Kubernetes High Availability

Webinar CERN : May 20th 2020

https://indico.cern.ch/event/915688/

Webinars

Webinars starting May 20th: Weekly on Wednesdays 2pm CET

Live on Zoom, open to everyone, recording available shortly after

https://clouddocs.web.cern.ch/containers/training.html#webinars

Webinars Agenda

20 May 2020 - Kubernetes and High Availability for Services and Control Plane

27 May 2020 - Kubernetes Cluster Auto Scaling

3 Jun 2020 - Use Case: Moving Rucio to Production in Kubernetes

10 Jun 2020 - Containerization and Image Best Practices

17 Jun 2020 - Monitoring Kubernetes Clusters with Prometheus

24 Jun 2020 - Managing Kubernetes Deployments with Helm and Flux

1 Jul 2020 - Debugging Kubernetes Services and Nodes

8 Jul 2020 - Managing OpenStack with Kubernetes

About

Computing Engineer in the CERN cloud team

Focusing on containers, kubernetes and networking

Accelerators and ML

Previous work in storage and the WLCG (worldwide LHC computing grid)

@ahcorporto

ricardo.rocha@cern.ch

High Availability

“ A characteristic of a system which aims to ensure an agreed level of operational performance, usually uptime, for a higher than normal period. “

“ Can my control plane survive losing one master ? “

“ Can my control plane survive losing one AZ ? “

“ Can my application survive losing one node ? “

“ Can my application survive losing one AZ ? “

“ Can my control plane survive losing one master ? “

“ Can my control plane survive losing one AZ ? “

Multi Master, Split AZs

“ Can my application survive losing one node ? “

“ Can my application survive losing one AZ ? “

Replica Set, Service, Anti Affinity, Topology Spread

“ Can my control plane survive a bad upgrade ? “

“ Does my application stay up during upgrades ? “

“ Is my application running as expected ? “

“ Is my application ready to serve traffic ? “

“ Can my cluster survive a bad upgrade ? “

Multiple Clusters, External Load Balancer

“ Does my application stay up during upgrades ? “

“ Is my application running as expected ? “

“ Is my application ready to serve traffic ? “

Pod Disruption Budgets, Liveness Probes, Readiness Probes

Infrastructure

3 Availability Zones : cern-geneva-a / b / c

Split of compute nodes, racks

Redundancy on the networking layer

Redundancy on storage

Infrastructure

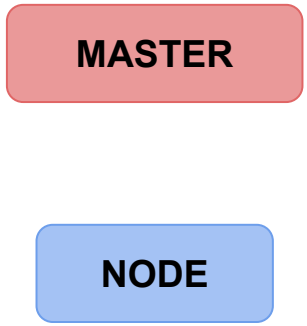
3 Availability Zones : cern-geneva-a / b / c

Split of compute nodes, racks

Redundancy on the networking layer

Redundancy on storage

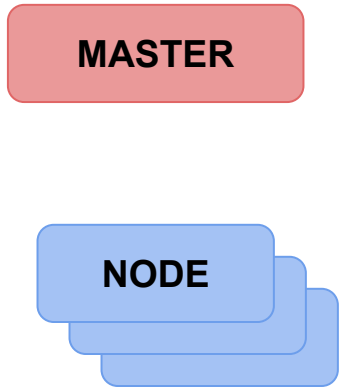
```
openstack coe cluster create \
    --cluster-template kubernetes-1.18.2-2 \
    --master-count 1 \
    --node-count 1 \
    …
```



No high availability of the control plane

No high availability of the applications

```
openstack coe cluster create \
    --cluster-template kubernetes-1.18.2-2 \
    --master-count 1 \
    --node-count 3 \
    …
```



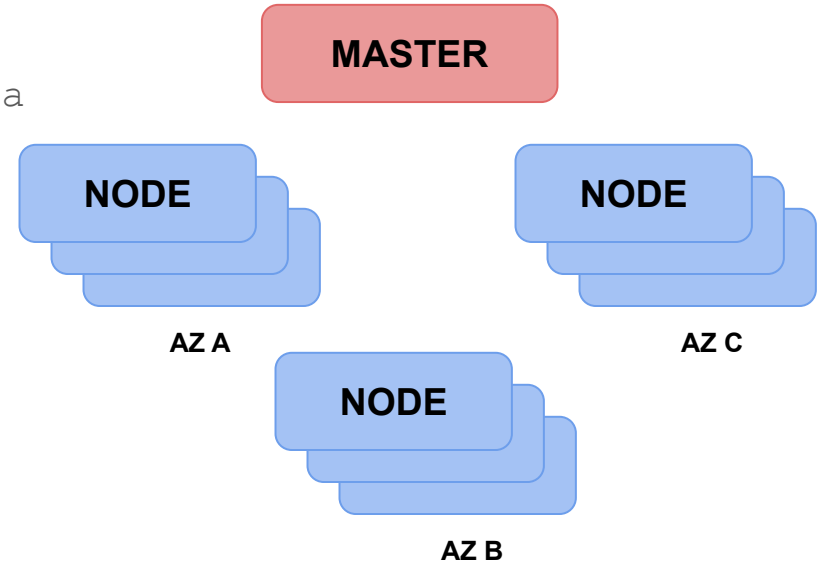
No high availability of the control plane

Improved availability of the applications

- Can survive one or more nodes dying
- Cannot survive a full AZ becoming unavailable

```
openstack coe cluster create \
    --cluster-template kubernetes-1.18.2-2 \
    --master-count 1 ...

openstack coe nodegroup create \
    --label availability_zone=cern-geneva-a
    --node-count 3 ...
```



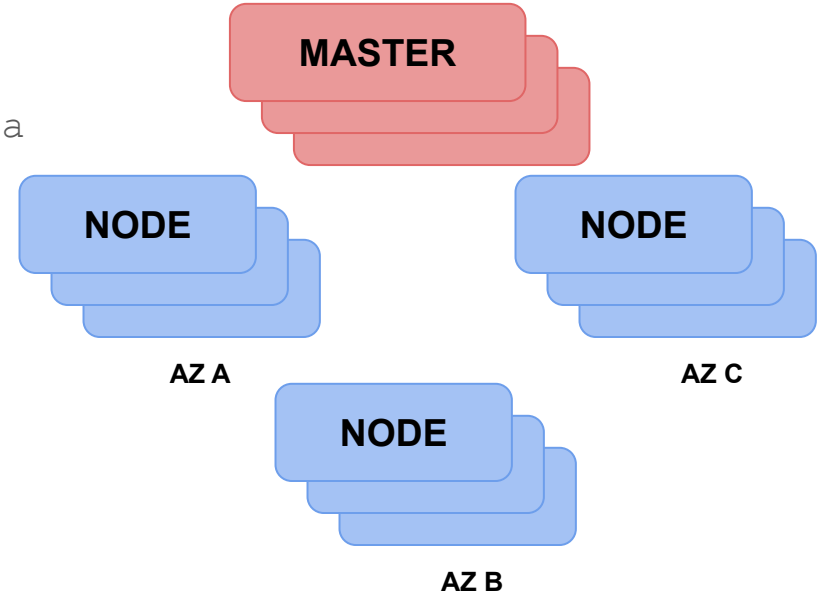
No high availability of the control plane

High availability of the applications

- Can survive one or more nodes dying
- Can survive a full AZ becoming unavailable

```
openstack coe cluster create \
    --cluster-template kubernetes-1.18.2-2 \
    --master-count 3 ...

openstack coe nodegroup create \
    --label availability_zone=cern-geneva-a
    --node-count 3 ...
```



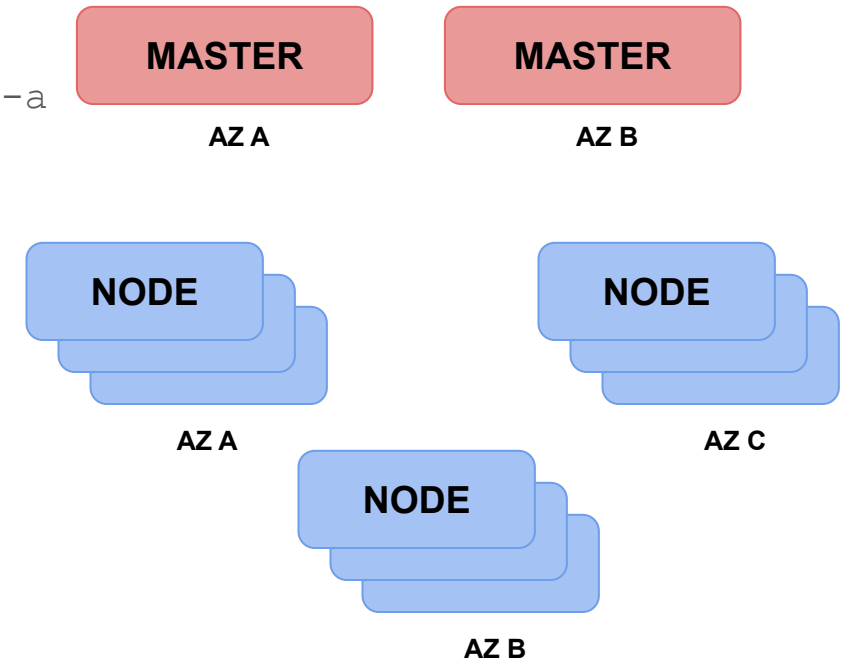
Improved availability of the control plane

High availability of the applications

- Can survive one or more nodes dying
- Can survive a full AZ becoming unavailable

```
openstack coe cluster create \
    --cluster-template kubernetes-1.18.2-2 \
    --master-count 1 ...

openstack coe nodegroup create \
    --label availability_zone=cern-geneva-a
    --node-count 3 ...
```



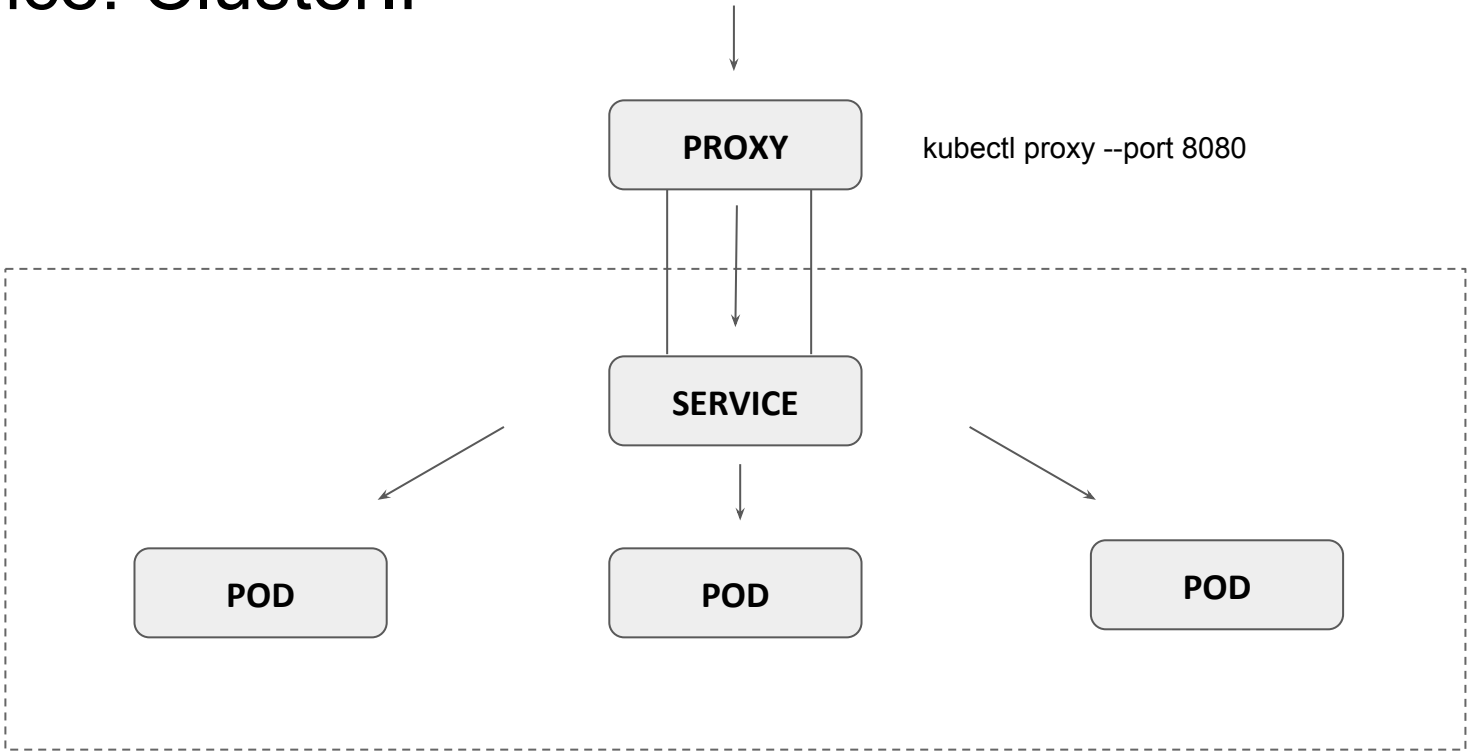
High availability of the control plane

High availability of the applications

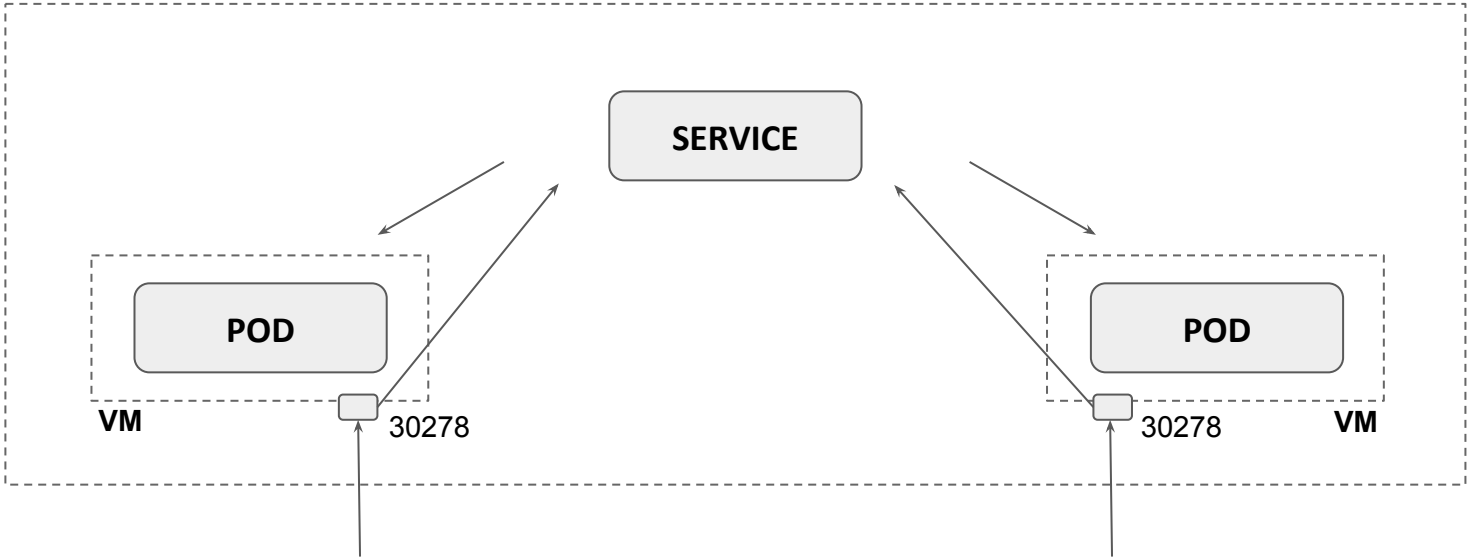
- Can survive one or more nodes dying
- Can survive a full AZ becoming unavailable

Demo

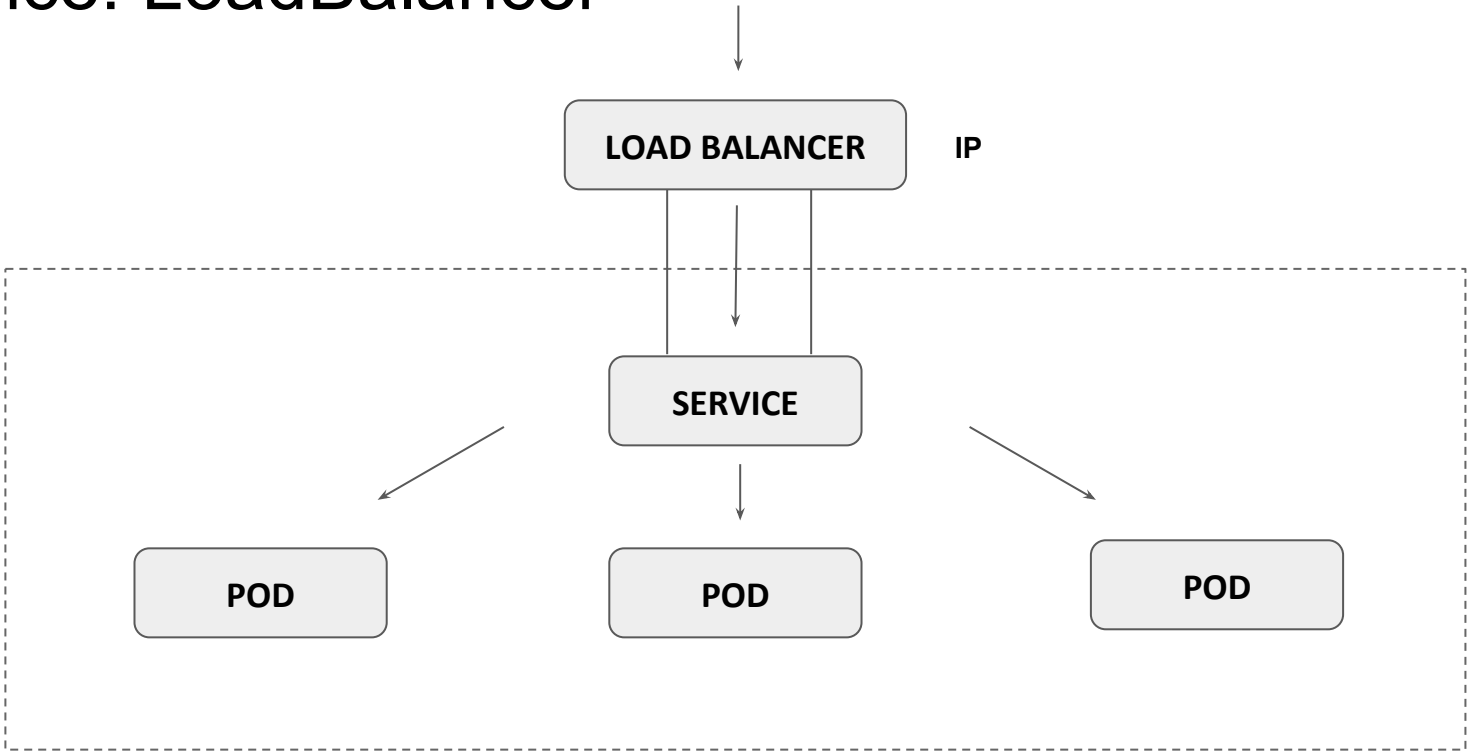
Service: ClusterIP



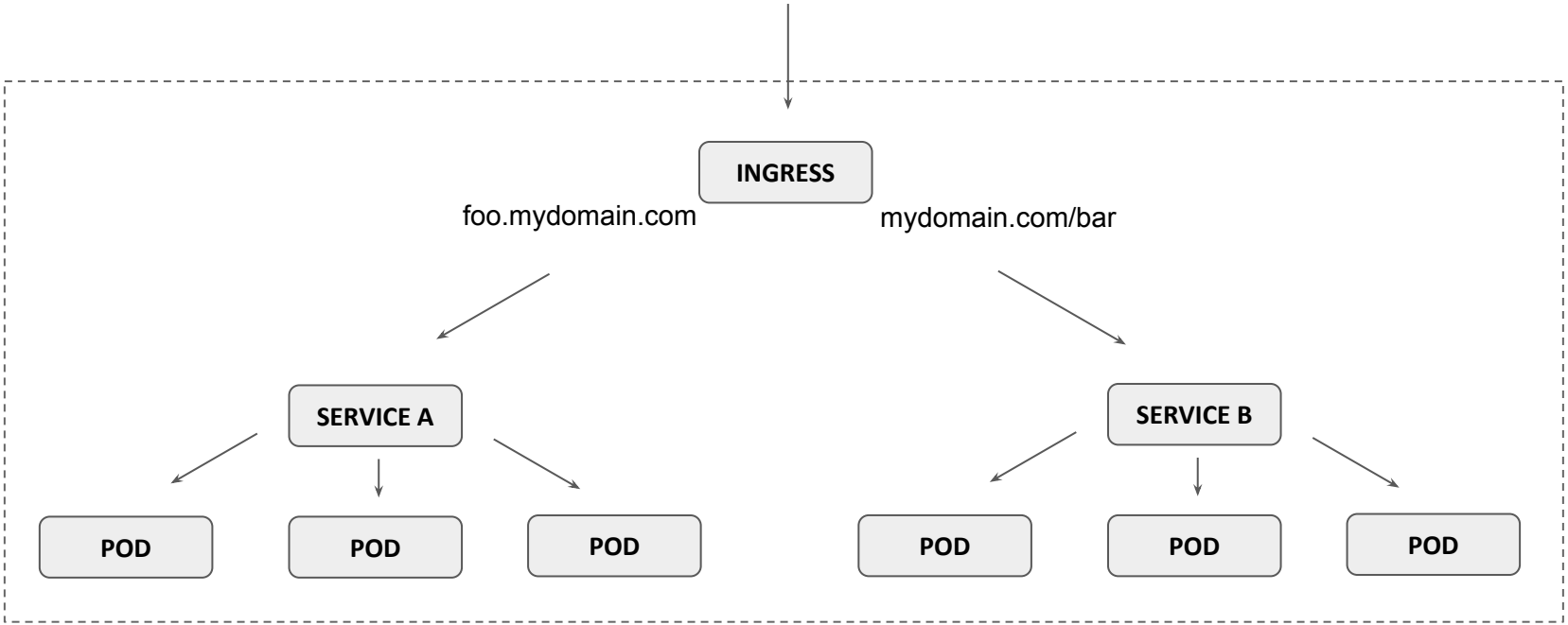
Service: NodePort



Service: LoadBalancer



Ingress


INGRESS
foo.mydomain.com
mydomain.com/bar
SERVICE A
SERVICE B
POD
POD
POD
POD
POD
POD

Demo

Future Topics

There’s a lot more to cover for HA on Kubernetes

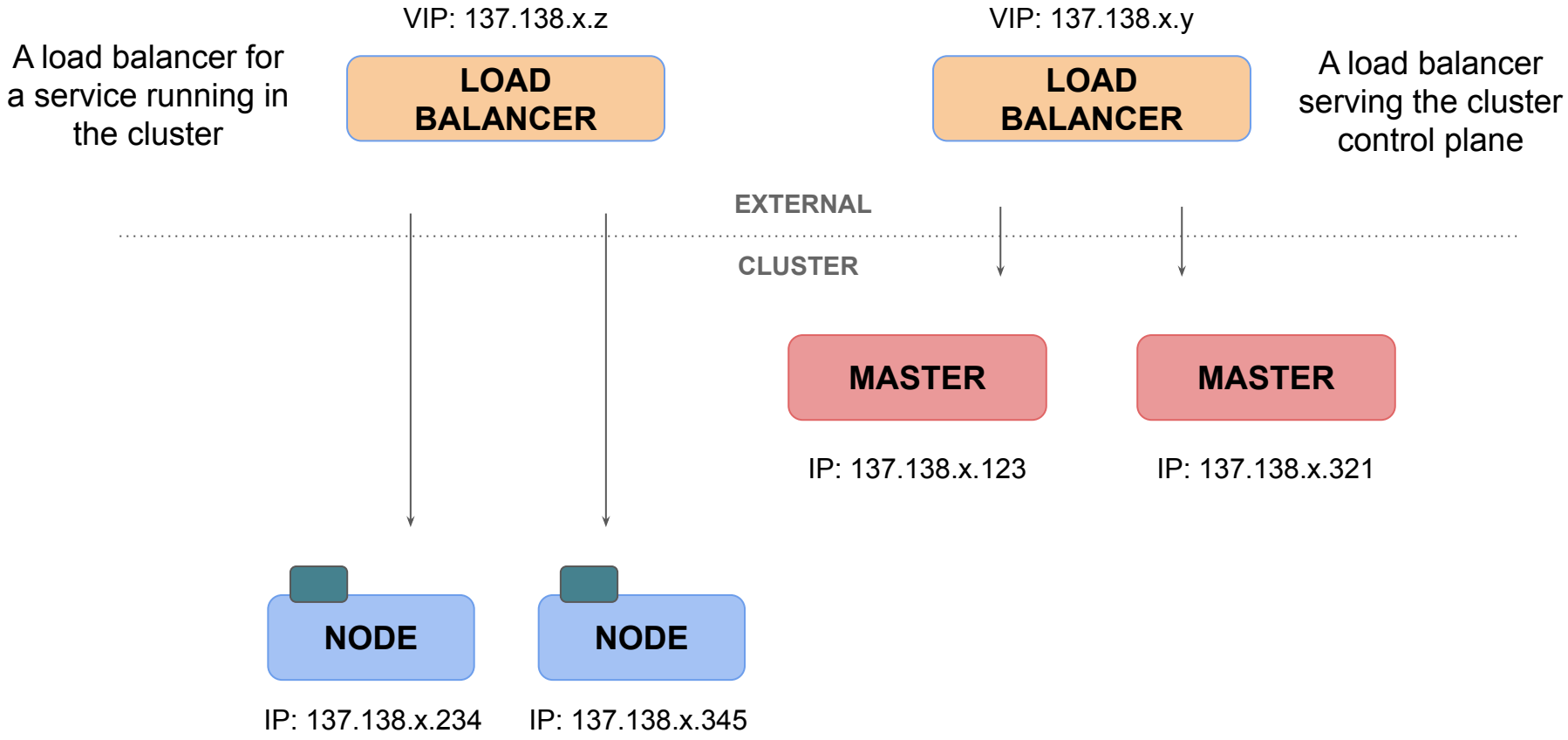
Stateful Workloads

Storage Provisioning and Access

Multi Cluster Services

Gateway, GatewayClass, Route

Questions?


A load balancer for a service running in the cluster
VIP: 137.138.x.z
LOAD BALANCER
VIP: 137.138.x.y
LOAD BALANCER
A load balancer serving the cluster control plane
EXTERNAL
CLUSTER
MASTER
MASTER
IP: 137.138.x.123
IP: 137.138.x.321
NODE
NODE
IP: 137.138.x.234
IP: 137.138.x.345