AWS re:Invent

NOV. 29 – DEC. 3, 2021 | LAS VEGAS, NV

CMP404

Running cost-optimized Spark workloads on Kubernetes

Pallami Bhattacharjee (she/her)
Sr. Software Dev Engineer,
EC2 Core Product Management
AWS

Kinnar Kumar Sen (he/him)
Sr. Specialist Solutions Architect, EC2 Spot
AWS

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Agenda

Why are we here today?

What are the AWS services involved?

How does it all work? (whiteboard)

Demo and statistics

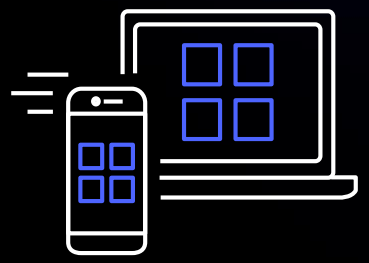
Key takeaways and Q&A

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

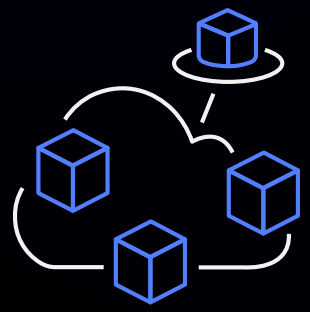
Why are we here today?

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

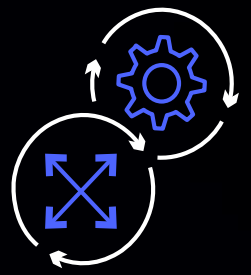
What customers want



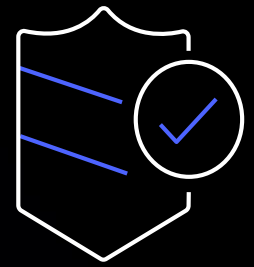
Build applications, not infrastructure



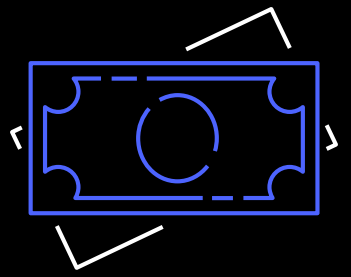
Manage infrastructure to their requirements



Scale quickly and seamlessly



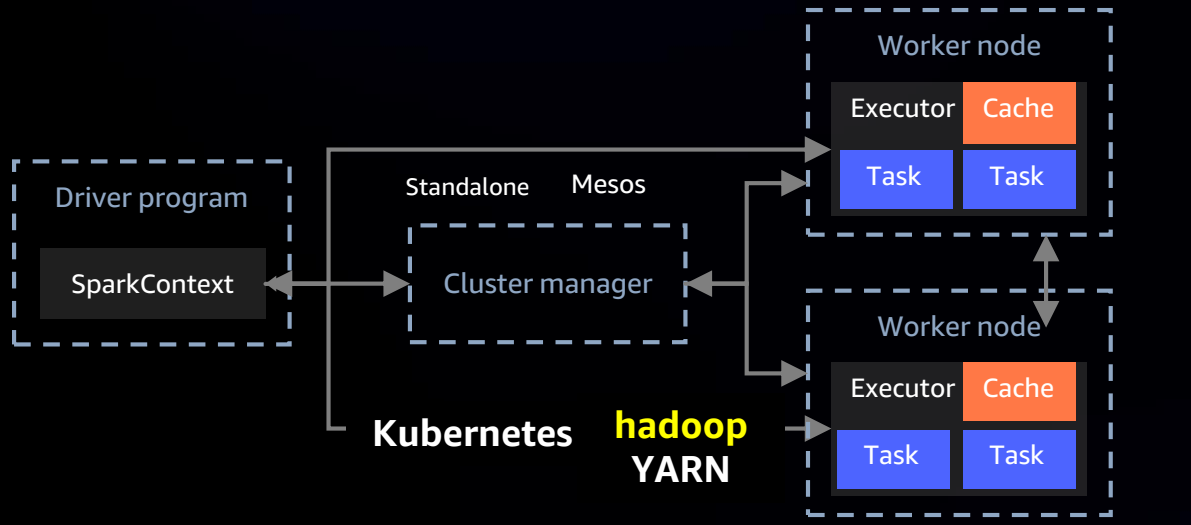
Security and isolation by design



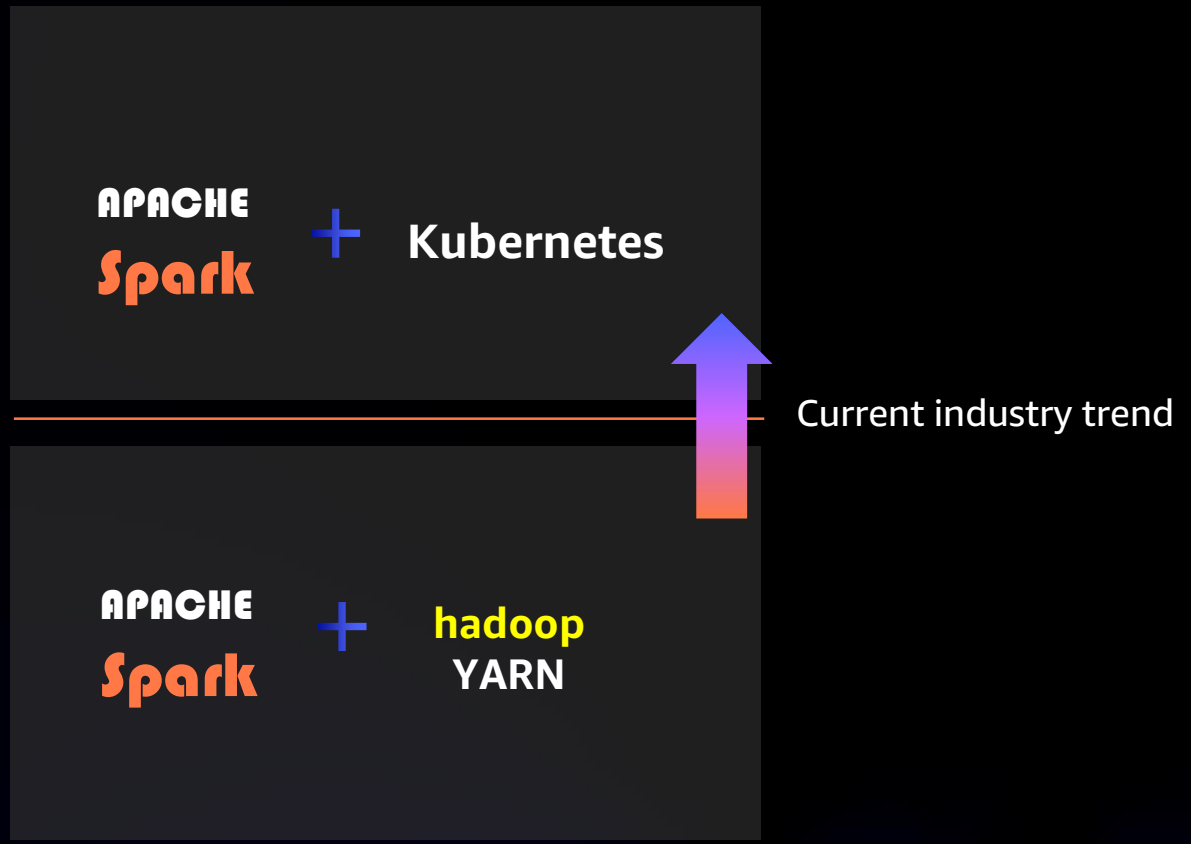
Optimize cost of operation

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Back to basics

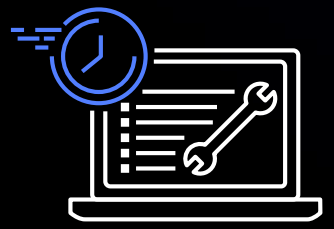


Spark deployment



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Why Spark on Kubernetes



Deployment and portability



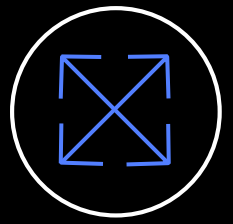
Resource sharing



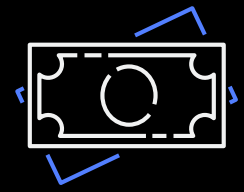
Secure and resilient



Observability



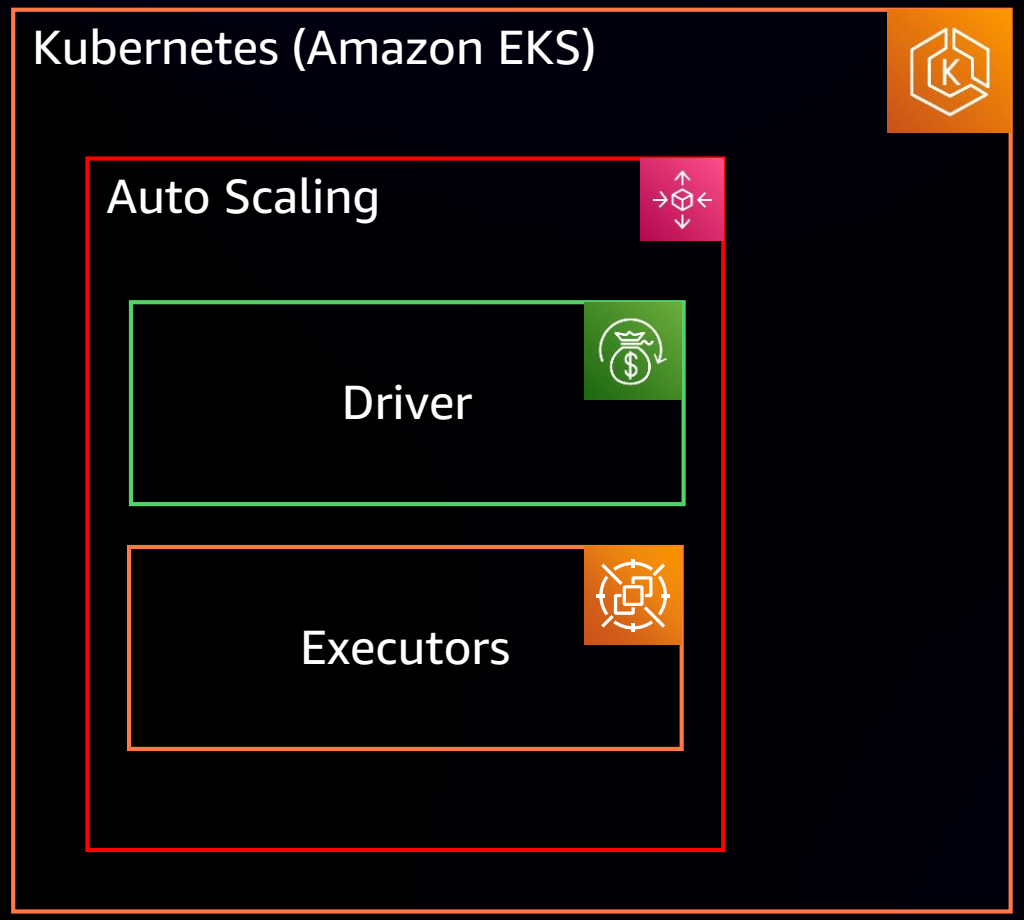
Auto Scaling



Cost optimization

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Logical architecture for cost optimization



Amazon EC2 Spot

up to 90% off compared to on-demand pricing

AWS Savings Plans

up to 72% for committed compute usage

ASG

on-demand scaling

Amazon EKS

optimal resource sharing and managed service

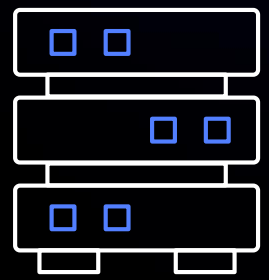
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

What are the AWS services involved?

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

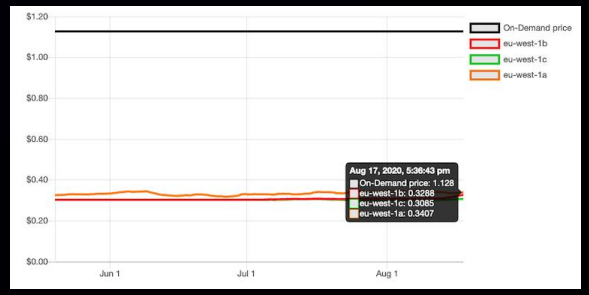
Amazon EC2 Spot

LEVERAGE SPARE CAPACITY ON DEMAND WITH STEEP DISCOUNTS, PAID BY THE SECOND



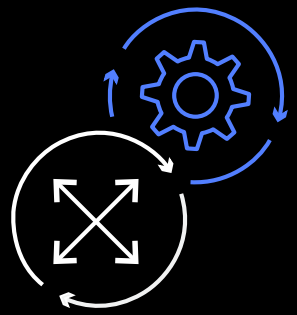
Spare Amazon EC2 capacity

Same infrastructure as on-demand



Up to 90% off compared to on-demand pricing

Spot price is based on long-term supply and demand trends of Amazon EC2 Instances (no bidding!)



Interruptible

Spot Instances can be interrupted with a two-minute warning if Amazon EC2 needs the capacity back

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

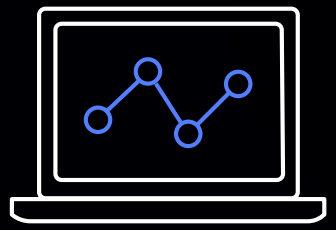
AWS Savings Plans

A FLEXIBLE PRICING MODEL THAT HELPS YOU SAVE UP TO 72% FOR COMMITTED COMPUTE USAGE



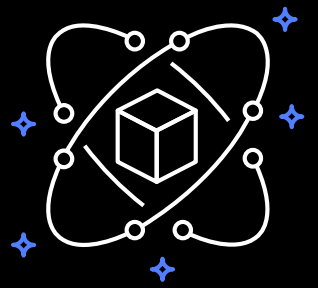
Significant savings

Maximize **cost optimization** efforts by broadly applying discounts throughout your portfolio



Easy to use

Dramatically improves **operational efficiency** by simplifying operational overhead for infrastructure management



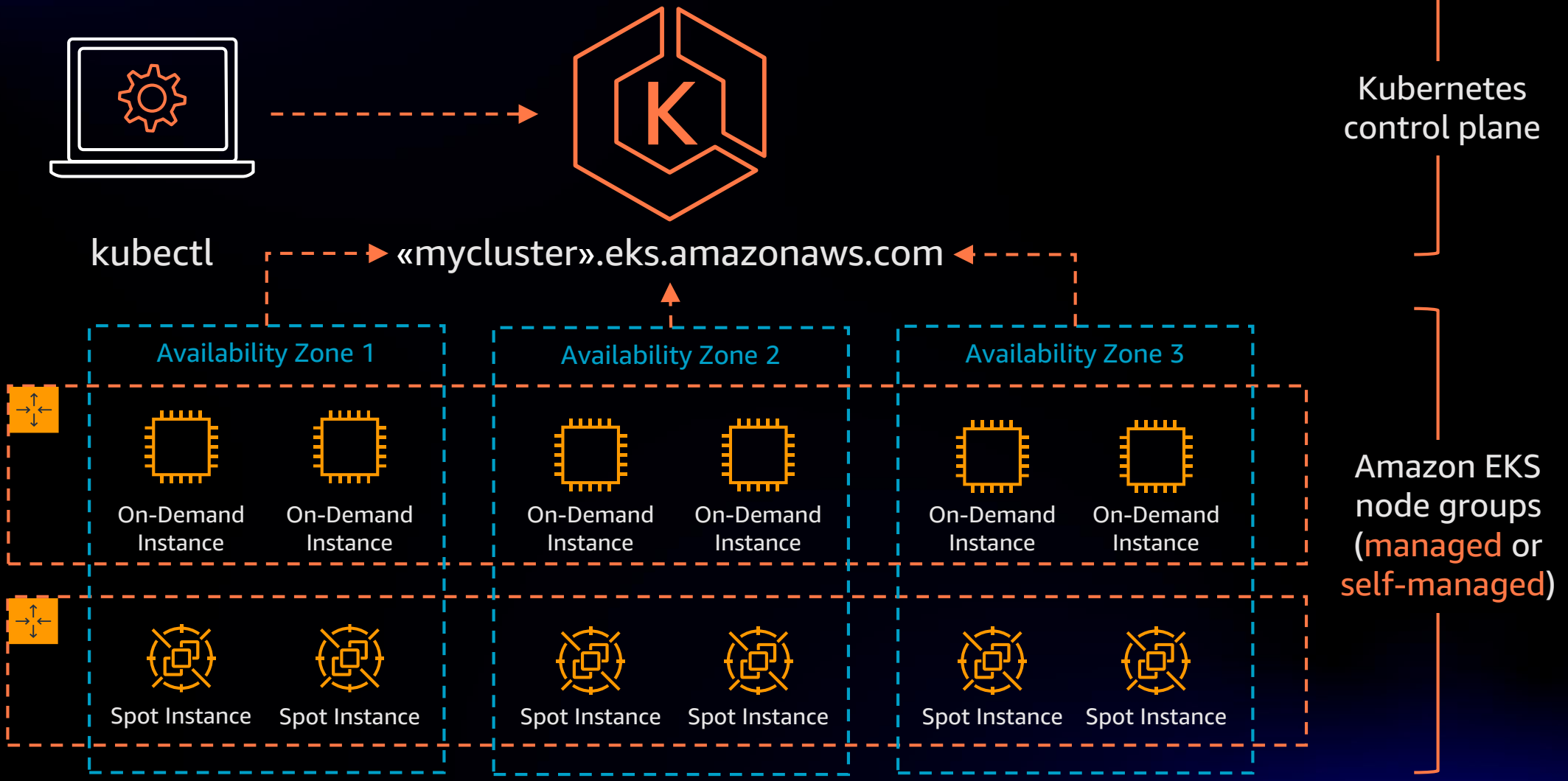
Flexible

Supports **modernization** by seamlessly integrating applications across instances and serverless offerings

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Amazon EKS



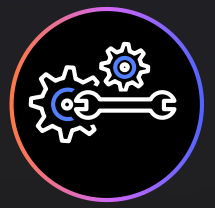
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Amazon EKS managed node groups with Amazon EC2 Spot Instances



Automatic node provisioning

- Support for mixed instance types
- Amazon EC2 Auto Scaling group provisioning with capacity-optimized allocation strategy
- Automatically configures Kubernetes labels and AWS tags (e.g., cluster autoscaler)



Managed node update and termination

- Managed node draining during node termination and updates



Spot Instance lifecycle management

- Managed Spot Instance draining integrated with Amazon EC2 Auto Scaling Capacity Rebalancing



Launch template support

- Custom configuration and custom AMIs for your Kubernetes nodes

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

How does it all work?

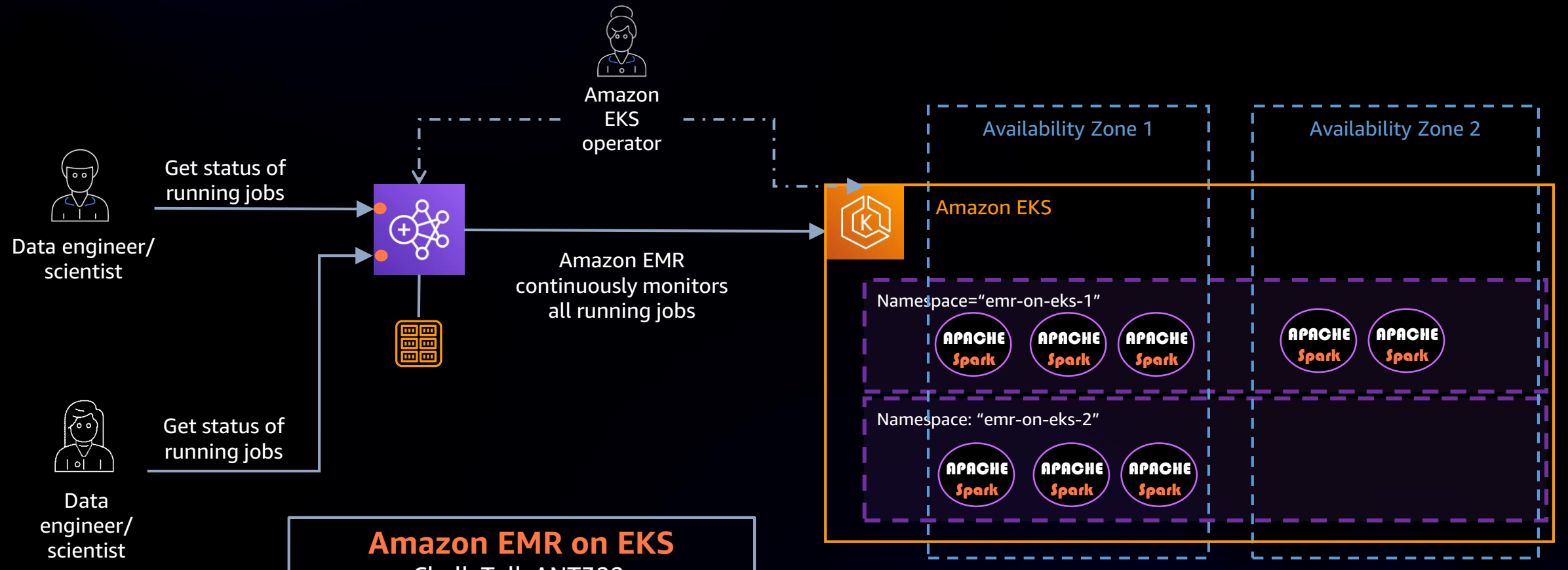
Whiteboard

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Where is this applicable?

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

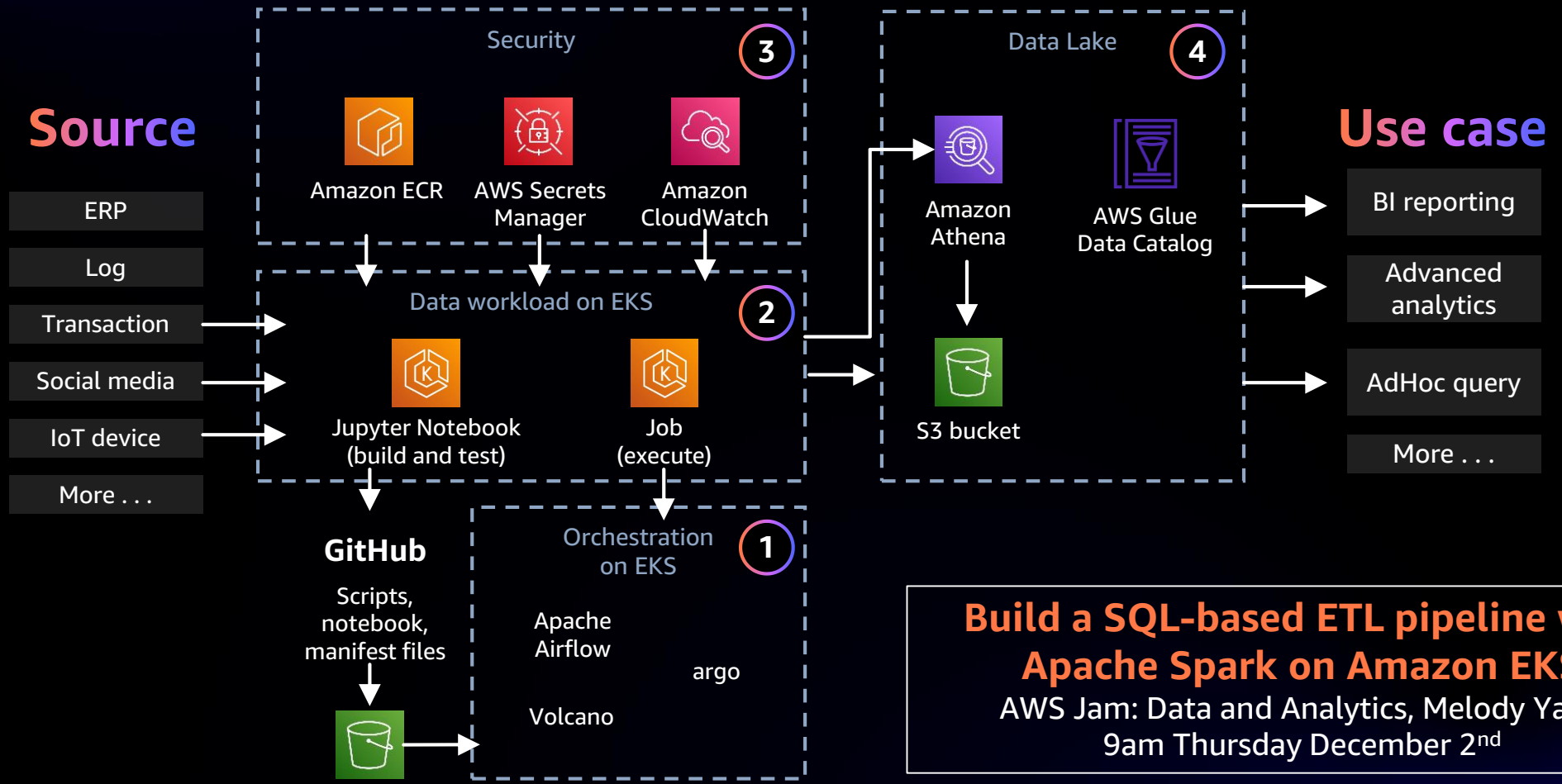
Amazon EMR on Amazon EKS: Job-centric workflow



Amazon EMR on EKS
Chalk Talk ANT322,
Vincent Gromakowski
5 p.m., Tuesday November 30th

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Submit Spark jobs using SQL



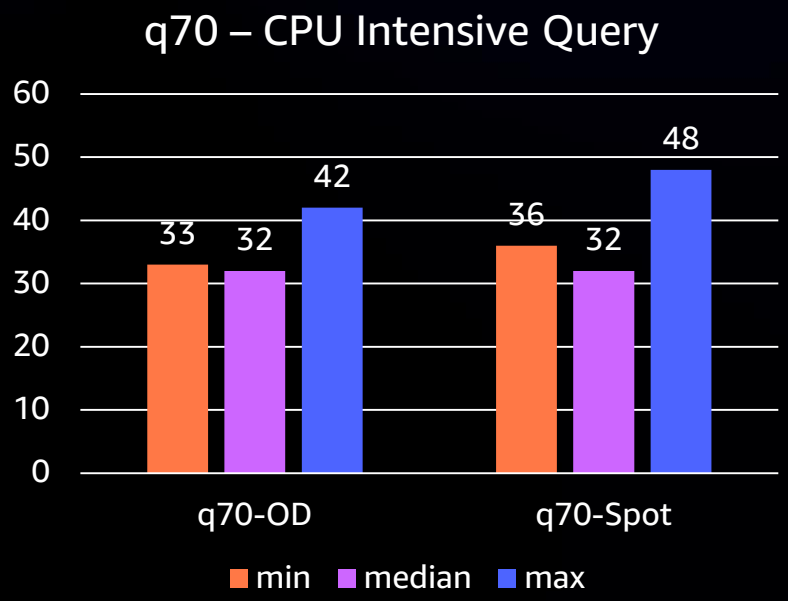
Build a SQL-based ETL pipeline with Apache Spark on Amazon EKS

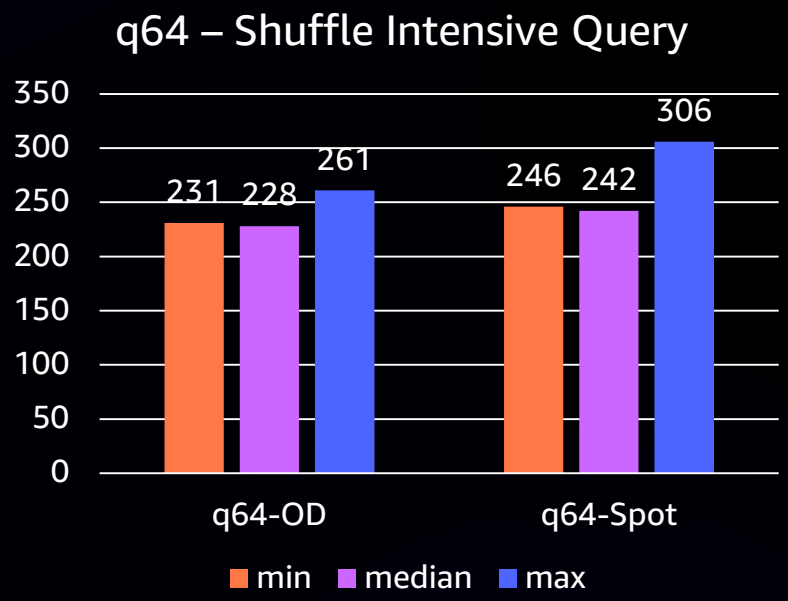
AWS Jam: Data and Analytics, Melody Yang

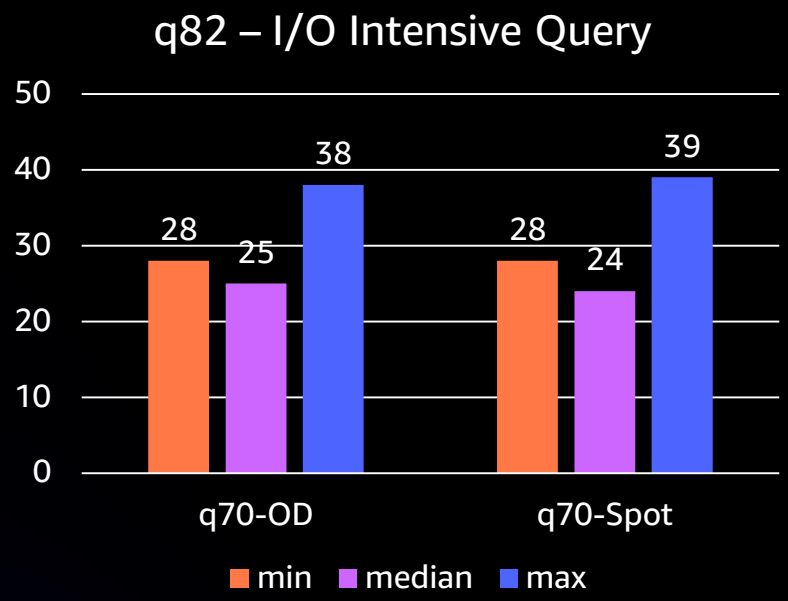
9am Thursday December 2nd

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Statistics (TPC-DS 1T)







Spark 3.1.2 | Driver :1 Executor :50 | Cores - 2 Memory - 8000m

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Key takeaways

- Use EC2 Spot for Spark executors and On-Demand Instances with AWS Savings Plans for Spark drivers
- Use cluster Auto Scaler for scaling
- Use managed nodegroups (Spot and On-Demand)
- Diversify EC2 Spot Instances with similar vCPU/Mem ratio in a node group
- Use ASG with capacity optimized allocation strategy (for self-managed node groups)
- Use Spark decommission and dynamicAllocation
- Decouple compute and storage
  - S3 for input/output
- Reduce shuffle size and blast radius

EC2 spot labs GitHub: Cost-optimized Spark on K8

EMR on EKS workshop

SQL-based ETL implementation

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Thank you!

Pallami Bhattacharjee
pallami@amazon.com

Kinnar Kumar Sen
kinnarse@amazon.com

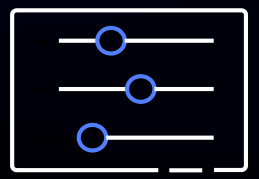
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AWS re:Invent

NOV. 29 – DEC. 3, 2021 | LAS VEGAS, NV

Flexibility is key to successful adoption

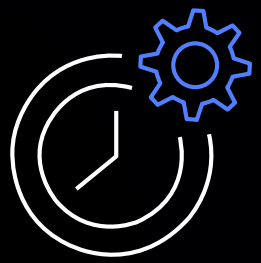
Instance flexible



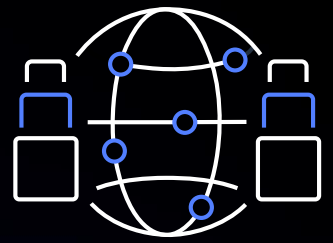
- More than one instance type can get the job done
- Instance weighting gives you more flexibility on instance types
- Multiple instance types are key to resilient clusters

OR

Time flexible



Region flexible



Be time flexible to account for interruptions and/or location flexible to maximize application uptime

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

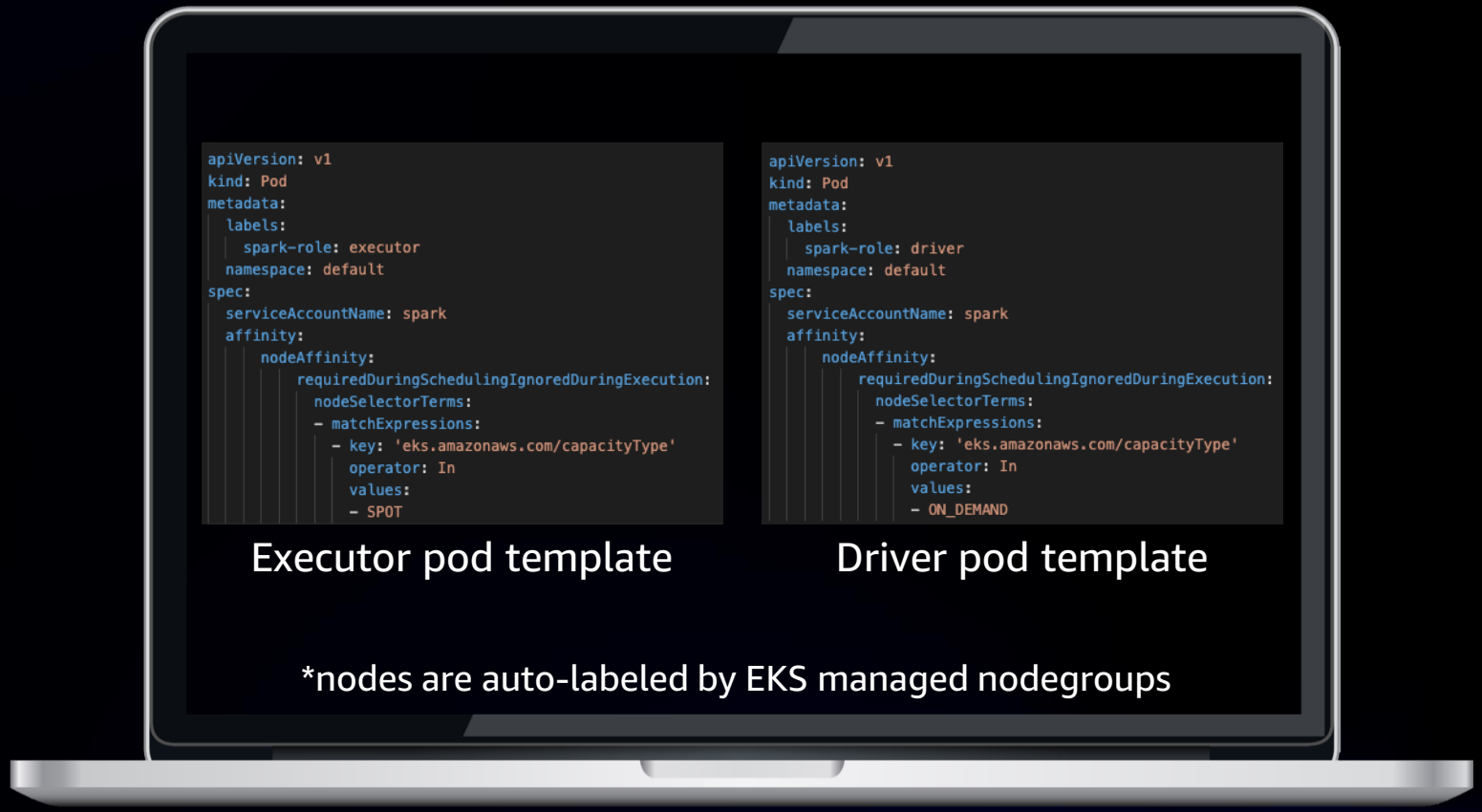
Scheduling pods

How?

- nodeSelector
- Affinity and anti-affinity

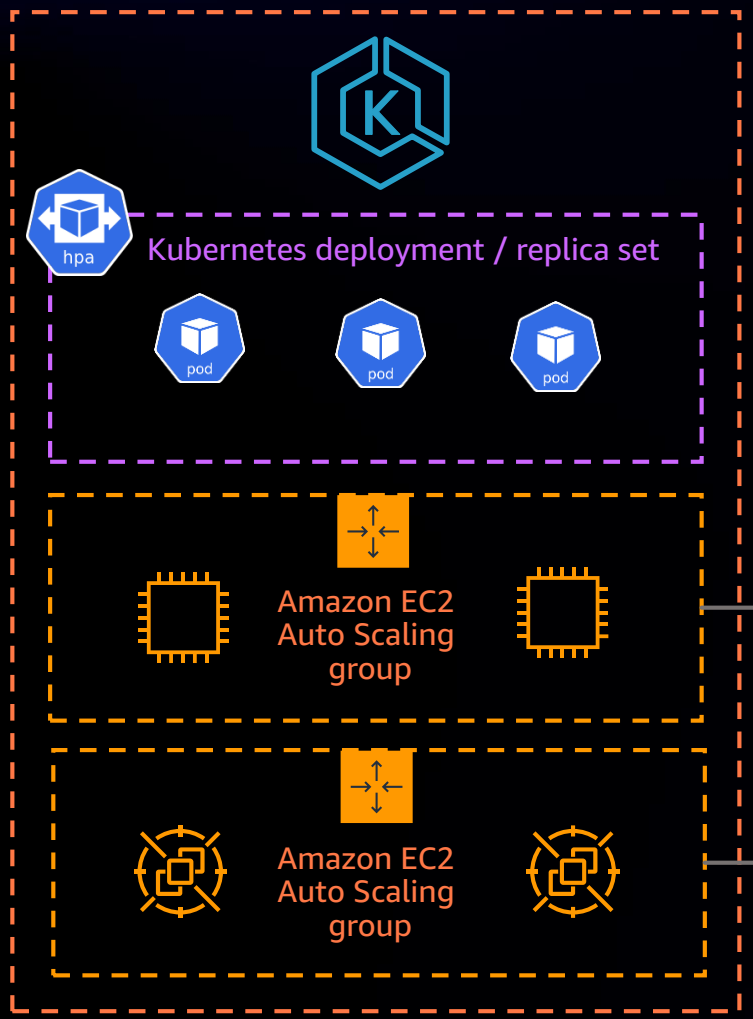
Pod templates (Spark 3.0.0)

- Specifications for creating pods



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Scaling



Spark dynamic allocation

- Allow dynamic allocation without an external shuffle service (SPARK-27963)

Kubernetes cluster autoscaler

- Scale out Amazon EC2 Auto Scaling groups of nodes matching pod scheduling requirements when pods are in a pending state because of insufficient resources
- Scale in when nodes are underutilized and important pods can be rescheduled somewhere else

Auto Scaling group tags:
k8s.io/cluster-autoscaler/enabled: true
k8s.io/cluster-autoscaler/${cluster_name}: owned

aws

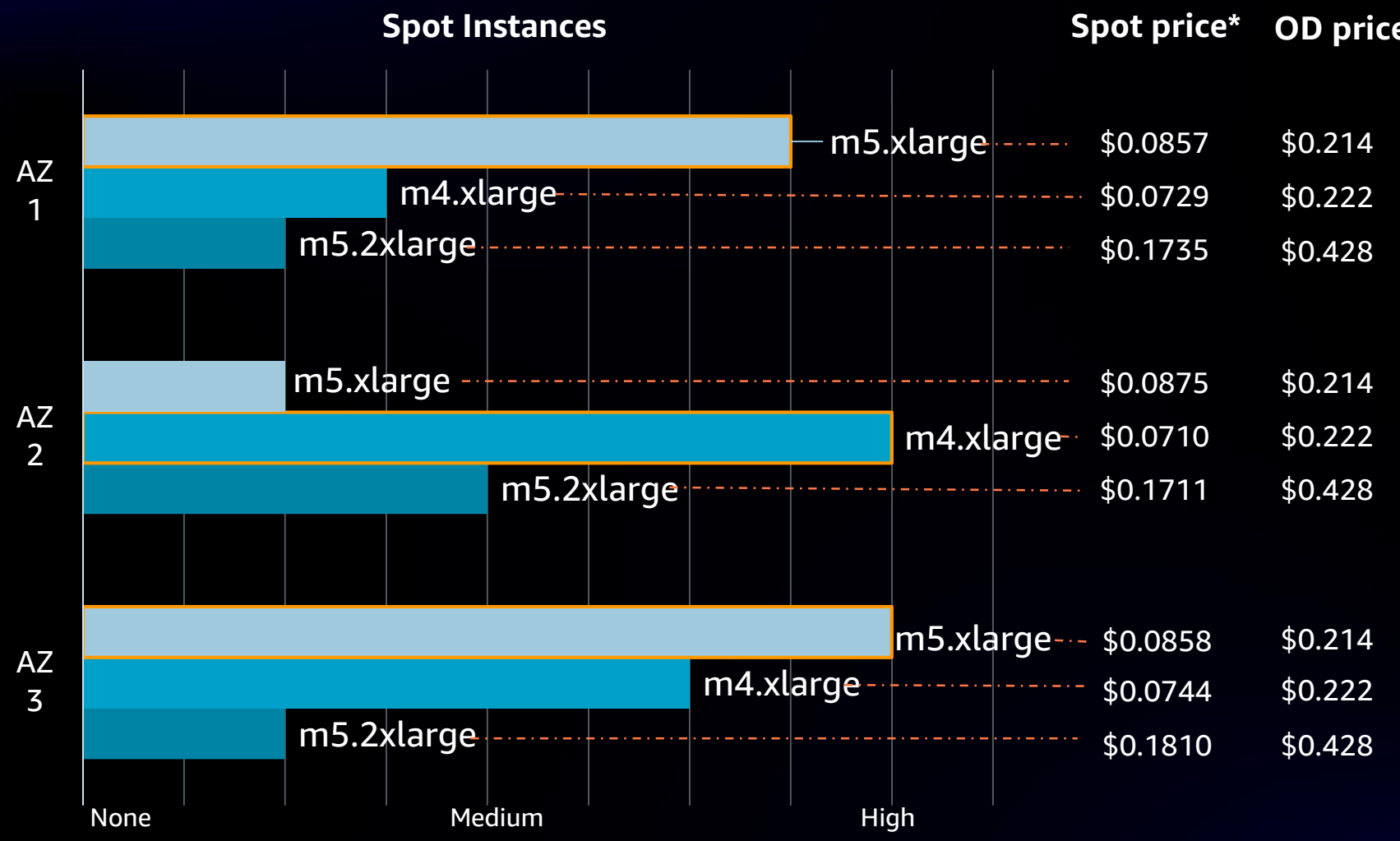
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Karpenter (experimental)

https://github.com/awslabs/karpenter

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Understanding Spot capacity pools



Each instance type

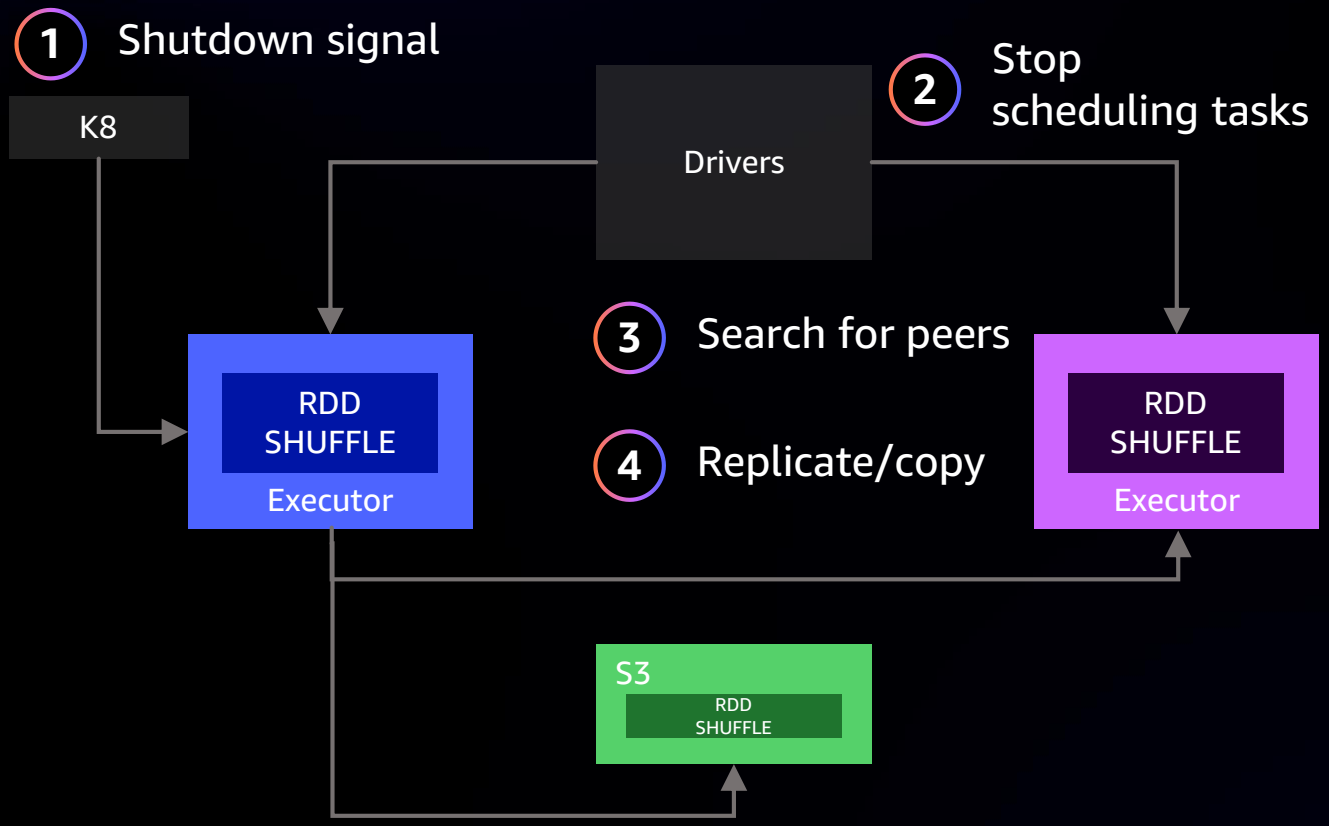
In each Availability Zone

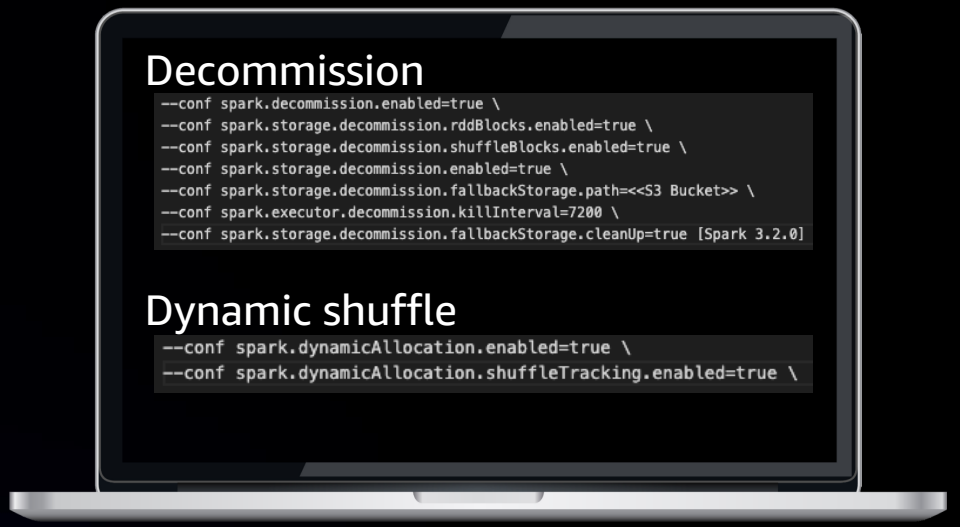
In every Region

Is a separate Spot capacity pool

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Node decommission feature





*reduces the time taken to complete Spark jobs on EC2 Spot

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Configuring your Spot-managed node group



my-cluster-definition.yaml

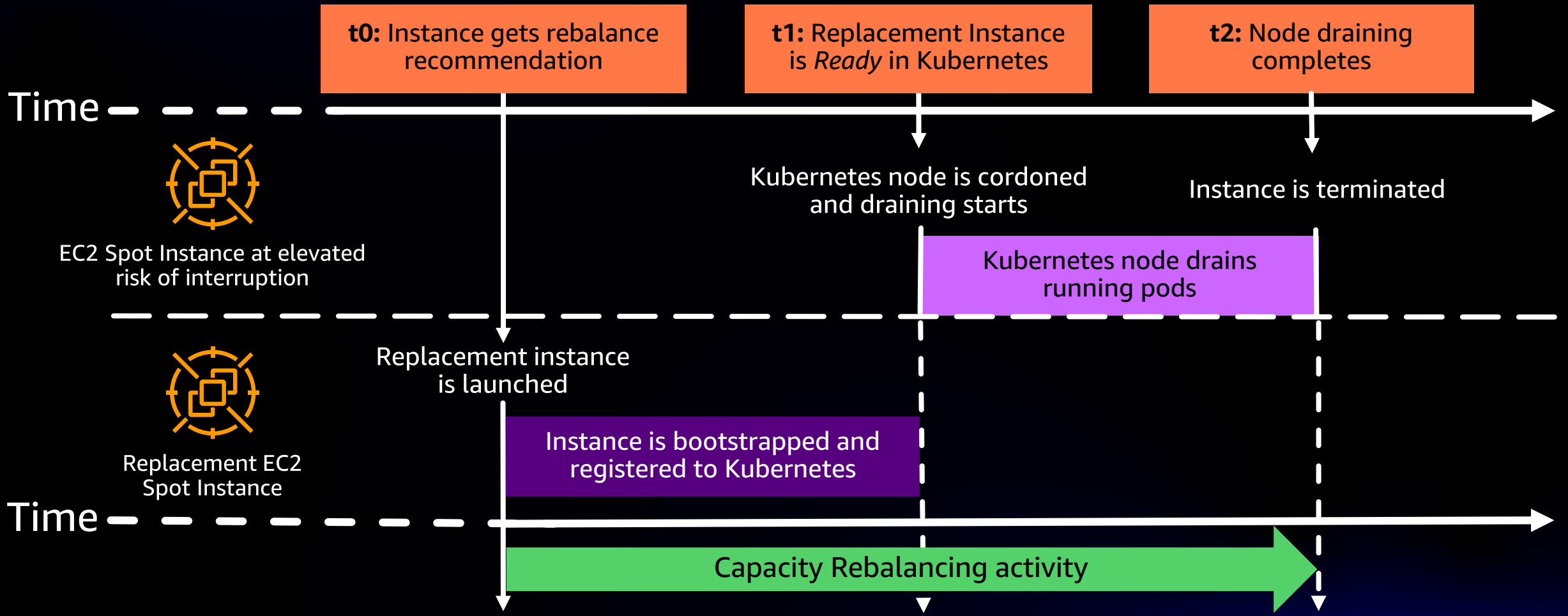
```
apiVersion: eksctl.io/v1alpha5
kind: ClusterConfig
metadata:
    name: my-eks-cluster
    region: us-west-2
managedNodeGroups:
        - name: my-spot-managed-node-group
          desiredCapacity: 2
          spot: true
          instanceTypes:
                - m5.xlarge
                - m5d.xlarge
                - m5a.xlarge
                - m5ad.xlarge
                - m4.xlarge
```

Select same-size instance types (4 vCPUs and 16 GiB memory) for consistent scaling performance with Kubernetes Cluster Autoscaler

```
$ eksctl create nodegroup -f my-cluster-definition.yaml
```

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

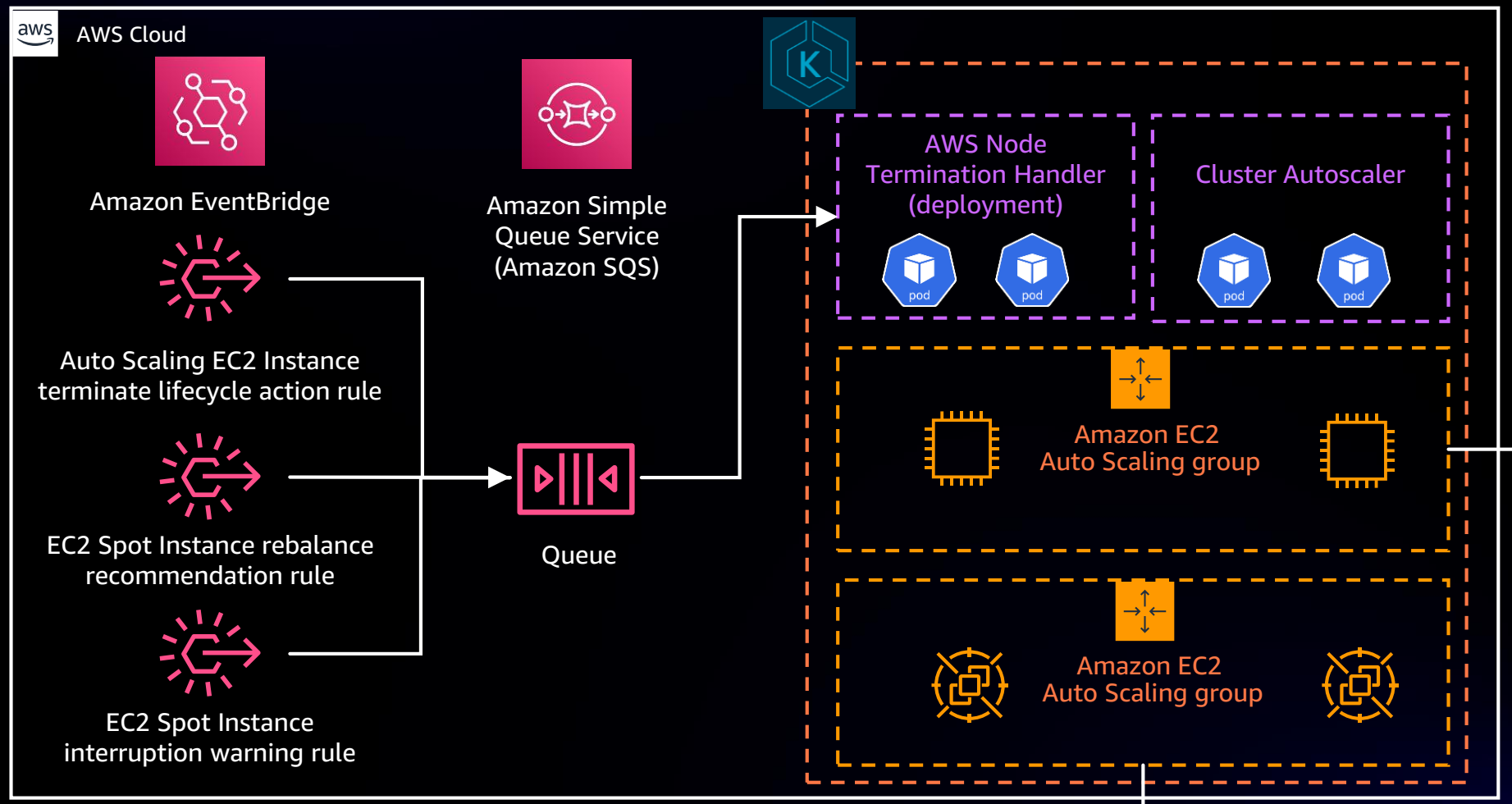
Managed node group Spot Capacity Rebalancing activity



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Handling interruptions on (self-managed) Kubernetes

AWS node termination handler queue processor for self-managed node groups/DIY Kubernetes



All ASGs configuration:

EC2_INSTANCE_TERMINATING lifecycle hook: enabled

- Scale-in events
- AZ Rebalance events
- Capacity Rebalance Recommendations

Spot Instances ASGs:

Capacity Rebalancing: enabled

- Attempts launch of replacement capacity
- Upon succesful replacement, triggers instance termination lifecycle hook

https://github.com/aws/aws-node-termination-handler

Auto Scaling group tags:
`aws-node-termination-handler/managed:`

es. All rights reserved.

AWS re:Invent

NOV. 29 – DEC. 3, 2021 | LAS VEGAS, NV

AWS re:Invent

NOV. 29 – DEC. 3, 2021 | LAS VEGAS, NV