Containers at AWS

More options and power than ever before

Jason Hoog, AWS Solution Architect

Agenda

Why are companies moving to containers

What are customers building

How are customers building

Why are companies moving to containers?

The only constant is change

Customers today face unprecedented business challenges

BUT they also have incredible opportunities to reinvent themselves

What customers ask for

Build applications, not infrastructure

Manage infrastructure to their requirements

Scale quickly and seamlessly

Security and isolation by design

Why customers adopt containers

Reduced risk

Uniform security across environment, maintained with automation

Operational efficiency

Reduced operational burden by removing undifferentiated heavy lifting

Speed

Consistent environment improves developer velocity

Agility

Automation increases speed and ease of testing and iterating

What are our customers building?

Applications

Mobile & web applications

Back-end web services

IoT

Data processing

Shared services platform

CI/CD

IaaS

Management, security,
& governance

Logging & monitoring

Enterprise app migration

.NET Classic Windows apps

Linux apps

Third-party applications

Machine learning

Autonomous vehicle

Recommendation engines

Fraud detection

Chatbots

AWS Container Service customers

ACQUIA AETION affirm ALERT LOGIC Amenity AMGEN ancestry App Annie BIRD Blackboard CLASS101

BuzzFeed Caltech Capital One cheetahmobile chime ClearDATA Clever CommonwealthBank cookpad CASTING NETWORKS

datree.io carsales DOORDASH DATADOG coursera CyberAgent. CyberCube EBSCO edmunds experian 강남언니 freee RAC

Genesys GoPro GoDaddy hulu heybit lyft MONSANTO Helix HUMAN LONGEVITY, INC. iRobot KAPLAN KDDI Kingdee 金蝶

komodohealth instacart Lamudi Lime Linden Lab mapbox matomy METEOR MORNINGSTAR movinga

nab nauto NAVITIME Nextdoor okta og OASIS GAMES OMNIOUS Prezi PIC NIC QMENTA remind

realtor.com Segment RICOH SEEDRS sportsbet .com.au shutterstock shippable TIBCO TRACKTICS Travelex TUNE IN TRUECar

SAMSUNG UBISOFT Vanguard WattzOn Upserve verizon VIACOM Trimble weever apps zynga zuora vingle vend

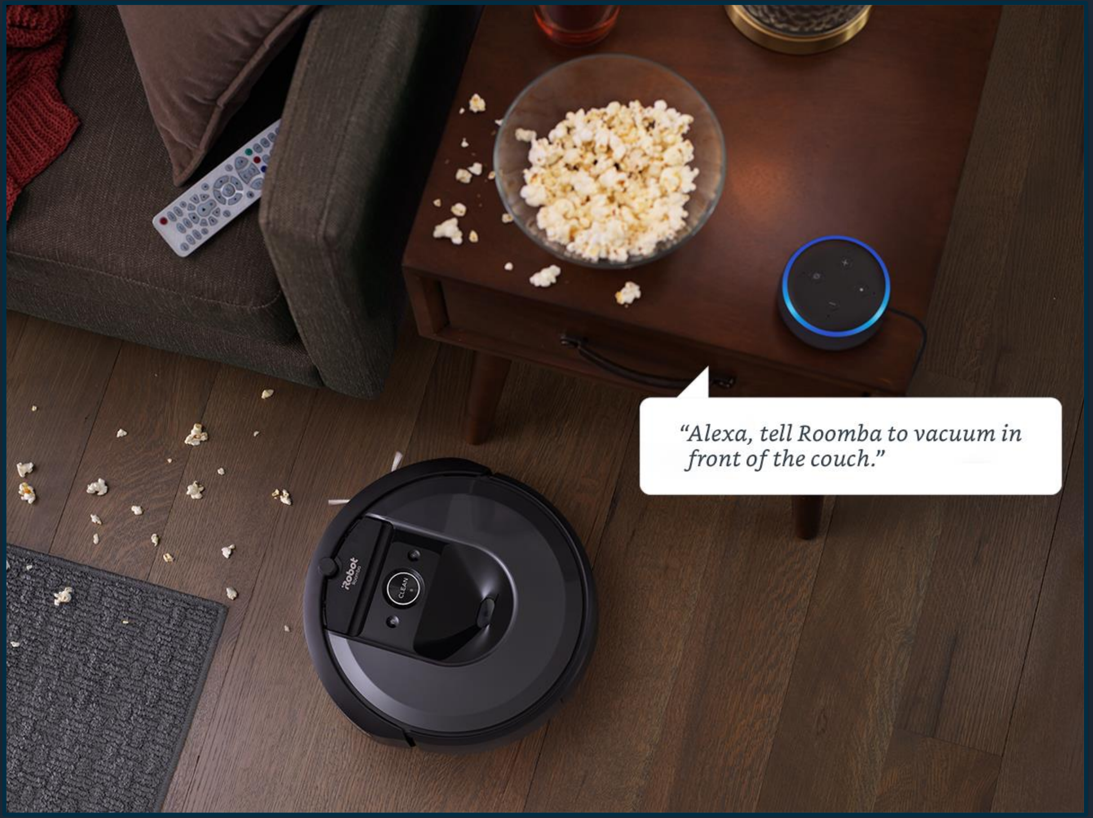
xero Uniplaces Workiva inflearn typecast wp palringo RIDECELL CloudHealth INSIDER Skyscanner

iRobot and Amazon EKS

Running Machine Learning workloads using Kubeflow

Challenge:

iRobot needed a flexible machine learning platform that can run across on-prem and AWS cloud environments



Solution and results:

"Running Kubeflow on Amazon EKS gave us a scalable machine learning platform that integrated seamlessly with AWS, and abstracted away infrastructure complexity so ML engineers could perform rapid experimentations that leveraged powerful AWS GPU based instances"

Danielle Dean, PhD

Technical Director of Machine Learning

iRobot

Ubisoft and Amazon ECS

Managing multi-player gaming service

UBISOFT

Challenge:

Ubisoft needed a way to build a highly scalable peer-to-peer relay service that reduced their costs and improved player experience

Solution and results:

”We operated 120 servers with a team of three people here, because we didn't have to spend our time managing the backend. Instead of installing and overseeing a management and orchestration tool ourselves, which would have taken us weeks, we configured our service to support the traffic spike in a few days using Amazon ECS”

Eric Fortin

Technical Architect

Ubisoft

AWS Container Competency Partners

Foundation

- docker
- HashiCorp
- D2 IQ
- VMware Tanzu
- Red Hat
- CANONICAL
- spot
- SUSE

Monitoring & Logging

- DATADOG
- dynatrace
- epsagon
- splunk>
- sumo logic
- New Relic.

DevOps & CI/CD

- ATLASSIAN
- cloudbees
- CODESHIP by CloudBees
- GitLab
- pulumi
- Travis CI
- XebiaLabs Enterprise DevOps
- weaveworks

Security & Networking

- alcide
- ALERT LOGIC
- aqua
- paloalto NETWORKS
- StackRox
- sysdig
- threat stack
- TIGERA
- TREND MICRO

AWS Consulting Partners for Containers

accenture

ALTOROS™

Capgemini

clearscale

Cloudreach™

Cognizant

mission

nClouds℠

slalom

wipro

ONICA
a rackspace technology company

How are customers building?

The containers stack on AWS

CONTAINERS

CONTAINER SERVICE
MANAGED KUBERNETES
STORE & RETRIEVE DOCKER IMAGES

COMPUTE

AUTO SCALING
BATCH JOBS
EVENT-DRIVEN SERVERLESS COMPUTING
INSTANCE TYPES
MANAGED VIRTUAL PRIVATE SERVERS
MANAGED REPOSITORY FOR SERVERLESS APPS
RUN & MANAGE WEB APPS
SERVERLESS COMPUTE
VIRTUAL SERVERS
ISOLATED COMPUTE ENVORNMENTS (FOR NITRO ENCLAVES)

NETWORKING & CONTENT DELIVERY

APPLICATION DELIVERY
DEDICATED NETWORK CONNECTION
DOMAIN NAMING SERVICE
LOAD BALANCING
MONITOR APIS
MONITOR MICROSERVICES
NETWORK TOPOLOGY
NETWORKING HUG
PRIVATE CONNECTION TO APPS
SCALE VPS & ACCOUNT CONNECTIONS
SERVICE DISCOVERY
VIRTUAL PRIVATE CLOUD

HYBRID ARCHITECTURE

AWS-SERVICES ON-PREMISES
DATA INTEGRATION
INTEGRATED DEVICES & EDGE SYSTEMS
INTEGRATED IDENTITY & ACCESS
INTEGRATED NETWORKING
INTEGRATED RESOURCE &
DEPLOYMENT MANAGEMENT
VMWARE CLOUD ON AWS

INFRASTRUCTURE

AVAILABILITY ZONES
CUSTOM HARDWARE
DATA CENTER INFRASTRUCTURE
GLOBAL NETWORK BACKBONE
POINT OF PRESENCE
POWER INFRASTRUCTURE
REGIONS

STORAGE

ARCHIVE STORAGE
BACKUP & RESTORE
BLOCK STORAGE
DATA TRANSFER
EDGE PROCESSING & COMPUTING
FILE STORAGE
HIGH PERFORMANCE FILE SYSTEM
HYBRID CLOUD STORAGE
OBJECT STORAGE
WINDOWS FILE SYSTEM

SECURITY, IDENTITY, & COMPLIANCE

ACCESS CONTROL
ASSESSMENT & REPORTING
CONFIGURATION COMPLIANCE
DATA PROTECTION
DDOS PROTECTION
IDENTITY MANAGEMENT
KEY MANAGEMENT & STORAGE
MONITORING & LODGING
RESOURCE MANAGEMENT
THREAT DETECTION
WEB APPLICATION FIREWALL
THREAT DETECTION & INVESTIGATION FOR AMAZON DETECTIVE
AUTOMATED SECURITY POSTURE CHECKS FOR AWS SECURITY HUB

MANAGEMENT & GOVERNANCE

ACTIVITY & API USAGE TRACKING
CHATBOT
CONFIGURATION TRACKING
GOVERNANCE
INVENTORY TRACKING
LICENSE MANAGER
MANAGE POLICIES
MANAGE RESOURCES
MONITORING
PROVISIONING
RESOURCE TEMPLATES
SECURITY RECOMMENDATIONS
SERVER MANAGEMENT
SERVICE CATALOG
SYSTEM MANAGER

DEV TOOLS

ANALYSE & DEBUG
APPLICATION LIFECYCLE MANAGEMENT
AUTHORING
BUILD & TEST
CONTAINERS
DEVOPS RESOURCE MANAGEMENT
ONE CLICK APP DEVELOPMENT
PATCHING
PIPELINE ORCHESTRATION
RESOURCE TEMPLATES
TRIGGERS

MOBILE

API GATEWAY
DEVELOPMENT FRAMEWORK
IDENTITY
MOBILE ANALYTICS
MOBILE APP TESTING
SINGLE INTEGRATED CONSOLE
SYNC
TARGETED PUSH NOTIFICATIONS

APPLICATION INTEGRATION

EMAIL
MESSAGE BROKER
QUEUING & NOTIFICATIONS
SEARCH
TRANSCODING
WORKFLOW

Container technology

Amazon EKS or Amazon ECS?

ECS

Powerful simplicity

EKS

Open flexibility

Powerful simplicity

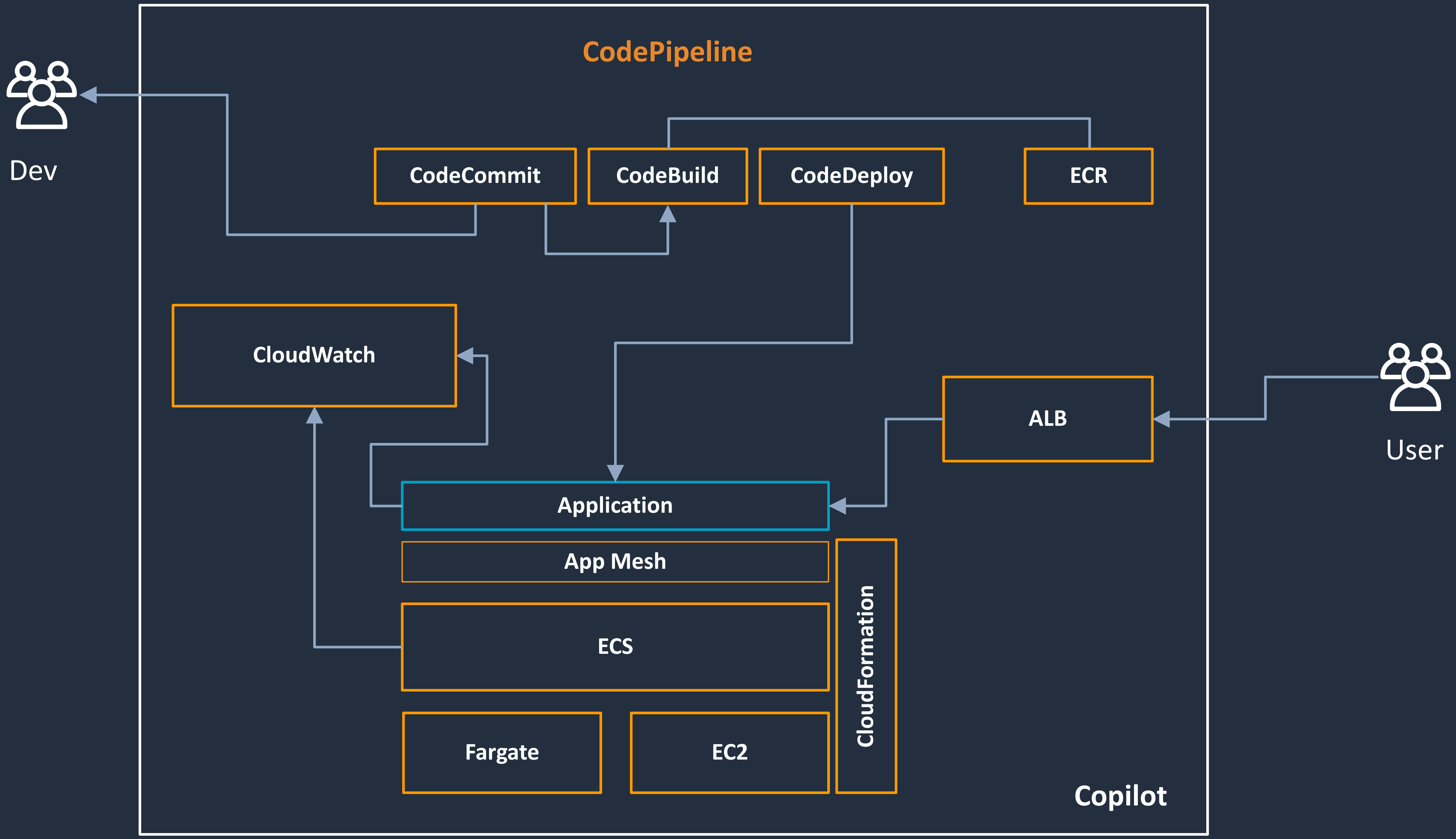
ECS

AWS-opinionated way to run containers at scale

Reduce decisions without sacrificing scale or features

Reduce time to build, deploy, and migrate applications

Powerful simplicity



Open flexibility

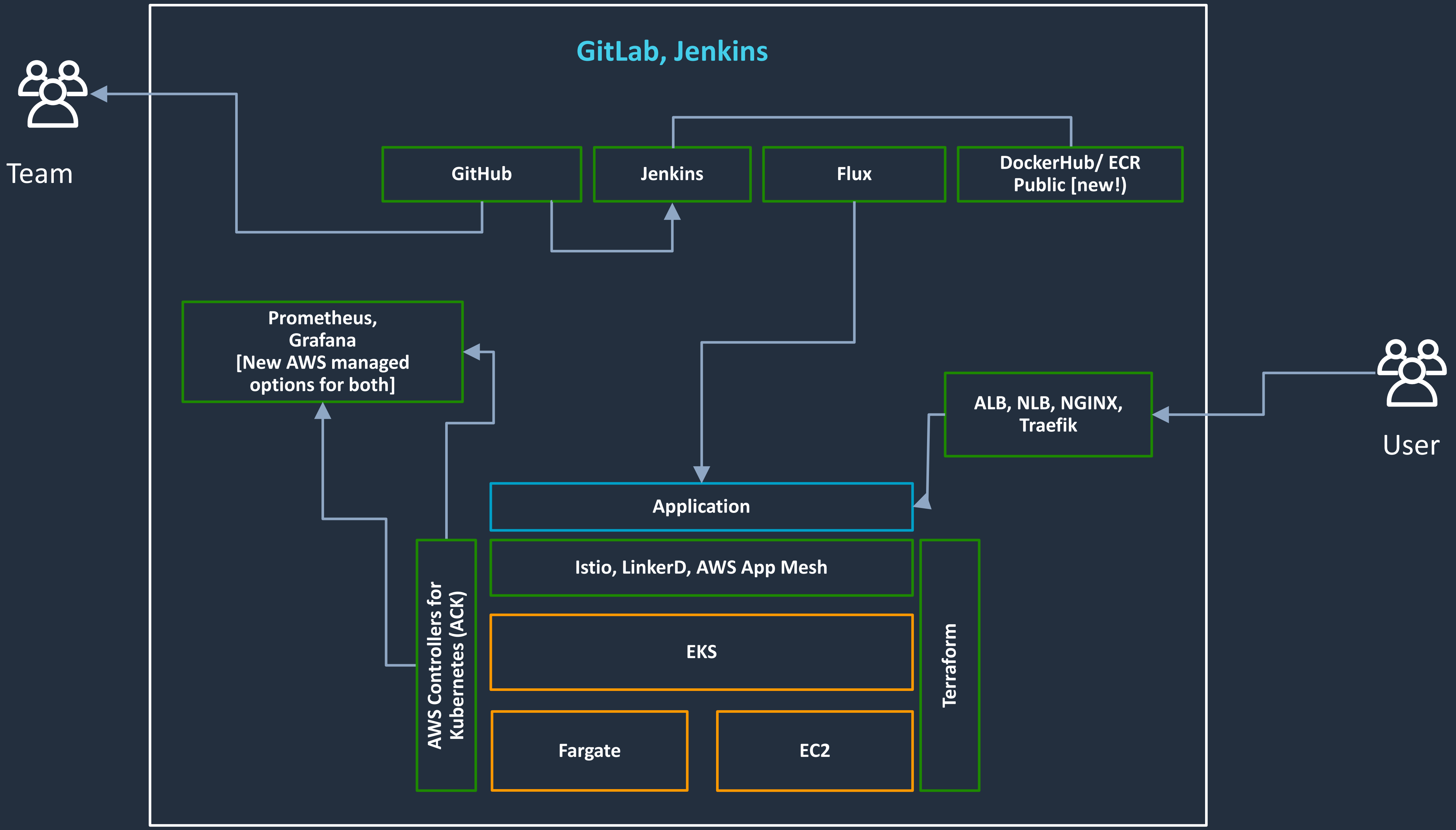
EKS

Gain agility and efficiency with AWS-optimized Kubernetes, and standardize operations everywhere

Secure, highly available, with observability across all Kubernetes deployments

Build with choice of solutions from the broader community around Kubernetes

Open flexibility


GitLab, Jenkins
Team
GitHub
Jenkins
Flux
DockerHub/ ECR Public [new!)
Prometheus, Grafana [New AWS managed options for both]
ALB, NLB, NGINX, Traefik
User
Application
Istio, LinkerD, AWS App Mesh
EKS
Fargate
EC2
AWS Controllers for Kubernetes (ACK)
Terraform

Amazon Elastic Kubernetes Service (EKS) is a fully managed Kubernetes service. EKS runs upstream Kubernetes and is certified Kubernetes conformant.

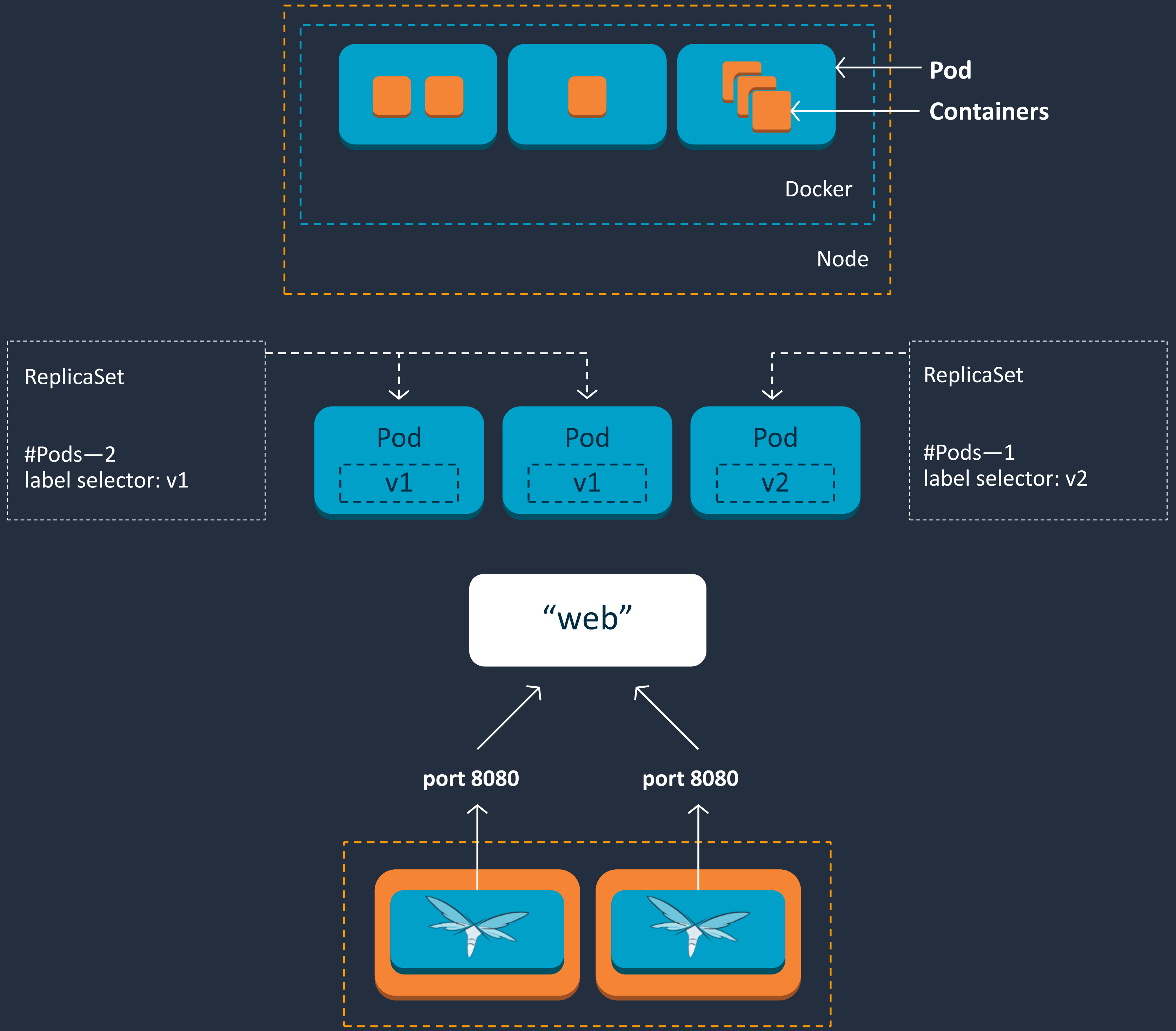
Kubernetes Concepts

Pods: Co-located group of containers that share an IP, namespace, storage volume

Replica Set: Manages the lifecycle of pods and ensures specified number are running

Service: Single, stable name for a set of pods, also acts as LB

Label: Used to organize and select group of objects



Kubernetes Concepts

Namespaces: “Virtual” clusters for users/projects

Ingress controller: L7 load balancing

Deployments: Declarative version updates

Jobs: Run to completion

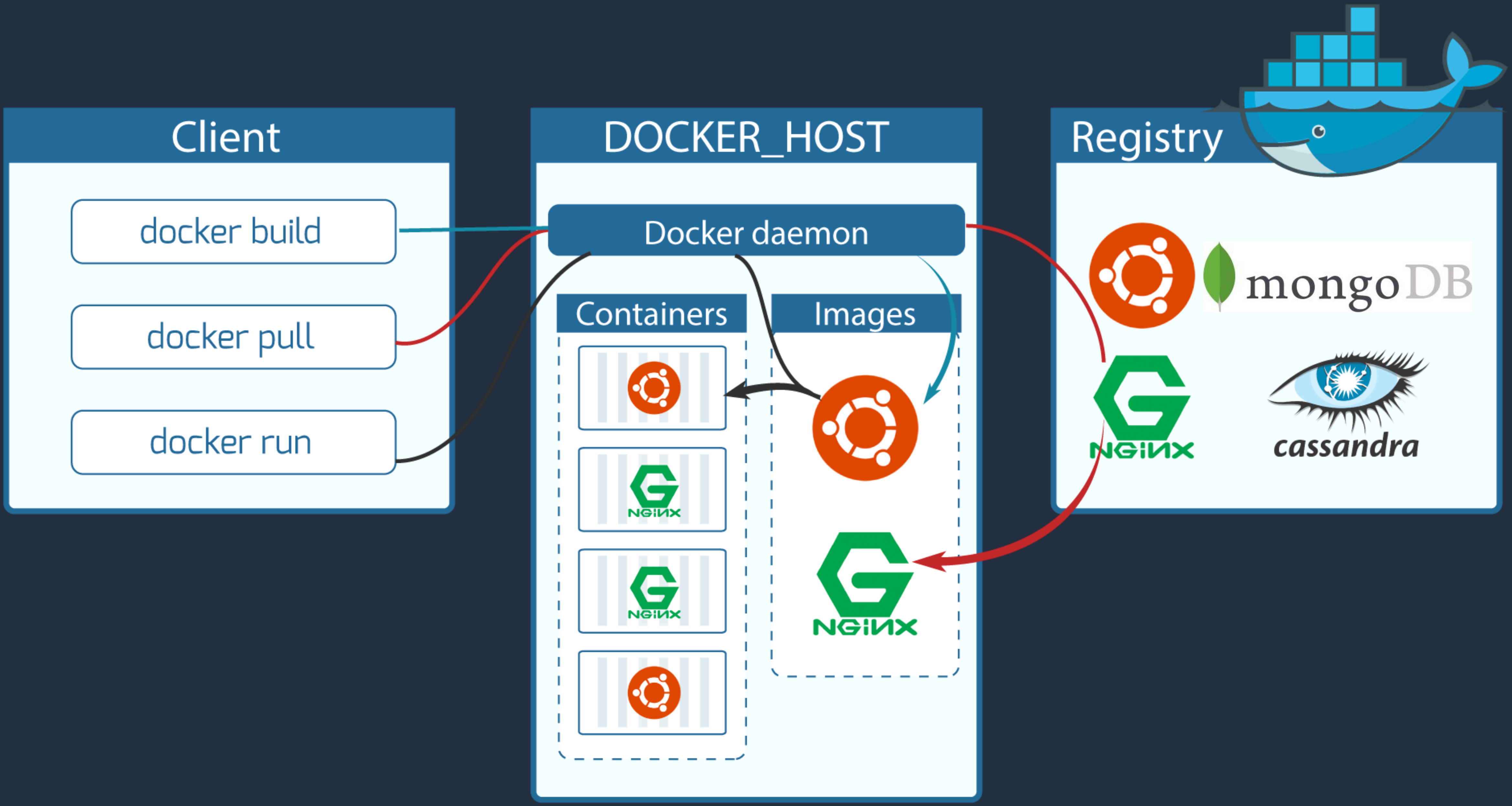
Autoscale: Automatically adjust number of Pods

Network Policies: AKA Security Groups for Pods

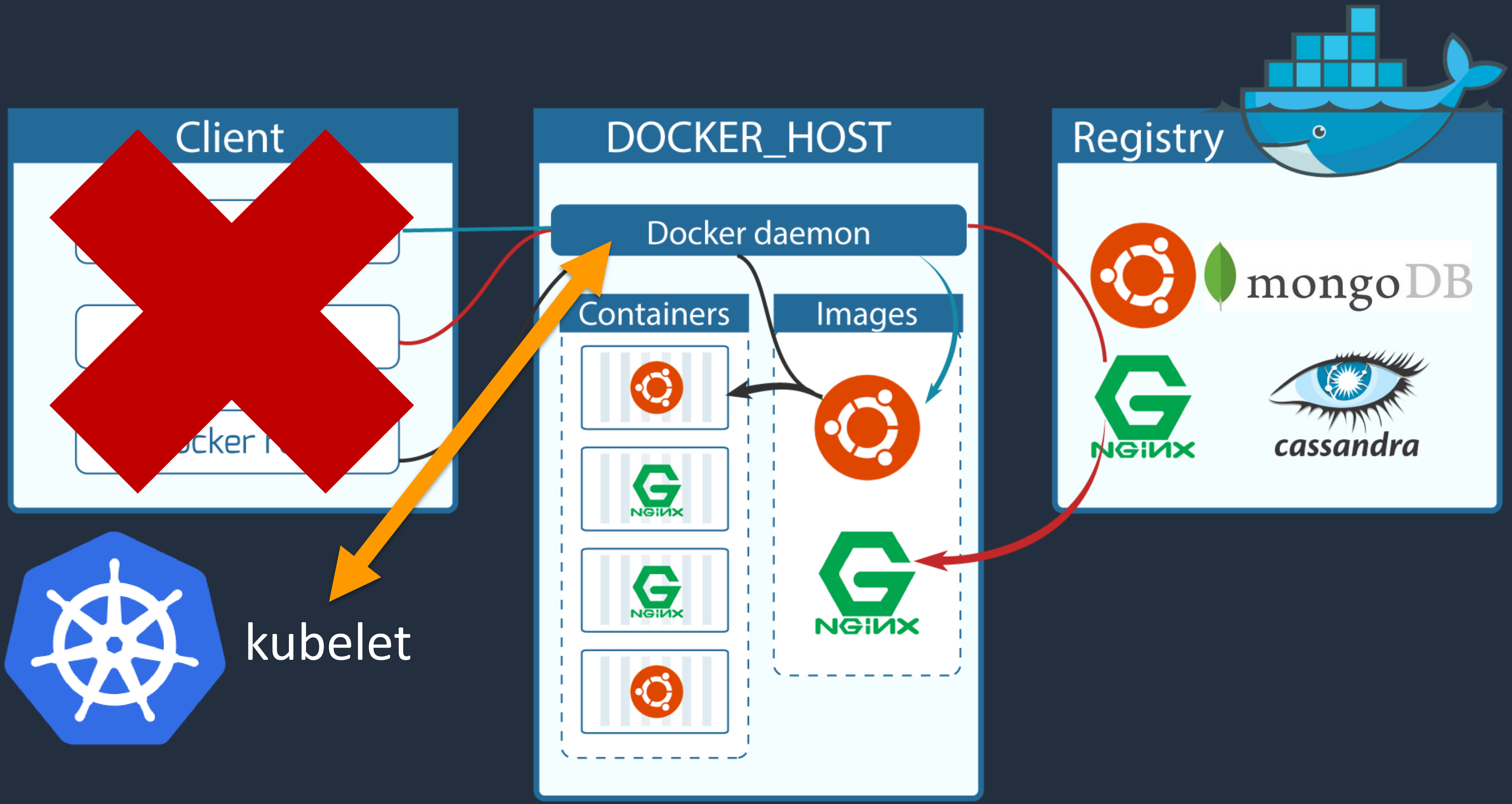
StatefulSet: Support for long-term stateful distributed systems

More…

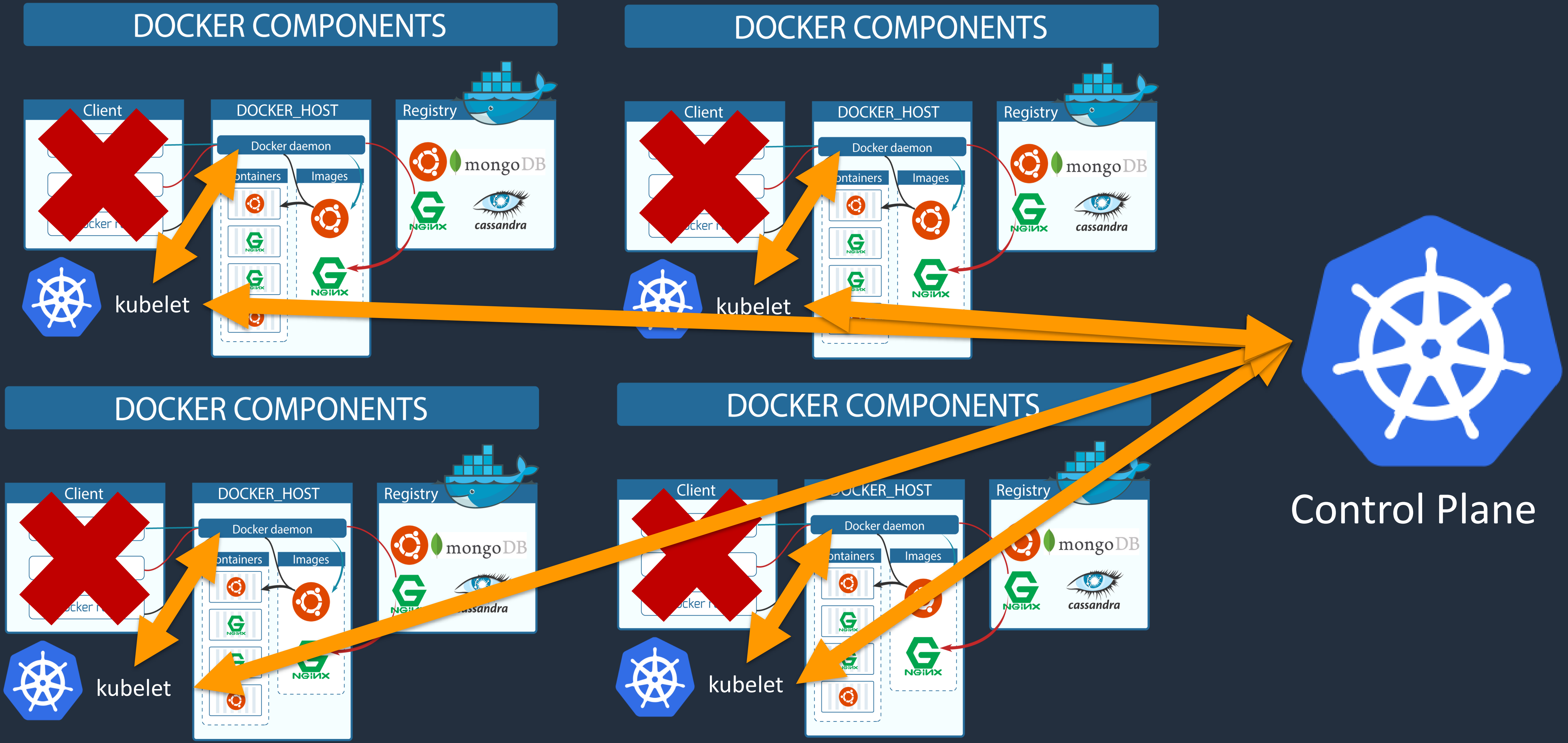
DOCKER COMPONENTS



DOCKER COMPONENTS



aws


DOCKER COMPONENTS
Client
DOCKER_HOST
Registry
Docker daemon
Containers
Images
mongoDB
cassandra
NGINX
kubelet
DOCKER COMPONENTS
Client
DOCKER_HOST
Registry
Docker daemon
Containers
Images
kubelet
DOCKER COMPONENTS
Client
DOCKER_HOST
Registry
Docker daemon
Containers
Images
kubelet
DOCKER COMPONENTS
Client
DOCKER_HOST
Registry
Docker daemon
Containers
Images
kubelet
Control Plane

What does EKS do for you over running Kubernetes on EC2?

- We deploy the Kubernetes Control Plane and etcd in a highly-available configuration across 3 AZs
- We manage that control plane for you in a similar way to our managed relational database service RDS
- We provide a network (CNI) plugin that integrates Pod networking natively with AWS VPC
- We integrate/federate user access to the Kubernetes CLI (kubectl) and API with AWS IAM

EKS is Kubernetes Certified

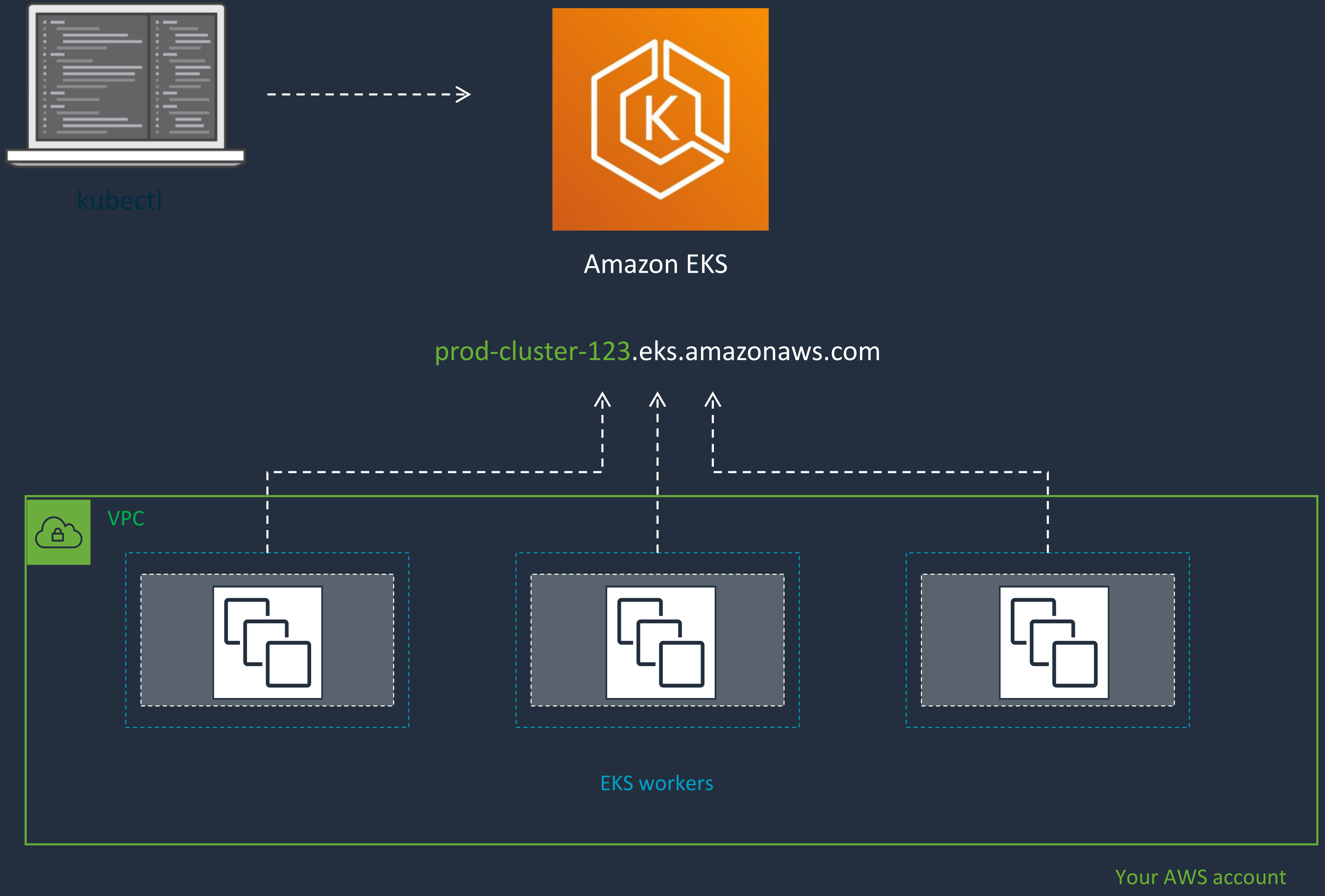
EKS runs upstream Kubernetes and is certified as Kubernetes conformant.

This means that applications managed by Amazon EKS are fully compatible with applications managed by any standard Kubernetes environment.

certified

kubernetes

Amazon EKS Architecture



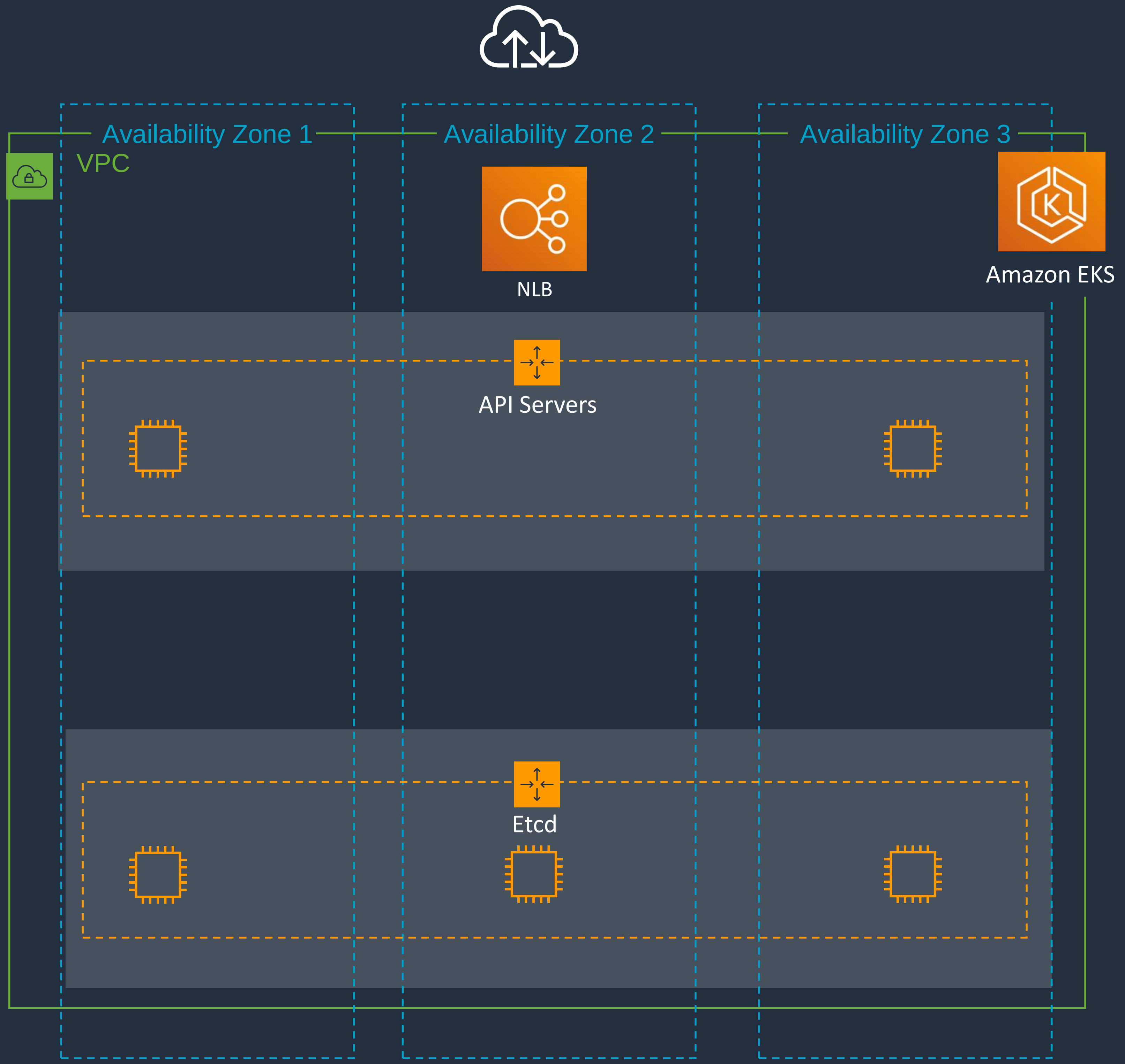
EKS Control Plane

Highly available and single tenant infrastructure

All “native AWS” components

Separate etcd to help with safe and seamless ops/upgrades

Fronted by an NLB - which enables Private endpoints into your VPC



Run your containers anywhere based on your workload needs

Serverless

AWS Fargate

EC2 options

Amazon EC2

Spot instance

Edge and 5G

AWS Local Zones

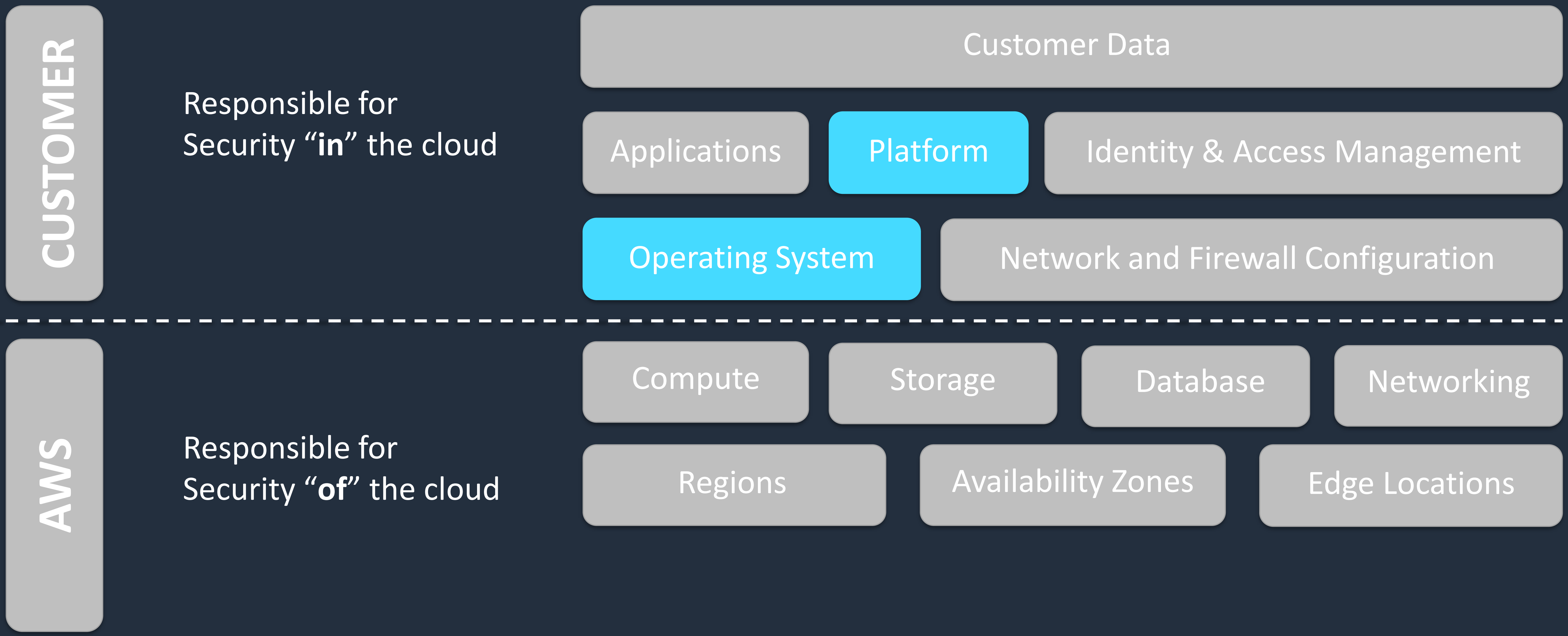
AWS Wavelength

On-premises

AWS Outposts

EKS Anywhere
ECS Anywhere

Shared Responsibility Model



CUSTOMER

Responsible for
Security “**in**” the cloud

Customer Data

Applications

Platform

Identity & Access Management

Operating System

Network and Firewall Configuration

AWS

Responsible for
Security “**of**” the cloud

Compute

Storage

Database

Networking

Regions

Availability Zones

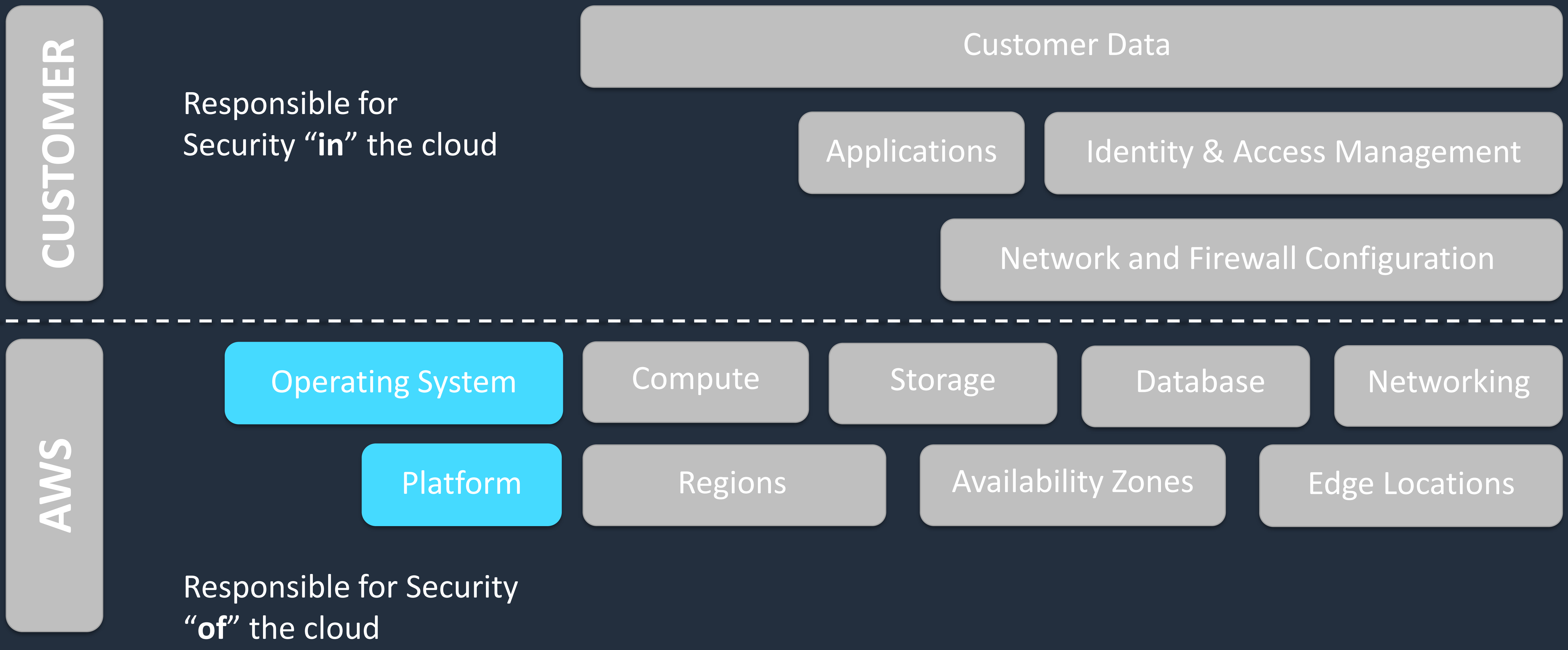
Edge Locations

EC2 Mode – Customer Responsibilities

- Instance type and quantity to choose?
  - What is the CPU to RAM ratio?
  - Excess capacity for scaling and availability?
- Which OS to choose?
  - If Amazon Linux we provide AMIs
- Hardening the OS (e.g. against CIS benchmark)
- The patching of the OS, Docker, ECS Agent or kubelet etc.

Photo & Licence

Shared Responsibility Model - Fargate



CUSTOMER

Responsible for
Security “**in**” the cloud

Customer Data

Applications

Identity & Access Management

Network and Firewall Configuration

AWS

Operating System

Compute

Storage

Database

Networking

Platform

Regions

Availability Zones

Edge Locations

Responsible for Security
“**of**” the cloud

Updating EKS

- Kubernetes has a new major version every quarter
- Kubernetes has a new minor version quite regularly
- Sometimes Kubernetes updates are security-related
- EKS has APIs to trigger an update of the control plane
- You then need to update the worker Nodes - both re: Kubernetes as well as Docker and OS
    - Often the workers are in an Autoscaling Group so this means building updating AMIs
    - We provide a regularly updated EKS Node AMI as well as scripts to build your own.

Many containerized applications need persistent storage

Long-running stateful applications

Developer tools

Jenkins
Jira
Git

Web & content management

WordPress
Drupal
nginx

Shared data sets

Machine learning

MXNet
TensorFlow
Kubeflow

Data science tools

Jupyter(hub)
Airflow

Address common networking challenges

Set up network

Amazon VPC

Discover services & secure traffic

AWS App Mesh

AWS Cloud Map

Balance incoming traffic

Elastic Load Balancing

Application load balancer

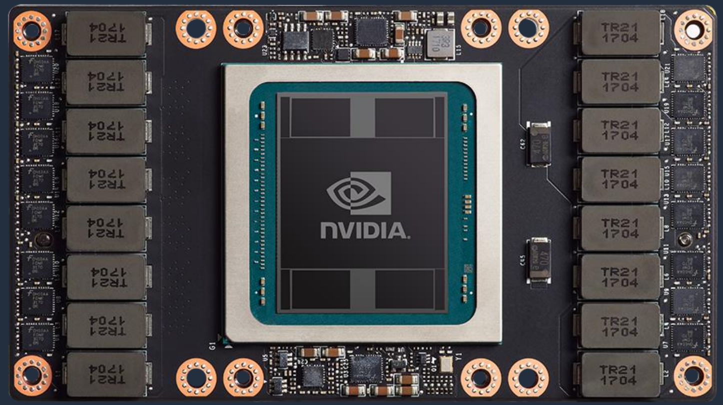
Network load balancer

You can use GPU-enabled instances in EKS

AWS has several instance types that have GPUs in them such as our new p3s.

These instance types are useful for things like Machine Learning and other computationally intensive things that can benefit from massive parallelisation.

It is possible to have some of them in a separate NodeGroup within the a mixed-use cluster and then use Kubernetes' taints and tolerations to ensure that only those workloads that benefit from the GPUs get scheduled on those instances.



P3 P2 G3

You can use spot instances with EKS

AWS sells instances that are currently sitting unused at a discount of up to 90%.

The catch is that if a customer comes along and wants to pay the OnDemand or Reserved pricing for that instance we'll take it back with a two-minute warning.

It is possible to have either a cluster totally running on Spot (and a few tricks to make that safe-ish) or a mixed cluster with a foundational capacity of OnDemand with bursting out to Spot when it is affordable.

You can now add Windows nodes to your EKS cluster

We’ve just announced that adding Windows nodes to your EKS cluster is now Generally Available.

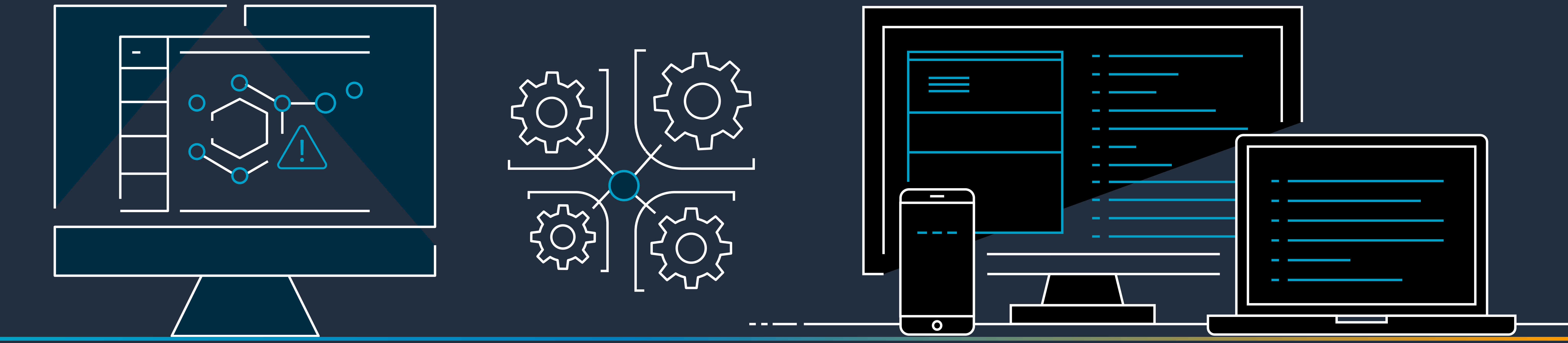
A couple provisos to Windows Server Containers generally that extend to Kubernetes:

- **Half of Windows is in each container**
- **The windows version in the container image must match the host OS**

Also all the containers/sidecars in the Pod need to run on the same host/OS.

Connecting it all together

Abstract away the developer pain



Building applications is hard

With the right automation, it gets a lot easier

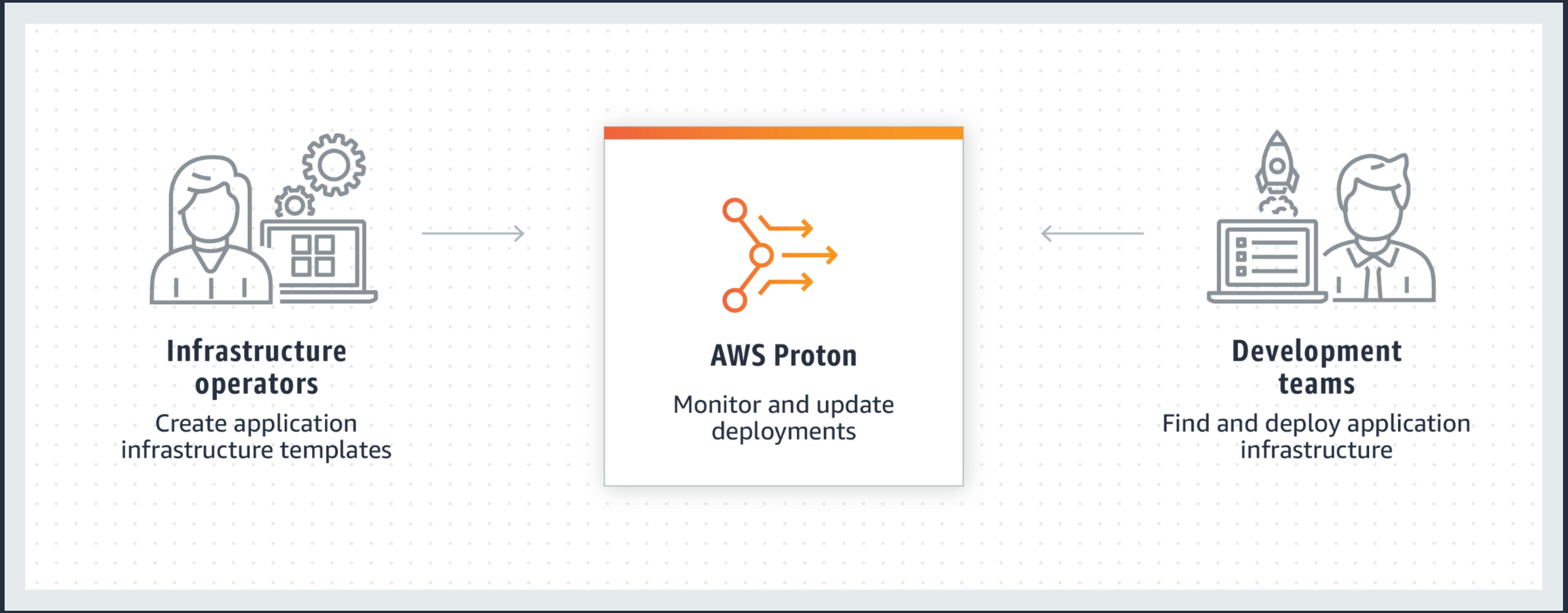
You can be more productive and agile, focus on creation, not admin work

New: Manage deployments with AWS Proton

PROBLEM

Deploy code pipelines with consistent standards and management

The first fully managed deployment service for container and serverless applications.



Infrastructure operators

Create application infrastructure templates

AWS Proton

Monitor and update deployments

Development teams

Find and deploy application infrastructure

Conclusion

Customers today face unprecedented business challenges and going faster than ever before

Customers are building their business critical applications with AWS container services

We have a full stack of services, technologies, and tools for every workload

Many more to come in 2021!

Thank you!

Container launch roundup

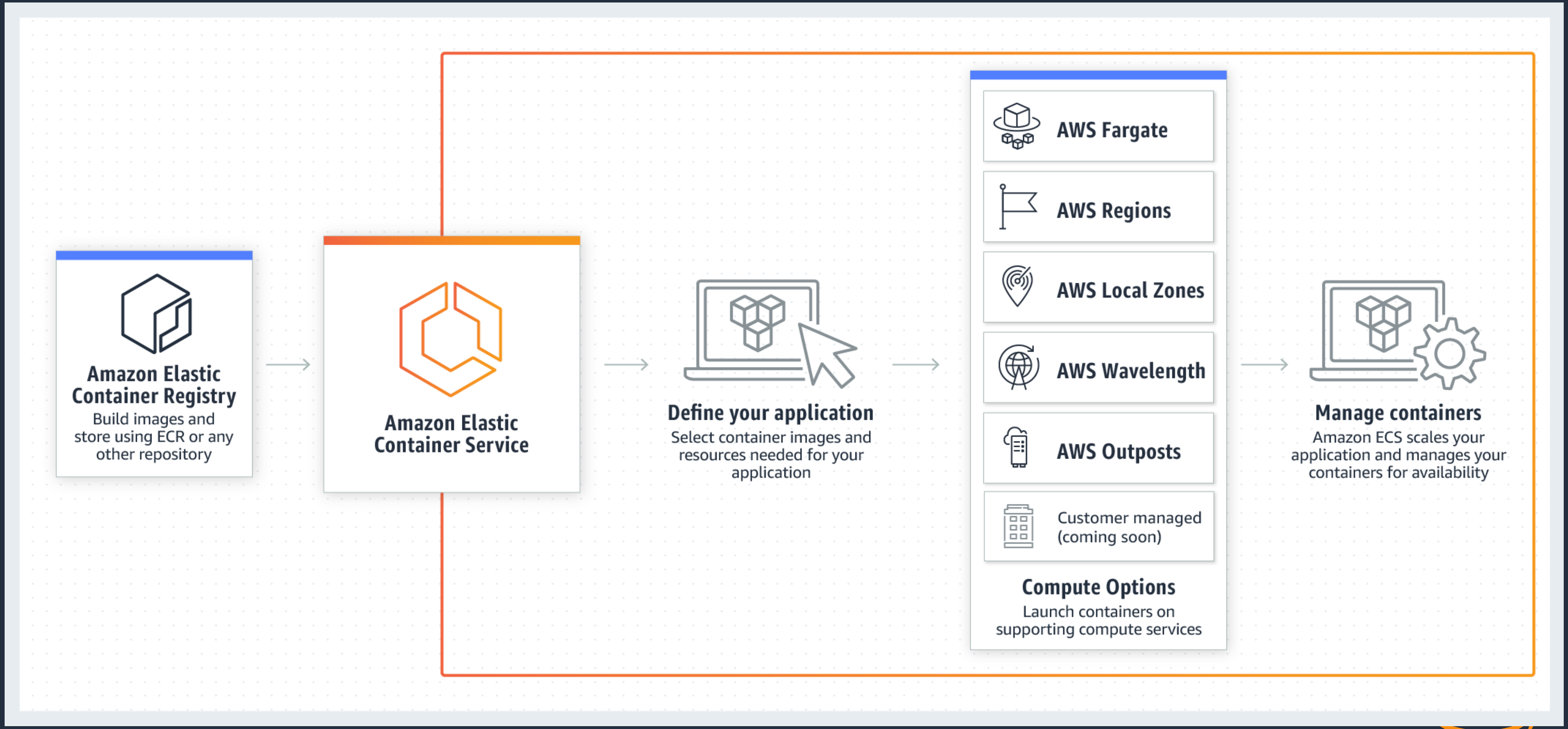
New: Amazon ECS Anywhere

PROBLEM

How to use ECS for application portfolios that span AWS, on-premises, and other customer managed infrastructure

Fully managed, and highly scalable

Powerful simplicity for your hybrid footprint



Amazon Elastic Container Registry
Build images and store using ECR or any other repository
Amazon Elastic Container Service
Define your application
Select container images and resources needed for your application
AWS Fargate
AWS Regions
AWS Local Zones
AWS Wavelength
AWS Outposts
Customer managed (coming soon)
Compute Options
Launch containers on supporting compute services
Manage containers
Amazon ECS scales your application and manages your containers for availability

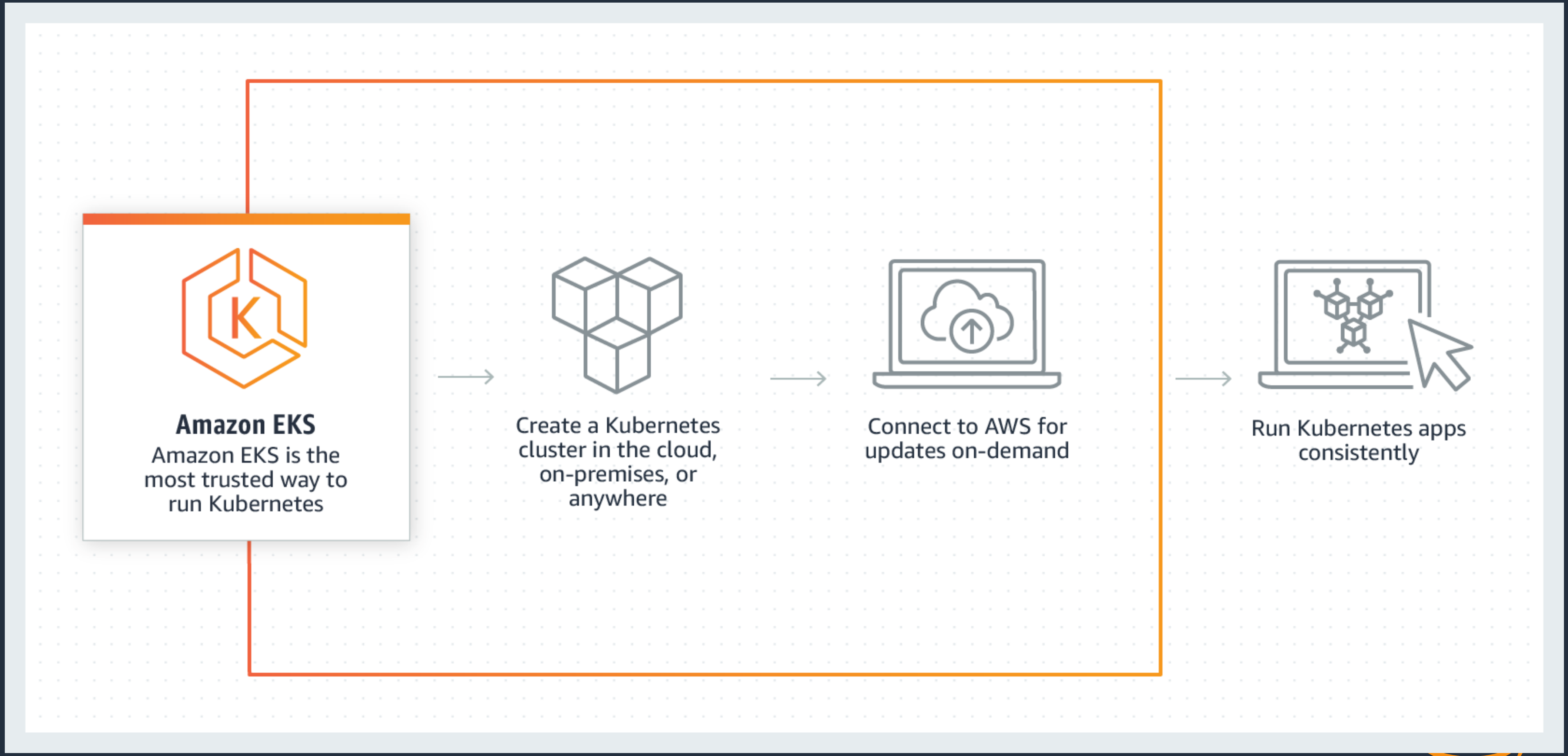
New: Amazon EKS Anywhere and EKS Distro

PROBLEM

How to run EKS outside of AWS regions. Customers have workloads, workflows, and application portfolios that span AWS, on-premises, and other clouds.

Manage K8s clusters on-premises

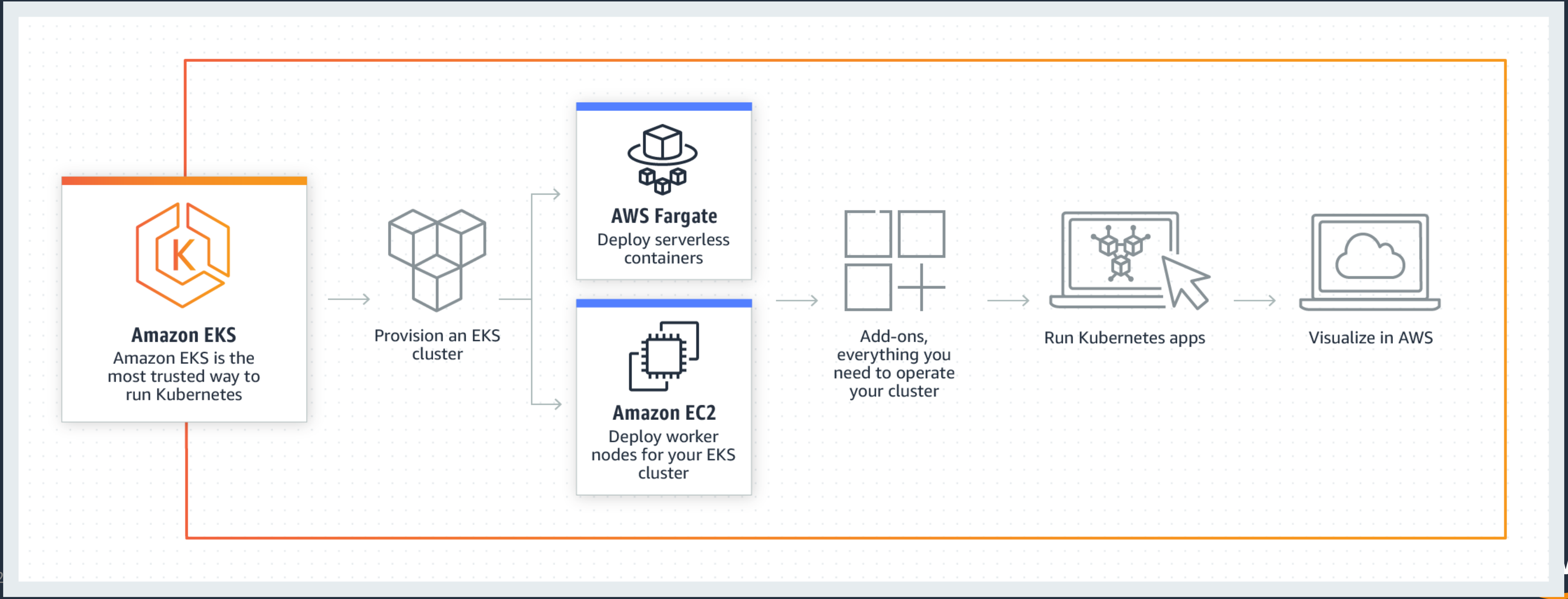
Amazon EKS Distro is the same one run on AWS cloud


Amazon EKS
Amazon EKS is the most trusted way to run Kubernetes
Create a Kubernetes cluster in the cloud, on-premises, or anywhere
Connect to AWS for updates on-demand
Run Kubernetes apps consistently

New: EKS Add-ons and dashboard

PROBLEMS

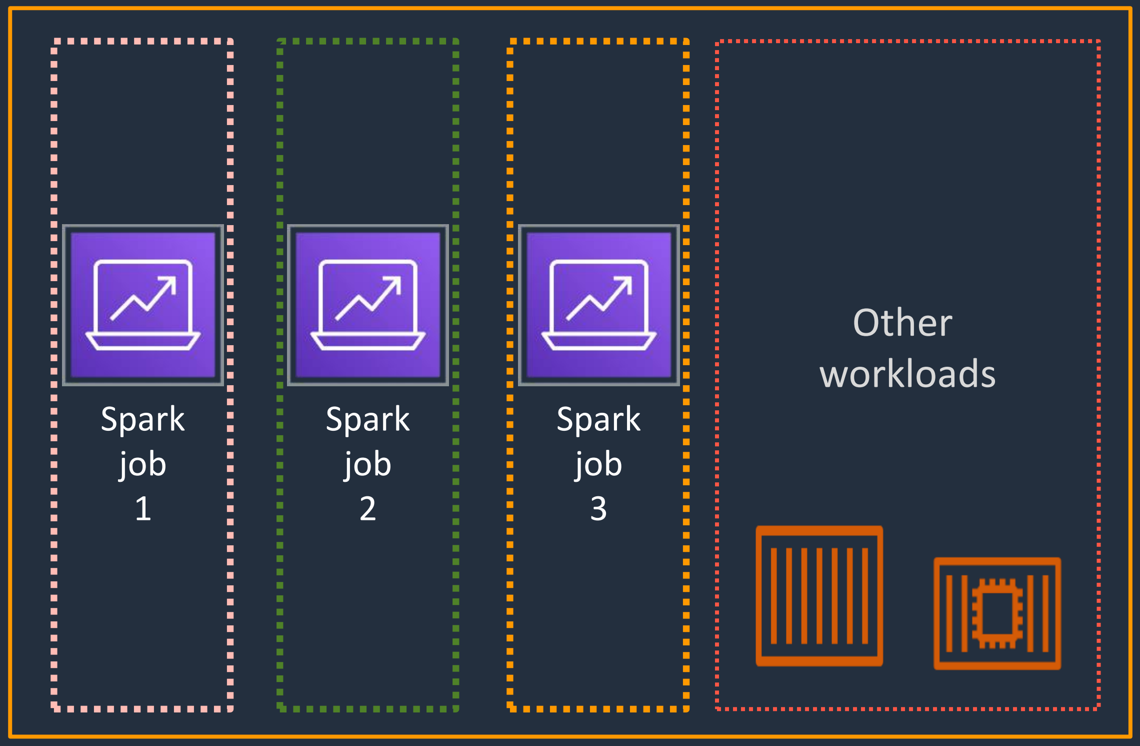
Limited observability to containers in different places. Complexity in managing and patching Kubernetes components.


Amazon EKS
Amazon EKS is the most trusted way to run Kubernetes
Provision an EKS cluster
AWS Fargate
Deploy serverless containers
Amazon EC2
Deploy worker nodes for your EKS cluster
Add-ons, everything you need to operate your cluster
Run Kubernetes apps
Visualize in AWS

Amazon EMR on Amazon EKS

Run Apache Spark jobs on Amazon Elastic Kubernetes Service (EKS)

INTRODUCING
GA



Amazon EKS

Eliminate underutilized and over-provisioned resources

Automate management of big data frameworks

Optimize Apache Spark performance

A new public container registry from AWS

Meeting customer need, leveraging our experience operating Amazon ECR at scale

PROBLEM

How to innovate and collaborate with images and artifacts

Amazon ECR Public

Geo-replicated image storage

Amazon CloudFront cache

Single, global URL

Essentially free to use

No AWS account needed to pull

```
docker pull public.ecr.aws/ecs/amazon-ecs-agent
```

Amazon ECR Public Gallery

Search for public container artifacts

Image detail pages

Custom aliases

Verified accounts

Free for anyone to browse

https://gallery.ecr.aws

Red Hat OpenShift on AWS (ROSA)

Console service

- Create OpenShift clusters from the AWS console or CLI
- AWS integrated experience for cluster creation and management
- Foundation based on RHEL

Unified bill

- Leverage your existing AWS commitment to use OpenShift
- Get a single unified bill from AWS for both OpenShift and AWS consumption

Joint support

- Integrated support systems
- Contact Red hat or AWS support
- Built on Red Hat and AWS' decades of enterprise IT knowledge and experience

Integration with AWS

- Build containerized applications that integrate natively with the more than 170 AWS cloud-native services

Amazon Managed Service for Prometheus

IN PREVIEW TODAY

Highly available, secure, and managed monitoring and alerting for containers

Ingest, query, and store Prometheus metrics at scale, easily and securely

Monitor containers on AWS and on-premises

Get started quickly monitoring EKS and ECS

Amazon Managed Service for Grafana

IN PREVIEW TODAY

Powerful, interactive data visualizations for builders, operators, and business leaders

Run Grafana at scale, easily and securely

Visualize, analyze, and correlate metrics, logs, traces, and IoT data securely across multiple data sources

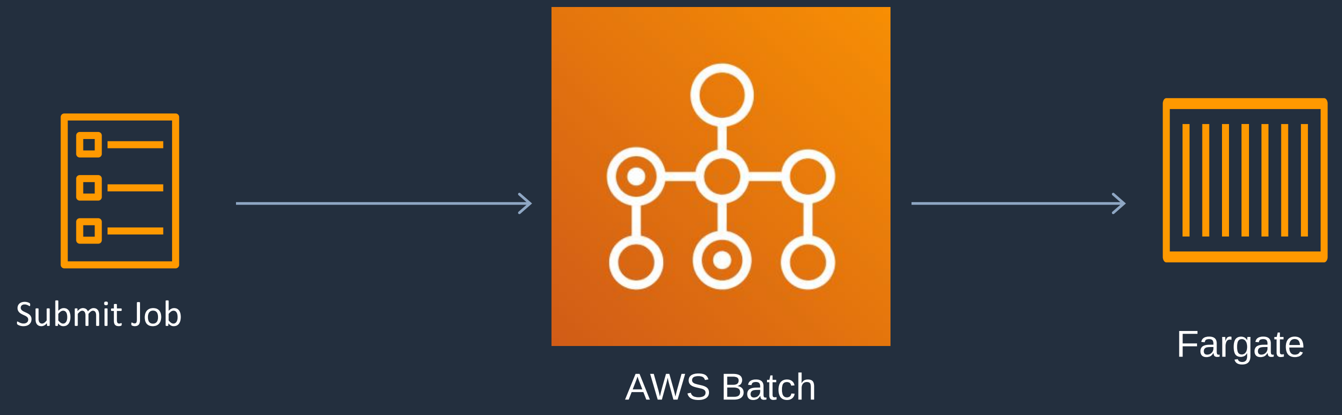
Upgrade to Grafana Enterprise to connect to 3rd party ISVs such as Datadog and Splunk

Introducing AWS Batch Jobs for Fargate

Fully Serverless Batch computing with AWS-owned compute resources; no need to specify instance type or manage machine images

AWS Batch provides you with a managed batch queue, complete with the ability to specify priority, dependencies, and retries.

- Use Fargate Spot for savings of up to 70%
- Batch handles queueing, submission, and lifecycle management



New launches and features—highlights

EKS

New EKS Console

EKS Anywhere/EKS Kubernetes Distro

EKS Add-Ons

Managed Node Groups—Spot

Control Plane Auto-tuning

OIDC Authentication API

Load Balancer Controller

App Mesh Controller

Amazon Controllers for Kubernetes (ACK)

Open-source cluster scaler—Karpenter

ECS

New Amazon ECS Console

ECS Anywhere

Proton for ECS

AWS Copilot

Docker Compose with ECS GA

CDK Extensions (FireLens, AppMesh)

AWS Distro for OpenTelemetry

Deployment circuit breaker

Capacity Providers enhancements

Amazon FSx for Windows support

New launches and features—highlights

Fargate

AWS Batch on Fargate

Persistent Storage w/ EFS

Bigger ephemeral storage

Ephemeral storage default encryption

Dual-stack for IPv4/IPv6

NLB Support

Increased default task/pod quotas

Usage against quotas in CloudWatch

Enhanced network perf metrics

Logging via FireLens (EKS)

Savings Plan (EKS)

ECR

ECR Public

ECR Public Gallery

Cross region replication

Increased layer sizes

Lambda image support

Helm chart support

Additional artifact types

Encryption at rest with KMS

Multi-Arch Support

App Mesh

mTLS—SPIRE or bring your cert

TLS—ACM or bring your cert

Circuit breaking

App Mesh K8S Controller GA

Ingress—Virtual gateways

Cross Account with AWS RAM