AWS re:Invent

CON421-R

Amazon EKS under the hood

Eswar Bala
Sr. Software Development Manager
Amazon Web Services

Richard Sostheim
Principal Engineer
Amazon Web Services

Ahmed El Baz
Software Engineer
Snap Inc

AWS re:Invent

aws

Agenda

Amazon EKS architectural overview

Amazon EKS under the hood

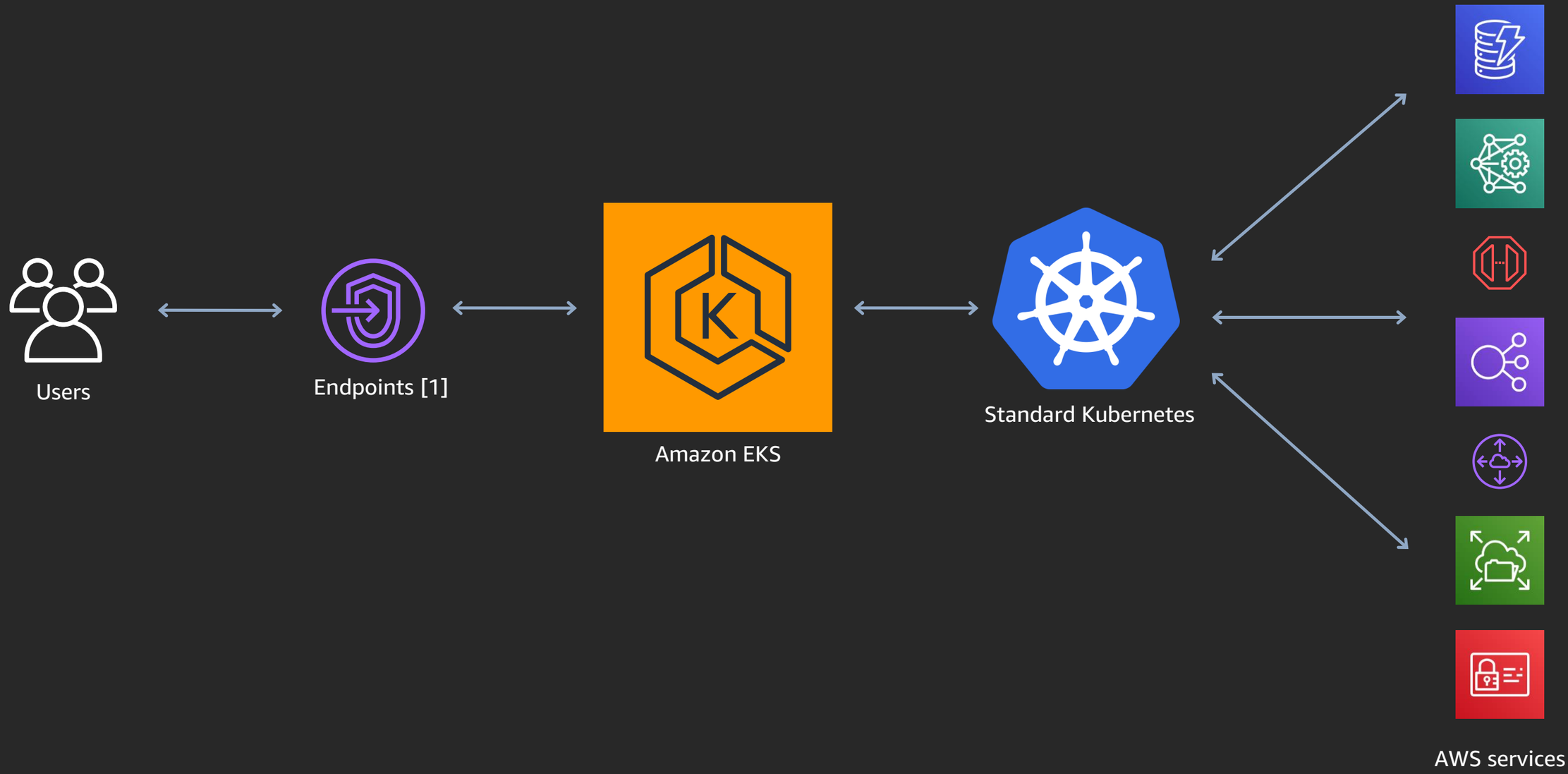
Amazon EKS operations

Amazon EKS enhancements

Snap Service Mesh

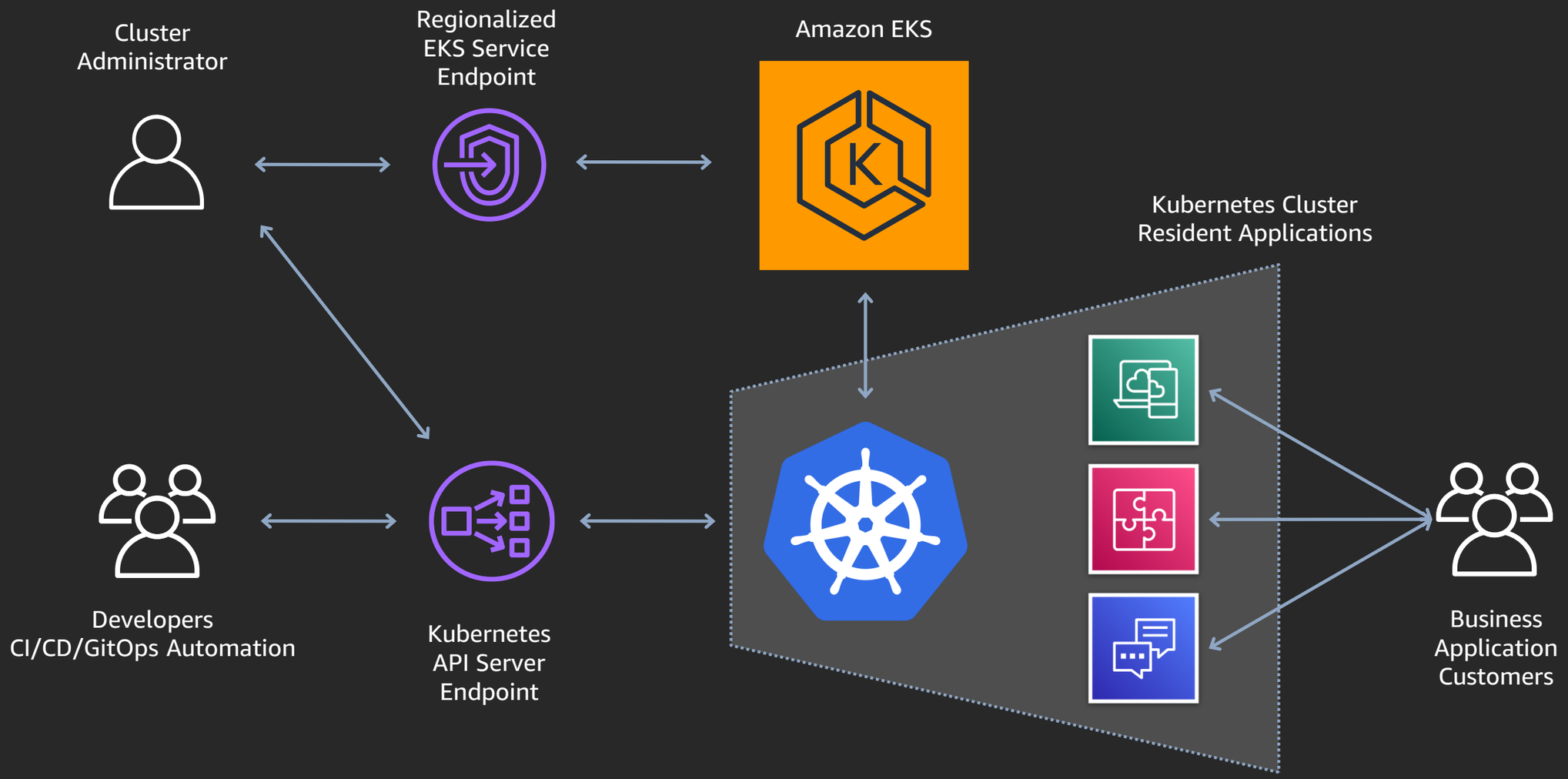
Amazon EKS architectural overview

Amazon Elastic Kubernetes Service (Amazon EKS)

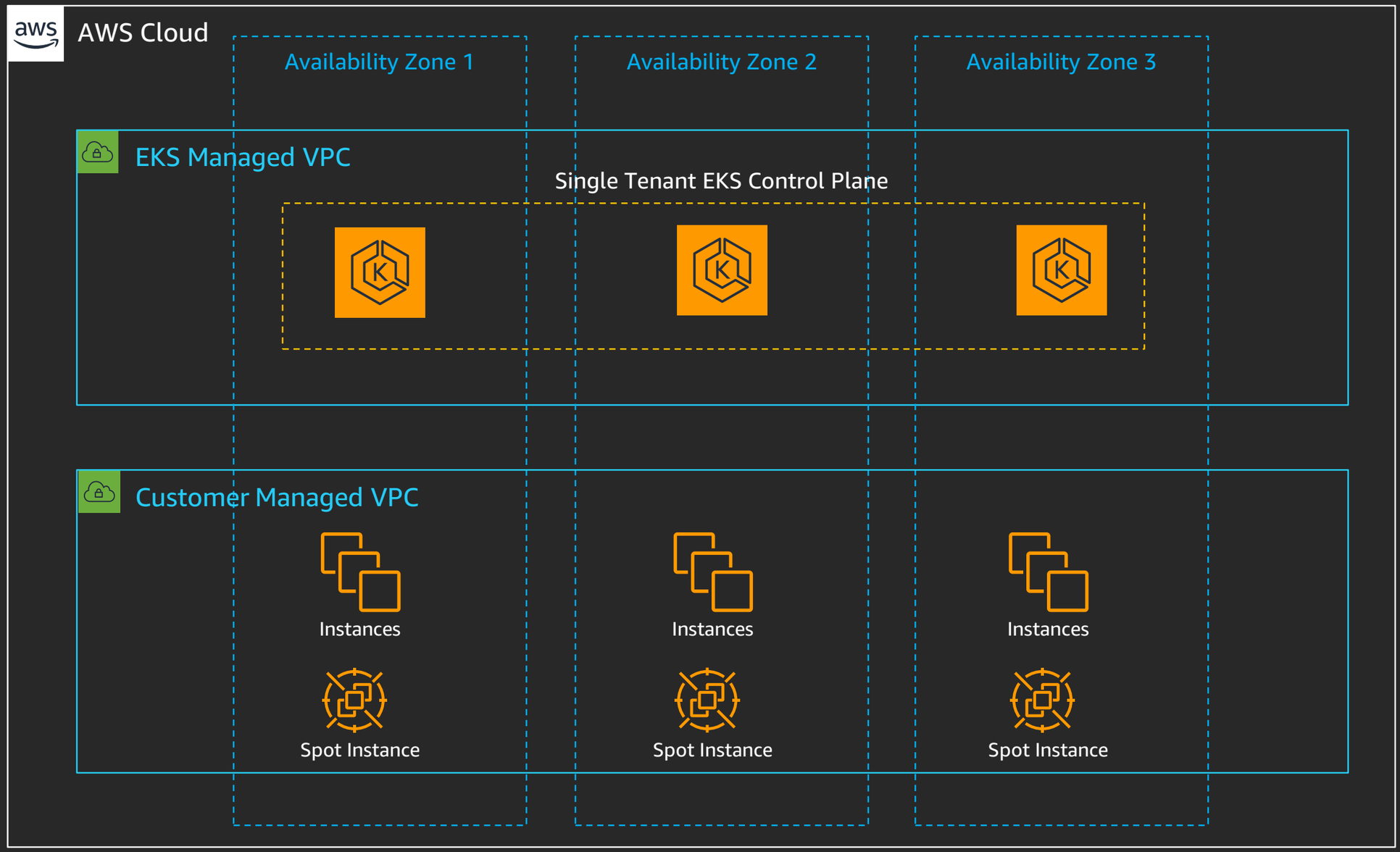


[1] https://docs.aws.amazon.com/general/latest/gr/rande.html#eks_region

EKS Service / Kubernetes Logical Overview

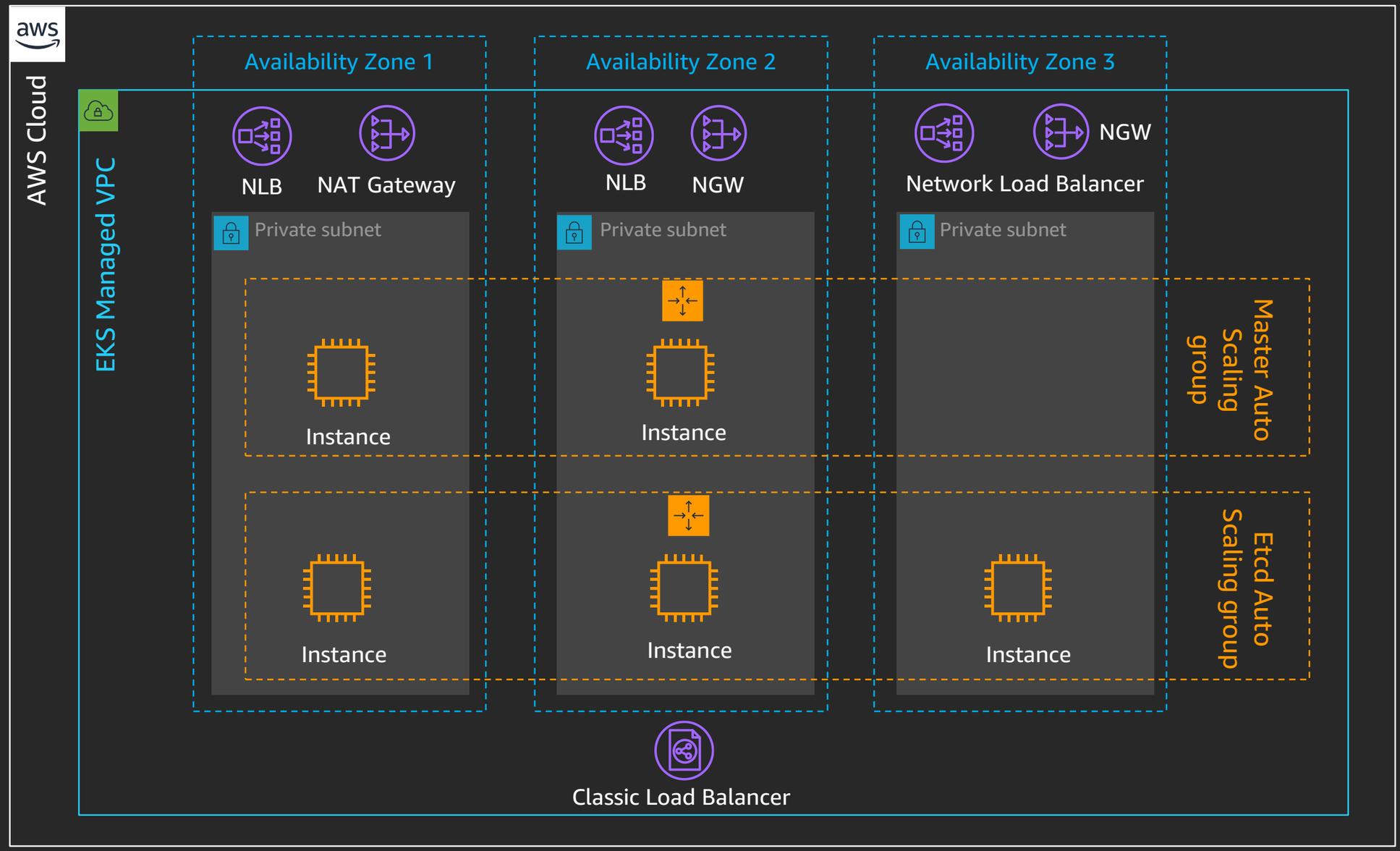


Amazon EKS Cluster



Amazon EKS

Single Tenant EKS Control Plane



EKS Under the Hood

EKS Cellular Architecture

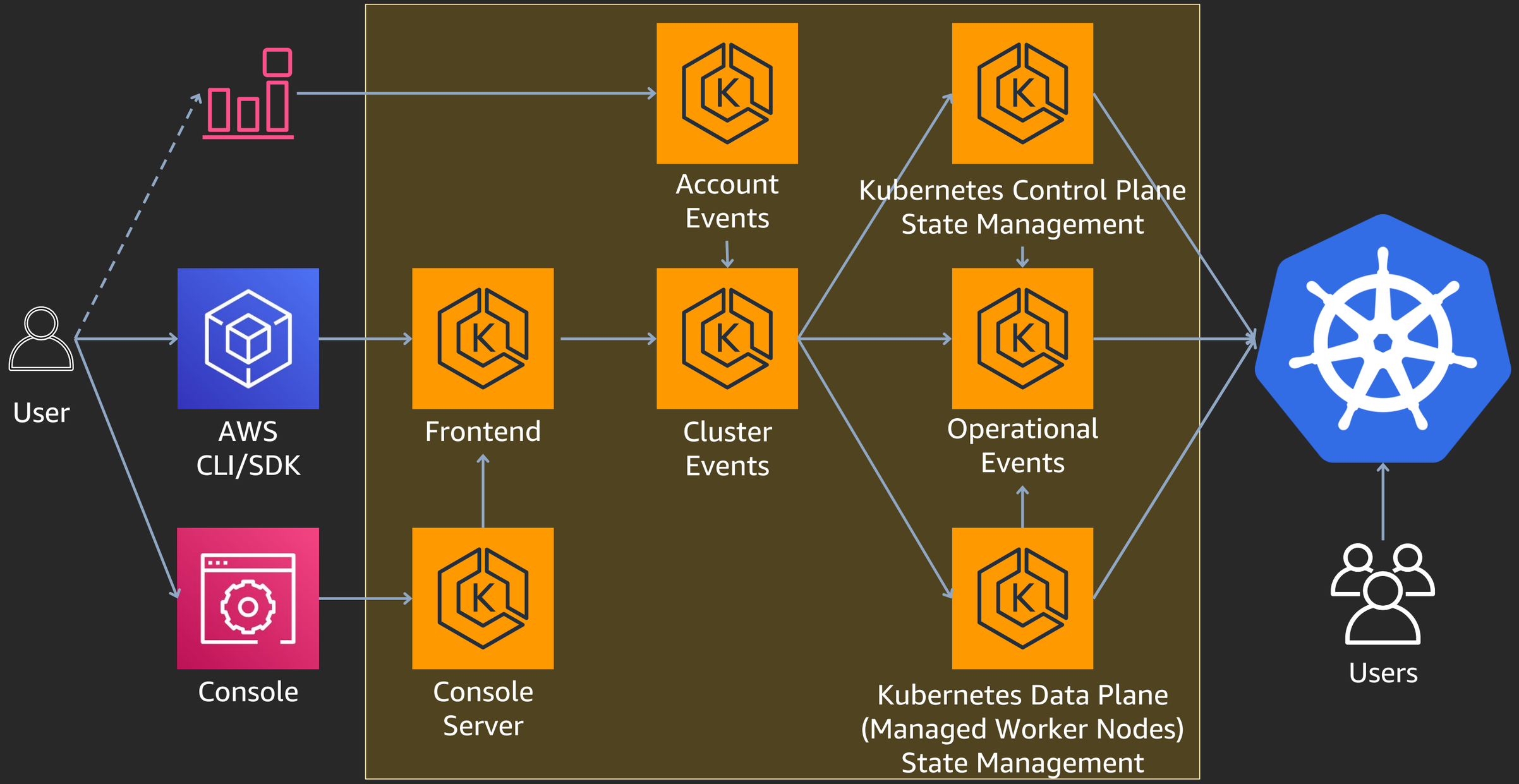
EKS Service Failure Domains – isolated failure domains designed to limit the blast radius of events

- Region - top level of isolation
  - Force majeure, hurricane, asteroid (space junk), earthquake, other significant event
- Availability Zone – subdivides region geographically
  - Localized event, natural disasters, lightning, tornado, power grid failure, civil unrest
- AWS Account – subdivides region by resource ownership
  - Security isolation, limit management, load partitioning (shard)

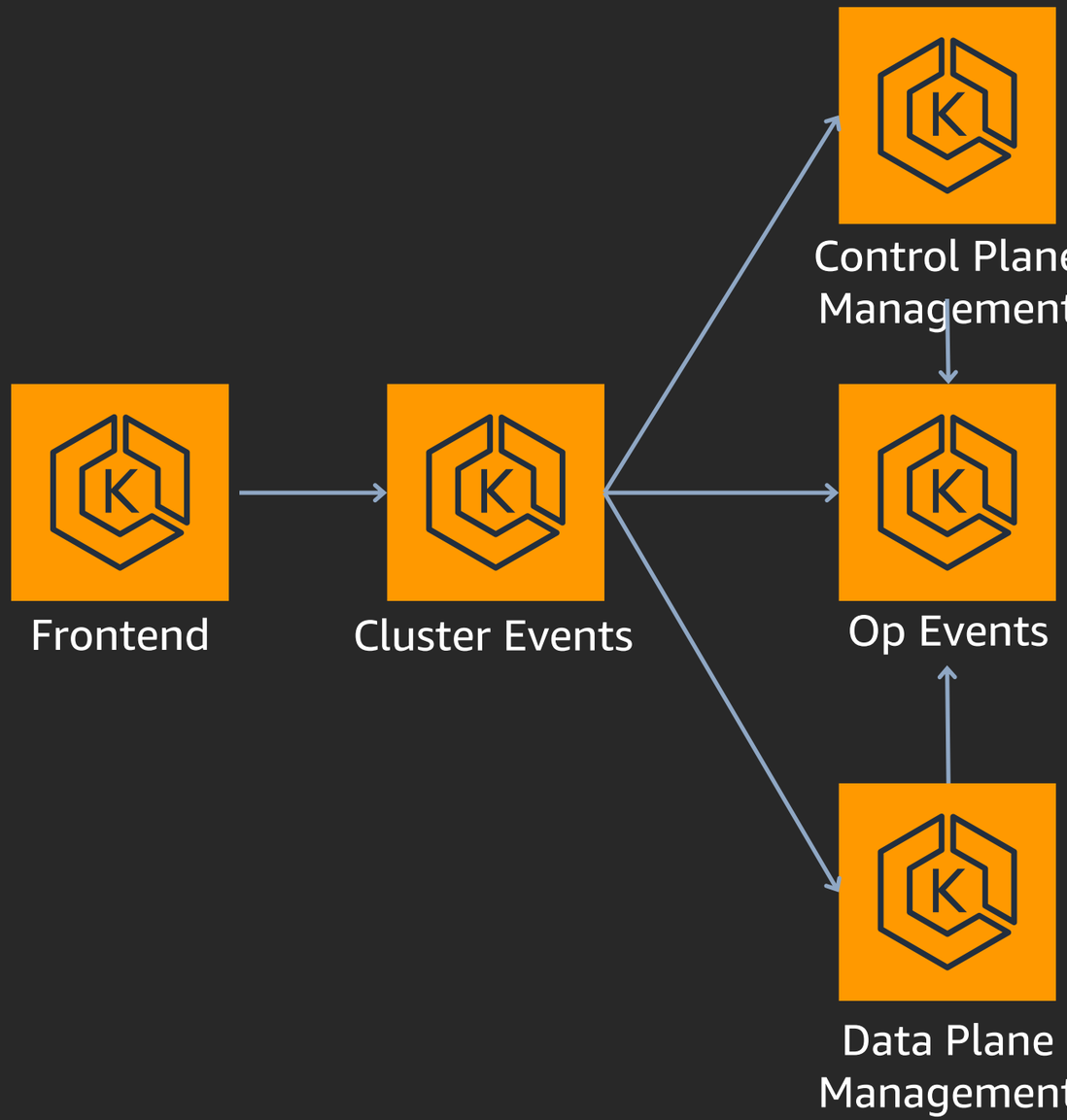
1 cell = 1 AWS account

EKS Logical Single Highly Available Service

EKS Scalable Architectural Components



EKS Scalable Architectural Components



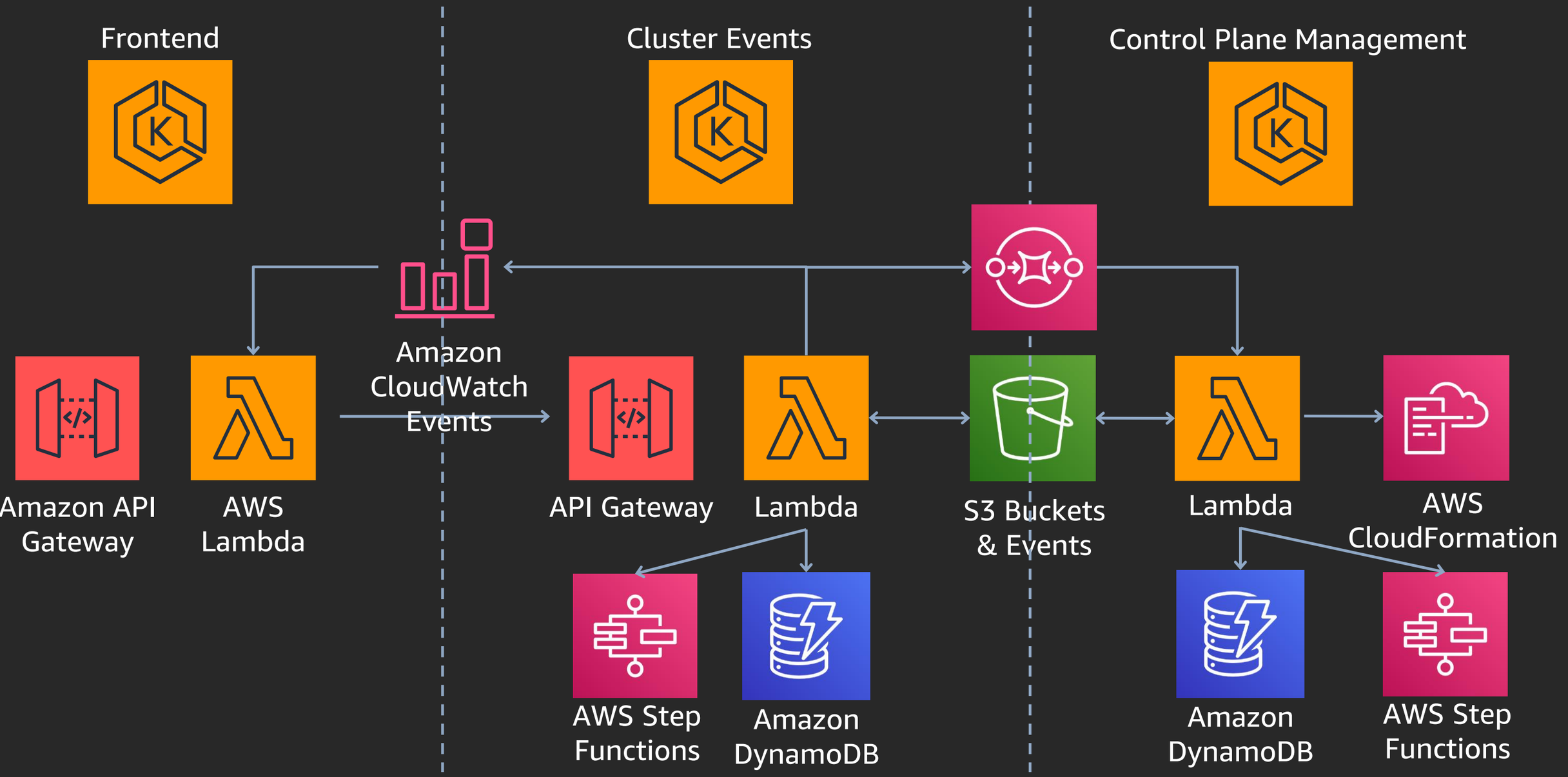
EKS Scalable Architectural Components

Frontend

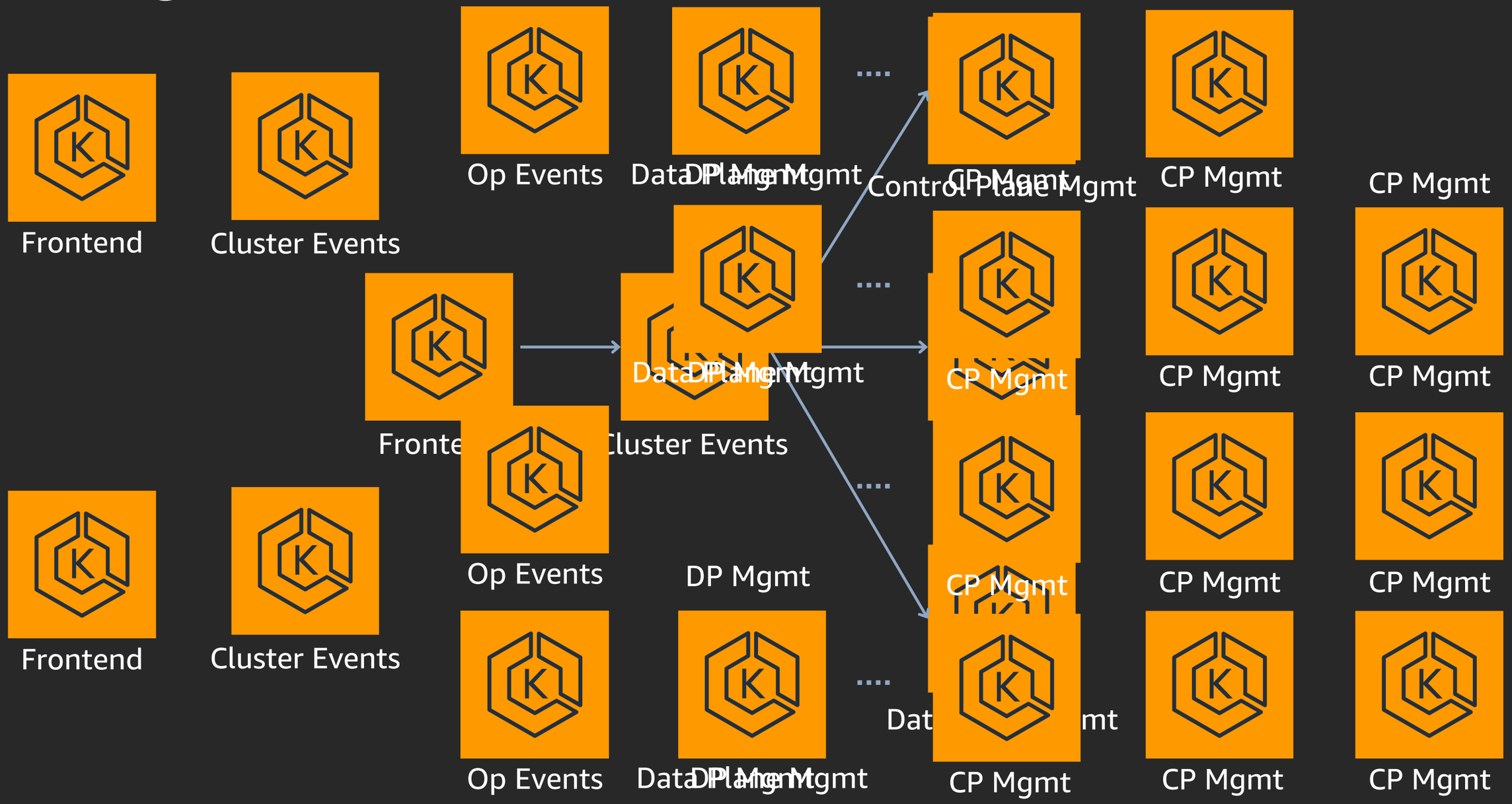
Cluster Events

Control Plane Management

EKS Scalable Architectural Components

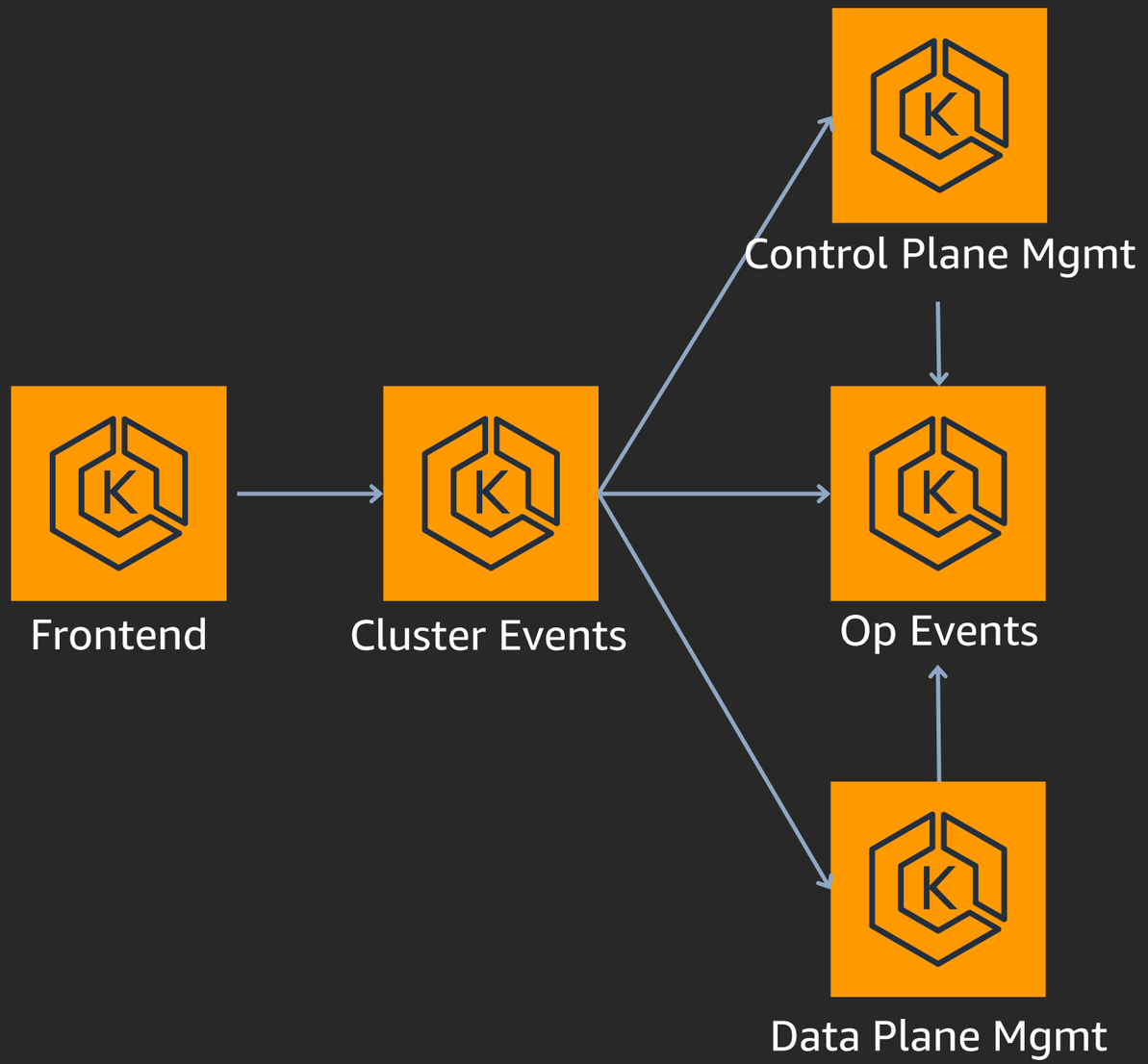


EKS Regional Cellular Architecture

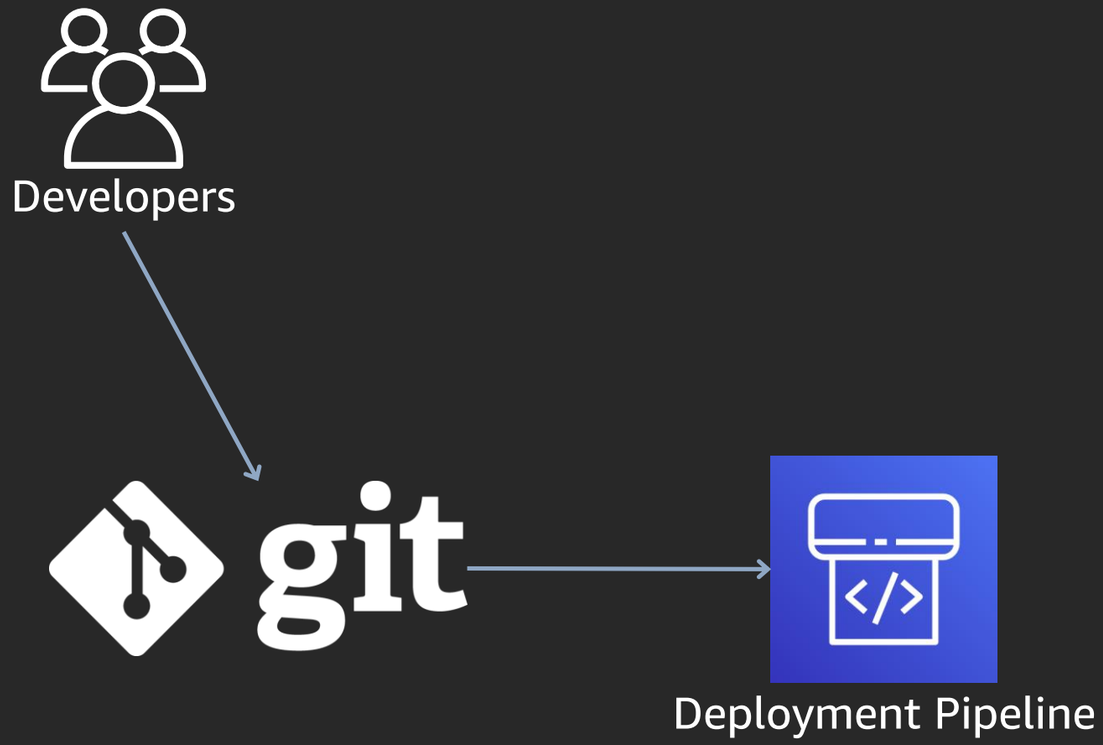


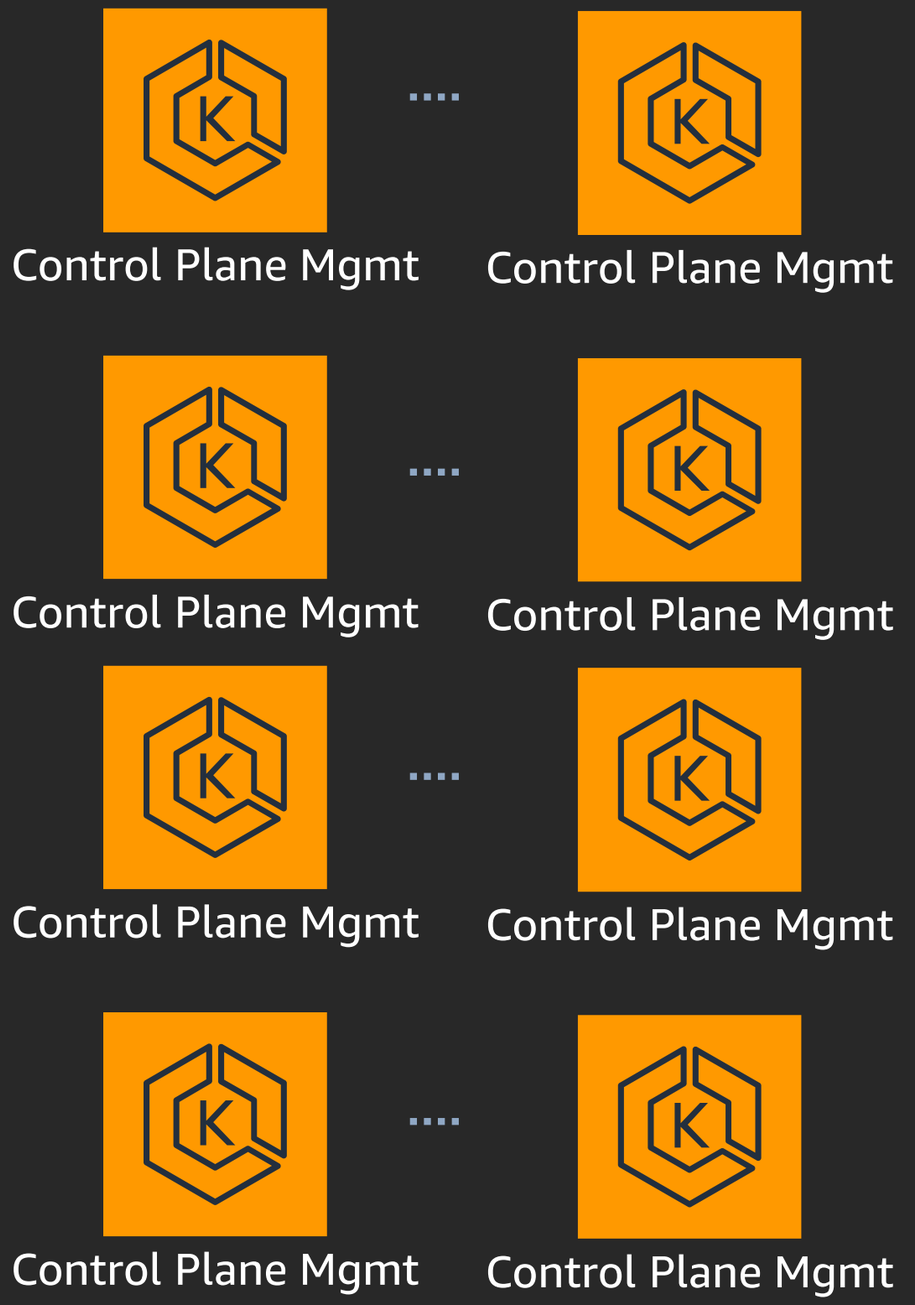
EKS Operations

EKS Region Deployment Safety

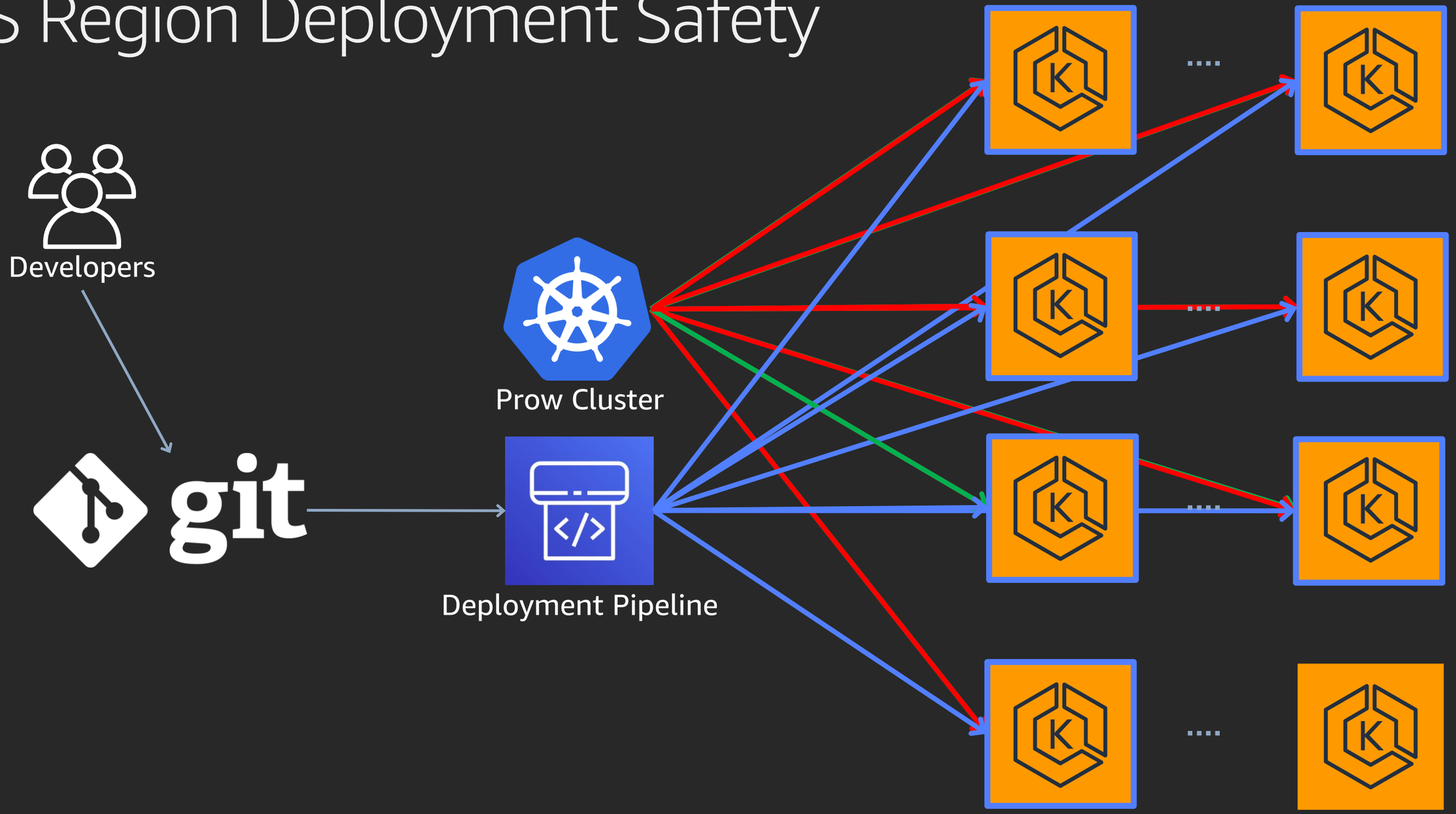


EKS Region Deployment Safety

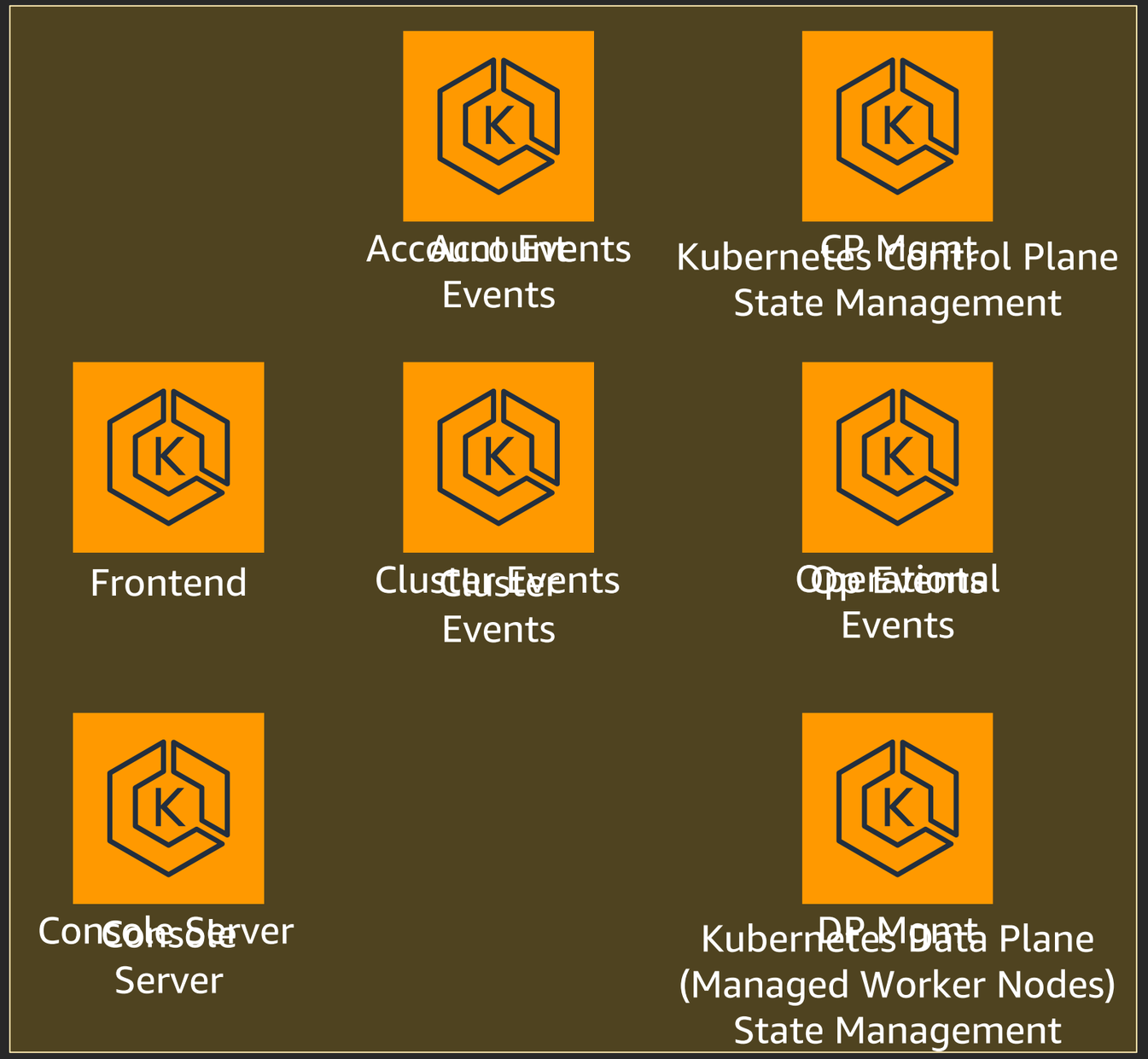




EKS Region Deployment Safety



EKS Components



EKS Components

Account Events

CP Mgmt

Frontend

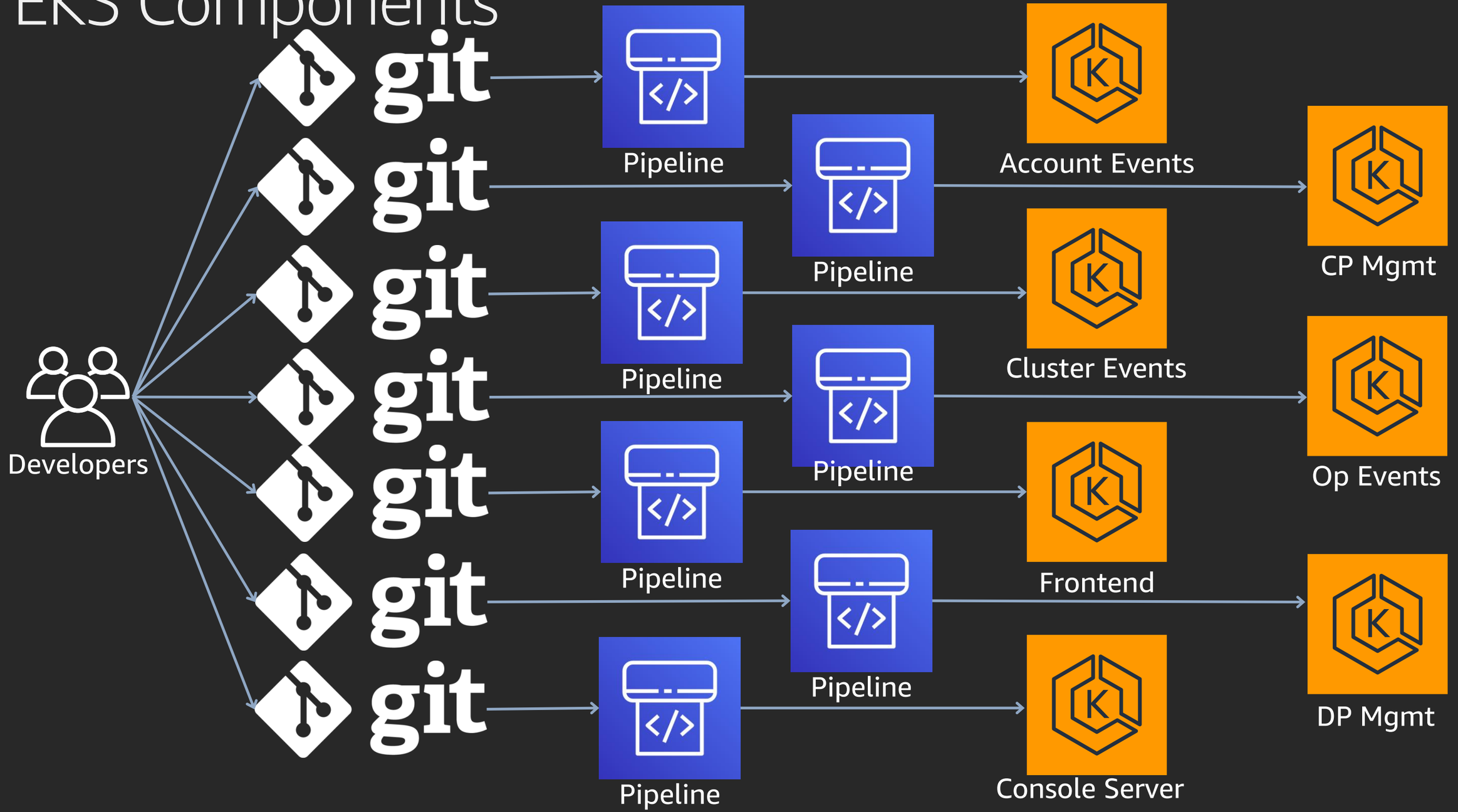
Cluster Events

Op Events

Console Server

DP Mgmt

EKS Components



EKS Enhancements: What we've been up to

AWS re:Invent

aws

The year in review

Security & Reliability
ISO, SOC 123, and PCI compliance
99.9% Service Level Agreement
Cluster creation limit raised to 50 per region
API Server Endpoint Access Control
Control Plane Logs in Amazon CloudWatch
AWS IAM authenticator integration
EKS v1.10 and 1.11 end of life
Amazon ECR PrivateLink support
Kubernetes pod security policies
AWS IAM for Service Accounts
Cluster tagging

Regions & Versions
Seoul, Mumbai, London, Paris, Ohio, Frankfurt, Singapore, Sydney, Tokyo, Hong Kong, São Paulo, Bahrain
Support for Kubernetes versions 1.11, 1.12, 1.13, and 1.14

Nodes
Windows Node Support (GA)
A1 (ARM) instance support (preview)
EKS-Optimized AMI AWS Systems Manager parameters

Storage & Networking
Alpha CSI Driver for Amazon FSx for Lustre
Beta CSI Drivers for Amazon EBS and Amazon EFS
Support for Public IP Addresses Within Cluster VPCs
AWS ALB Ingress Controller
Amazon VPC CNI plugin v1.3, 1.4, 1.5

Tooling
AWS App Mesh controller
Managed Cluster Version Updates
CloudWatch Container Insights
eksctl as the official EKS CLI
AWS Node Termination Handler
Mixed instance policy support and GPU-provider for Cluster Autoscaler

Machine Learning
Deep Learning Benchmark Utility
AWS in official Kubeflow documentation
Support for P3dn and G4dn instances
Escalator autoscaler one-click capacity

All since re:Invent 2018

The year in review

Security & Reliability
ISO, SOC 123, and PCI compliance
99.9% Service Level Agreement
Cluster creation limit raised to 50 per region
API Server Endpoint Access Control
Control Plane Logs in Amazon CloudWatch
AWS IAM authenticator integration
EKS v1.10 and 1.11 end of life
Amazon ECR PrivateLink support
Kubernetes pod security policies
AWS IAM for Service Accounts
Cluster tagging

Regions & Versions
Seoul, Mumbai, London, Paris, Ohio, Frankfurt, Singapore, Sydney, Tokyo, Hong Kong, São Paulo, Bahrain
Support for Kubernetes versions 1.11, 1.12, 1.13, and 1.14

Nodes
Windows Node Support (GA)
Managed Node Groups
A1 (ARM) instance support (preview)
EKS-Optimized AMI AWS Systems Manager parameters

Storage & Networking
Alpha CSI Driver for Amazon FSx for Lustre
Beta CSI Drivers for Amazon EBS and Amazon EFS
Support for Public IP Addresses Within Cluster VPCs
AWS ALB Ingress Controller
Amazon VPC CNI plugin v1.3, 1.4, 1.5

Tooling
AWS App Mesh controller
Managed Cluster Version Updates
CloudWatch Container Insights
eksctl as the official EKS CLI
AWS Node Termination Handler
Mixed instance policy support and GPU-provider for Cluster Autoscaler

Machine Learning
Deep Learning Benchmark Utility
AWS in official Kubeflow documentation
Support for P3dn and G4dn instances
Escalator autoscaler one-click capacity

All since re:Invent 2018

AWS IAM Roles for Service Accounts

Secure

- IAM policy restrictions can restrict roles to Service Accounts or Namespaces
- Enables isolated AWS permissions per Service Account
- Credentials are automatically rotated
- The cluster's signing key is automatically rotated

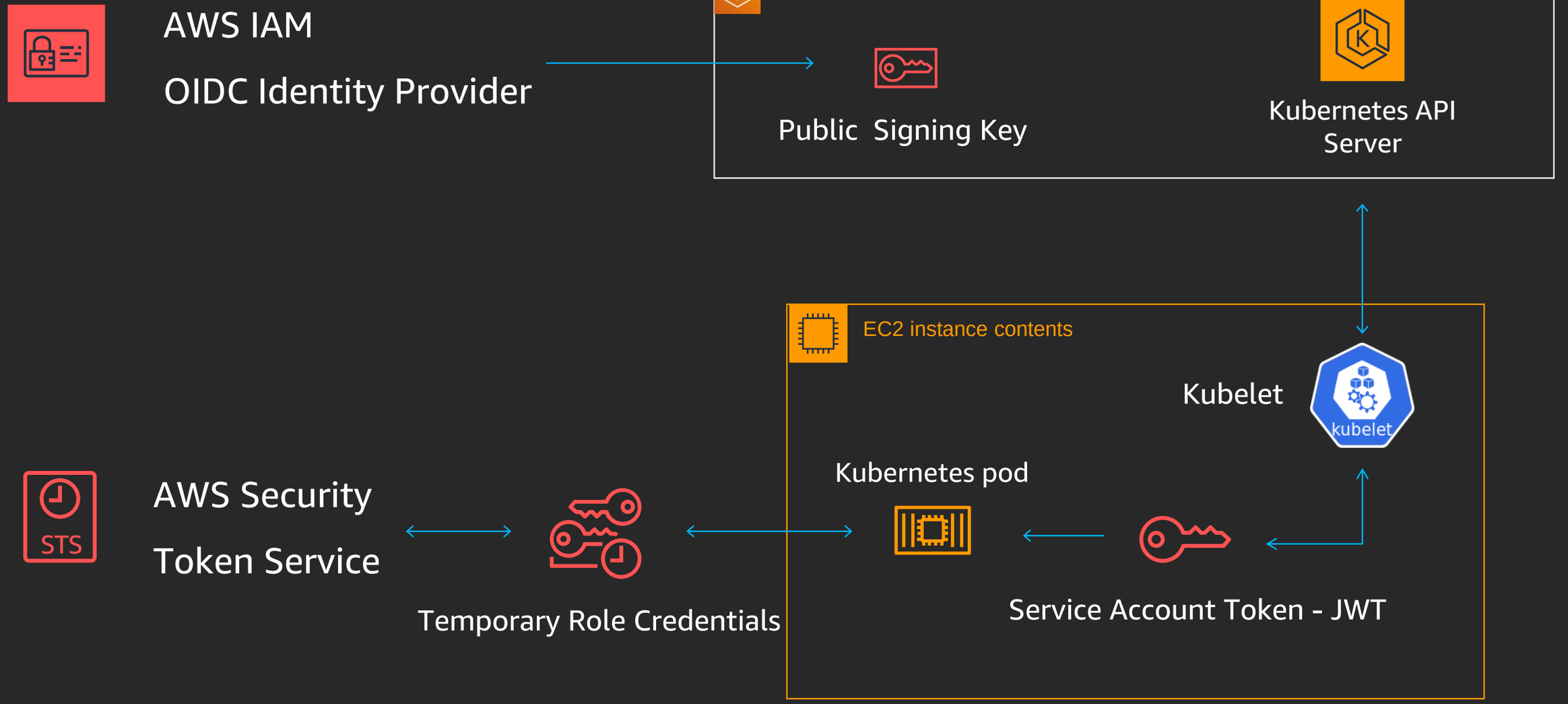
Easy Integration

- Annotate the Service Account
- Built into the default credential chains in the AWS SDKs and CLI

Auditable

- Service Account names are logged in AWS CloudTrail

AWS IAM Roles for Service Accounts



p0

Security

p1

Reliability

Investments in security and reliability

- Cellular Architecture
- Version qualification and release
- Security Patching
- Operations tooling

EKS Enhancements: Things you're gonna love

AWS re:Invent

aws

AWS Fargate for Amazon EKS

Fargate is a serverless compute platform for containers on AWS

The differences between using EKS and ECS with Fargate are driven by the orchestration system.

AWS Fargate for Amazon EKS

Bring existing pods

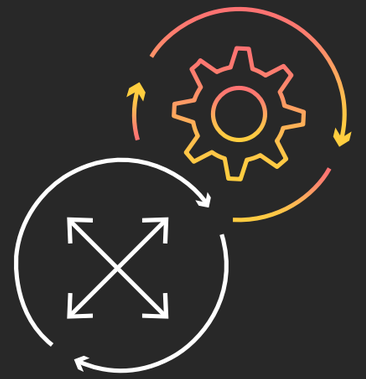
You don’t need to change your existing pods.

Fargate works with existing workflows and services that run on Kubernetes.

Production Ready

Launch pods quickly. Easily run pods across multiple AZs for high availability.

Each pod runs in an isolated VM compute environment.

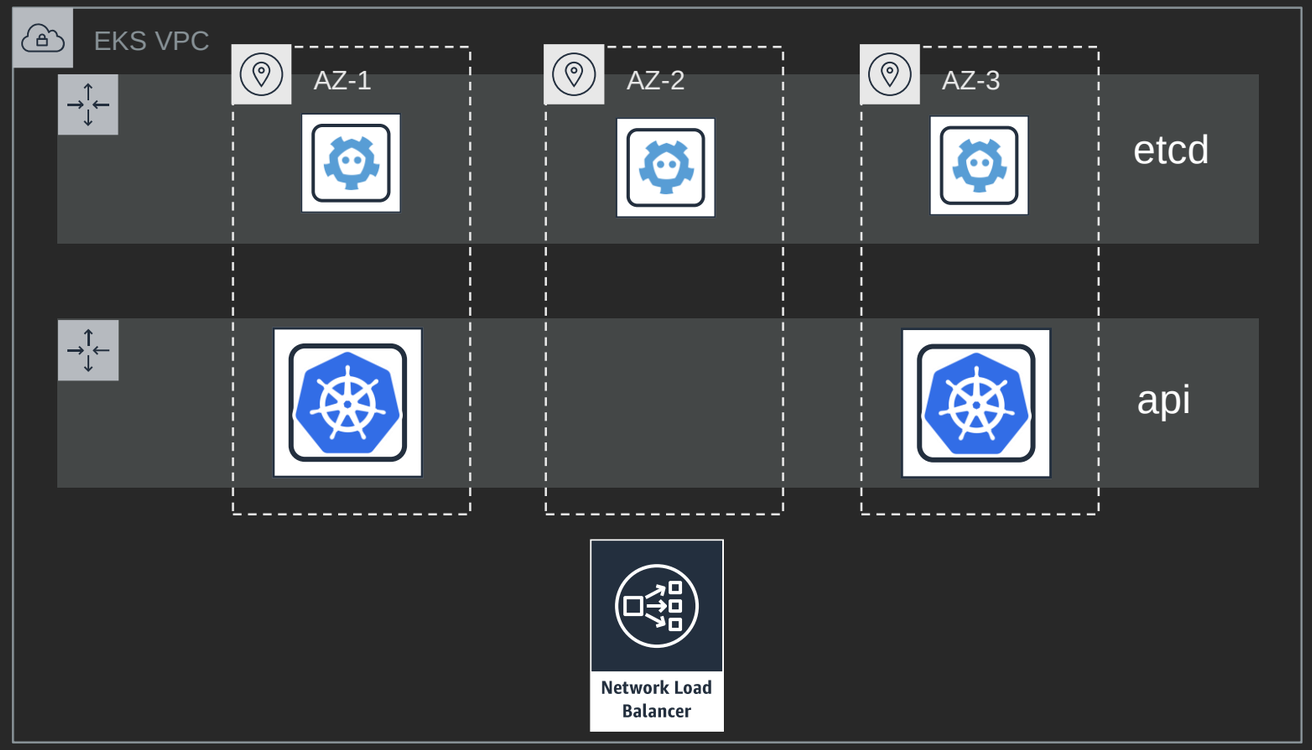


Right-Sized and Integrated

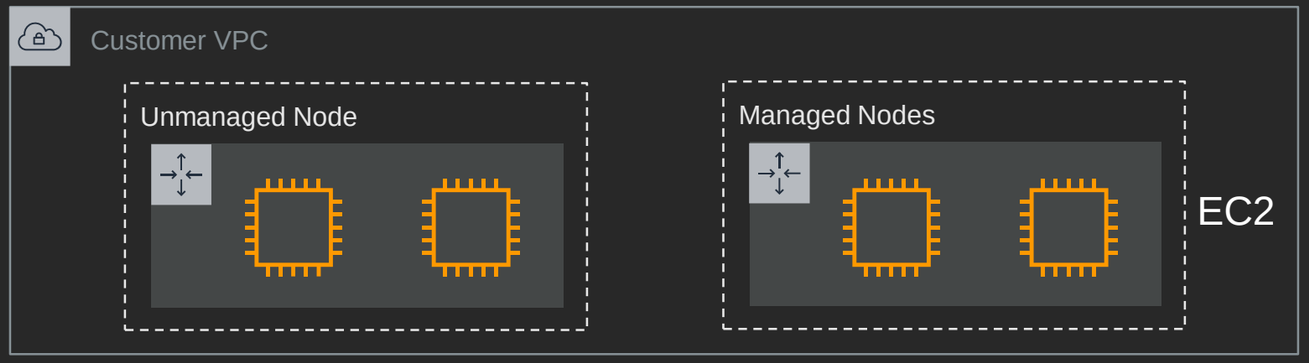
Only pay for the resources you need to run your pods.

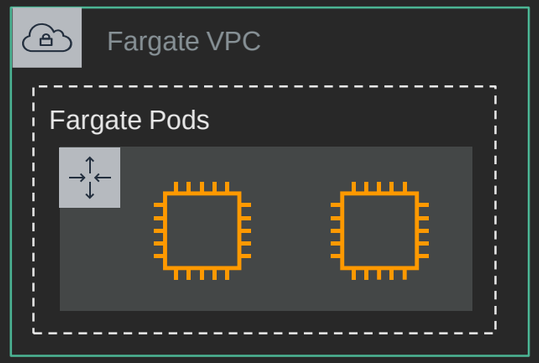
Includes native AWS integrations for networking and security.

EKS Cluster Architecture



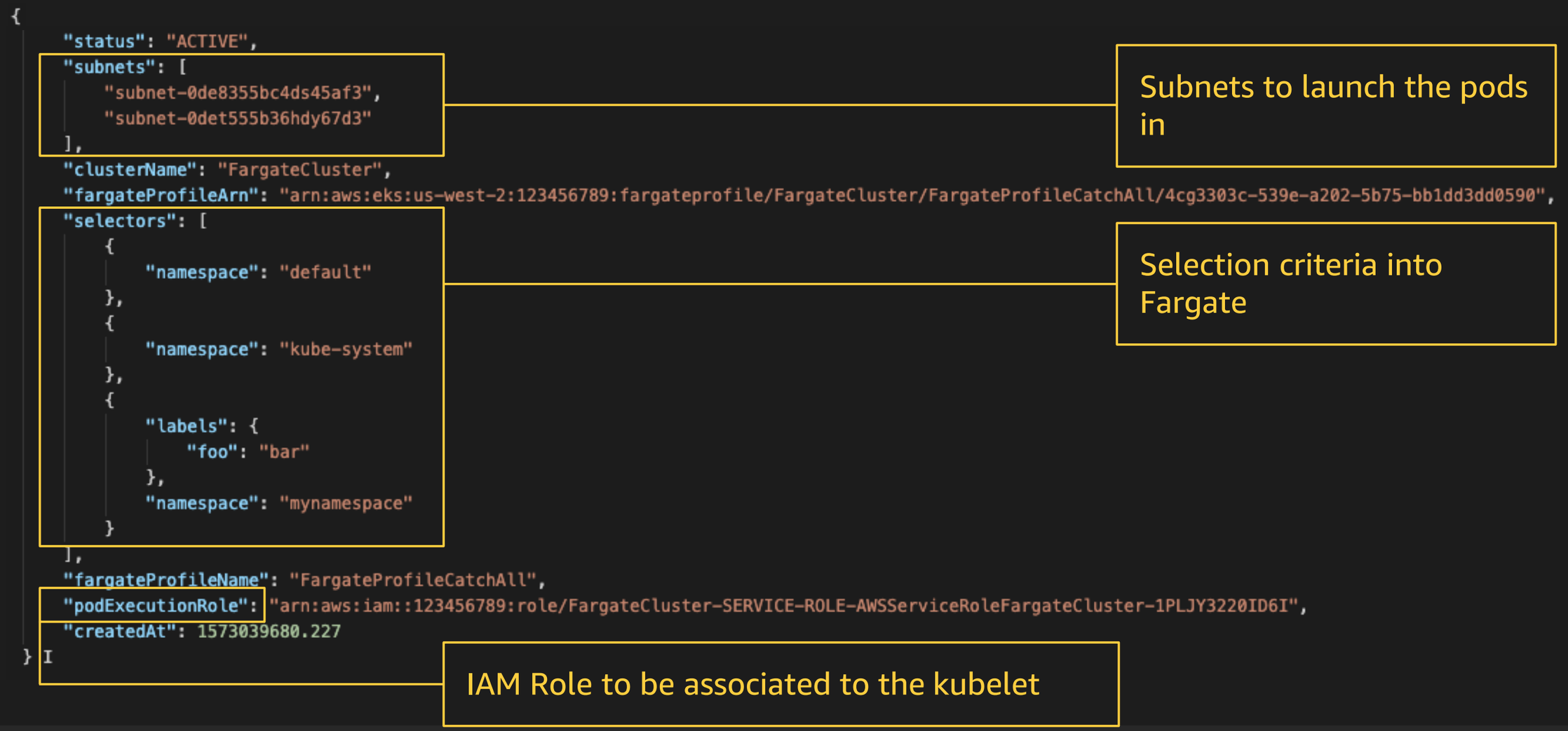
EKS Managed Control Plane





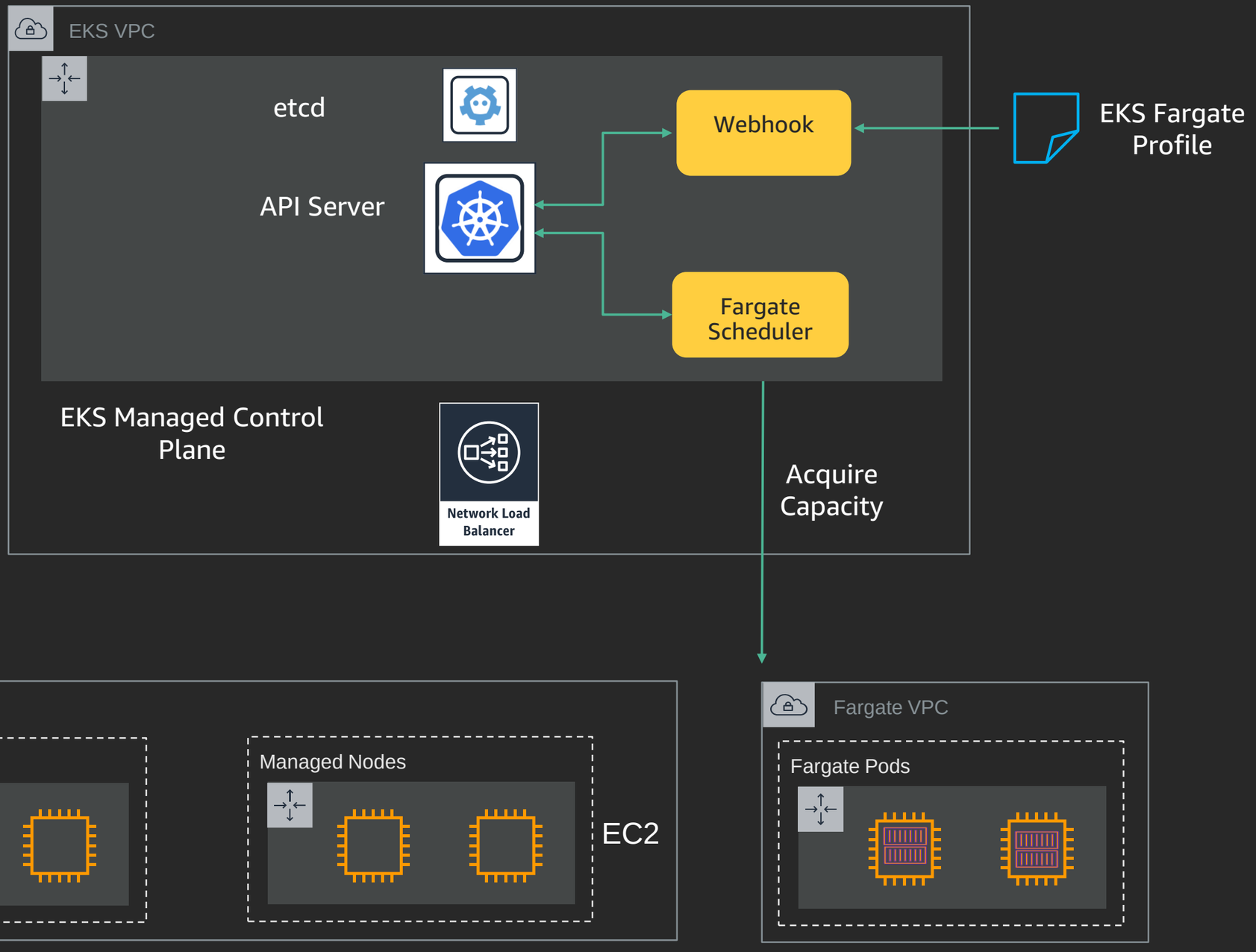
EKS Data Plane

EKS Fargate profile template

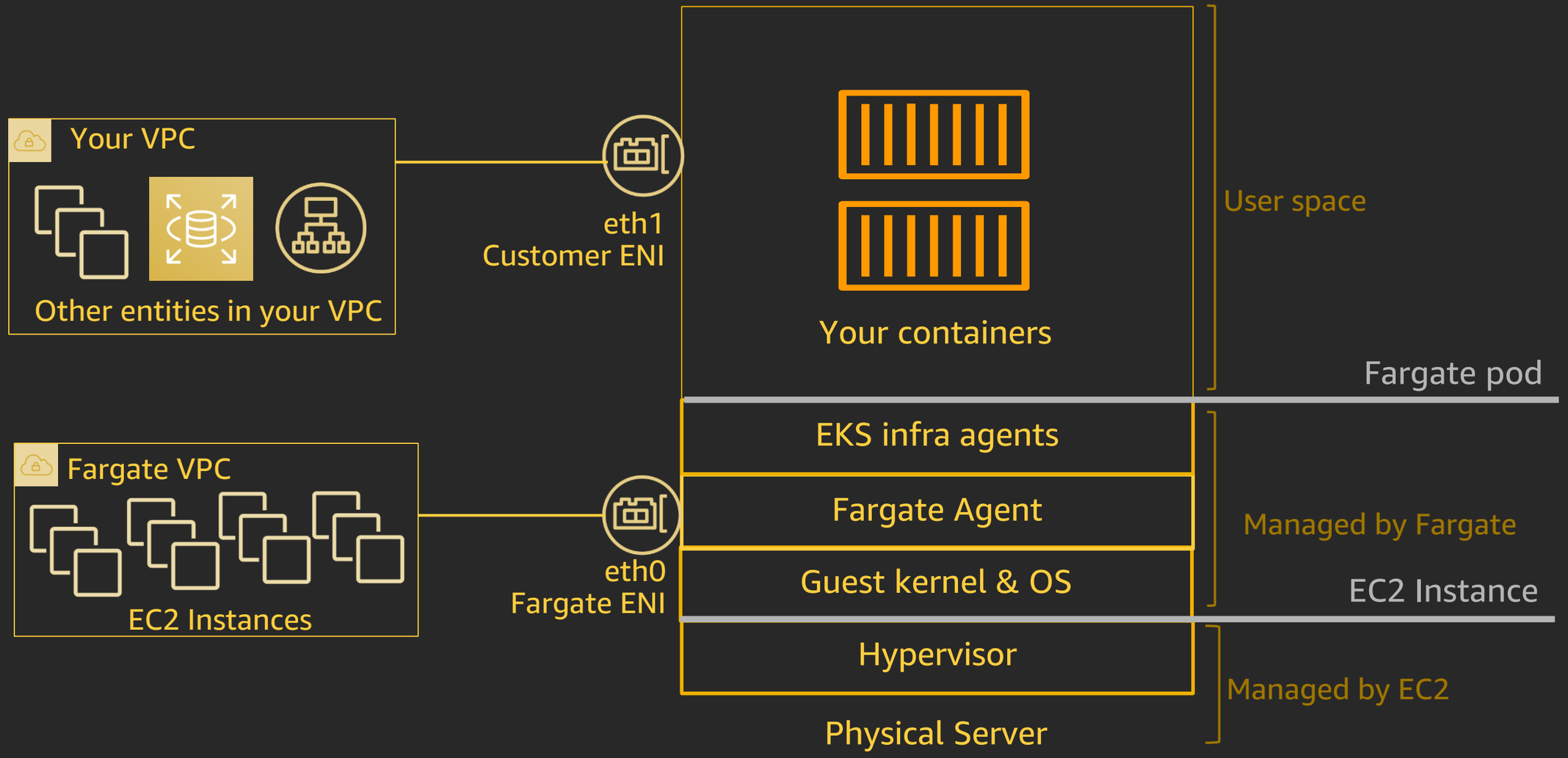


EKS Fargate flow at 33,000 feet

EKS Fargate Architecture



EKS Fargate Data Plane



Recap: EKS Fargate UX changes

Things you no longer need to do

- Manage Kubernetes worker nodes
- Pay for unused capacity
- Use K8s Cluster Autoscaler (CA)

Things you get out of the box

- VM isolation at pod level
- Pod level billing
- Easy chargeback in multi tenant scenarios

Things you can't do (for now)

- Deploy Daemonsets
- Use service type LoadBalancer (CLB/NLB)
- Running privileged containers
- Run stateful workloads

EKS Fargate Availability

Available today for all new 1.14 clusters

- Create a new cluster
- Update a 1.13 cluster to 1.14

Use EKS with Fargate in

- Virginia (us-east-1)
- Ohio (us-east-2)
- Dublin (eu-west-1)
- Tokyo (ap-northeast-1)

EKS Enhancements: What's Next?

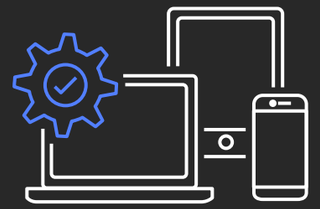
AWS re:Invent

aws

Our vision for EKS

Globally available

Easy to use



Production ready

Cost-effective

High-performance


All the building blocks of Kubernetes
in one place

Snap Service Mesh on EKS

AWS re:Invent

aws

Snap service mesh …
Infrastructure layer providing foundation for SOA enables core capabilities by default at the platform level

- Security by default
- Standardized traffic management and routing policies:
  - Service discovery—Just call `<service>.snap`
  - Zonal affinity and regional proximity to favor closest endpoints
  - Traffic splitting, mirroring, and failover
  - Automatic resilience and circuit breakers
- Observability by default

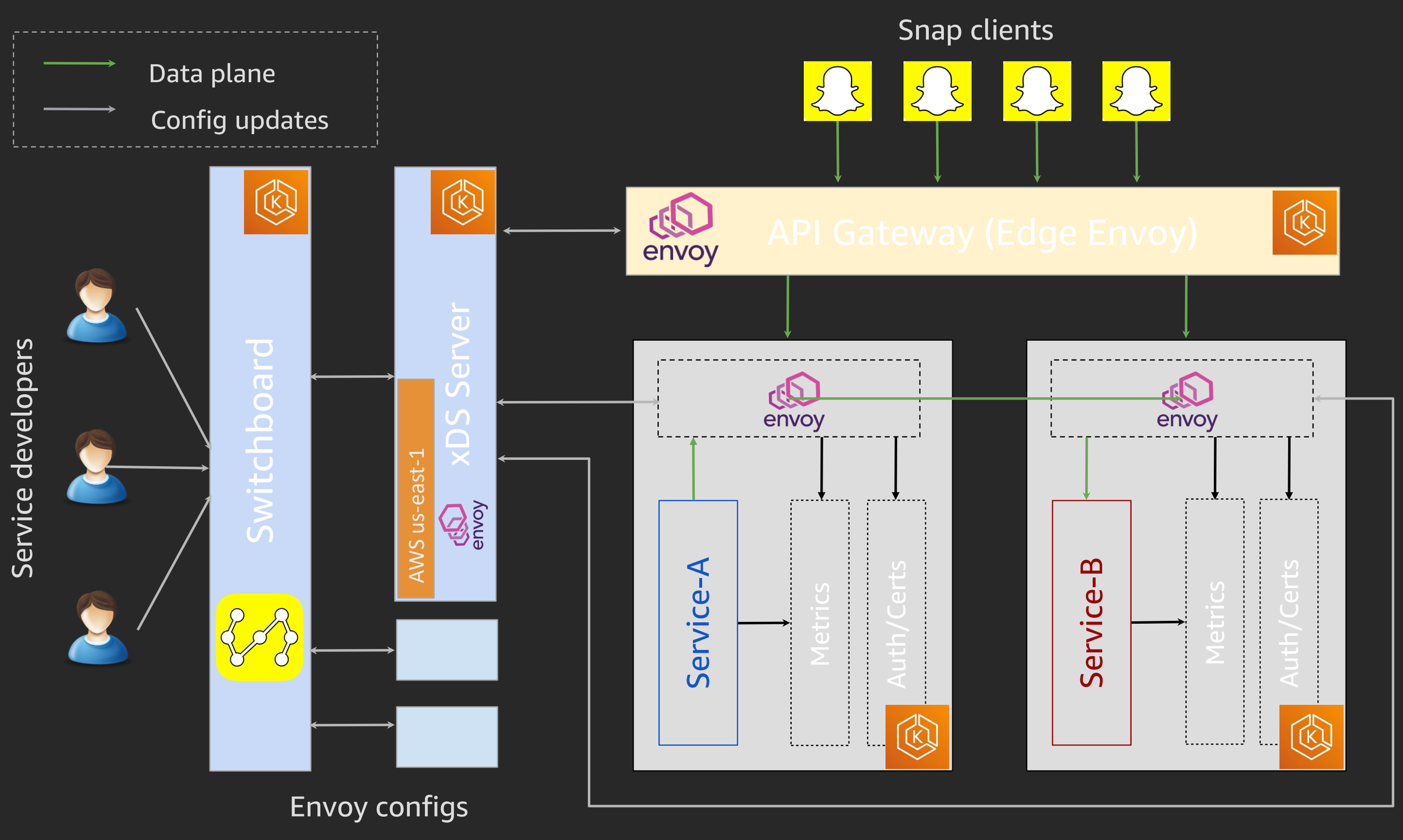
Standardizing service infrastructure across clouds

Amazon EKS: Compute, application, and sidecar management

Envoy Data plane operations:
Load-balancing, traffic routing, observability, and security controls

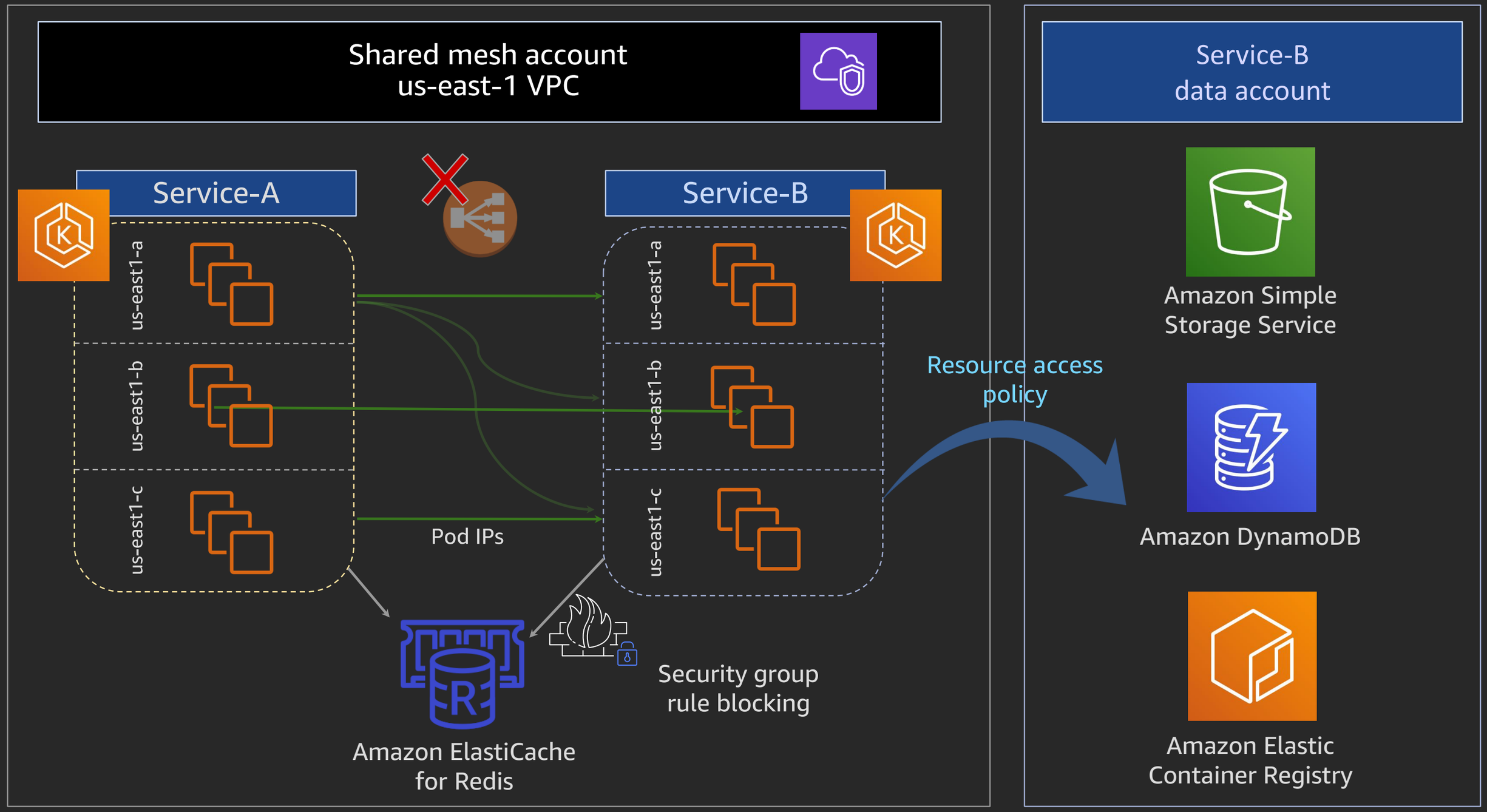
Switchboard:
In-house control plane for managing services, routes, and security policies

Spinnaker: Deployment orchestration and safe rollouts


Data plane
Config updates
Snap clients
envoy
API Gateway (Edge Envoy)
Service developers
Switchboard
xDS Server
AWS us-east-1
envoy
envoy
envoy
Service-A
Metrics
Auth/Certs
Service-B
Metrics
Auth/Certs
Envoy configs

Architectural design choices for AWS

- Accounts
  - One shared account for compute and network
  - Service data is isolated into separate accounts
- Compute: Amazon EKS
  - One EKS cluster per group of correlated services
  - ~300 EKS clusters in 4 mesh regions (as large as ~3K nodes)
- Network: > 4M QPS in AWS Regions
  - One VPC/Region, with subnets in 3 AZs
  - Security perimeter at the edge
  - Network-level protection: Security groups, network ACLs, resource access policy


Shared mesh account
us-east-1 VPC
Service-A
Service-B
us-east1-a
us-east1-b
us-east1-c
Pod IPs
Amazon ElastiCache
for Redis
Security group
rule blocking
Resource access
policy
Service-B
data account
Amazon Simple
Storage Service
Amazon DynamoDB
Amazon Elastic
Container Registry

Tooling for common service requirements

- Resource management:
  - Automate Amazon EKS cluster provisioning, and version upgrades
  - Standardize cluster add-ons: Cluster Auto Scaler, CoreDNS, and CNI
  - Per-service AWS Identity and Access Management (IAM) roles and granular access controls

Switchboard | Services | Gateway Routes | More | Search | File an issue | Help

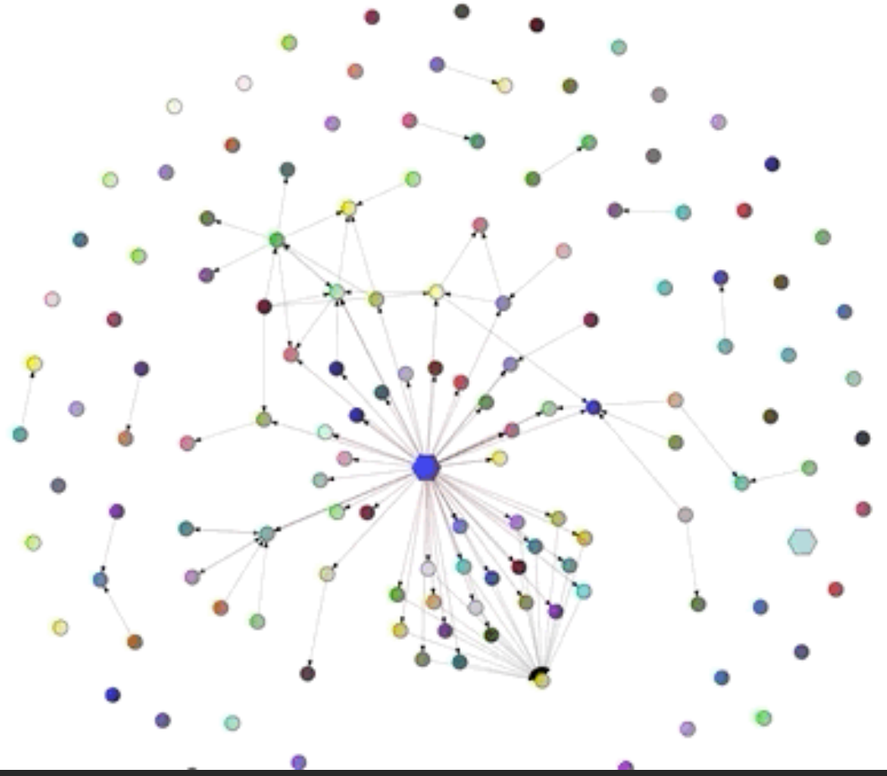
Home / All Services

All Services

New Service

Service Graph

Service Gateway



Tooling for common service requirements

- Standardize service deployments
  - Injection and upgrades of common sidecar containers
  - Default best practices through Spinnaker pipelines:
    - Uniform pod distribution per zone
    - Safe rollouts with integrated health checks

Switchboard | Services | Gateway Routes | More | Search | File an issue | Help

Home / All Services / Example

☆ Example

Edit ···

example

spinnaker.demo@snapchat.com (members) / INF

demo service for Spinnaker Summit slides

Clusters

Register New Cluster

Clusters represent a service's deployments across different stages, regions, and clouds.

Staging

Operator lease | Manage AWS IAM Roles

default.aws-us-west-2.STAGING AWS us-west-2 Security Updates Needed Edit

Dependencies

Declaring a dependency adds Example's LCA issuers to the dependency's auth whitelist. Call dependencies using `http://<service-id>.snap` URLs that route through the Envoy sidecar.

Request New Service Dependency

Consumers

Service mesh consumers can call this service via `http://example.snap`. Read more.

This service has no consumers yet.

Cloud Resources

New Kubernetes Cluster

Kubernetes clusters associated with this service.

Looking ahead—Amazon EKS features to consume

- IAM roles for service accounts:
  - Least privilege: Scope permissions at the pod level instead of worker nodes
  - Access isolation between pods
- Managed worker node groups:
  - Node draining and graceful node shutdown
  - Integrated cluster Auto-Scaling (with multi-AZ node group)
  - Simplified cluster upgrade experience
- Managed cluster add-ons
  - Metrics server
  - CoreDNS auto-scaling

More Information at re:Invent

Related breakouts

CON203 - Getting started with Kubernetes on AWS

CON205 - Deploying applications using Amazon EKS

CON206 - Management and operations for Amazon EKS

CON212 - Running Kubernetes at Amazon scale using Amazon EKS

CON306 - Building ML infrastructure on Amazon EKS with Kubeflow

CON310 - Achieving zero-downtime deployments with Amazon EKS

CON316 - Adopting CSI for stateful workloads on Amazon EKS

CON317 - Securing your Amazon EKS cluster

CON327 - Oversubscription at scale: Running tons of containers with Kubernetes

CON334 - Running high-security workloads on Amazon EKS

CON411 - Advanced network resource management on Amazon EKS

CON413 - Move your machine learning workloads to Amazon EKS


IT’S STILL DAY ONE . . .

Thank you!

Eswar Bala
Sr. Software Development Manager
Amazon Web Services
Twitter: @bala_eswar

Richard Sostheim
Principal Engineer
Amazon Web Services

Ahmed El Baz
Software Engineer
Snap Inc

Please complete the session survey in the mobile app.