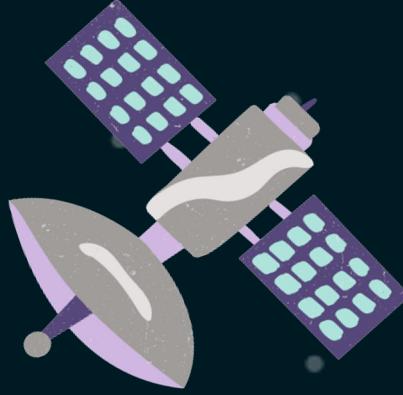
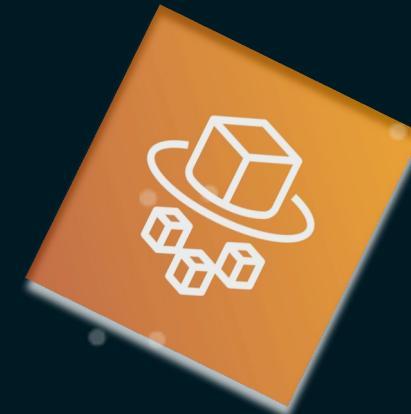
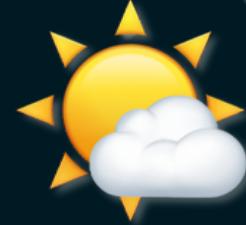


# Kubernetes Autoscalers





# What is K8s Autoscaler ?



Automatically adjusts the number of pods or nodes



Optimal resource utilization



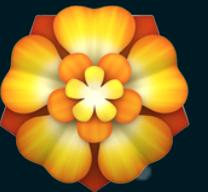
Maintains APP performance



# Overview of K8s AutoScalers



Pod Level



Horizontal Pod Autoscaler



~~Vertical Pod Autoscaler~~



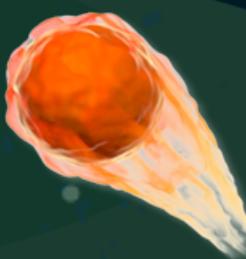
Node Level



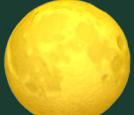
Cluster Autoscaler

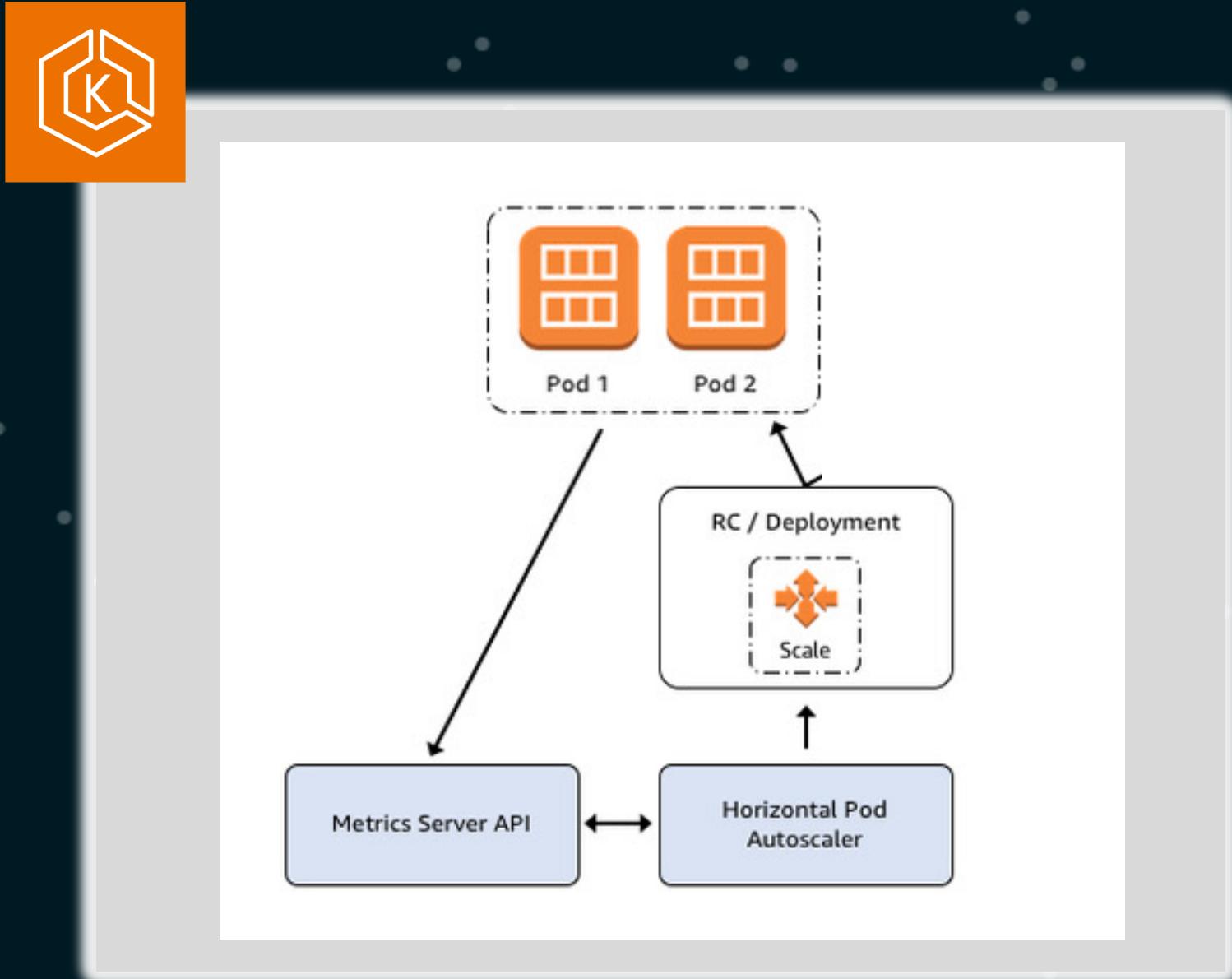


Karpenter



# Horizontal Pod Autoscaler

-  Automatically scales the number of pods
-  Autoscaler rule is based on CPU/Memory (Metrics Service) or Customer Metrics (Prometheus)





# Cluster Autoscaler

- UP! Adjusts add/remove nodes based on pending pods and resource requirements
- Ensures efficient cluster resources
- Optimal performance

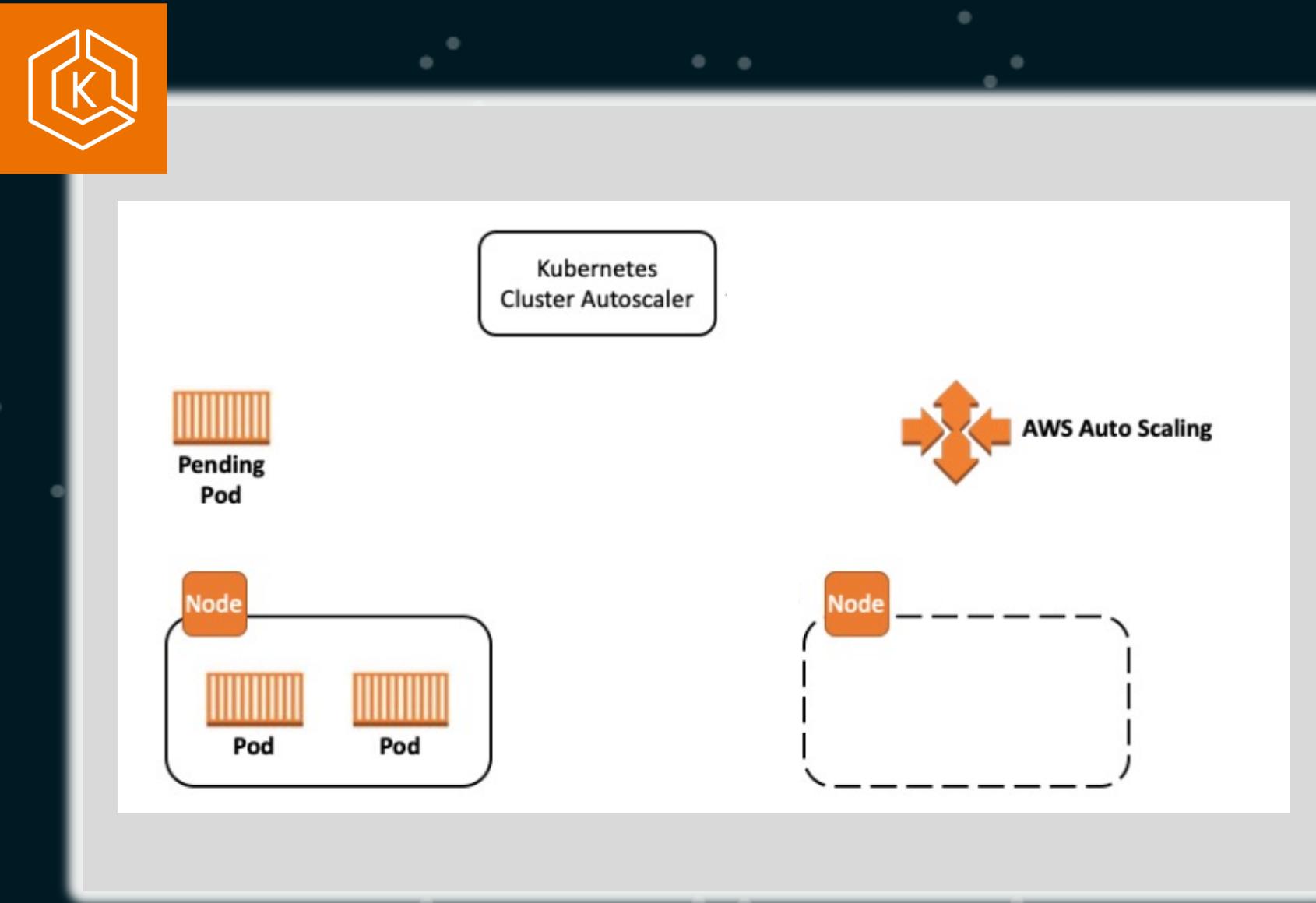


Image is from <https://aws.github.io/aws-eks-best-practices/cluster-autoscaling/architecture.png>

# Karpenter

- 🙌 Provides just-in-time compute resources
- 🚧 Instance lifecycle management
- 🏛️ Launches right-sized Node
- 🎯 Support Spot Instances

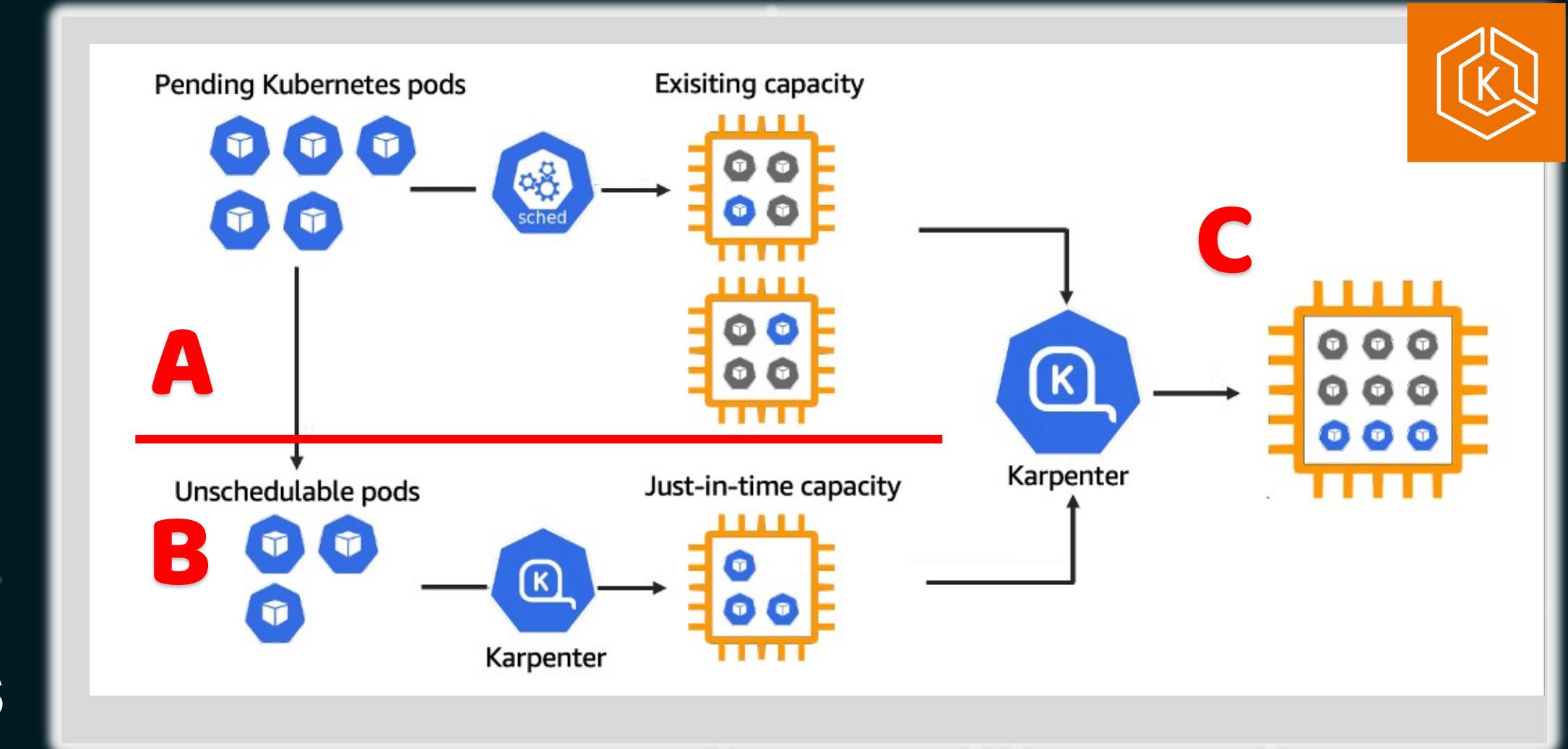


Image is from <https://d2908q01vomqb2.cloudfront.net/da4b9237bacccdf19c0760cab7aec4a8359010b0/2021/11/23/2021-karpenter-diagram.png>

# Thanks

Any questions?

You can find me at:  
[https://www.linkedin.com/in](https://www.linkedin.com/in/lian-duan-23b81059/)

