Optimize AI/ML, Kubernetes and Big Data Workloads at Your Startup

Ran Sheinberg Principal Specialist Solutions Architect, EC2 Spot

Scott Horsfield Sr. Specialist Solutions Architect, EC2 Spot

© 2020, Amazon Web Services, Inc. or its Affiliates. All rights reserved. Amazon Confidential

Enabling lean startups with AWS

Zero upfront cost

With AWS's infrastructure-on-demand, startups can pay only for the resources they use instead of investing in servers upfront

Launch faster

Startups can have new IT resources available in just a few clicks, increasing agility

Focus on core business value

Startups can focus on growing their business rather than on infrastructure

Experiment often at lower risk

Being able to deprovision servers as needed enables startups to experiment often and fail fast if an idea doesn't work

aws

Once you start using Amazon EC2 you want more

More of the same

Increase application throughput

More workloads

Workloads not possible w/o the cloud

How can I do more without increasing costs?

aws

EC2 Spot Instances is how you can do more for less

- Spot Instances are spare EC2 capacity offered at up to a 90% discount
- Spot Instances offer the same features as EC2
  - ✓ Compliant with 89 security standards
  - ✓ 290+ instances and AWS Nitro System
  - ✓ 22 Availability Zones

Spot helps you migrate new workloads and also cost optimize existing ones

More Compute + Faster Results = Accelerated Innovation

aws

Accelerate innovation at your startup

Reduce Time-to-Market
Reinvest the cost savings back into your company. Hire more developers or access more compute

Extend your Runway
Significantly lowering your overall infrastructure costs, allows you to extend the life of your funding.

Build for Scale
Quickly ramp up massive data analysis jobs or optimize the resilience of your web based application.

Spot Instances help you reduce time-to-insight, build applications faster, and save on infrastructure costs you can invest back into your company.

aws

The Simple Rules of Spot

Spot infrastructure

Is **same** as On-Demand and RIs

70 - 90% off

Spot pricing

Smooth, infrequent changes no spikes, more predictable

(no bidding)

Interruptions

Happen when OD needs capacity (**no bidding**)

Diversify

Choose different instance types, size and AZ in a single fleet

© 2020, Amazon Web Services, Inc. or its Affiliates.

Diversification increases resiliency to interruption

C4	1a	1b	1c	On Demand
8XL	$0.27	$0.27	$0.29	$1.76
4XL	$0.21	$0.30	$0.16	$0.88
2XL	$0.08	$0.07	$0.08	$0.44
XL	$0.04	$0.05	$0.04	$0.22
L	$0.01	$0.01	$0.04	$0.11

Each instance family

Each instance size

Each Availability Zone

In every region

Is an independent ***Spot pool***

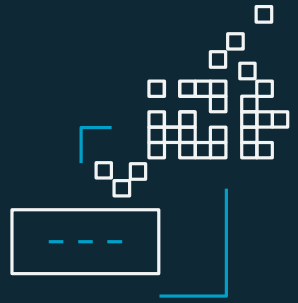
R5 R4 I3 C5 R5d M5 M4 i3en R5a

© 2018, Amazon Web Services, Inc. or its Affiliates. All rights reserved. Amazon Confidential and Trademar

Spot Instances are perfect for flexible fault-tolerant workloads

Containers

Increase application resilience and save costs



Big data

Accelerate time-to-insight and time-to-market

AI/ML

Access more compute to train models faster

Web services

CI/CD

Batch & HPC

Lean on Spot for these workloads!

aws

What happens when AWS reclaims an instance?

Minimal interruptions

Less than 5% of Spot instances were interrupted in the last 3 months

Handling Options

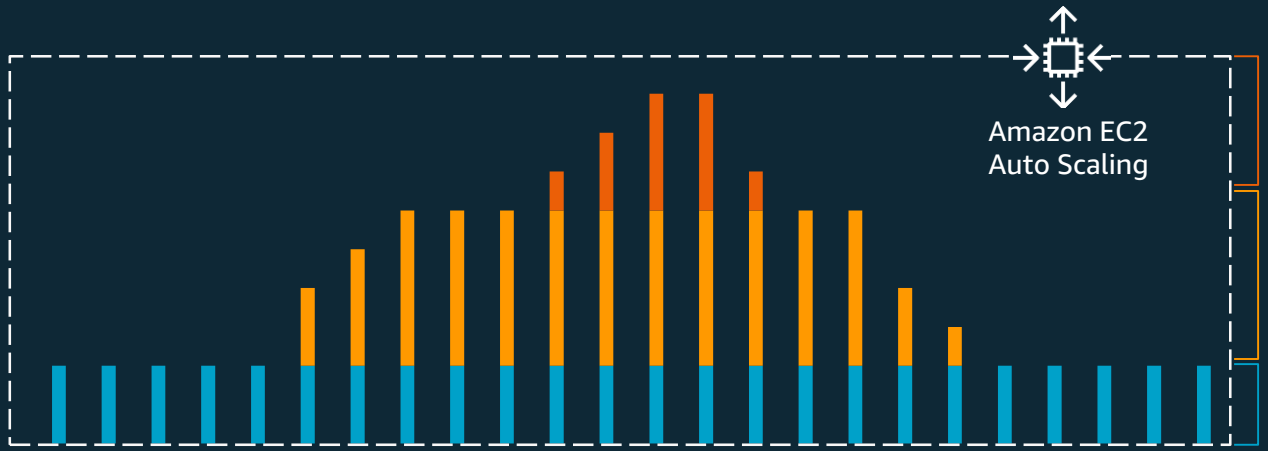
Alerts → Automation

Spot Capacity

AWS customers are using more compute capacity on Amazon EC2 Spot Instances than customers in 2014 were running across all of Amazon EC2

aws

Simplifying capacity and cost optimization



Scale using Spot for flexible, fault-tolerant workloads

Scale using On-Demand for new or stateful spiky workloads

Use Ris and Savings Plans for known/steady-state workloads

AWS SERVICES MAKE THIS EASY AND EFFICIENT

- Amazon Elastic Container Service (Amazon ECS)
- Amazon Elastic Kubernetes Service (Amazon EKS)
- AWS Thinkbox
- Amazon EMR
- AWS CloudFormation
- AWS Batch
- Amazon SageMaker

Delivery Hero saves 70% running Kubernetes on Amazon EC2 Spot Instances

Challenge

Delivery Hero is among the largest food delivery networks worldwide with delivery fleets in 39 countries. The company is always looking for ways to lower its infrastructure costs

Solution

The company transitioned its Kubernetes clusters to run only on Amazon EC2 Spot Instances to take advantage of unused Amazon EC2 capacity at a discount

Benefits

Saves 70% on infrastructure costs

Transitions Kubernetes workloads to discounted cloud instances in 6 months

Scales to manage complex deployments spread over 39 countries

> "Our experience **running Amazon EKS on Amazon EC2 Spot Instances was eye-opening.** It has become a big cost saver and freed our time and energy to focus on business growth instead."
>
> —**Vojtech Vondra,** Senior Director of Engineering, Logistics, Delivery Hero

Delivery Hero

Company:	Delivery Hero
Industry:	E-commerce and Retail
Country:	Germany
Employees:	19,000
Website:	deliveryhero.com

About Delivery Hero

Delivery Hero is a leading online food ordering and delivery marketplace. The company operates delivery fleets in 39 countries, transporting more than 1 million food orders a day

aws

Mobile app startup Moovit saves by incorporating Spot Instances into EMR clusters

Challenge

With over 750M users, Moovit generates between 2-4 billion data points per day. Running EMR clusters with On-Demand instances was beginning to become too expensive.

Solution

The company transitioned 95% of its production EMR clusters to run on Spot Instances

Benefits

Moovit saved significantly on their EC2 costs by moving to Spot Instances

Even after doubling their compute capacity, running EMR clusters on Spot was still cheaper than On-Demand

> "The best practice as of today is to use EMR clusters [with On-Demand] which are 3-4X more expensive compared to Spot Instance. The straight forward thing [about Spot] is that it's much less expensive."
>
> —**Anan Kenig, Chief Architect**, Moovit

Company: Moovit

Industry: Software

Country: Israel

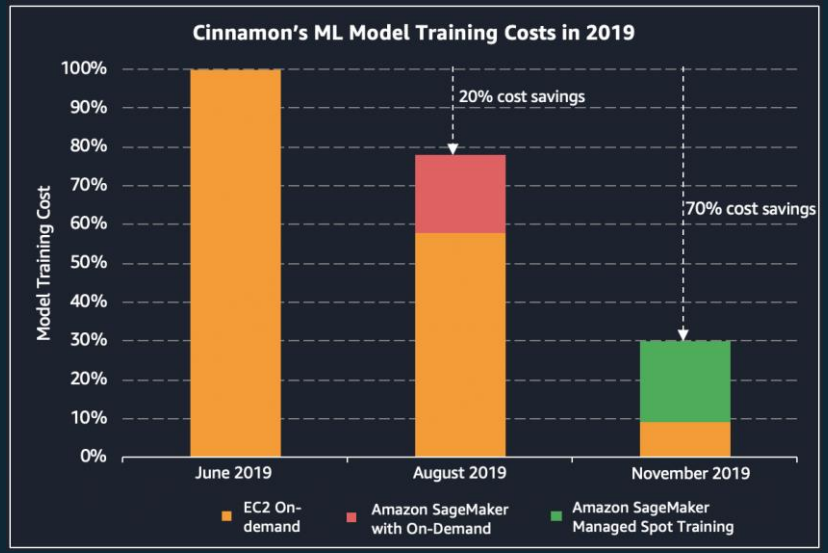
Employees: 200+

Website: https://company.moovit.com/

About Moovit

Moovit guides 750M+ transit riders across 100 countries to the best, most efficient routes, and makes it easy for locals and visitors alike to navigate the world's cities.

CINNAMON AI SAVES 70% IN MODEL TRAINING COSTS



Cinnamon AI is a Japan-based startup with ML-powered products like Flax Scanner (a NLP powered document reader), Rossa Voice (a real-time voice recognition service), and Aurora (an information extraction service).

Cinnamon moved all their ML workloads onto AWS and used EC2 Spot via the Managed Spot Service to save on model training costs.

"The Managed Spot Training feature of Amazon SageMaker has had a profound impact on our AWS cost savings. Our AWS EC2 costs reduced by up to 70% after using Managed Spot Training," said Tetsuya Saito, GM of Infrastructure at Cinnamon AI. "In addition, Managed Spot Training does not require complicated methods and can be used simply from the SageMaker SDK."

© 2019 Amazon Web Services, Inc. or its affiliates. All rights reserved

aws machine learning

WHY AWS FOR ML?

Broadest and deepest set of AI and ML services

200 new features and services launched this last year alone

Unmatched flexibility

Accelerate your adoption of ML with SageMaker

70% cost reduction in data-labeling

10x faster performance

75% lower inference cost

Built on the most comprehensive cloud platform

AWS Named as a Leader in Gartner's Infrastructure as a Service (IaaS) Magic Quadrant for the 9th Consecutive Year

© 2019 Amazon Web Services, Inc. or its affiliates. All rights reserved

THE AWS ML STACK

Broadest and deepest set of capabilities

AI Services

VISION			SPEECH		LANGUAGE		CHATBOTS	FORECASTING	RECOMMENDATIONS
REKOGNITION IMAGE	REKOGNITION VIDEO	TEXTRACT	POLLY	TRANSCRIBE	TRANSLATE	COMPREHEND & COMPREHEND MEDICAL	LEX	FORECAST	PERSONALIZE

ML Services

Amazon SageMaker

Ground Truth	Notebooks	Algorithms + Marketplace	Reinforcement Learning	Training	Optimization	Deployment	Hosting

ML Frameworks + Infrastructure

FRAMEWORKS	INTERFACES
TensorFlow, mxnet, PYTORCH	GLUON, Keras

INFRASTRUCTURE

EC2 P3 & P3DN	EC2 G4 EC2 C5	FPGAS	DL CONTAINERS & AMIs	ELASTIC CONTAINER SERVICE	ELASTIC KUBERNETES SERVICE	GREENGRASS	ELASTIC INFERENCE	INFERENTIA

© 2019 Amazon Web Services, Inc. or its affiliates. All rights reserved

aws machine learning

AMAZON SAGEMAKER

REDUCE COSTS

70%
cost reduction for data labeling using Ground Truth

75%
cost reduction for inference with Elastic Inference

90%
cost reduction with managed spot training

INCREASE PERFORMANCE

10x
better algorithm performance

2x
performance increases from model optimization with Neo

90%
AWS-optimized TensorFlow

EASE-OF-USE

One-click
model training and deployment

Train once
run anywhere

SECURITY & COMPLIANCE

SOC, PCI, ISO, HIPAA, C5, OSPAR, HITRUST CSF

© 2019 Amazon Web Services, Inc. or its affiliates. All rights reserved

aws machine learning

AMAZON SAGEMAKER

Bringing machine learning to all developers

Pre-built notebooks for common problems	Built-in, high performance algorithms	One-click training	Optimization	One-click deployment	Fully managed with auto-scaling, health checks, automatic handling of node failures, and security checks
Collect and prepare training data	Choose and optimize your ML algorithm	Set up and manage environments for training	Train and tune model (trial and error)	Deploy model in production	Scale and manage the production environment

© 2019 Amazon Web Services, Inc. or its affiliates. All rights reserved

aws machine learning

AMAZON SAGEMAKER

Supported algorithms

Built-in, high performance algorithms

Choose and optimize your ML algorithm

- K-Means Clustering
- Principal Component Analysis
- Neural Topic Modelling
- Factorization Machines
- Linear Learner (Regression)
- BlazingText
- Reinforcement learning
- XGBoost
- Topic Modeling (LDA)
- Image Classification
- Seq2Seq
- Linear Learner (Classification)
- DeepAR Forecasting

- Optimized frameworks – Apache MxNet, TensorFlow, Chainer, & PyTorch containers.
- Custom training algorithm
- ML market place with 200+ model listings

© 2019 Amazon Web Services, Inc. or its affiliates. All rights reserved

aws machine learning

WHAT IS MANAGED SPOT TRAINING?

Spot-only instances for ML model training

90% cost reduction in model training costs *(compared to OD)*

Ideal for training workloads with flexibility *(non-time sensitive)*

Single-node and Distributed model training supported

Works with SageMaker Reinforcement Learning

Automated Spot management

SageMaker manages provisioning Spot instances and interruptions

Training jobs pause and resume reliably as Spot capacity becomes available

Built-in frameworks and algorithms checkpoint automatically*

Integrates with other AWS services

S3 – stores training data, checkpoints, and results of the training jobs

CloudWatch – training metrics and logs

CloudTrail – log of all API calls

© 2019 Amazon Web Services, Inc. or its affiliates. All rights reserved

aws machine learning

** Computer vision algorithms supported, support for others coming up*

KEY CONSIDERATIONS

Training with only Spot

Interrupted jobs resume if checkpointed and if Spot instances become available; Jobs restart if not checkpointed*

Works with Automated Model Tuning

Training jobs can run only with a single instance-type in a single-AZ

Does not integrate with Spot Fleet and Spot Block today

Checkpoint

Built-in algorithms checkpoint automatically

For custom models, checkpointing should be enabled

Checkpoints are saved to S3

Models that don't checkpoint are subject to *MaxWaitTime* of 60 mins

Pricing

View savings on AWS console or use DescribeTrainingJob API

Charged for the run duration before completion or termination; not charged for idle time, billing starts when instances are ready

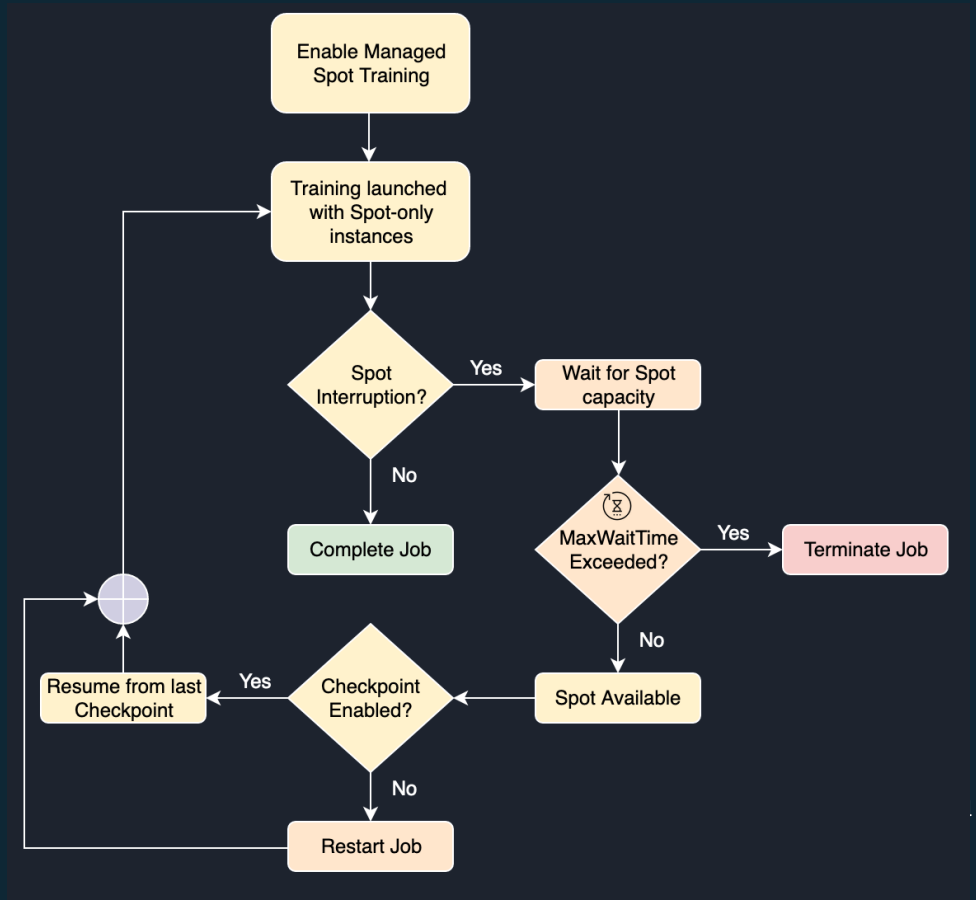
Charged for data download time only once even if the job is interrupted multiple times

© 2019 Amazon Web Services, Inc. or its affiliates. All rights reserved

aws machine learning

** Checkpointing is a best practice and is highly recommended*

MANAGED SPOT TRAINING WORKFLOW

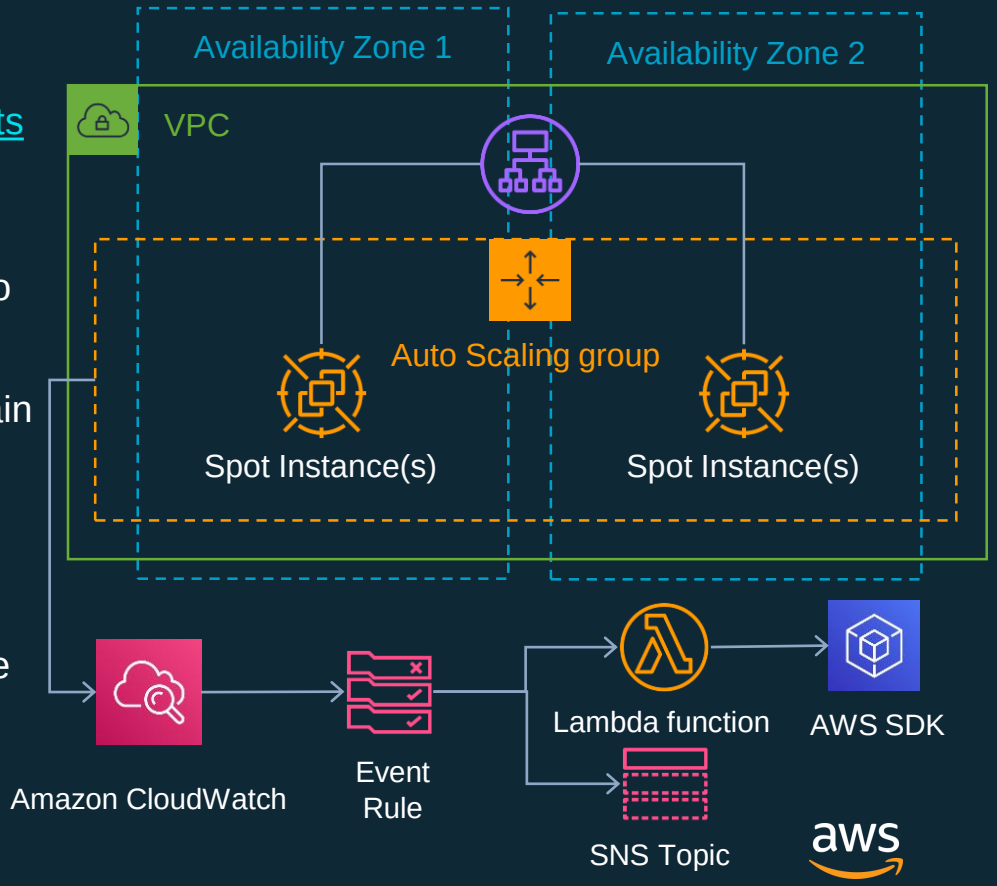


. rights reserved

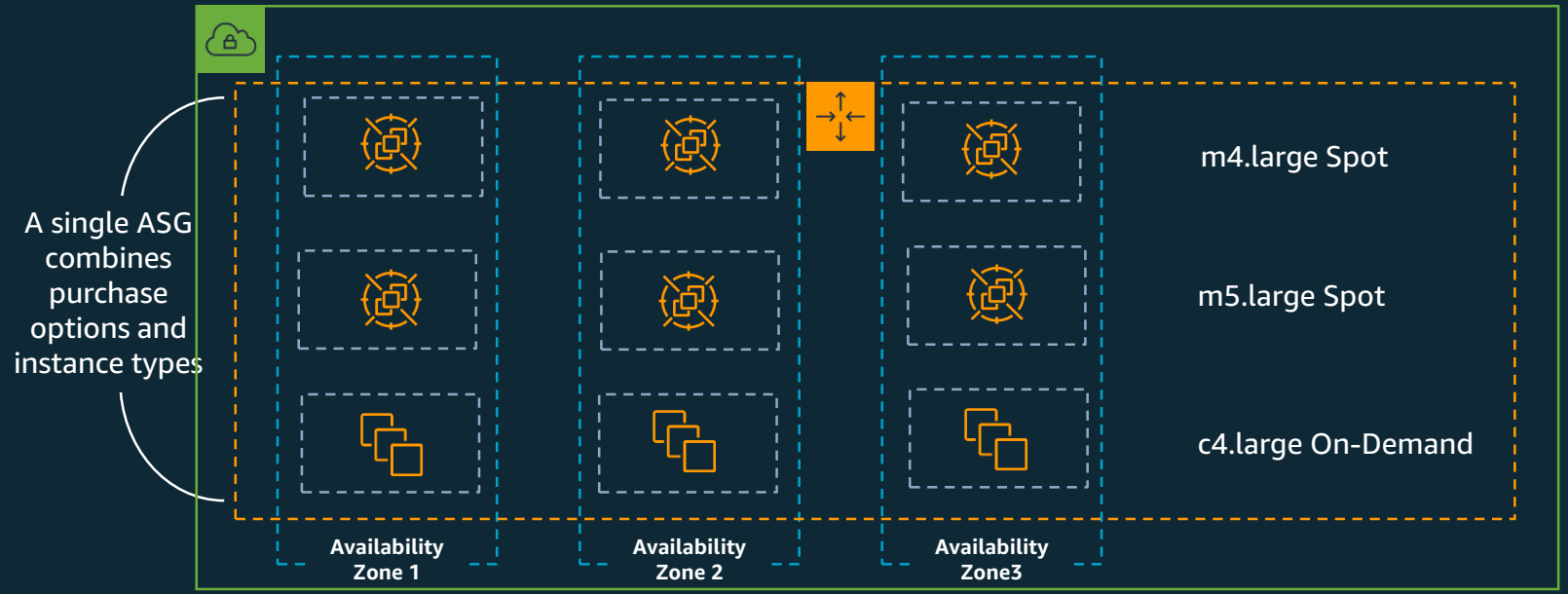
Demo

Stateless web application or API frontend

- Maintain a diversified fleet of instances
- ALB can now use Least Outstanding Requests load balancing algorithm to ensure fair load across different instance types
- Adjust the TargetGroup deregistration delay to 120 sec (connection draining)
- Lambda calls AutoScaling Detach API will drain ELB connections and launch replacement capacity. OS/Application level actions can be invoked via Systems Manager RunCommand API
- Ensure a short bootstrap time and allow some overprovisioning



Combining Instance types and purchase options on ASG



Auto Scaling groups allocation strategies

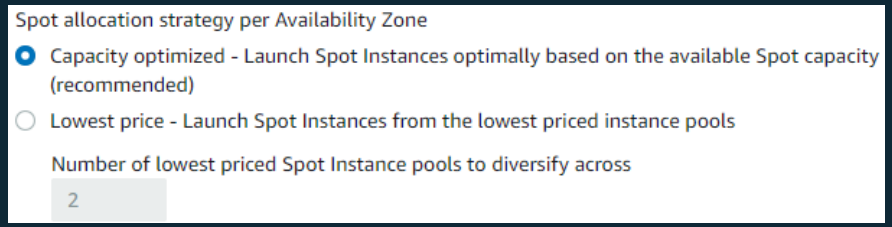
Allocation strategies determine which instance types from your selection are launched and terminated.

On-demand instances:

- On-Demand: prioritized list of instance types

Spot Instances:

- N lowest priced instance pools
- Capacity Optimized

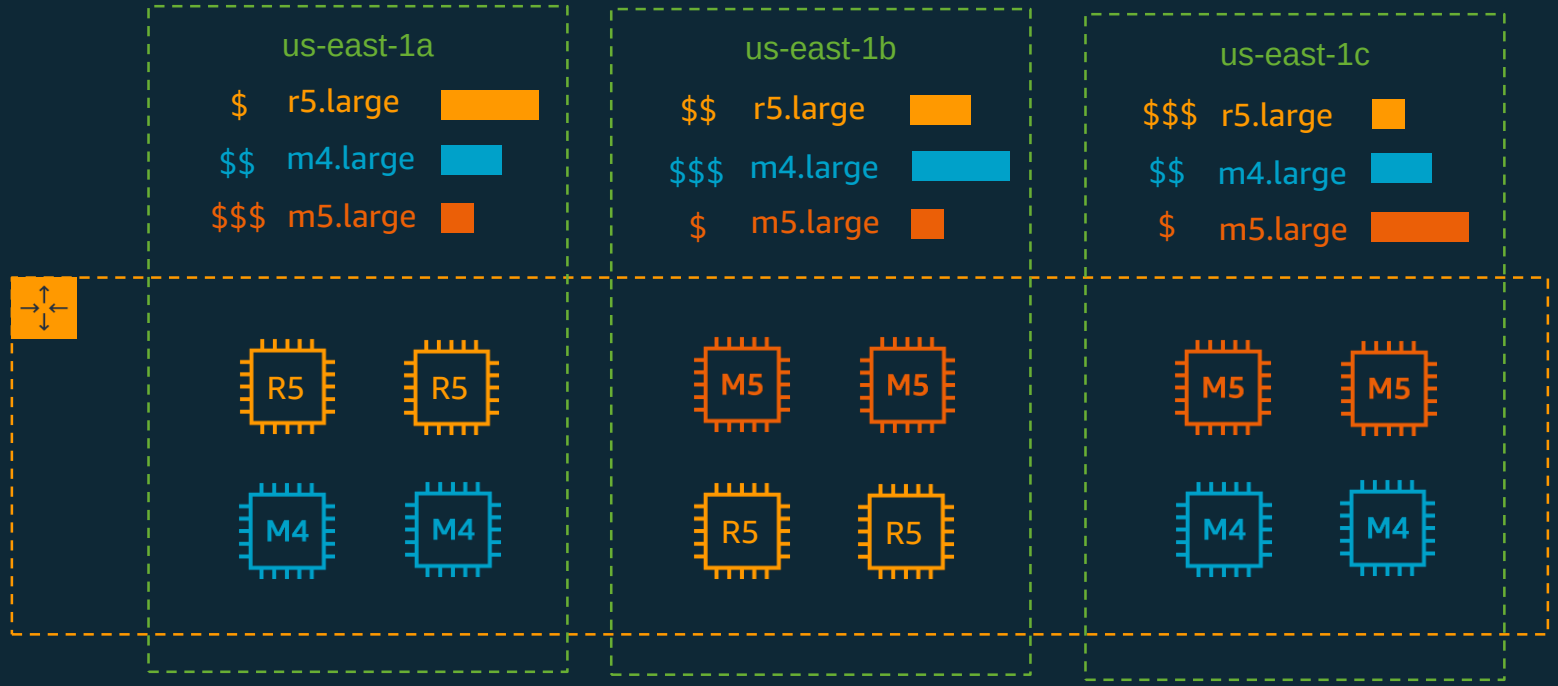


ASG lowest-price across N pools allocation strategy

Desired capacity: 12 OnDemandBaseCapacity: 0 OnDemandPercentageAboveCapacity: 0

Overrides: [“r5.large”, “m4.large”, ”m5.large”] SpotAllocationStrategy: lowest-price

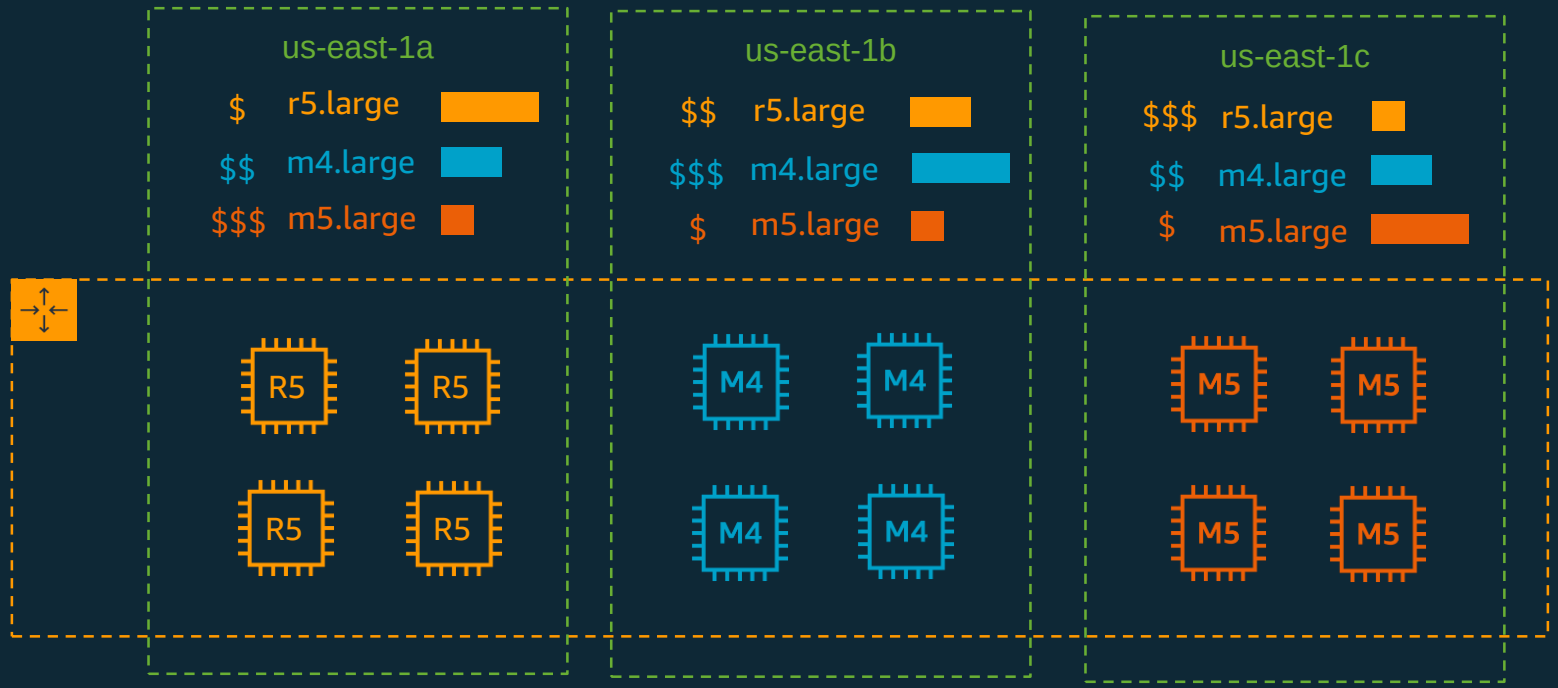
SpotInstancePools: 2



ASG capacity-optimized allocation strategy

Desired capacity: 12 OnDemandBaseCapacity: 0 OnDemandPercentageAboveCapacity: 0

Overrides: [“r5.large”, “m4.large”, ”m5.large”] SpotAllocationStrategy: capacity-optimized



EC2 Auto Scaling groups demo

aws

Containers + Spot = match made in heaven

- ✓ Containers are often stateless, fault-tolerant, and a great fit for Spot Instances
- ✓ Deploy containerized workloads and easily manage clusters at any scale at a fraction of the cost with Spot Instances
- ✓ Spot instances can be used with ECS or Kubernetes to run any containerized workload

skyscanner

Skyscanner is a travel fare aggregator website and travel metasearch engine based in Edinburgh, Scotland

"We are currently tracking **74% saving** over all regions."

—Paul Gillespie,
Principal Architect/Tribe Lead

lyft yelp MapBox Expedia Caltech

aws

EC2 Spot with Amazon EKS

kubernetes

aws

How do I provision EKS Worker Nodes?

What about
Spot Fleet /
EC2 Fleet?

eksctl

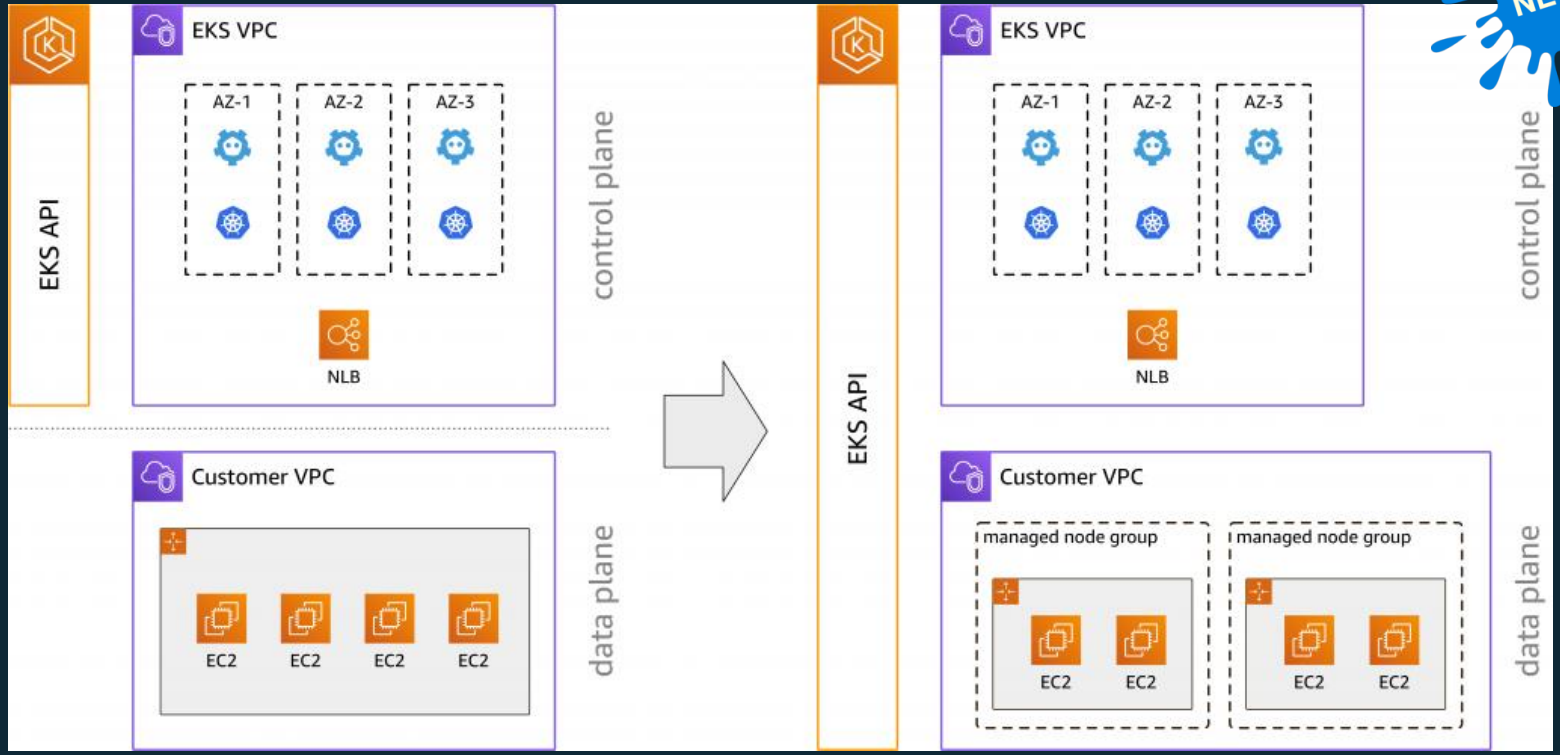
AWS CloudFormation

Amazon EC2
Auto Scaling groups

HashiCorp
Terraform

aws

EKS managed worker nodes



NEW!

Using EC2 Spot instances with Amazon EKS

1. Handling interruptions with AWS Node Termination Handler

2. Acquiring Diversified EC2 Spot capacity

3. Autoscaling Kubernetes apps and clusters

4. Tools & Techniques

AWS node termination handler

NEW!

Current functionality

Drain/cordon worker node upon EC2 Spot Interruption Notification,
Maintenance notifications

Coming soon

Drain / Terminate EKS worker node on
ASG scale-in event

```
log.Println("Kubernetes Spot Node Termination Handler has started successfully!")
waitForTermination(nthConfig)
drainer.Drain(nthConfig)
log.Printf("Node %q successfully drained.\n", nthConfig.NodeName)
```

```
helm repo add eks https://aws.github.io/eks-charts
helm install --name aws-node-termination-handler \
  --namespace kube-system eks/aws-node-termination-handler \
  --set nodeSelector.lifecycle=Ec2Spot
```

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

https://github.com/aws/eks-charts/tree/master/stable/aws-node-termination-handler

aws

Acquiring Diversified EC2 Spot Capacity

```yaml
apiVersion: eksctl.io/v1alpha5
kind: ClusterConfig
metadata:
    name: test-cluster
    region: us-west-2
nodeGroups:
    - name: dev-4vcpu-16gb-spot
      availabilityZones: ["us-west-2a","us-west-2b","us-west-2c"]
      ami: ami-03a55127c613349a7
      amiFamily: AmazonLinux2
      minSize: 1
      maxSize: 100
      instanceType: mixed
      instancesDistribution:
        instanceTypes: ["m5.xlarge", "m5d.xlarge", "m4.xlarge","m5a.xlarge",...]
        onDemandBaseCapacity: 0
        onDemandPercentageAboveBaseCapacity: 0
        spotInstancePools: 4
      labels:
        lifecycle: Ec2Spot
      taints:
        spotInstance: "true:PreferNoSchedule"
      tags:
        k8s.io/cluster-autoscaler/node-template/label/lifecycle: Ec2Spot
        k8s.io/cluster-autoscaler/node-template/taint/spotInstance: "true:PreferNoSchedule"
```

capacity-optimized

NEW!

nodegroup-4vCPUs-16GB

ge,...

nodegroup-8vCPUs-32GB

m5.2xlarge, m5d.2xlarge, m4.2xlarge, m5a.2xlarge,..

nodegroup-4vCPUs-32GB

rge,..

aws

Autoscaling the applications and the cluster

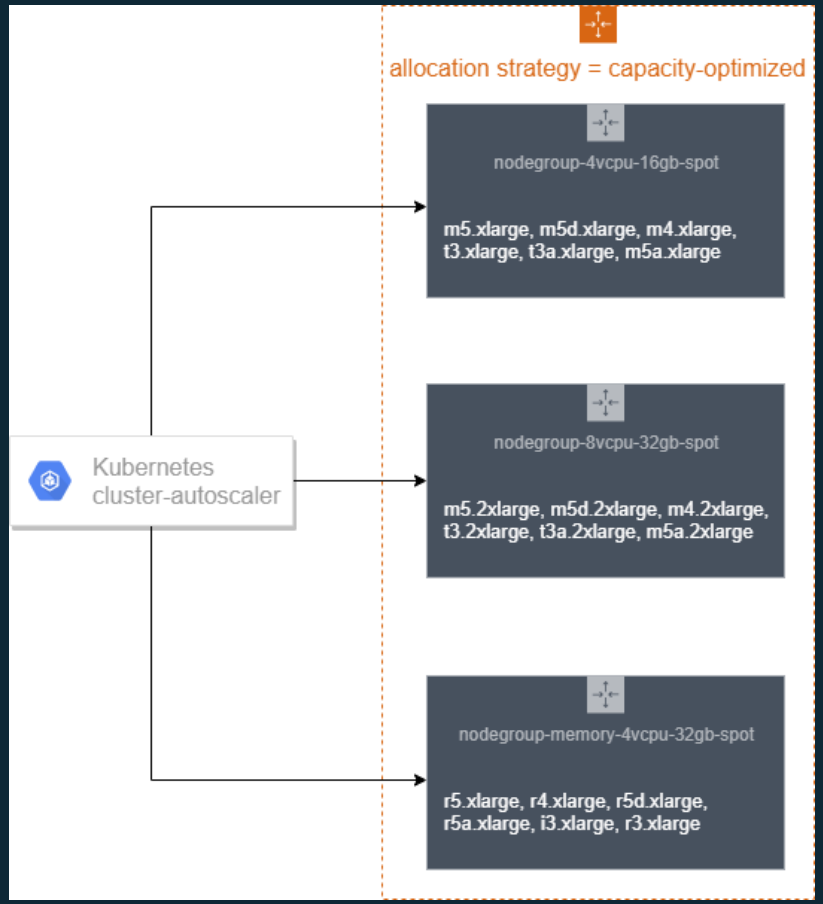
- Horizontal Pod AutoScaler (HPA)
  - AutoScales the number of pods in a Deployment/ReplicaSet

```
kubectl autoscale deployment hello-k8s --cpu-percent=$HPA_MIN_CPU --min=$HPA_MIN_PODS --max=$HPA_MAX_PODS
```

- Cluster Autoscaler (CA)
  - Autoscales the number of worker nodes in the cluster when:
    - Pods cannot be scheduled due to lack of resources (Pending state)
    - Nodes are underutilized and important pods can be rescheduled elsehere

aws

Cluster-autoscaler diversification example



Tools & Techniques

Affinity

```yaml
affinity:
  nodeAffinity:
    preferredDuringSchedulingIgnoredDuringExecution:
    - weight: 1
      preference:
        matchExpressions:
        - key: lifecycle
          operator: In
          values:
          - OnDemand
```

Taints & Tolerations

```yaml
tolerations:
- key: "spotInstance"
  operator: "Equal"
  value: "true"
  effect: "PreferNoSchedule"
```

Tools

- eksctl
- Cloudwatch Container Insights
- Descheduler
- K8s-node-drainer
- Overprovisioner

aws

Demo: K8s autoscaling with Spot

aws

EC2 Spot with Amazon EMR

aws

Amazon EMR + Spot best practices

- Use Spot Instances in EMR, unless heavily time-constrained
- Diversify and be instance type agnostic
- Decouple storage from compute (from HDFS to EMRFS)

Scenario	Master Node	Core Nodes	Task Nodes
Long-Running Clusters and Data Warehouses	On-Demand	On-Demand or instance-fleet mix	**Spot** or instance-fleet mix
Cost-Driven Workloads	**Spot**	**Spot**	**Spot**
Data-Critical Workloads	On-Demand	On-Demand	**Spot** or instance-fleet mix
Application Testing	**Spot**	**Spot**	**Spot**
	Manages the cluster, tracks status of tasks, monitors health	Run tasks and stores data in HDFS	Optional, only run tasks

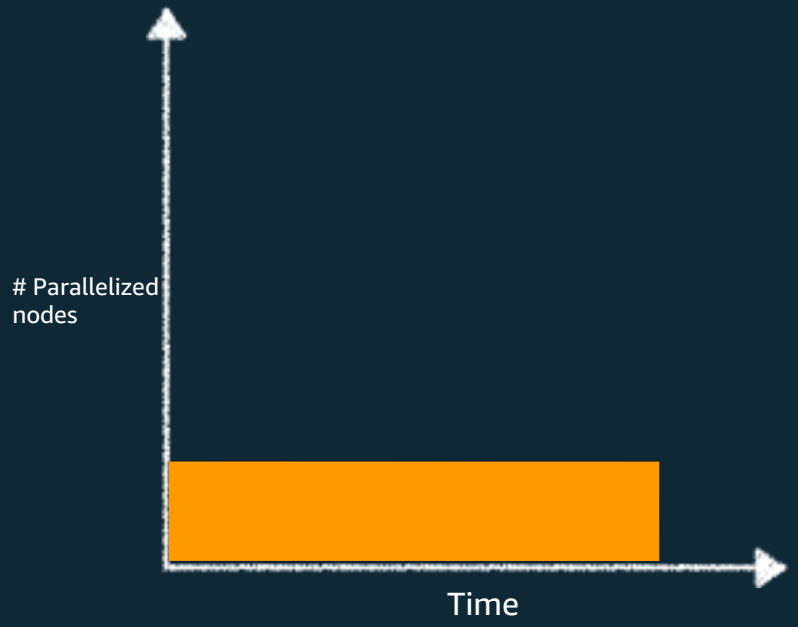
aws

Increased resilience with EMR Instance Fleets

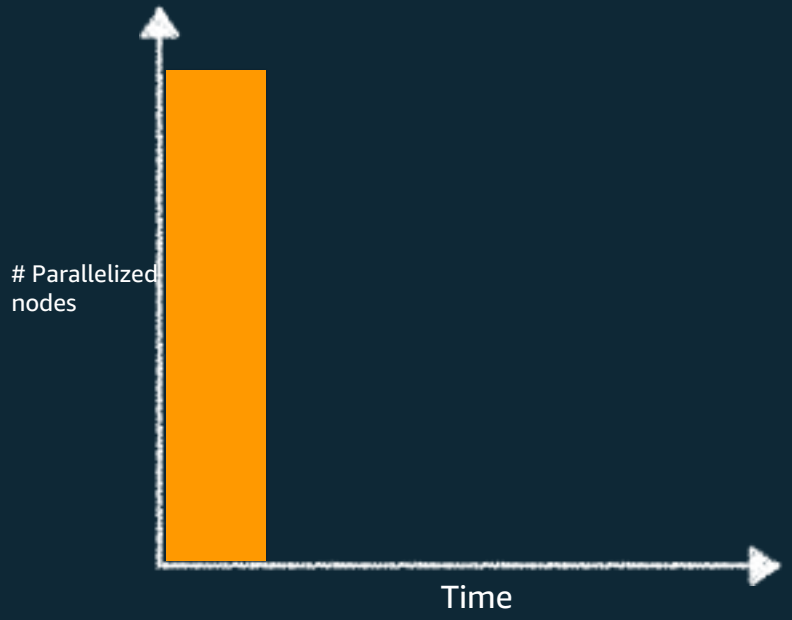
- Specify target capacity as a mix of instance types (up to 5)
- Amazon EMR will attempt to fulfill capacity from the most suitable pools
- Amazon EMR automatically replaces interrupted or failed instances with one of the instance types that you specified
- Spark is inherently resilient and applications continue running
- Coming soon: EMR Managed Resize

aws

Reduce time-to-insight with parallelization



Job running time: **10 hours**



Job running time: **1 hour**

Demo: EMR Instance Fleets

aws

Main takeaways for success with Spot

- SageMaker Managed Spot Training has super quick time to value
- EC2 Auto Scaling groups is the best fit for most of the workloads
- Be instance type agnostic and let ASG, Fleet or EMR provide the required capacity at the lowest price or from the deepest pools
- Architect for fault-tolerance to be Spot compatible and increase your availability

aws

More info:

https://aws.amazon.com/ec2/spot/startups

© 2020, Amazon Web Services, Inc. or its Affiliates. All rights reserved. Amazon Confidential