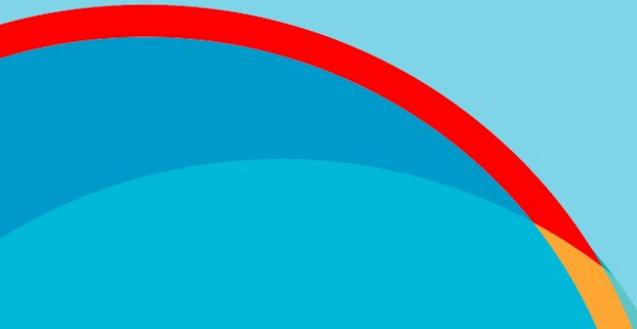




Red Hat  
**Summit**

**Connect**



---

Managed Services  
Cost Strategy  
Optimisation

# Who?

Managed OpenShift Black Belts (MOBB) Mission

*To remove customers' organizational, competitive, and technical blockers to enterprise-wide adoption of Managed OpenShift (ROSA, ARO, OSD)*



**Andrea Bozzoni**

Senior Cloud Services Black Belt



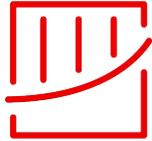
**Roberto Carratalá**

Senior Cloud Services Black Belt

# Agenda

- Overview of Cost Strategy Optimization
- Committed Spend Overview
- PAYGO and Reserved Instance Overview
- Spot Instances
- Machine right sizing & Autoscaling
- ROSA Hosted Control Plane (HCP)

# Cost optimisation levers



EDP/MACC  
Committed Spend



Machine right sizing



Reserved Instances and  
Savings Plans



Autoscaling



Use Included Open Source  
Tools



Spot Instances



Graviton Instances

\*coming soon



Hosted Control Planes

\*coming soon

# Committed Spend Overview

# Committed Spend Overview



## Enterprise Discount Program

(EDP) Enterprise Agreements give customers the option to tailor agreements that best suit their needs.



## Microsoft Azure Consumption Commitment

(MACC) is a contractual commitment that your organization may have made to Microsoft Azure spend over time.

# What is the Red Hat Hybrid Committed Spend (HCS) Buying Program?



**HCS allows customers to commit to a 3 year spend commitment to unlock additional benefits not available in our traditional subscription model**

The additional benefits include deployment and product **flexibility**:

- Customers get **flexibility** to purchase Red Hat products on-prem or in the cloud and receive volume discount considerations for total spend
- Product deployment **flexibility** across Red Hat's portfolio, allowing customers to shift from one product to another on a monthly (and in some cases an hourly) basis
- Maximum procurement and financing **flexibility** with option to merge traditional subscription (paid annually) with on-demand subscriptions (billed monthly in arrears)



Unlock greater flexibility with Hybrid Committed Spend

# PAYGO and Reserved Instance Overview

# PAYGO and Reserved Instance overview



## PAYGO



## Reserved Instance

<b>Price</b>	USD \$0.171/hr	One year = \$1,000 Three years = \$2,000 (\$667/year)
<b>Term</b>	Hourly; Paid at end of each month	Annual; Paid upfront, partial upfront + monthly, monthly
<b>Sizing</b>	Per 'worker node' = 4 vCPUs Eg. c5.4xlarge nodes (16 vCPU, 32 GB RAM) = Four 'worker nodes'	
<b>Discountable</b>	No	Yes, with Private Offer
<b>Use case</b>	<ul style="list-style-type: none"><li>▶ Pilot workloads</li><li>▶ Ephemeral environments</li><li>▶ Spikey/unknown workloads</li></ul>	<ul style="list-style-type: none"><li>▶ Production workloads with known usage profile</li><li>▶ Always-on workloads</li></ul>

# PAYGO and Reserved Instance example

## Production cluster example: 9 worker nodes - annual pricing

Estimated annual price for one multi-AZ ROSA cluster running nine (9) m5.xlarge worker nodes with a 1-year ROSA subscription and 1-year EC2 Instance Savings Plans:

ROSA service fee subtotal	\$9,263
AWS infrastructure fee subtotal	\$30,575
<b>Total estimated annual price***</b>	<b>\$39,838</b>

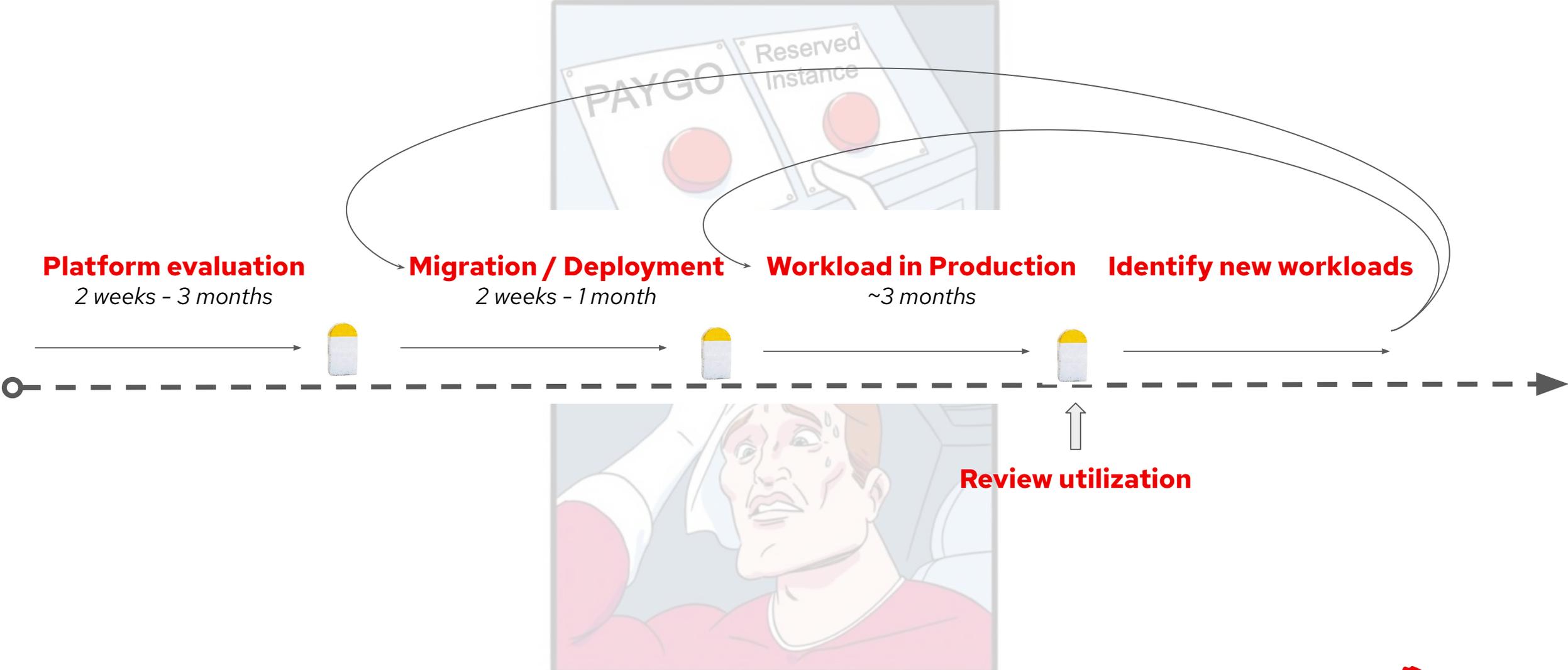
## Test cluster example: 3 worker nodes - on demand pricing

Estimated on-demand price for one multi-AZ ROSA cluster running three (3) m5.xlarge worker nodes for one month:

ROSA service fee subtotal	\$396
AWS infrastructure cost subtotal	\$2,501
<b>Total estimated monthly price***</b>	<b>\$2,897</b>

\*\*\* Does not include the cost of AWS services other than Amazon EC2 and Amazon EBS.

# When to go PAYGO or RI?



# Spot Instances

# Spot Instances Overview



## Amazon EC2 Spot Instances

Let you take advantage of unused EC2 capacity in the AWS cloud and are available at up to a 90% discount compared to On-Demand prices.



## Spot Virtual Machines

Let you buy unused compute capacity at deep discounts of up to 90 percent compared to pay-as-you-go prices.

# ROSA Spot Instances Example

- About machine pools
- Managing compute nodes**
- Configuring machine pools in Local Zones
- About autoscaling nodes on a cluster

Optional: Use Amazon EC2 Spot Instances if you want to configure your machine pool to deploy machines as non-guaranteed AWS Spot Instances:

- Select **Use Amazon EC2 Spot Instances**.
- Leave **Use On-Demand instance price** selected to use the on-demand instance price. Alternatively, select **Set maximum price** to define a maximum hourly price for a Spot Instance.

### Add machine pool ✕

A machine pool is a group of machines that are all clones of the same configuration, that can be used on demand by an application running on a pod.

Machine pool name \*

Worker node instance type \* ⓘ

4 vCPU 16 GiB RAM

Autoscaling ⓘ

Enable autoscaling

Worker node count (per zone) ⓘ

0

× 3 zones = 0 worker nodes

[> Edit node labels and taints](#)

Cost saving

Use Amazon EC2 Spot Instances

You can save on costs by creating a machine pool running on AWS that deploys machines as non-guaranteed Spot Instances. This cannot be changed after machine pool is created.

Use On-Demand instance price

The maximum price defaults to charge up to the On-Demand Instance price.

Set maximum price ⓘ

Specify the maximum hourly price for a Spot Instance.

# ARO Spot Instances Example

## Use Azure Spot Virtual Machines in an Azure Red Hat OpenShift (ARO) cluster

Article • 04/29/2023 • 2 contributors

[Feedback](#)

### In this article

[Before you begin](#)

[Add Spot VMs](#)

[Schedule interruptible workloads](#)

[Quota](#)

[Show 2 more](#)

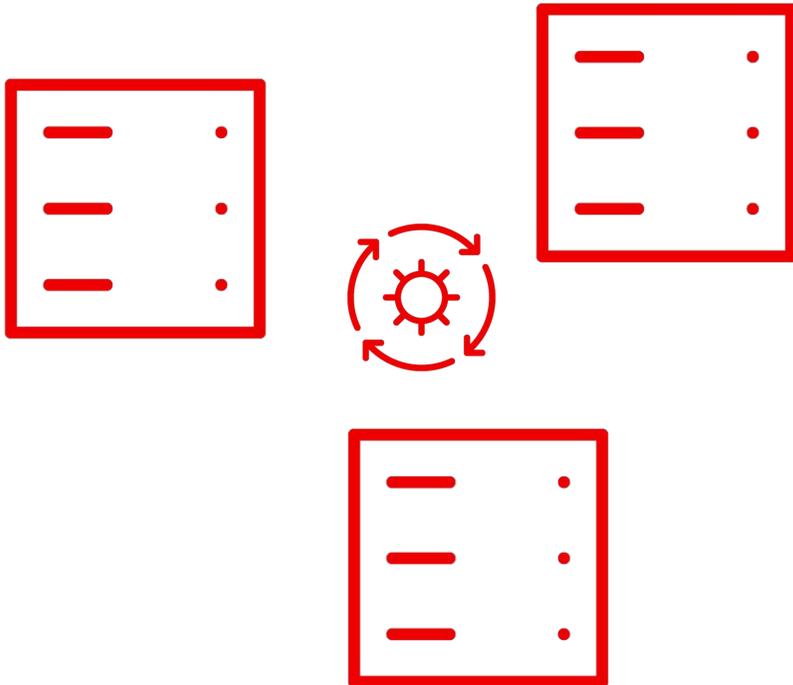
This article provides the necessary details that allow you to configure your Azure Red Hat OpenShift cluster (ARO) to use Azure Spot Virtual Machines.

Using Azure Spot Virtual Machines allows you to take advantage of our unused capacity at a significant cost savings. At any point in time when Azure needs the capacity back, the Azure infrastructure will evict Azure Spot Virtual Machines. For more information around Spot Instances, see [Spot Virtual Machines](#).

# Machine right sizing & Autoscaling

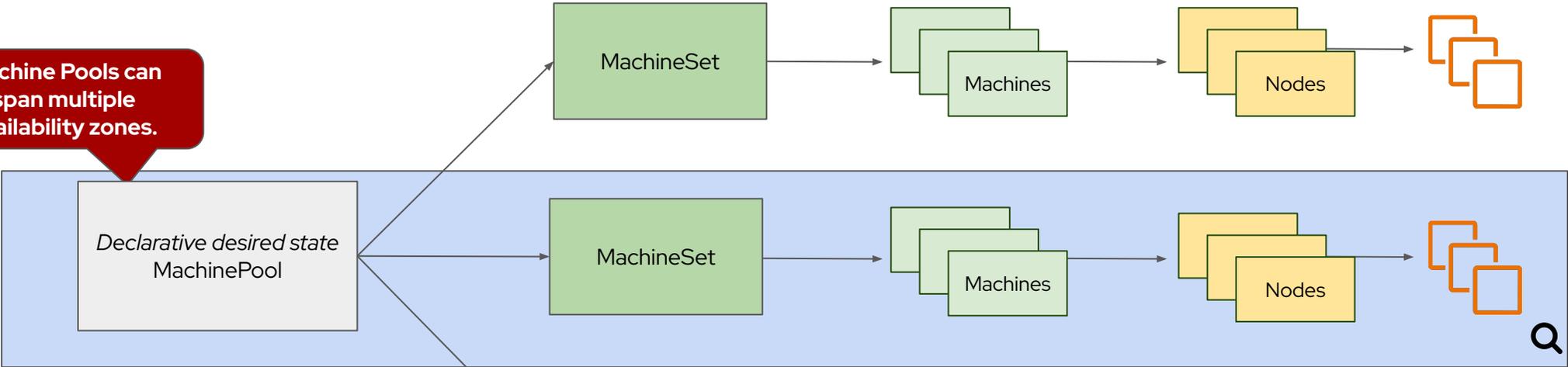
# Managing Worker Nodes

Providing highly available compute.



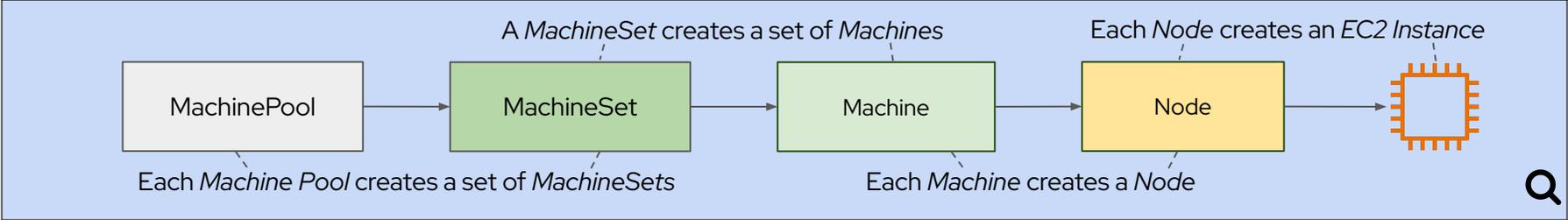
- ▶ MachinePools allows for worker nodes that span multiple availability zones (AZs).
- ▶ MachinePools provide a declarative desired state for worker nodes to ensure consistency across AZs.
- ▶ MachinePools can be scaled up or down manually or automatically.

Machine Pools can span multiple availability zones.



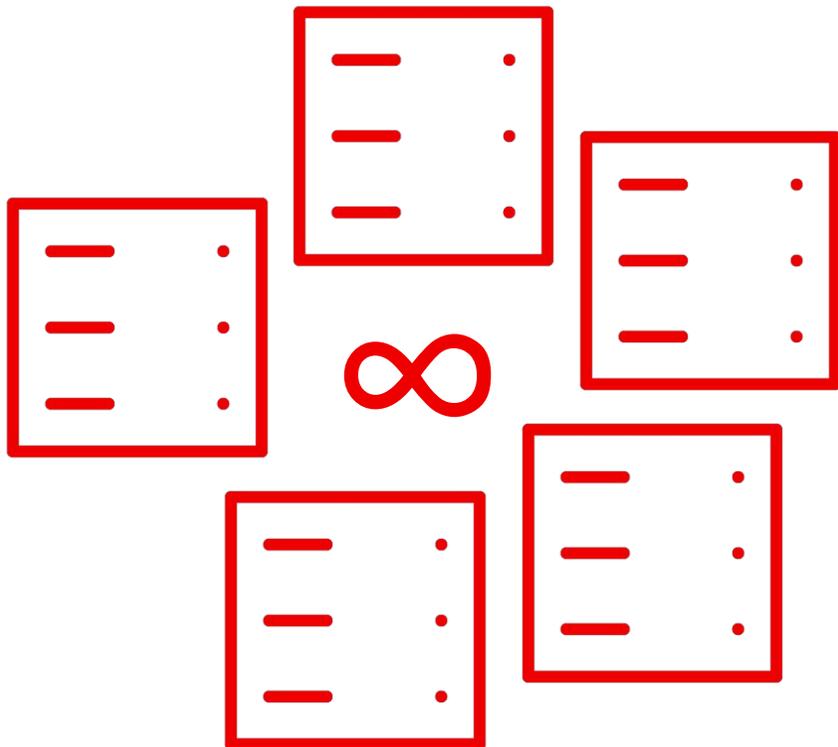
Machine Sets are specific to a single availability zone, which is why there are 3 in this diagram.

Machine Pools are managed by the *OpenShift Cluster Manager (OCM)*. The rest of the process is managed by the *Machine API Operator*. This operator interacts with the *AWS API* directly to provision *EC2 instances*.

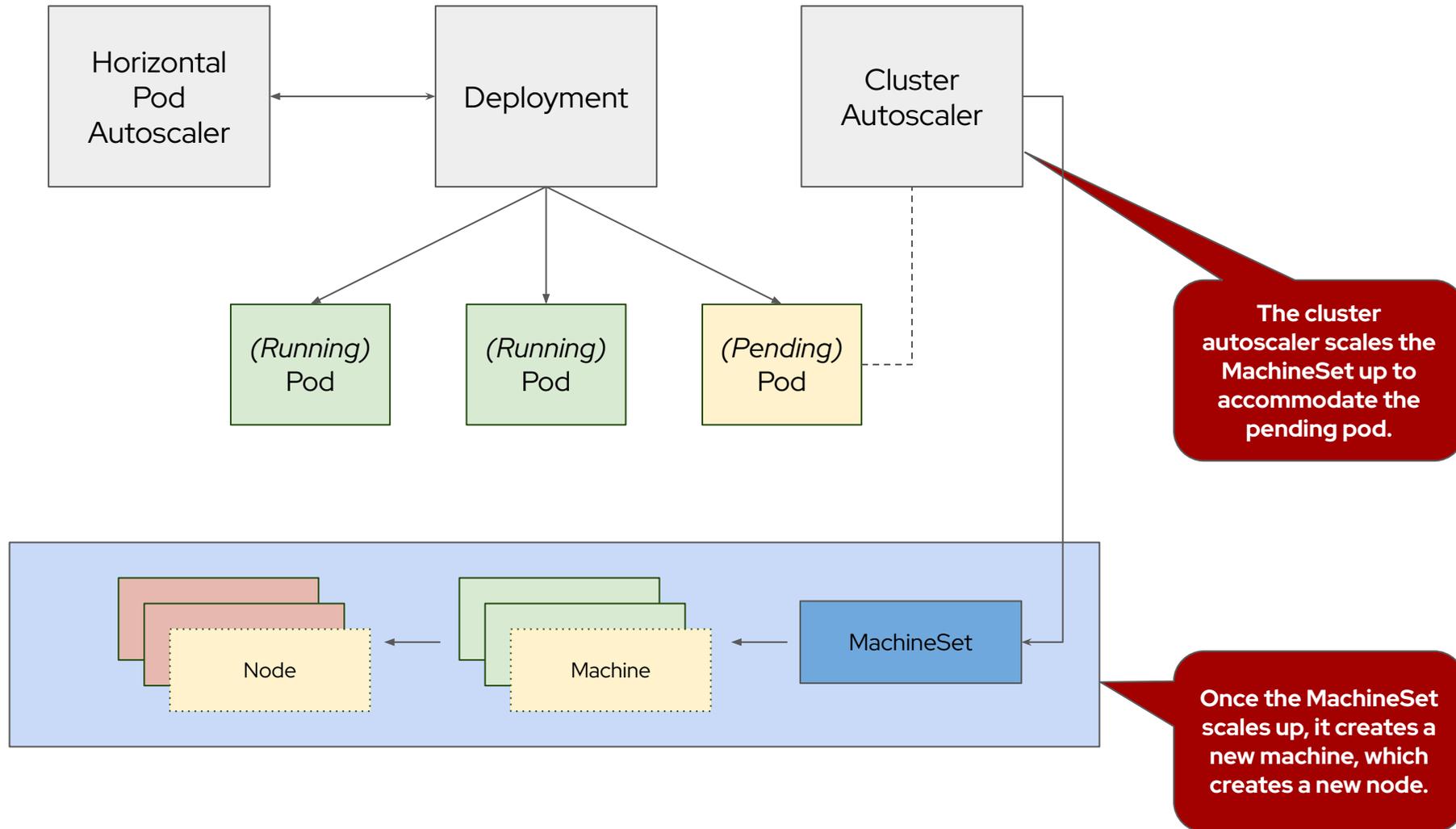


# Cluster Autoscaling

Automatically responding to cluster demand.



- ▶ MachinePools can be scaled to meet applications demands.
- ▶ Cluster AutoScaler will provision additional worker nodes when pods can not be scheduled due to resource constraints.
- ▶ Cluster AutoScaler will not scale beyond predefined limits.



# ROSA Hosted Control Plane (HCP)

# Why Hosted Control Plane (Cost Optimization)?



**Cost savings**

- ▶ Significantly reduced AWS / Azure infrastructure costs (typically \$8k/cluster/year)
- ▶ Quickly and easily spin up or tear down clusters when needed for efficiency and cost savings
- ▶ Smaller overall footprint (2 nodes vs 7)
- ▶ Scale worker nodes to 0 \* (post GA)

# ROSA with Hosted Control Planes Cost Example

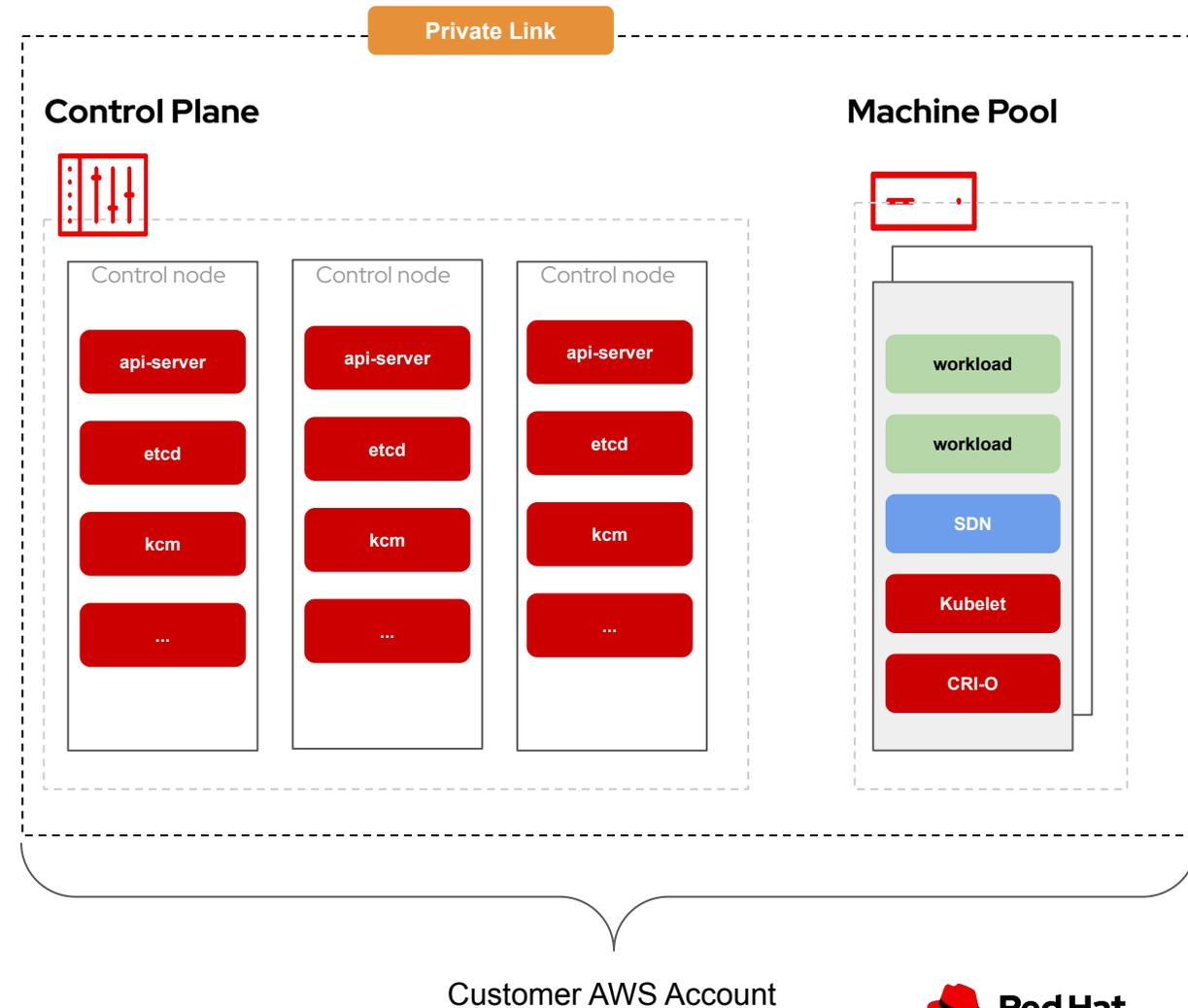
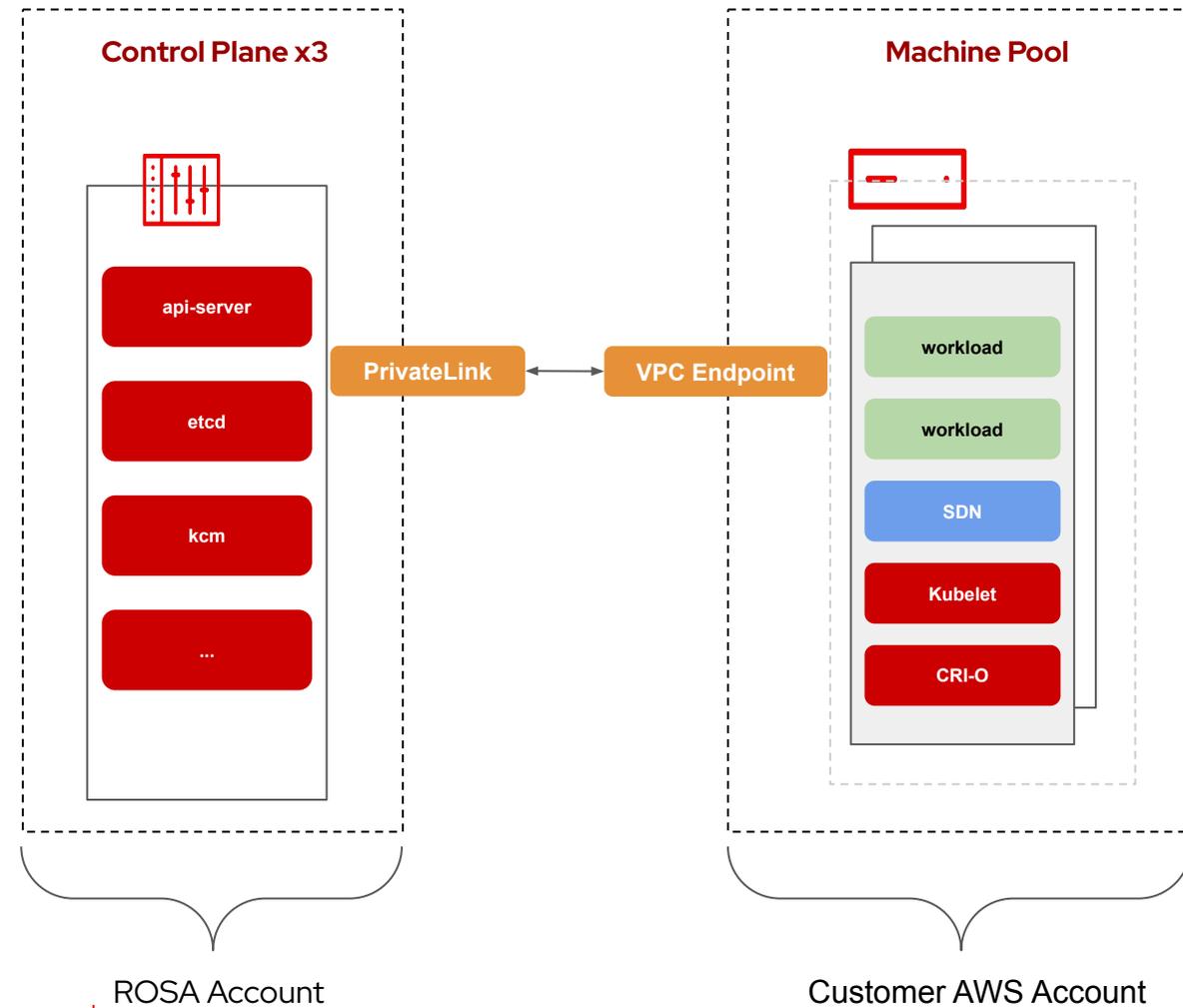
	<b>ROSA Classic</b>	<b>ROSA with HCP</b>
<b>ROSA subscription fees</b>		
Worker node subscription fee	\$9,000	\$9,000
Control plane subscription fee	\$263	\$2,191
<b>AWS infrastructure costs</b>		
<b>Worker nodes</b>		
Amazon EC2 (9 m5.xlarge nodes)	\$8,910	\$8,910
Amazon EBS (9 gp3 300GB SSDs)	\$2,592	\$2,592
<b>Infrastructure nodes</b>		
Amazon EC2 (3 r5.xlarge nodes)	\$3,891	\$0
Amazon EBS (3 gp3 300GB SSDs)	\$864	\$0
<b>Control plane nodes</b>		
Amazon EC2 (3 m5.2xlarge nodes)	\$5,940	\$0
Amazon EBS (3 gp3 350GB SSDs)	\$1,008	\$0
<b>Summary</b>		
ROSA subscription fee total	\$9,263	\$11,191
Amazon infrastructure costs total	\$23,205	\$11,502
<b>Total Costs</b>	<b>\$32,468</b>	<b>\$22,693</b>

# What is Hosted Control Plane (HCP) for ROSA

ROSA with HCP  
(Hosted Control Plane)

VS

ROSA Classic



Red Hat  
**Summit**

**Connect**

**Domande?**

Red Hat  
**Summit**

**Connect**

**Thank you**