Kubernetes(K8S)

Amir Mohammad Rezvaninia

Winter 1403 – 2025-02-27

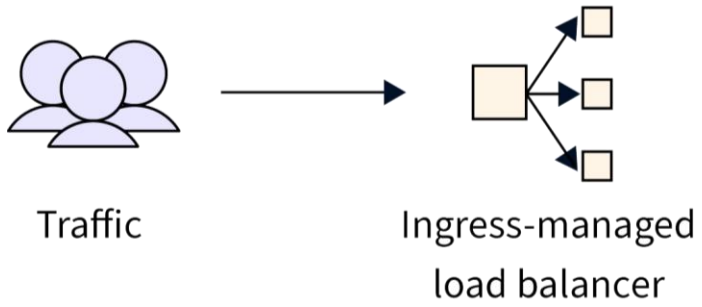
MashhadLUG

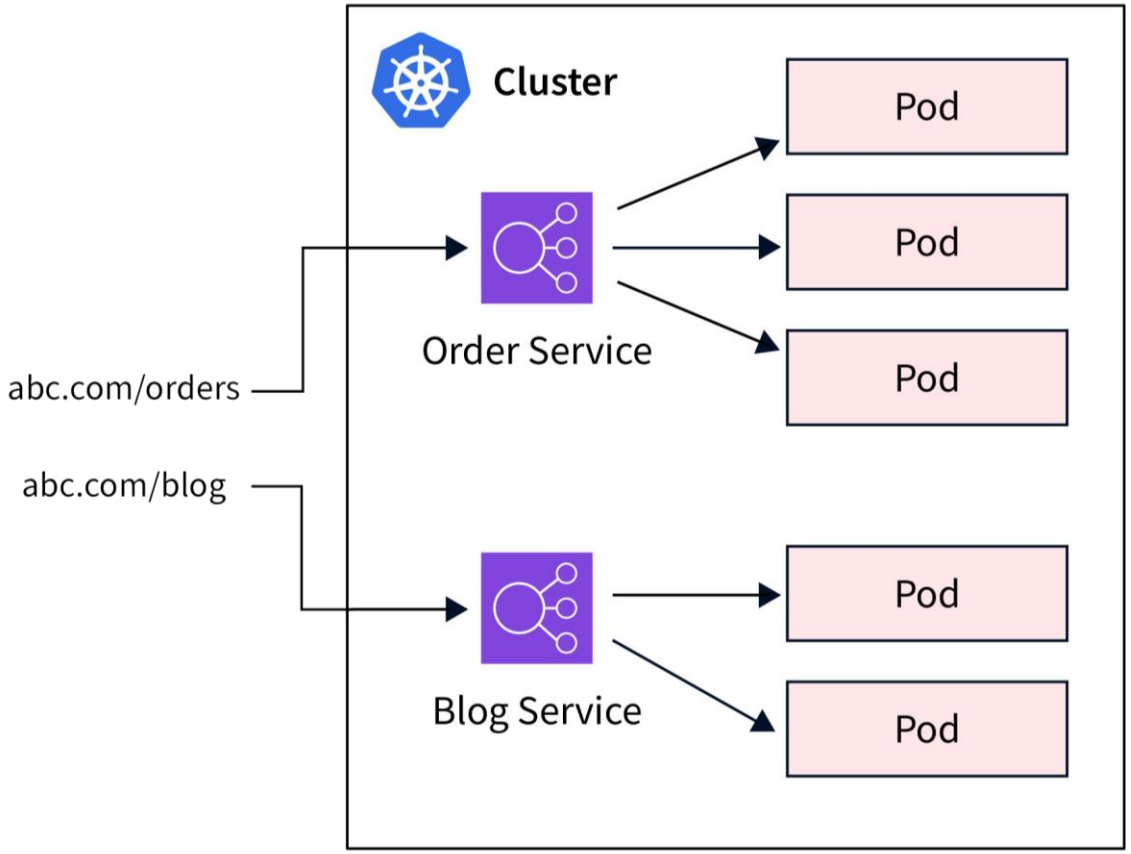
Main Challenges in Kubernetes

In a real-world environment, the workload on services is not always constant and fluctuates based on user requests.

✅ **During low load:** Running too many pods leads to resource wastage and increased costs.

✅ **During high load:** Insufficient resources can cause performance degradation, increased latency, or even service crashes.



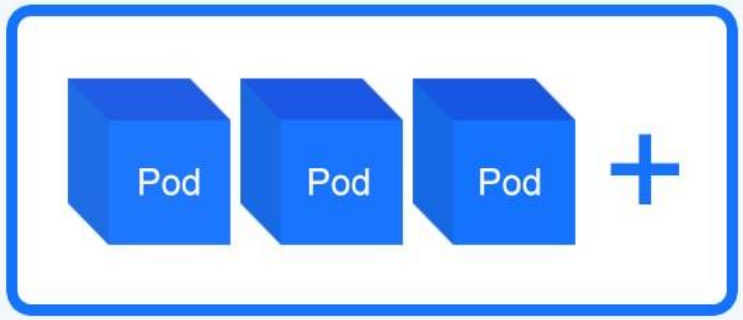


⏳ How can we solve this issue?

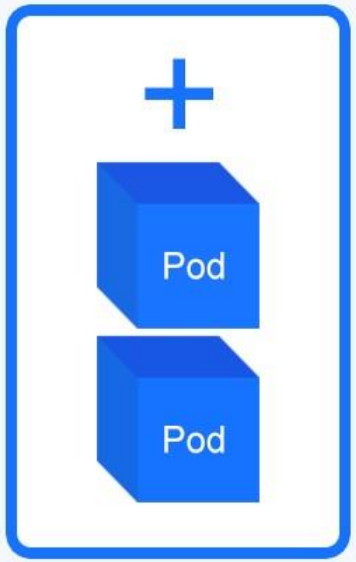
Kubernetes Auto Scaling helps dynamically adjust the number or size of pods based on workload demand.

Kubernetes Autoscaling

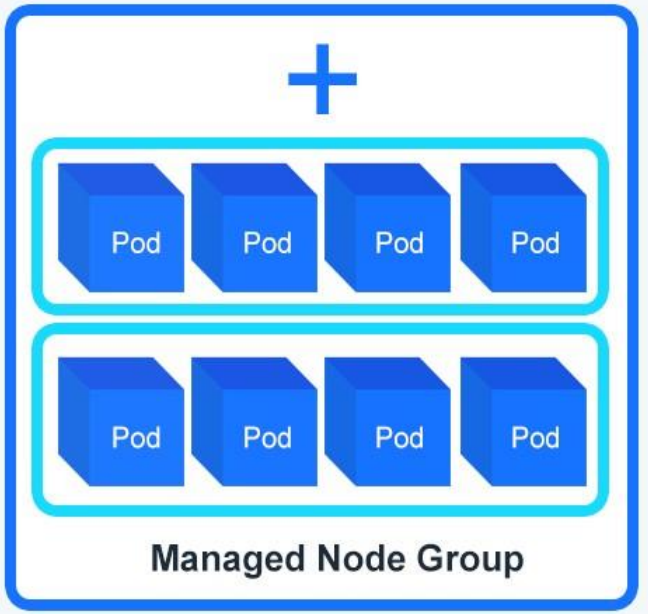
Different Methods for Autoscaling in Kubernetes



Horizontal Pod Autoscaler (HPA)



Vertical Pod Autoscaler (VPA)



Cluster Autoscaler

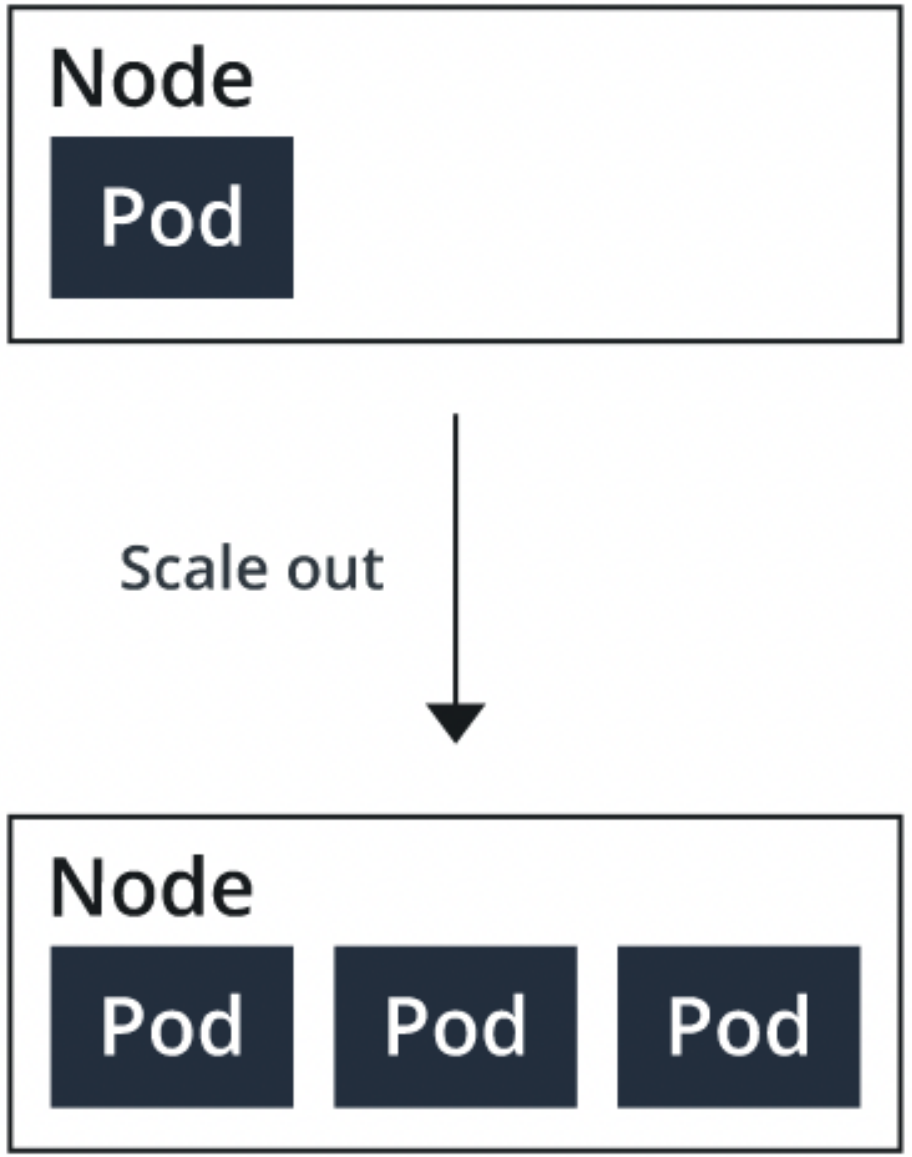
Horizontal Pod Autoscaler (HPA)

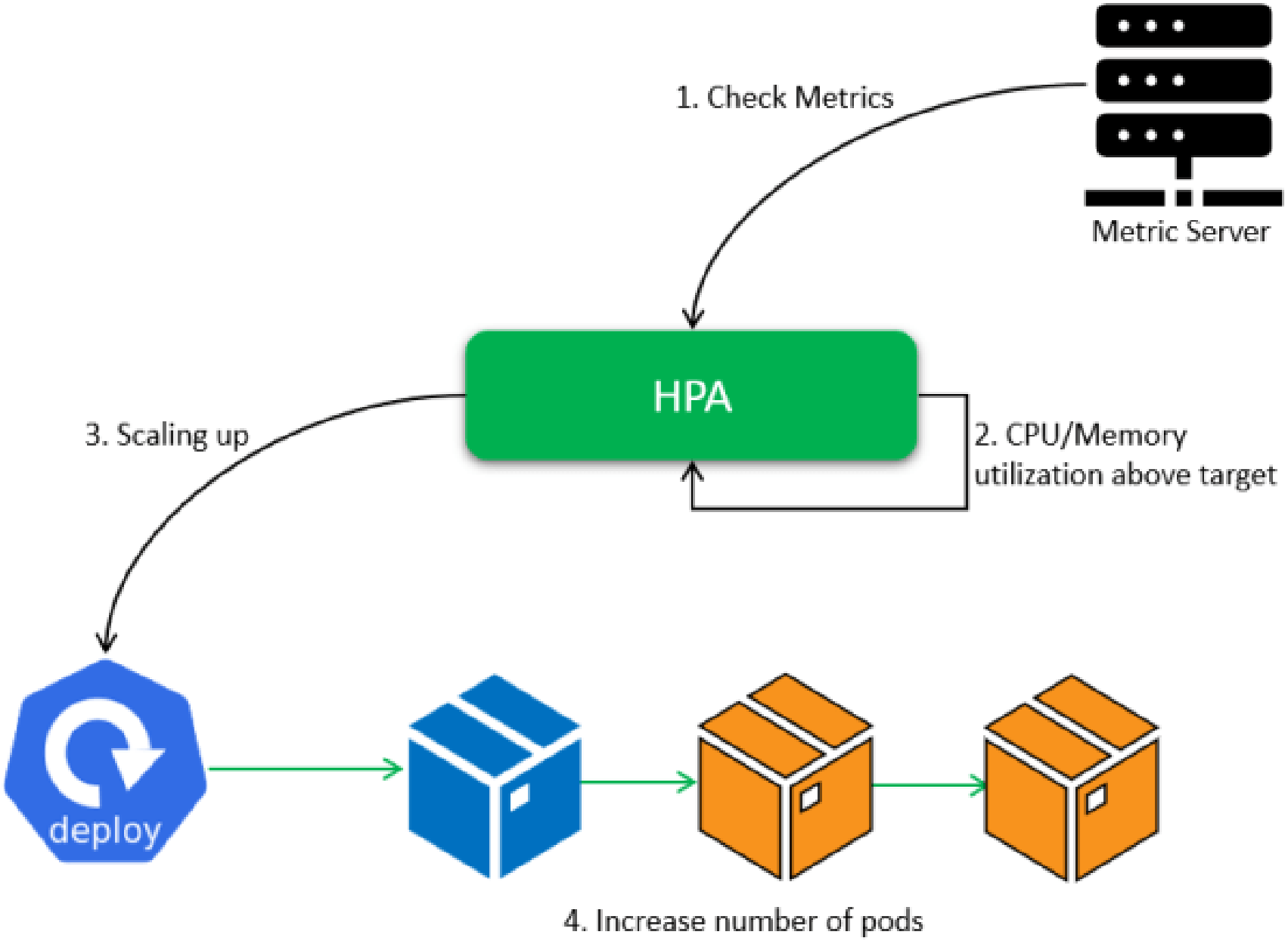
Implementation Steps and Explanation

- **HPA (Horizontal Pod Autoscaler)** automatically increases or decreases the number of pods in a Deployment or ReplicaSet based on CPU and RAM utilization.

👉 Practical Example:

Imagine you have a website that experiences high traffic during peak hours. If the load on the servers increases, the website may slow down. HPA dynamically scales the number of pods to distribute the load efficiently and maintain performance.

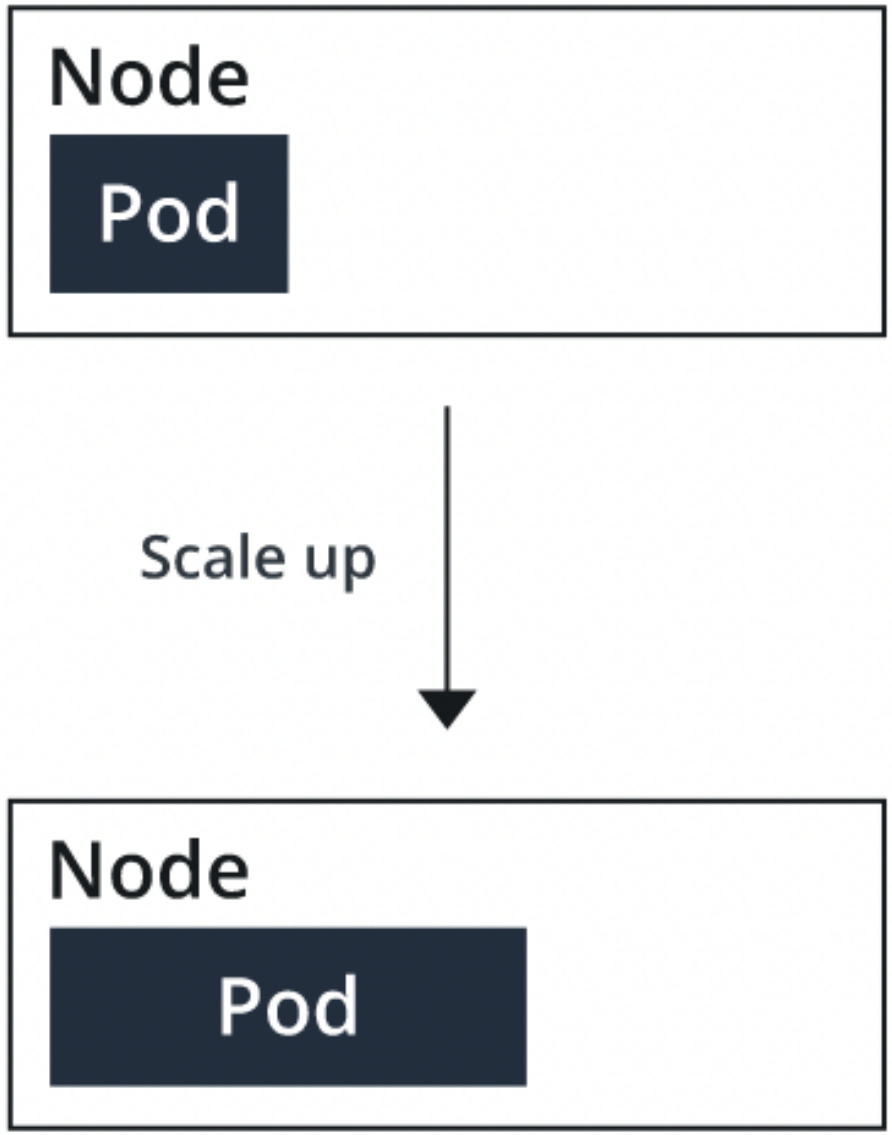



1. Check Metrics
Metric Server
HPA
2. CPU/Memory
utilization above target
3. Scaling up
deploy
4. Increase number of pods

Vertical Pod Autoscaler (VPA)

Implementation Steps and Explanation

- **VPA (Vertical Pod Autoscaler)** automatically increases or decreases the allocated CPU and RAM for a pod.



Practical Example:

Imagine you have a machine learning application running in Kubernetes. Over time, as the workload increases, the application requires more memory and CPU to process data efficiently. Instead of manually adjusting resource limits, VPA dynamically increases the allocated CPU and RAM, ensuring optimal performance without unnecessary resource wastage.

amirmohammadrezvaninia

amirmohammadre

@amirmohammadre

MashhadLUG