Red Hat Summit

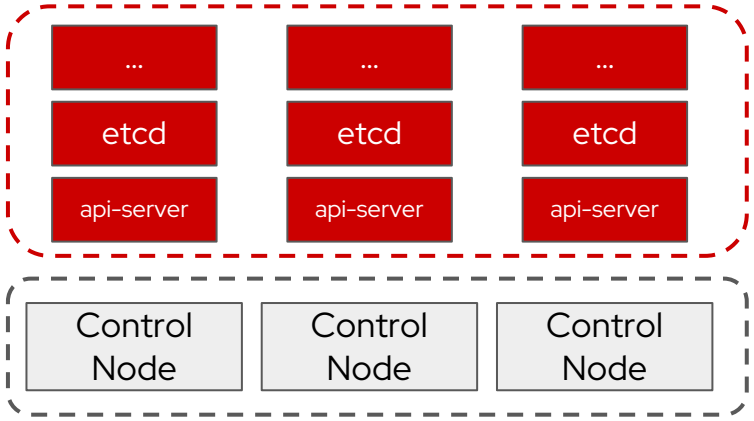
Connect

# OpenShift with Hosted Control Planes

Gökhan Göksu
Senior Solution Architect
gokhan@redhat.com
www.linkedin.com/in/ggoksu

# Standalone OpenShift

- Control Plane hosted across 3 machines



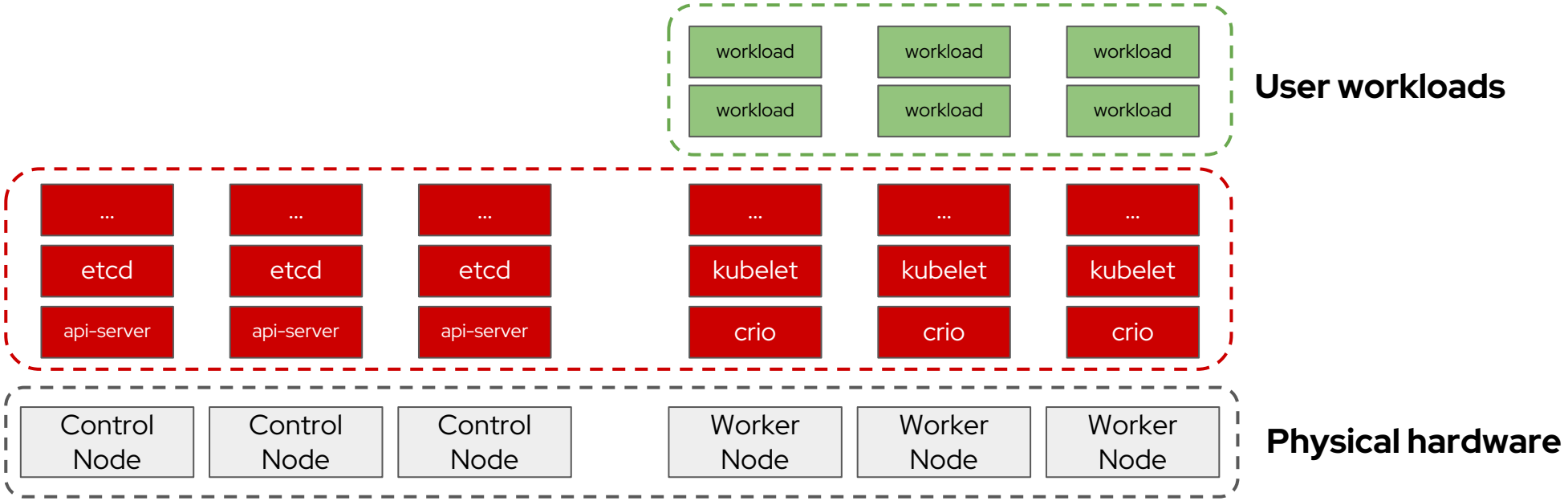

**Physical hardware**



Red Hat

# Standalone OpenShift

- Control Plane hosted across 3 machines
- Worker Nodes
- User Workloads

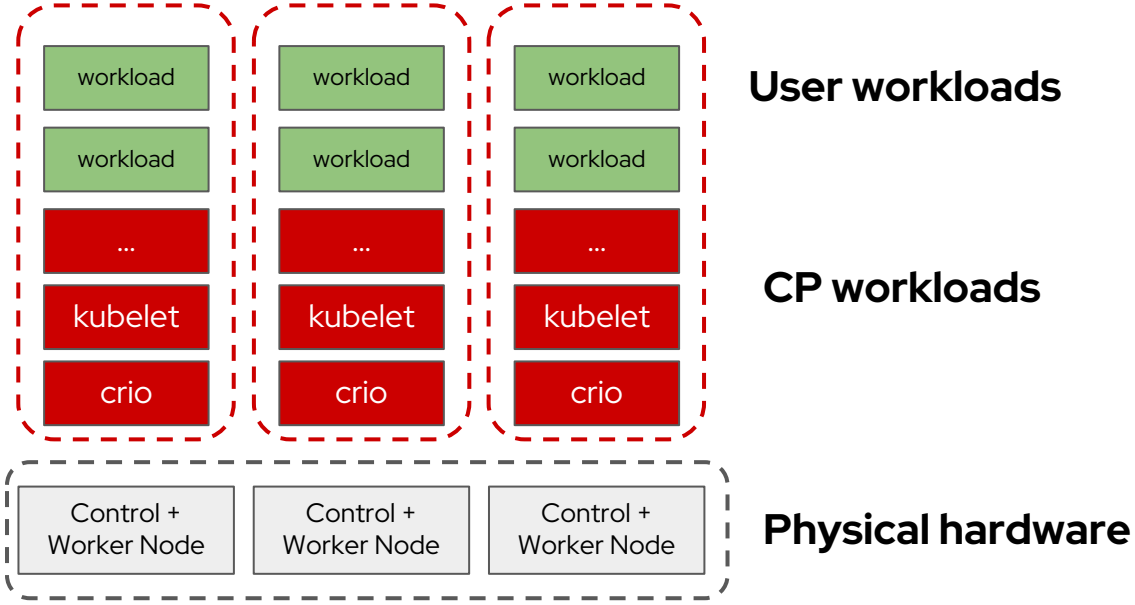





Red Hat

# Compact OpenShift

- Self-contained
- Highly available control-plane
- Less room for actual workloads
- Ideal for resource-constrained environments

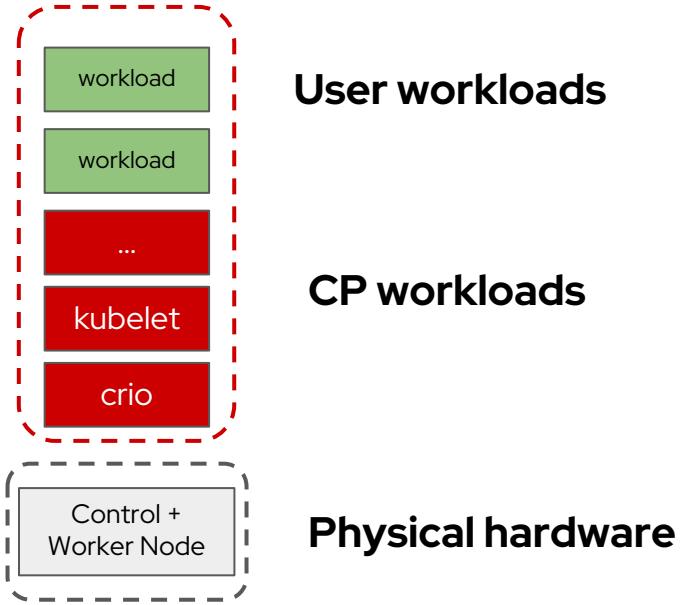





Red Hat

# Single-Node OpenShift (SNO)

- Self-contained
- No High Availability
- Resource Constraints
- Edge Locations

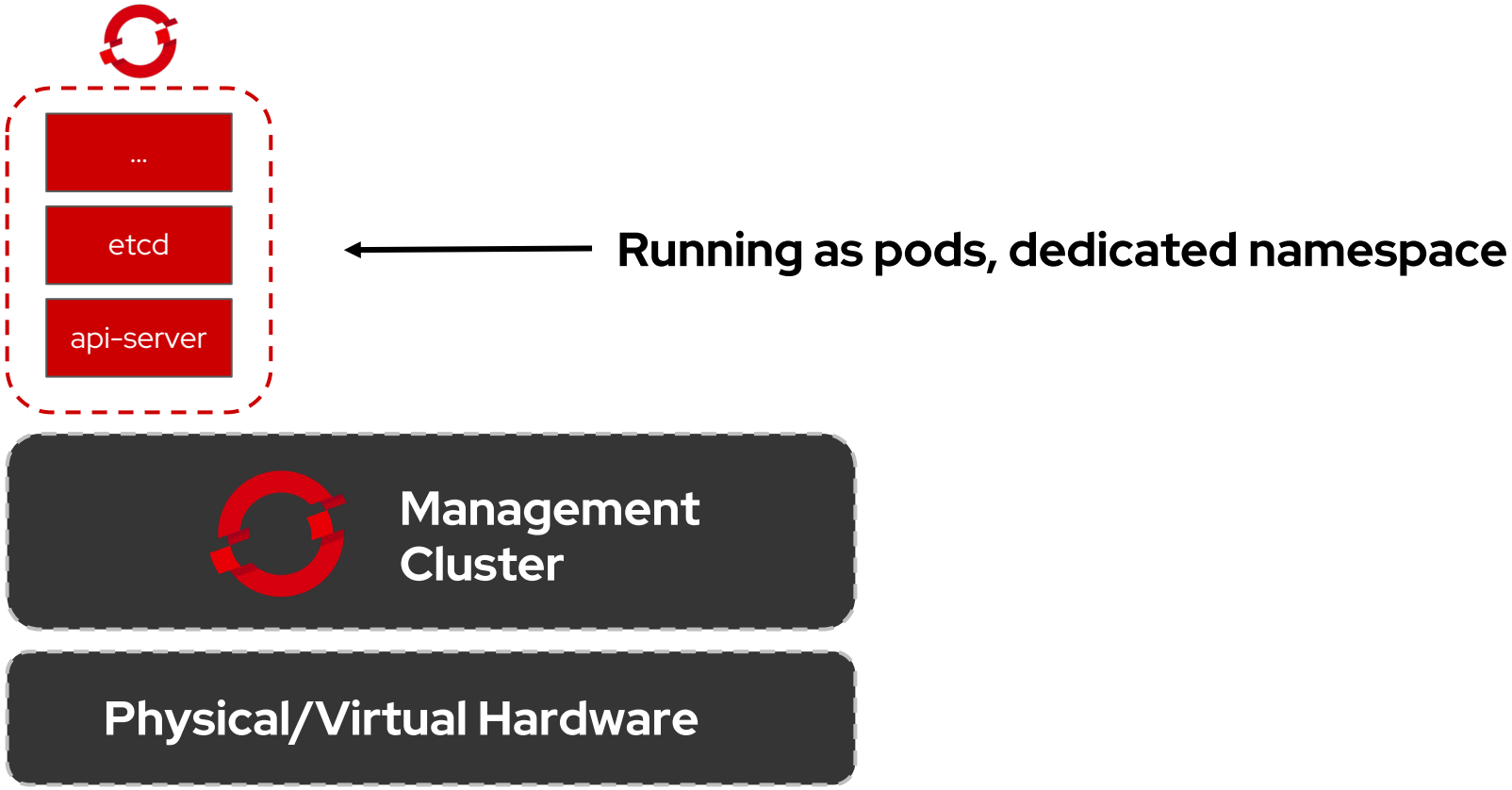





Red Hat

# Hosted Control Planes (HCP)

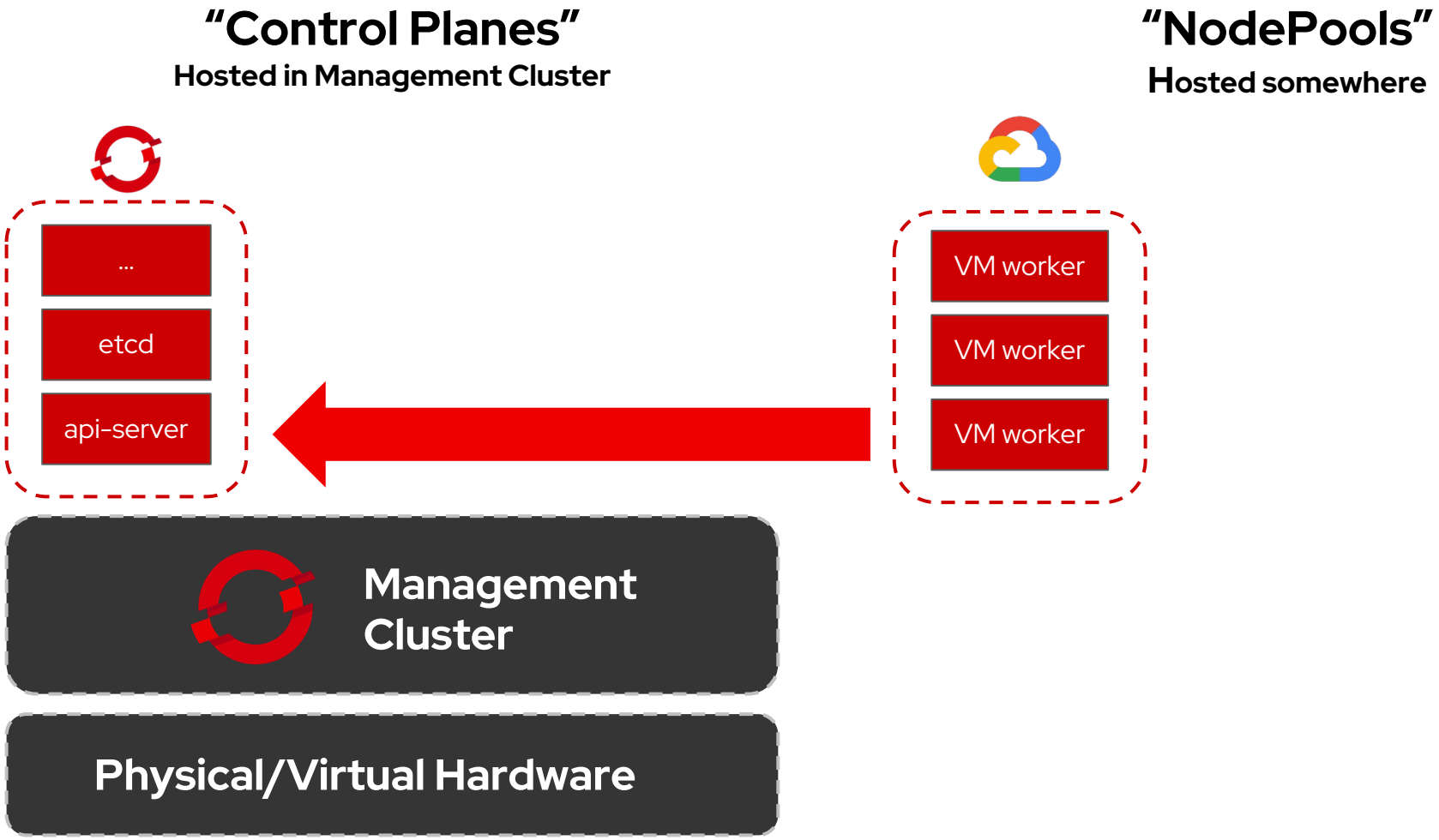
**“Control Planes”**
**Hosted in Management Cluster**






Red Hat

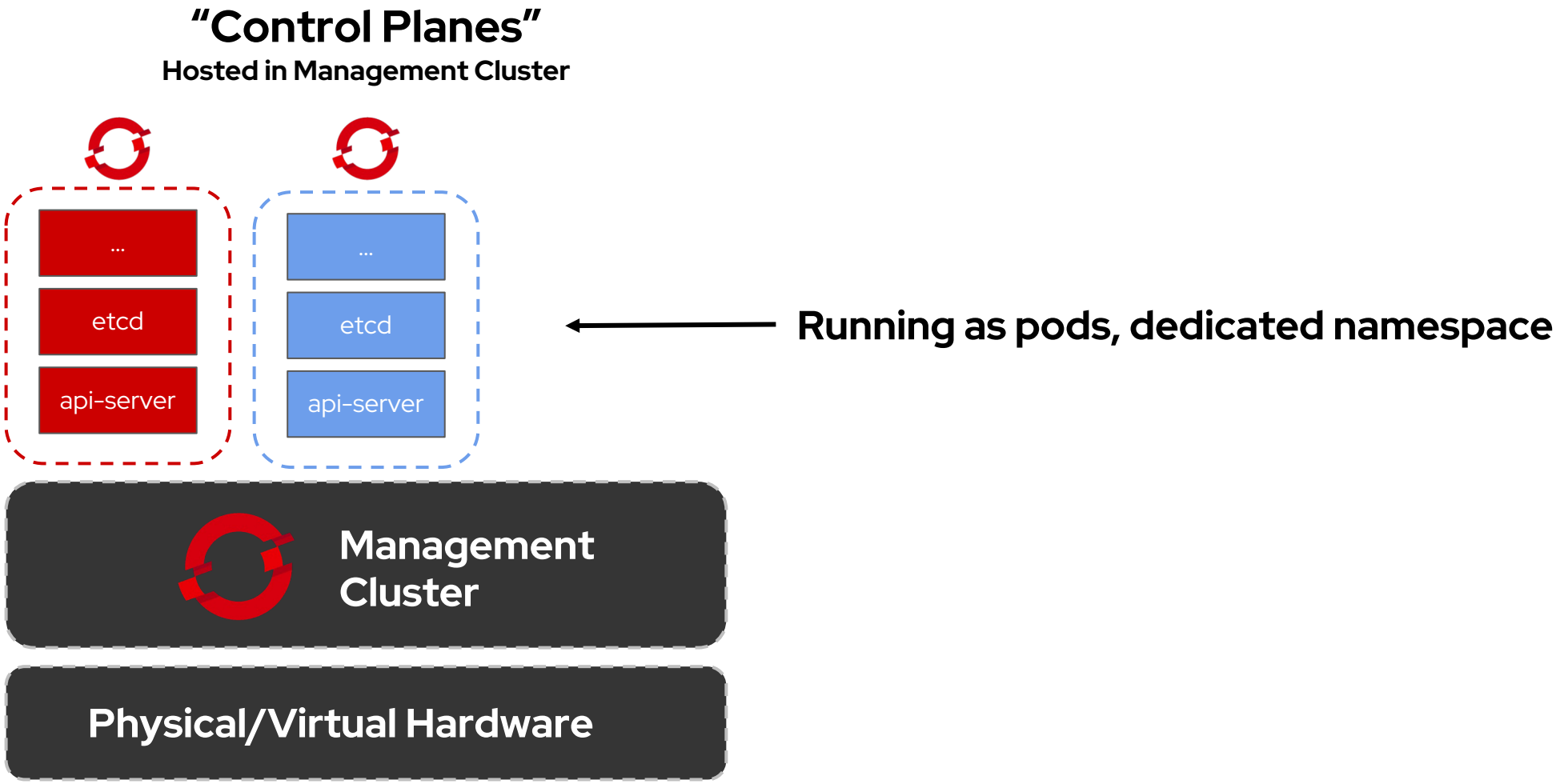
# Nodes Register with HCP






Red Hat

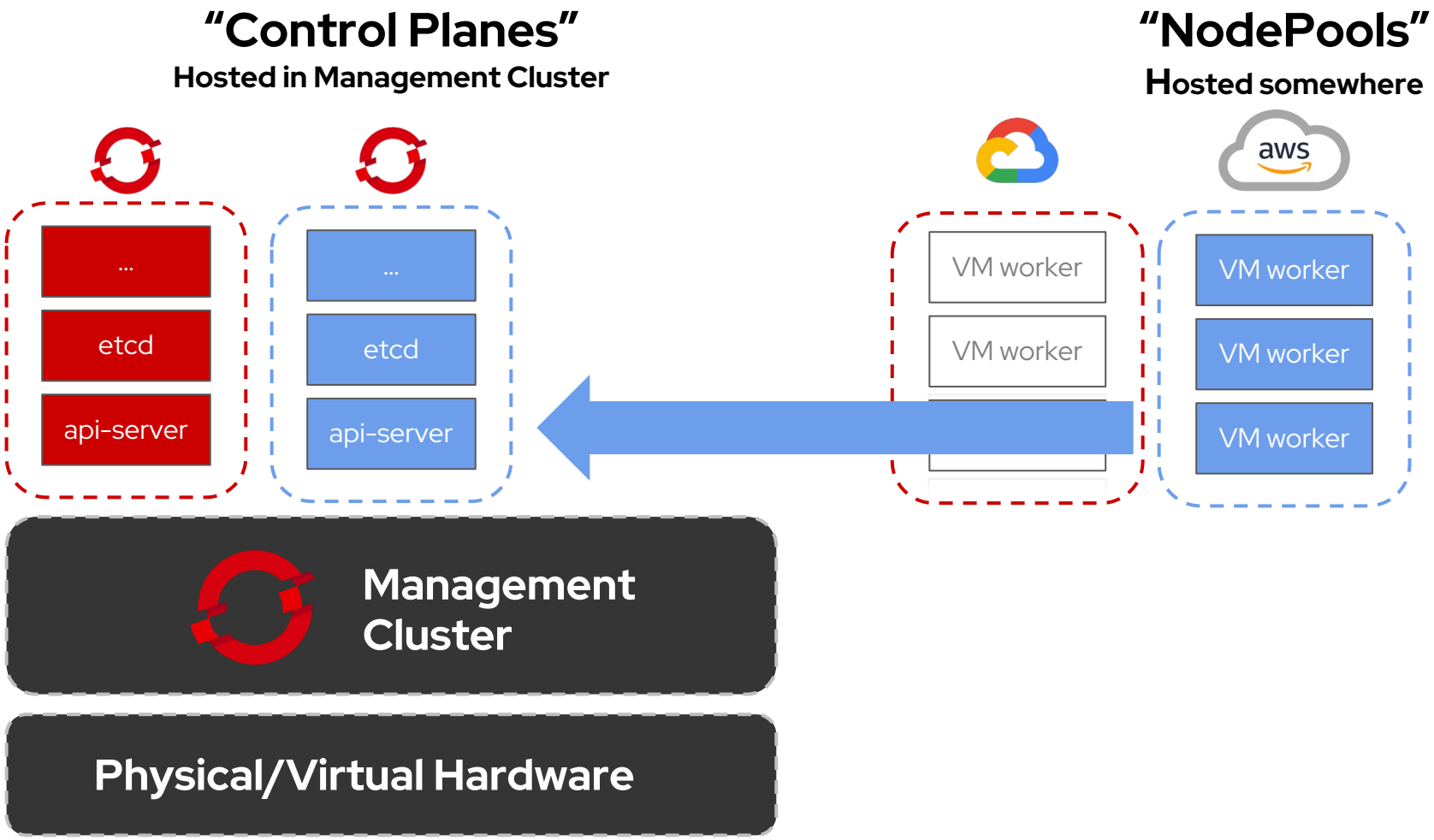
# HCP


"Control Planes"
Hosted in Management Cluster
...
etcd
api-server
...
etcd
api-server
Running as pods, dedicated namespace
Management Cluster
Physical/Virtual Hardware



Red Hat

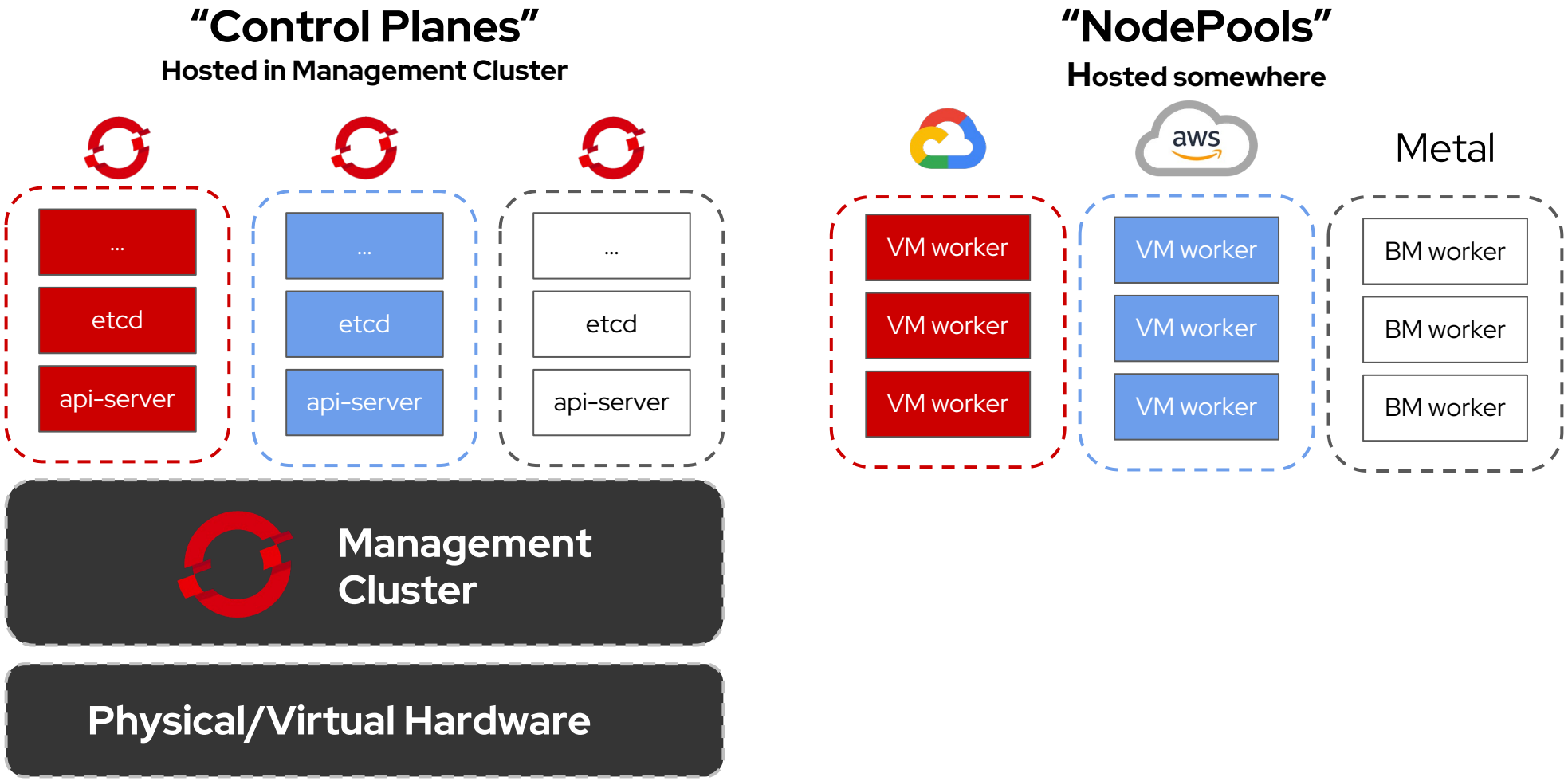
# Nodes Register with HCP






Red Hat

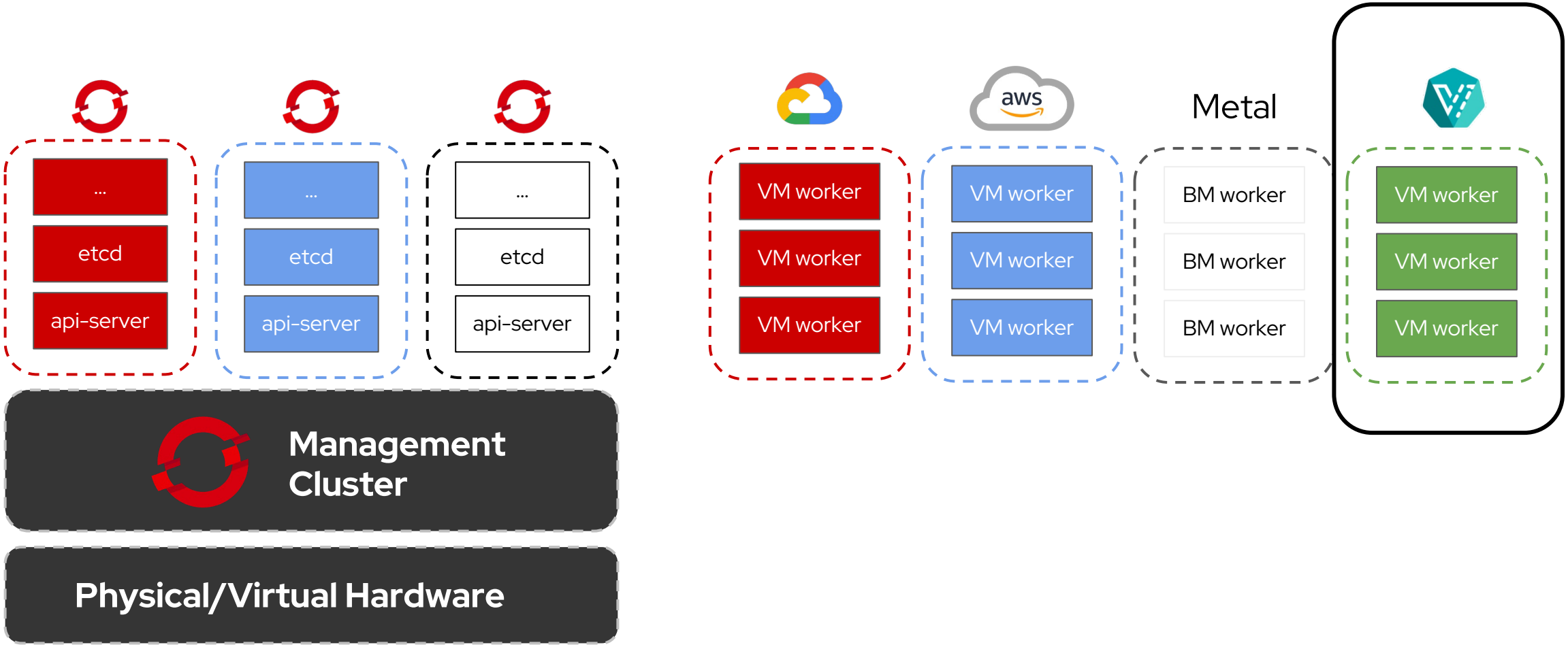
# HCP


"Control Planes"
Hosted in Management Cluster
...
etcd
api-server
...
etcd
api-server
...
etcd
api-server
Management Cluster
Physical/Virtual Hardware
"NodePools"
Hosted somewhere
aws
Metal
VM worker
VM worker
VM worker
VM worker
VM worker
VM worker
BM worker
BM worker
BM worker



Red Hat

# HCP


...
etcd
api-server
...
etcd
api-server
...
etcd
api-server
aws
Metal
VM worker
VM worker
VM worker
VM worker
VM worker
VM worker
BM worker
BM worker
BM worker
VM worker
VM worker
VM worker
Management Cluster
Physical/Virtual Hardware



Red Hat

# Use Cases

## Ephemeral Clusters

Quickly (< 10 min) spin up/destroy clusters for CI and developers.

## Clusters as a Service

On demand clusters driven by a declarative API

## Cheaper Control Planes

Multiple Control Planes per node vs. 3 nodes for 1 Control Plane

## Decoupled Life Cycle Management

Upgrade the consolidated control planes out of cycle from the segmented worker nodes



Red Hat

# Hosted Control Planes Enhancements

## Cluster AutoScaler API for Hosted Clusters

Customize the Cluster AutoScaler behaviour for hosted clusters such as CPU utilization thresholds

## OVN IPv4 subnet configuration

Customizable OVN IPv4 subnet configuration for HCP clusters to avoid CIDR conflicts and enable advanced networking features adding `ovnKubernetesConfig` to the HostedCluster API

## CNI Certification for HCP

New workflow certification guide including HCP (along with OCP Virt and Service Mesh). Cilium and Calico in the pipeline for certified CNIs for HCP

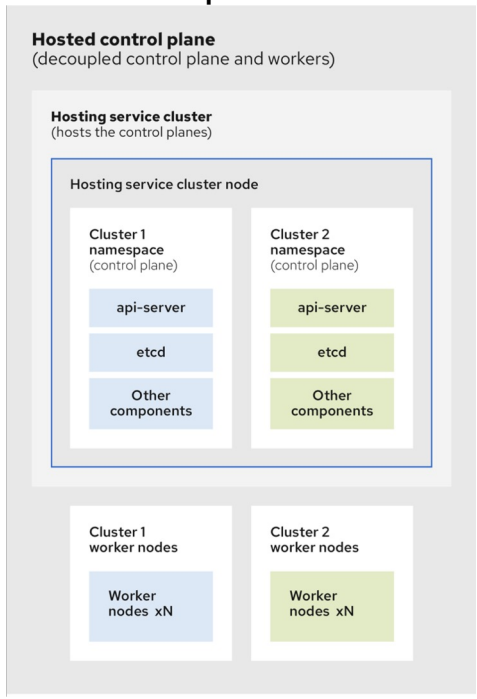
## OADP Integration for Hosted Control Planes

Backup and restore hosted control planes from the Management Cluster with the OpenShift API for Data Protection, including restoring a hosted control plane in a different Management Cluster

## Update Hosted Clusters API DNS on Day 2

Access your Hosted Clusters' API from a new name by updating at any time your DNS

HyperShift will regenerate your kubeconfig and console login command output







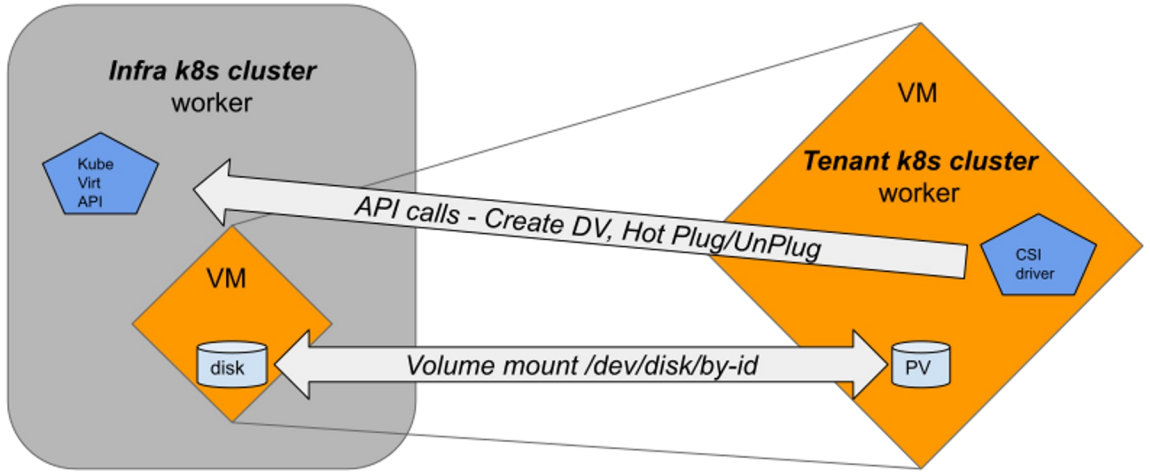

Red Hat

# Kubevirt-CSI

Maps storage from hosting infrastructure cluster into your tenant's virtualized OpenShift Clusters.

Agnostic of underlying CSI provider at the hosting/infra layer.

Included as part of a HCP deployment and is currently only supported for HCP deployment patterns




https://github.com/openshift/kubevirt-csi-driver

Red Hat

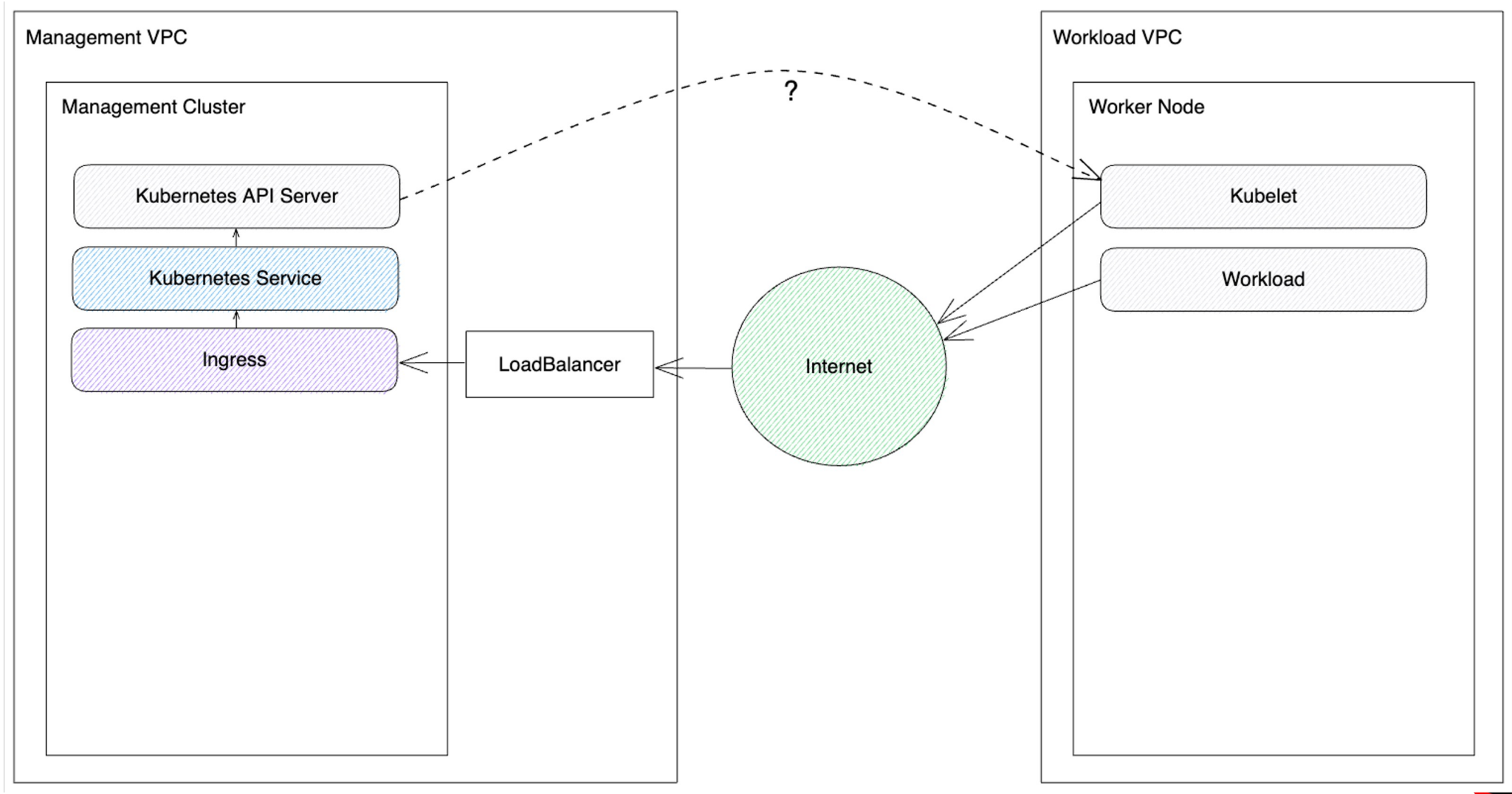
# CSI at the tenant cluster

Uses the native vendor provided CSI within the tenant cluster to directly access SAN backed storage.

For virtualized hosted clusters this will typically required network centric protocols like iSCSI / NFS / FCoE as native Fibre Channel devices will not be available at the VM level.

If you're using HCP for Bare Metal clusters all CSI SAN protocols can be supported.

# Networking

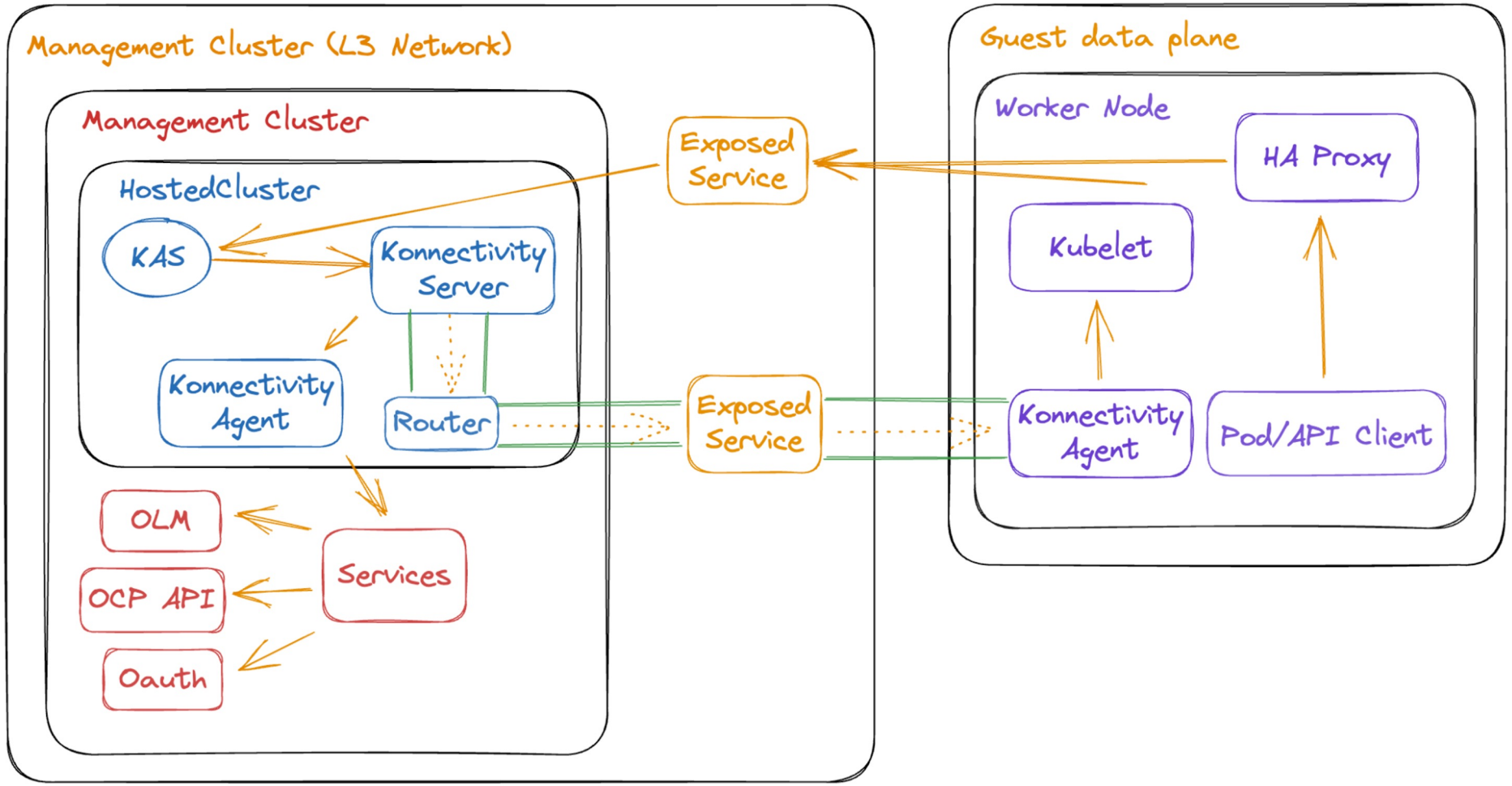





Red Hat

# Networking

## Networking between the management cluster and the hosted clusters

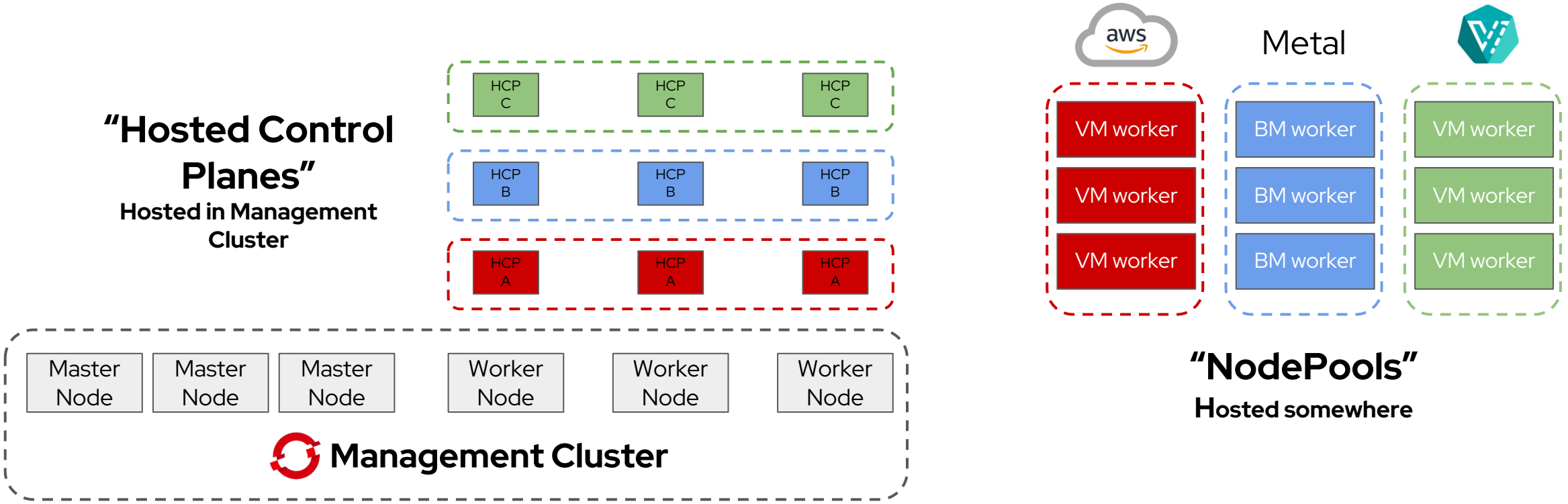





Red Hat

# High Availability

## High-level summary of different failure scenarios



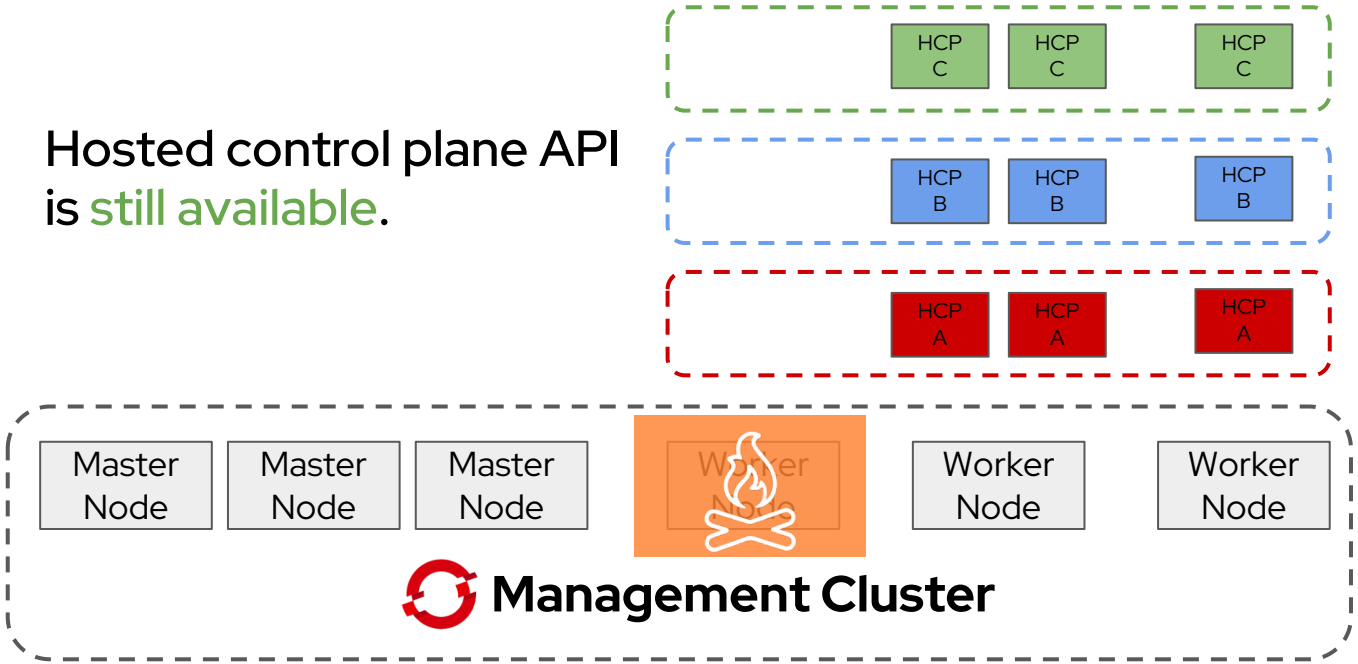



Red Hat

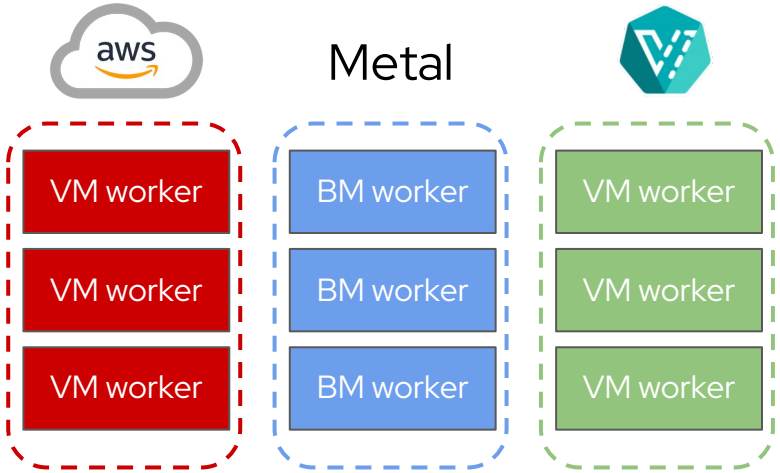
# High Availability

## Loss of management cluster worker

Hosted control plane API is still available.




Impacted hosted control plane components are rescheduled.




Hosted cluster data plane is still available.

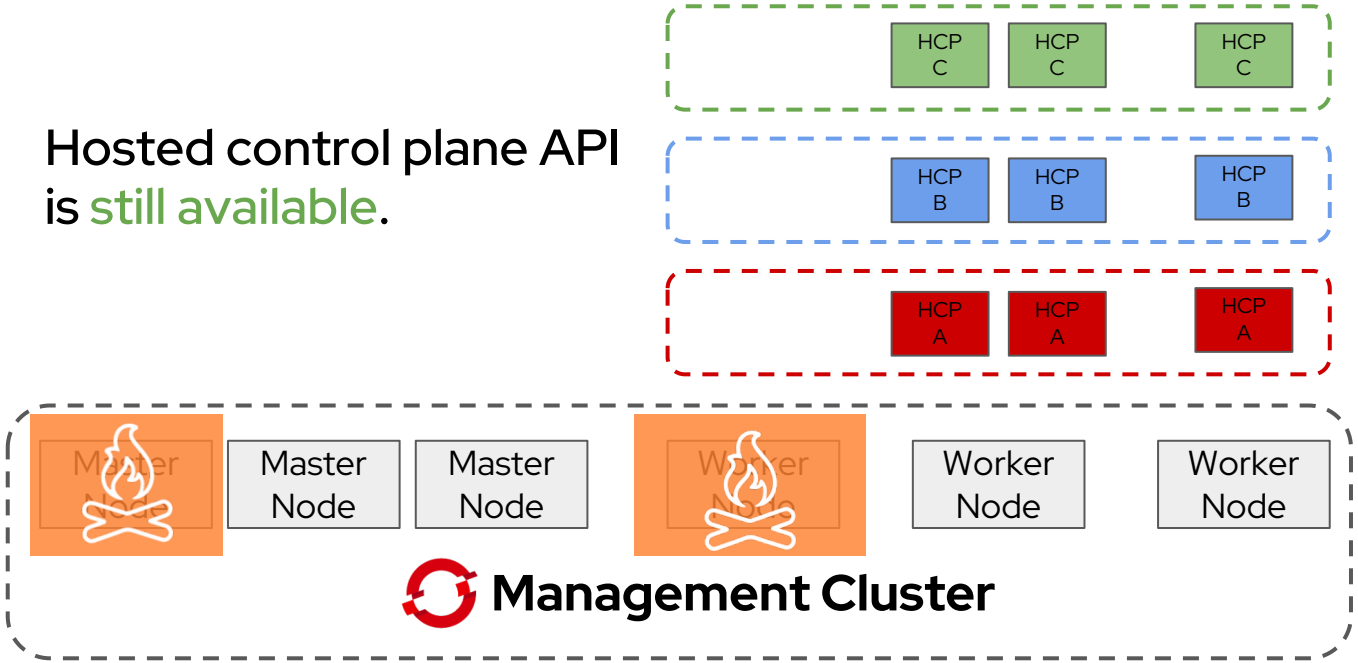


Red Hat

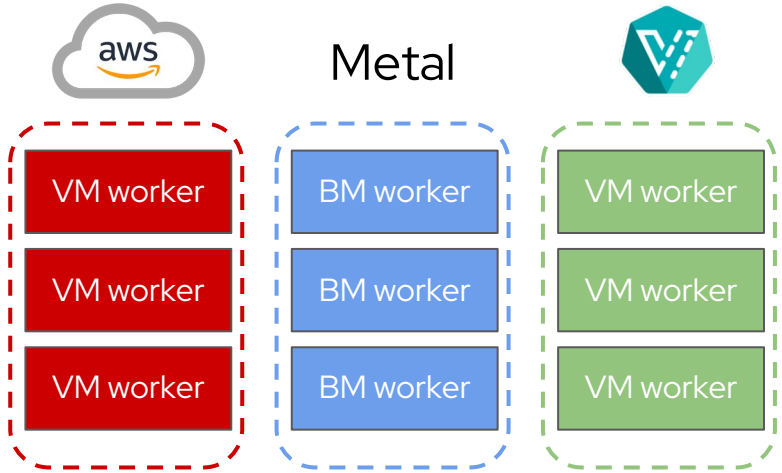
# High Availability

## Loss of management cluster availability zone

Hosted control plane API is still available.




Impacted hosted control planes maintain quorum.




Hosted cluster data plane is still available.

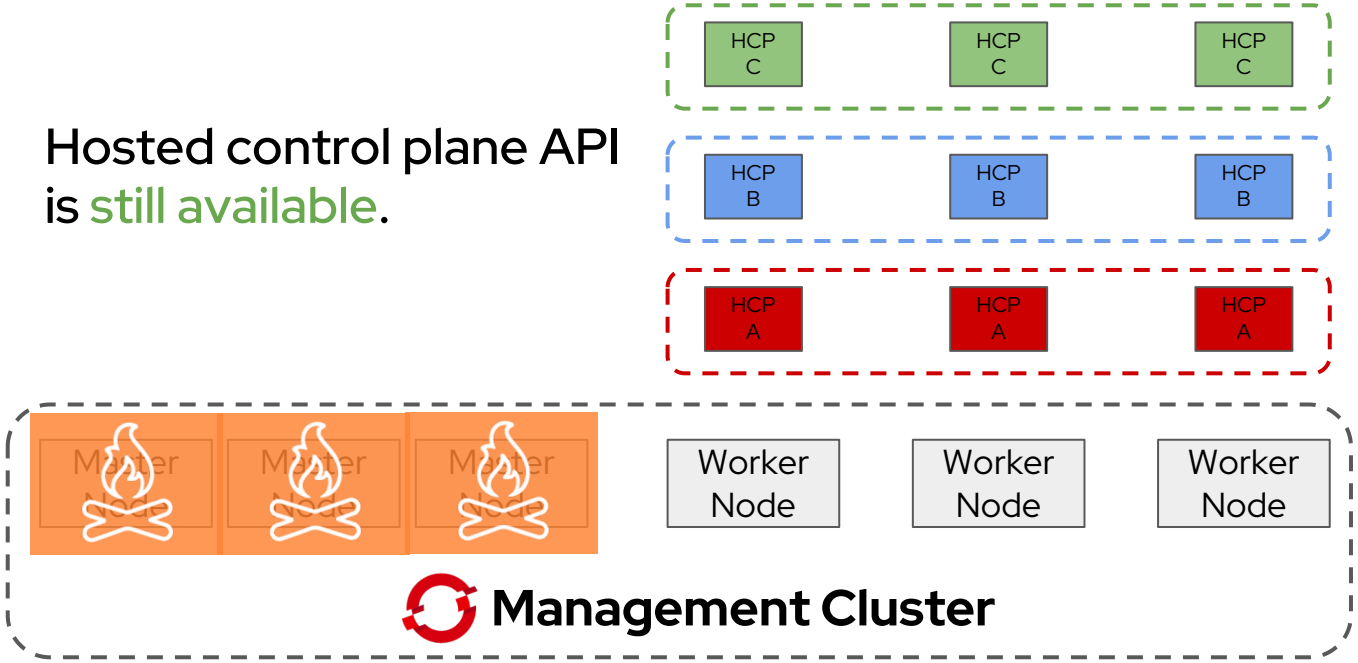


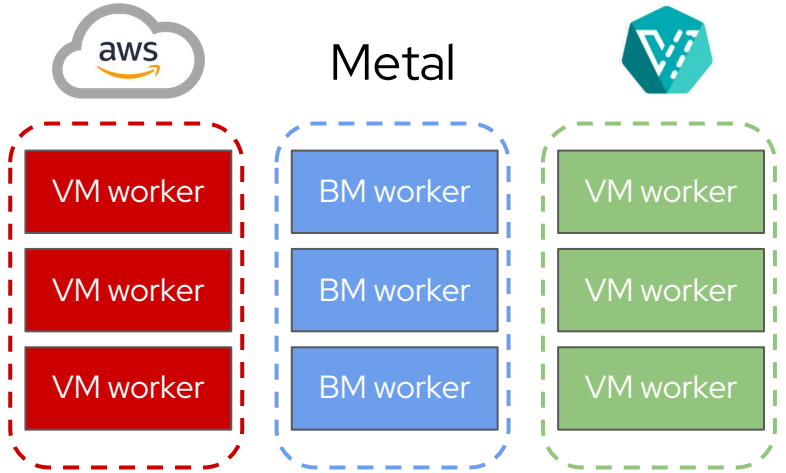
Red Hat

# High Availability

## Loss of management cluster control plane

Hosted control plane API is still available.



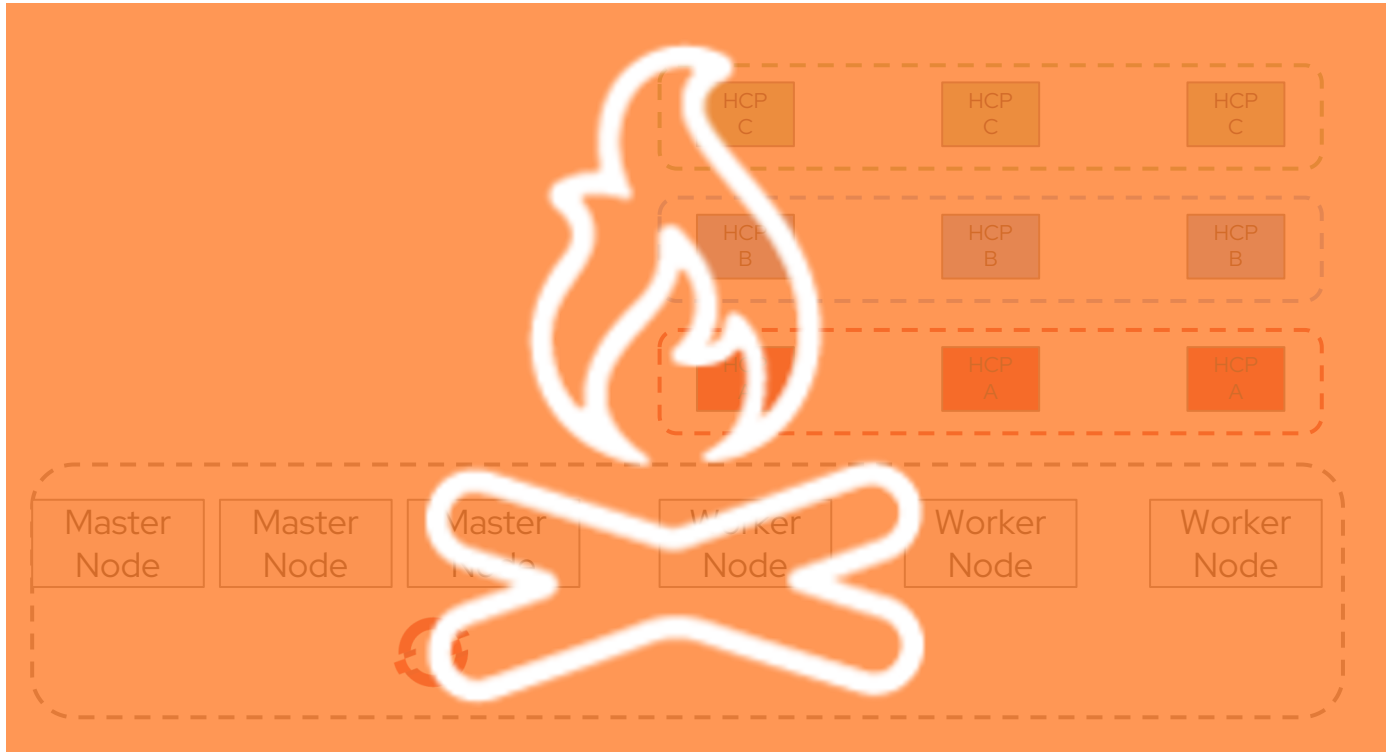

Hosted cluster data plane is still available.



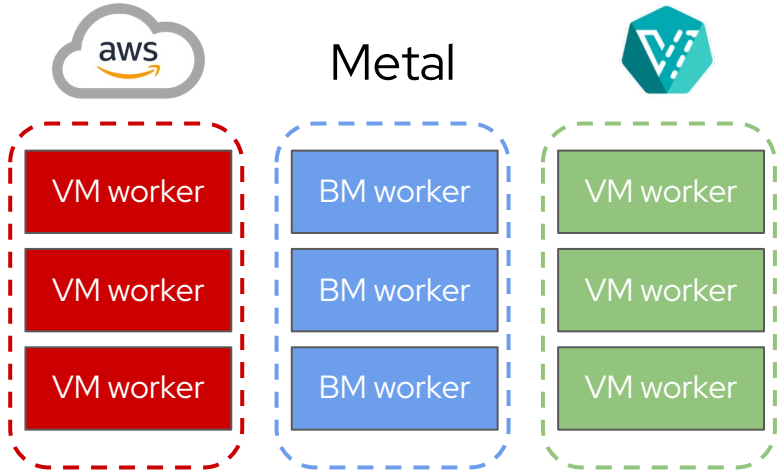
Red Hat

# High Availability

## Loss of management cluster control plane and workers




Hosted control plane API is not available.



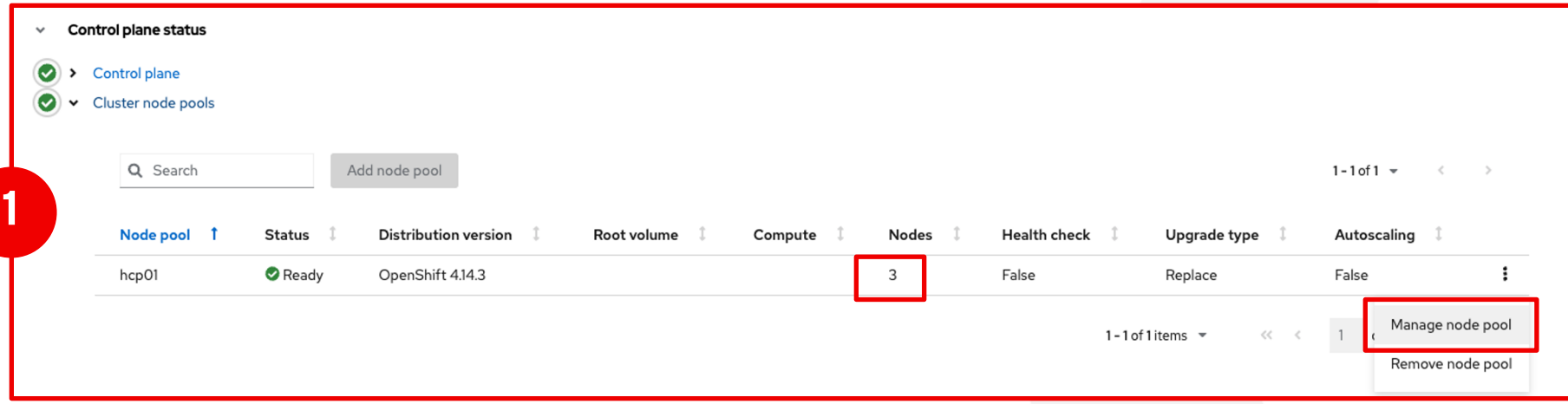

Hosted cluster data plane is still available.



Red Hat

# Scale UP / Scale DOWN

## Extend existing NodePool







Contributor: David Martini

Red Hat

# Scale UP / Scale DOWN

**Create new NodePool**

```
hcp create nodepool  kubevirt \
--cluster-name=hcp01 \
--name=new-nodepool \
--node-count=3 \
--cores=2 \
--memory=8Gi \
--root-volume-size=32 \
--node-upgrade-type=InPlace \
```

1

2

Search | Add node pool | 1 - 2 of 2

| Node pool | Status | Distribution version | Root volume | Compute | Nodes | Health check | Upgrade type | Autoscaling |
|---|---|---|---|---|---|---|---|---|
| hcp01 | Ready | OpenShift 4.14.3 | | | 3 | False | Replace | False |
| new-nodepool | Pending | OpenShift 4.14.3 | | | 3 | False | InPlace | False |

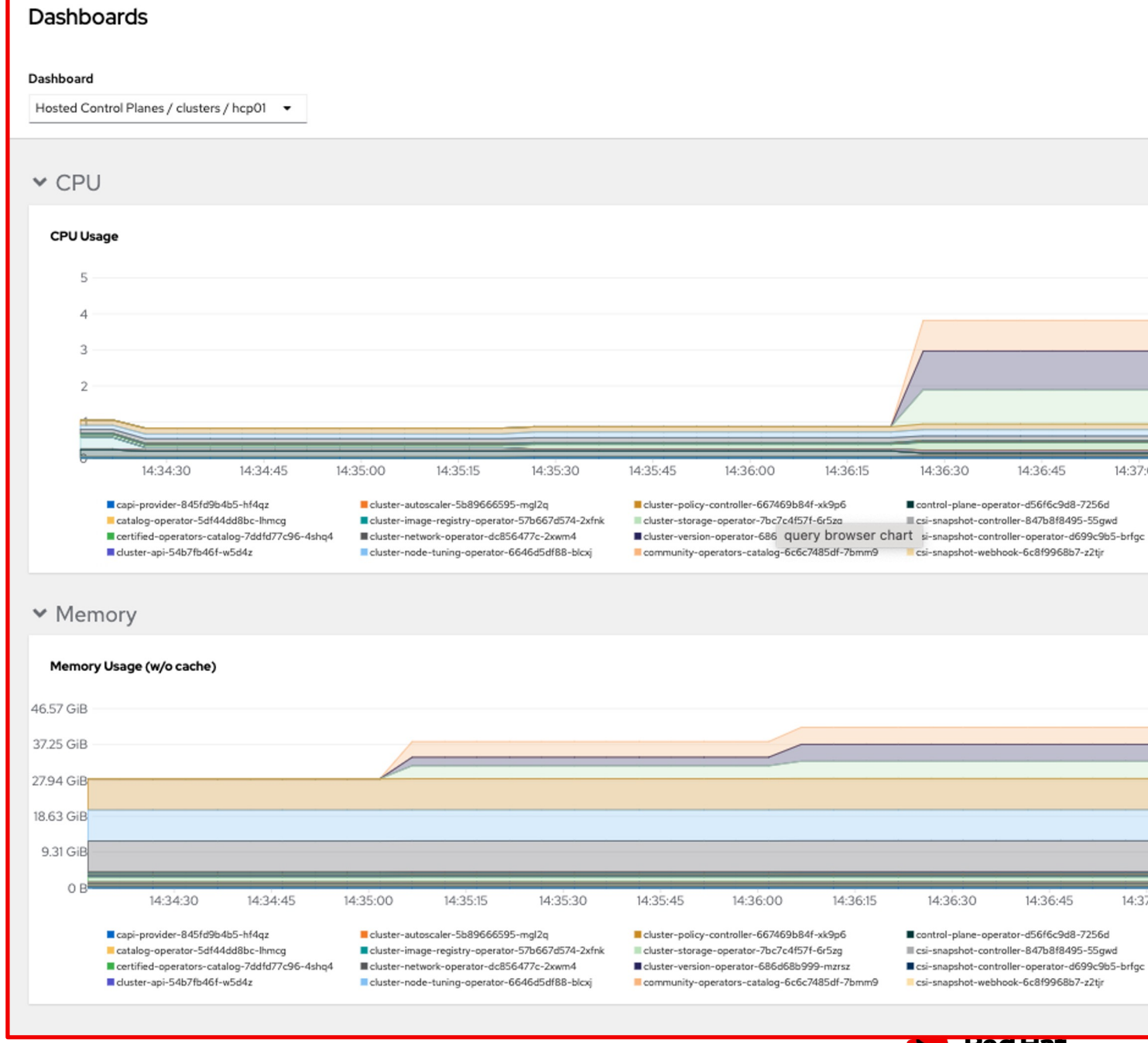


Contributor: David Martini

Red Hat

# HCP Monitoring

```
kind: ConfigMap
apiVersion: v1
metadata:
  name: hypershift-operator-install-flags
  namespace: local-cluster
data:
  installFlagsToAdd: "--monitoring-dashboards"
  installFlagsToRemove: ""
```

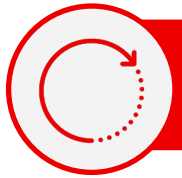




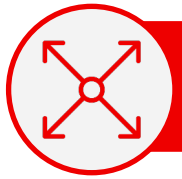

Contributor: David Martini

Red Hat

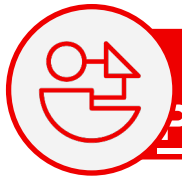
# HCP Upgrade – General aspects



Control plane and Data plane upgrades are decoupled



Each Node Pools can be upgraded separately



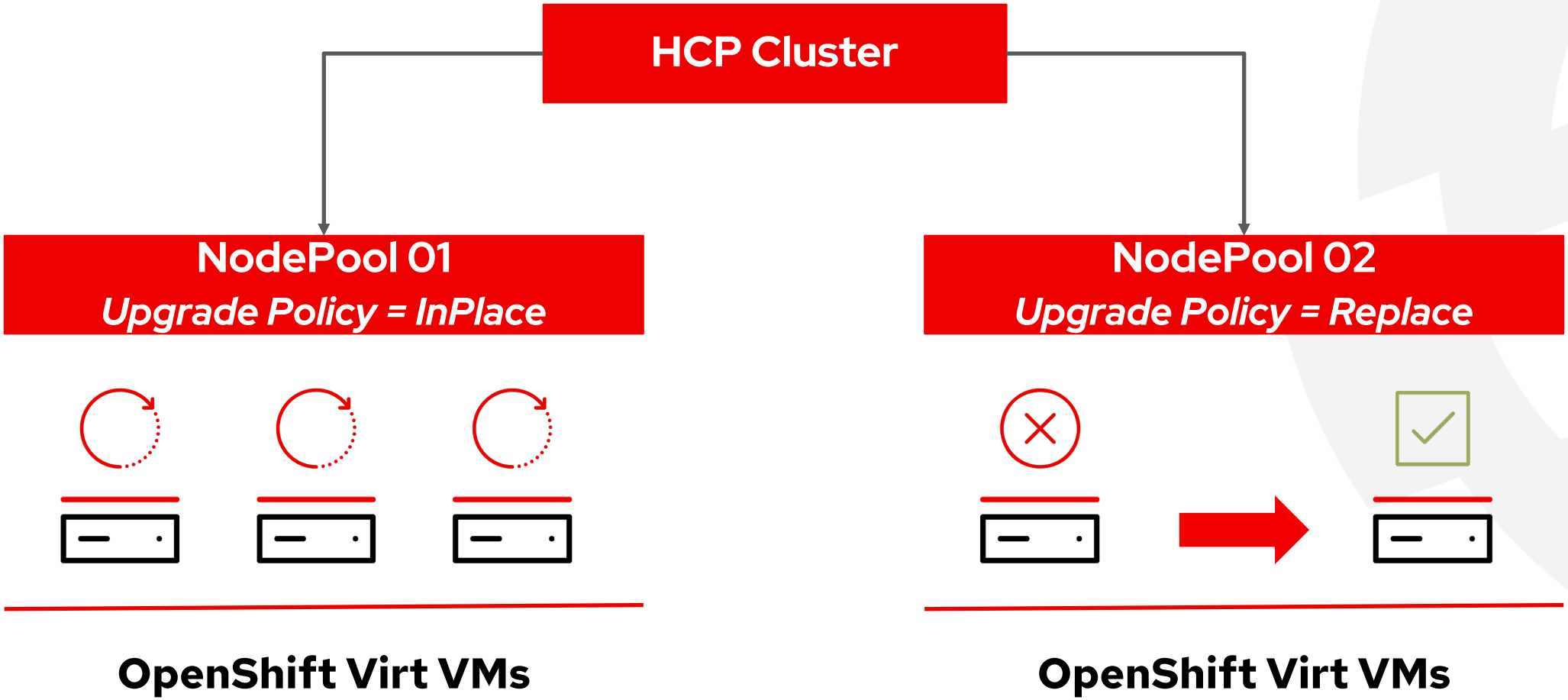
Different versions can be works in parallel following Version Skew Policy



Contributor: David Martini

Red Hat

# HCP Upgrade - Methods



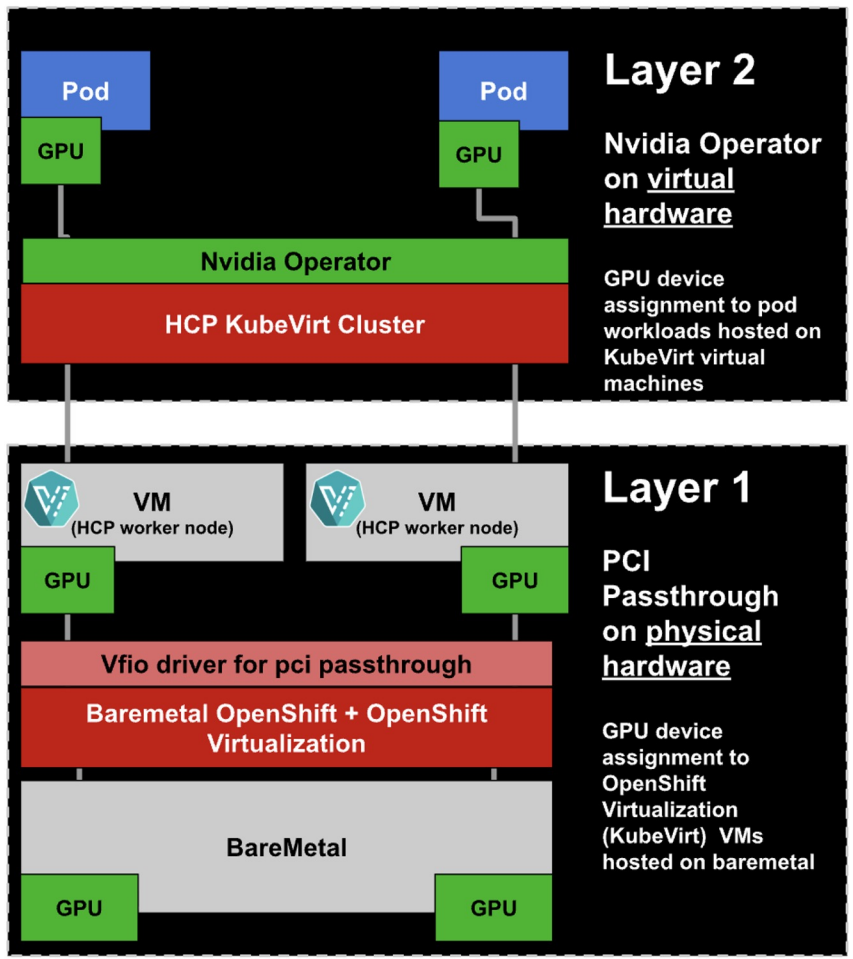



Contributor: David Martini

Red Hat

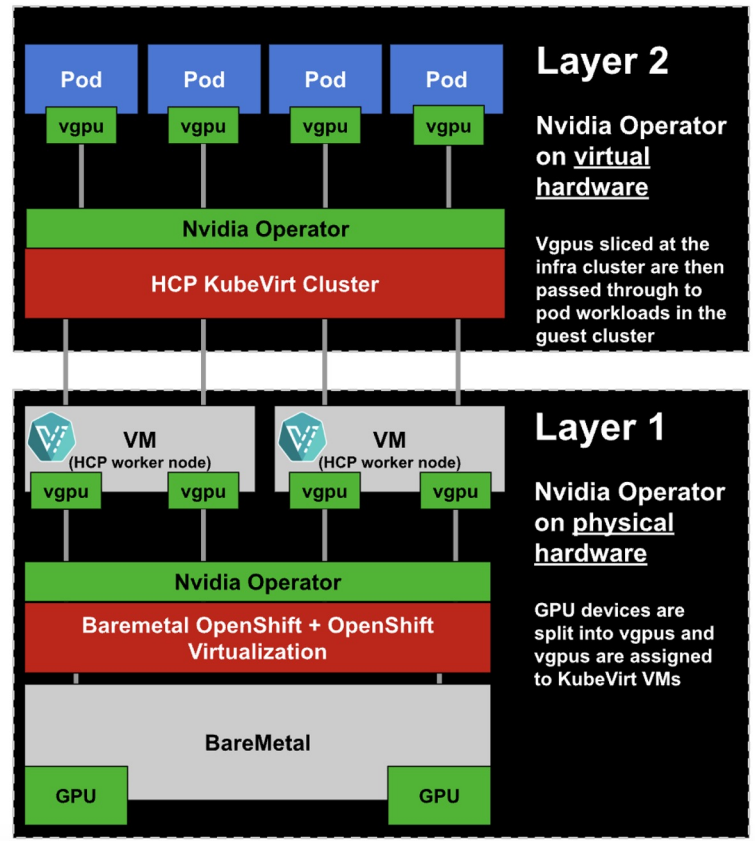
## Passthrough Architecture

**Layer 1** - Infra cluster has vfio driver installed and performs GPU passthrough to KubeVirt VMs
**Layer 2** - HCP KubeVirt cluster has nvidia operator installed and performs GPU passthrough to pod workloads




## Infra vGPU Slicing Architecture

**Layer 1** - Infra cluster has nvidia operator installed and performs vGPU slicing. VMs are provided with vGPU
**Layer 2** - HCP KubeVirt cluster has nvidia operator installed and provides vGPU to pod workloads



Red Hat Summit

Connect

# Thank you

linkedin.com/company/red-hat

facebook.com/redhatinc

youtube.com/user/RedHatVideos

twitter.com/RedHat