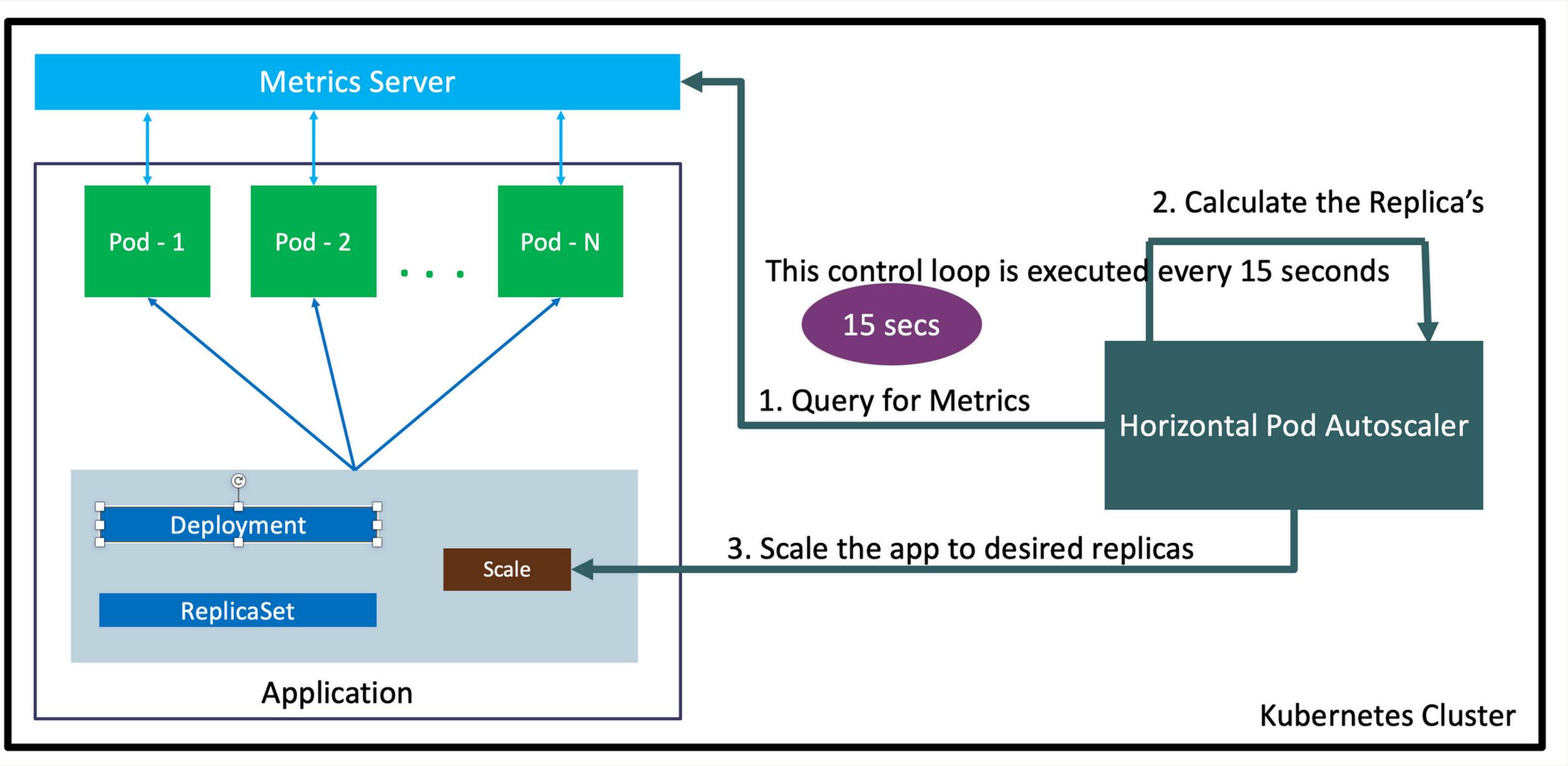
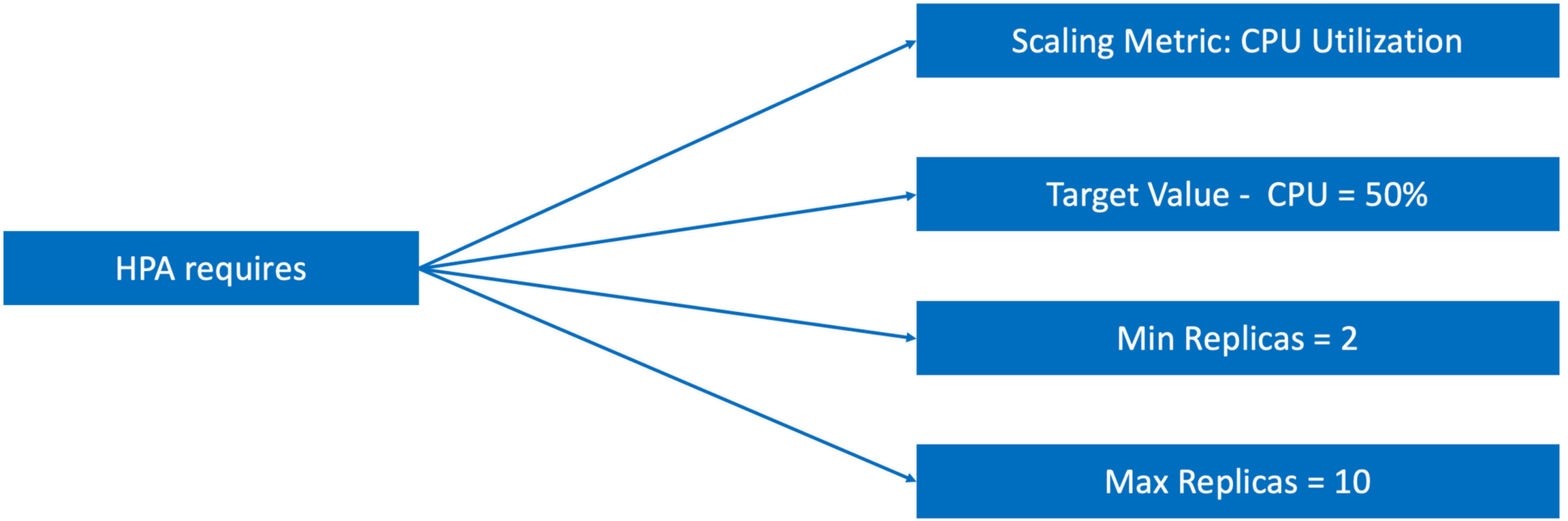
Kubernetes Horizontal Pod Autoscaler

-- Balakrishna

what we will learn today?

- *Ingress Demo*
- *HPA Architecture.*
- *HPA Demo*


Metrics Server
Pod - 1
Pod - 2
Pod - N
Deployment
ReplicaSet
Scale
Application
This control loop is executed every 15 seconds
15 secs
1. Query for Metrics
2. Calculate the Replica's
3. Scale the app to desired replicas
Horizontal Pod Autoscaler
Kubernetes Cluster




```
kubectl autoscale deployment demo-deployment  --cpu-percent=50 --min=1 --max=10
```

Thank You!!

Keep upskilling!!!