> with the name of your node
kubectl uncordon <node-to-drain>
```

Verify the Status of the Cluster

After upgrading the kubelet on all the Nodes, the status of the cluster must be checked. This is done to verify the availability of all the Nodes.

It can be done by running the following command :

Demo

```
kubectl get nodes
```

Note

The STATUS column should display **ready** for all your Nodes. The version number should also be updated.

Recovering from a Failure State

If an operation fails, for example, if kubeadm upgrade fails, kubeadm can be recovered from a bad state.

- Re-run **kubeadm upgrade**

- Run **kubeadm upgrade apply --force** without changing the version that the cluster is running

TECHNOLOGY

Managing Kubernetes Objects

Kubernetes Objects

Kubernetes objects are persistent entities that are used to represent the state of the cluster. They describe:

Available resources

Containerized applications that are running

Policies around how the applications behave

Object Spec and Status

A Kubernetes object usually includes two nested object fields, namely, object spec and object status. They govern the object's configuration.

spec

Describes the desired state of the object

status

Describes the current state of the object

Managing Kubernetes Objects

kubectl provides different ways to create and manage Kubernetes objects. Techniques that may be employed for managing Kubernetes objects are mentioned below:

Management technique	Operates on	Recommended environment	Supported writers	Learning curve
Imperative commands	Live objects	Development projects	1+	Lowest
Imperative object configuration	Individual files	Production projects	1	Moderate
Declarative object configuration	Directories of files	Production projects	1+	Highest

Managing Kubernetes Objects

Problem Statement:

Learn to manage Kubernetes objects.

Assisted Practice: Guidelines

Steps to demonstrate Managing objects in Kubernetes:

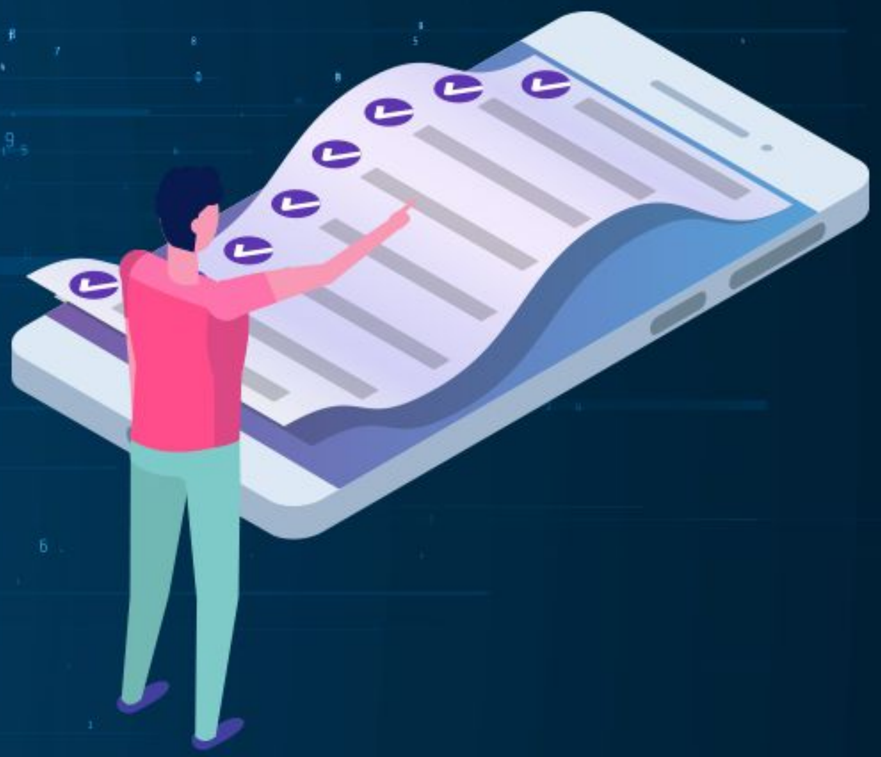
1. Get a list of Persistent Volumes
2. Describe the PersistentVolume output in yaml format
3. Sort PVs by storage capacity and store it in a text file

TECHNOLOGY

Demo: Cluster Networking

Key Takeaways

- Control Plane Node Communication catalogs the communication paths between the Control Plane and the Kubernetes cluster.
- Built-in controllers interact with the cluster API server and manage the state.
- Connections from the apiserver to a Node, Pod, or service default to plain HTTP connections are neither authenticated nor encrypted.
- When an object is created in Kubernetes, the object spec that describes its desired state needs to be created.



Knowledge Check

Knowledge Check 1

Which of the following is an authentication mechanism?

A. Bearer token

B. Client certificates

C. Authentication proxy

D. All of the above

Knowledge Check 1

Which of the following is an authentication mechanism?

A. Bearer token

B. Client certificates

C. Authentication proxy

D. All of the above

The correct answer is **D**

Bearer token, client certificates, and authentication proxy are authentication mechanisms.

Knowledge Check
2

Which of the following is NOT a type of Controller?

A. Node

B. Route

C. Service

D. Proxy

Knowledge Check
2

Which of the following is NOT a type of controller?

A. Node

B. Route

C. Service

D. Proxy

The correct answer is **D**

Proxy is not a type of controller.

Knowledge Check

3

________________is a type of heartbeat.

A. Updates of nodestatus

B. The Lease object

C. Both (A) and (B)

D. None of these

Knowledge Check
3

________________is a type of heartbeat.

A. Updates of nodestatus

B. The Lease object

C. Both (A) and (B)

D. None of these

The correct answer is **C**

Updates of nodestatus and the Lease object are the types of heartbeat.

Creating a Highly Available Cluster with Custom Role Binding

LESSON-END PROJECT

Problem Statement:

Your team lead has asked you to set up and configure a highly available Kubernetes cluster where role-based access controller can be created based on specific requirements. You are also required to create role and role bindings for a new user and check the access granted for it.

Steps to Perform:

1. Creating a cluster
2. Creating role-based access controller
3. Adding user credentials to the kubeconfig file
4. Creating role and role binding
5. Checking the access for the new user