AWS 中国（宁夏）区域由西云数据运营
AWS 中国（北京）区域由光环新网运营

aws INNOVATE
ONLINE CONFERENCE

分会场三：计算

在 AWS 上运行托管的 Kubernetes 集群

张浙，AWS 高级技术客户经理

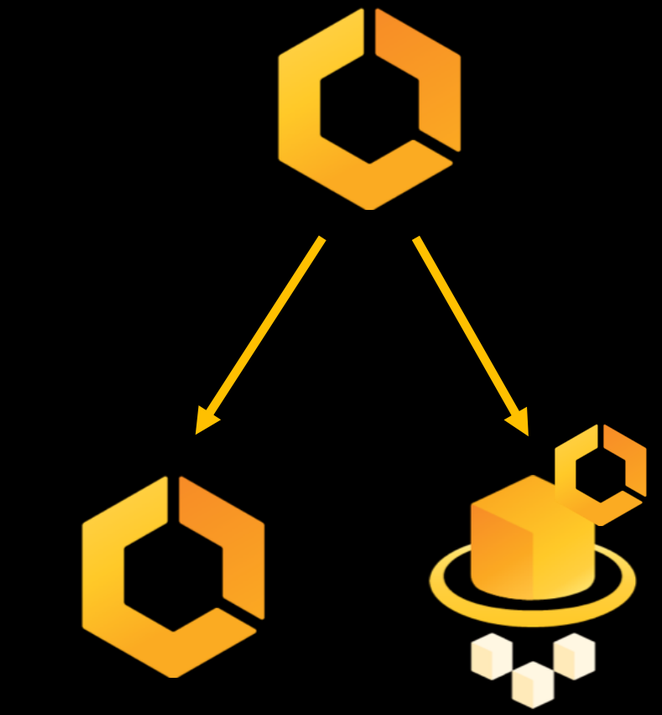
总览

AWS 中国（宁夏）区域由西云数据运营
AWS 中国（北京）区域由光环新网运营

我们为您提供选择:

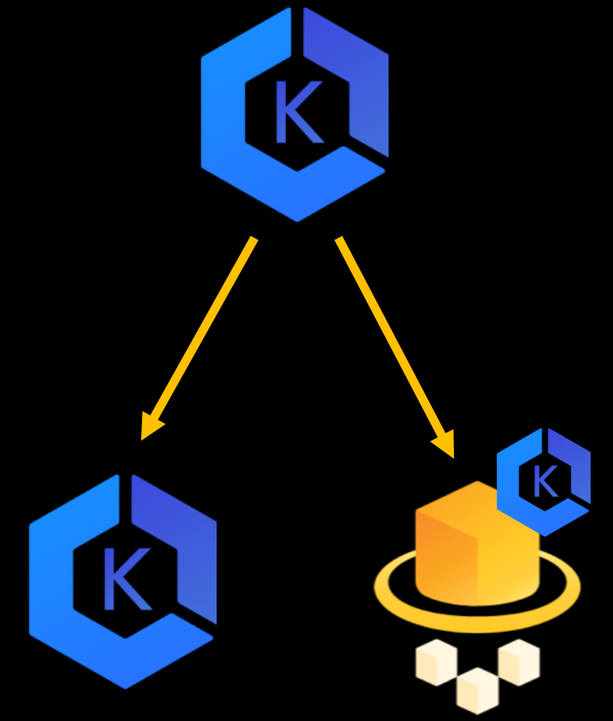
1. 选择您的排程工具
2. 选择您的运行方式

Amazon ECS



Amazon EC2　　AWS Fargate

Amazon EKS

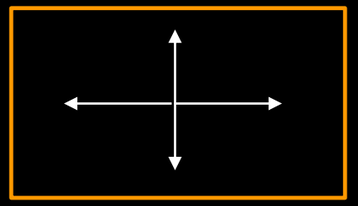


Amazon EC2　　AWS Fargate
(即将发布！)

Kubernetes 101

- 来源于 Google
- 受 Google Borg 设计思想影响
- 容器排程工具
- 托管在 CNCF 基金会

什么是 Kubernetes ?



开源的容器管理平台

应对各种规模的容器化应用

为构建现代应用架构提供基础平台

活跃的社区，大量的贡献者，未来的选择

kubernetes

AWS 中国（宁夏）区域由西云数据运营
AWS 中国（北京）区域由光环新网运营

但是，需要关注在哪里运行 Kubernetes

云平台的品质 **应用程序的品质** **使用者体验**

+

aws

51%

的 Kubernetes 工作负载当今部署在 AWS 平台上

—CNCF 调查

AWS 中国（宁夏）区域由西云数据运营
AWS 中国（北京）区域由光环新网运营

Amazon Elastic Container Service for Kubernetes

AWS 中国（宁夏）区域由西云数据运营
AWS 中国（北京）区域由光环新网运营

Amazon EKS 信条

原生的
Kubernetes

Upstream

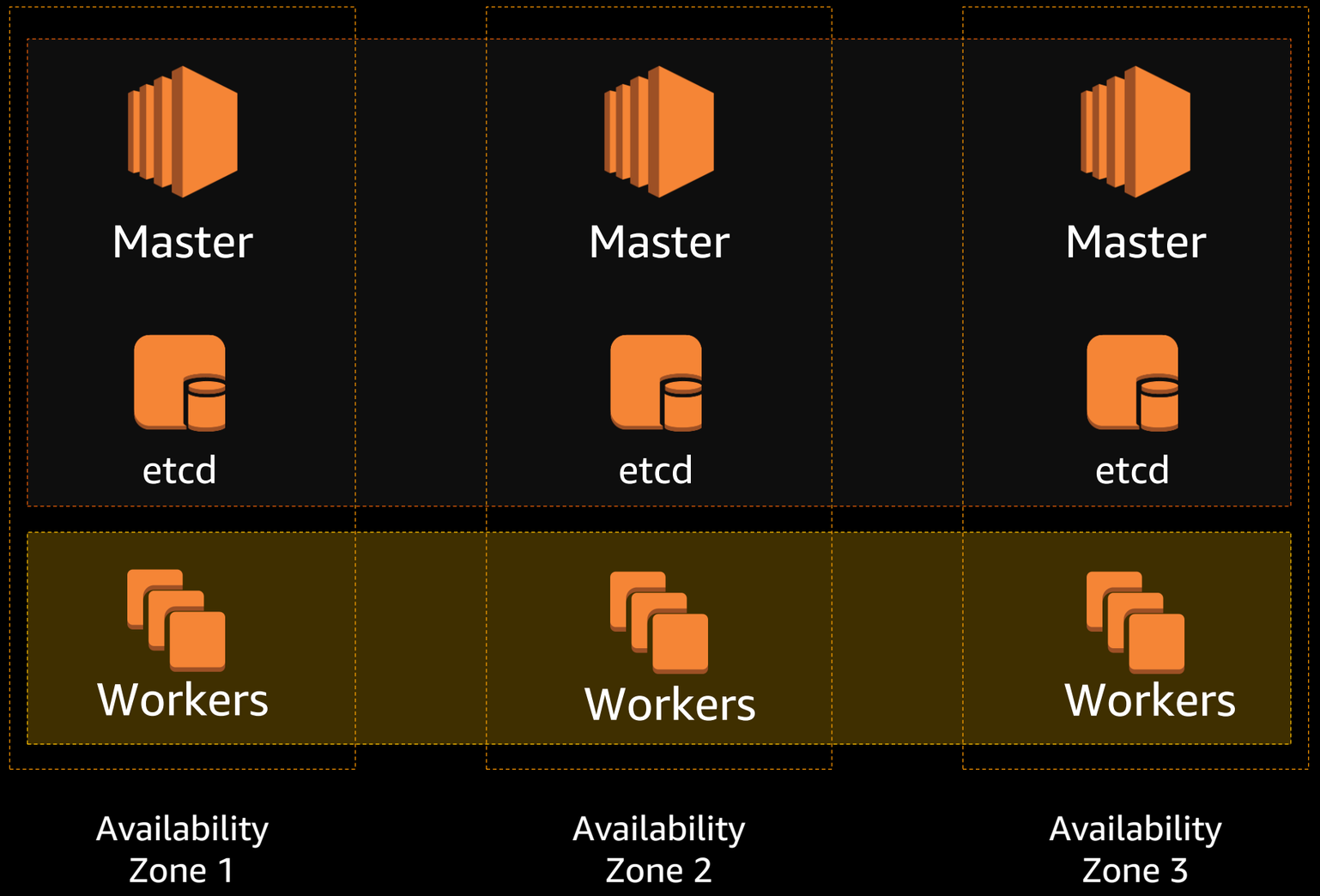
AWS 服务集成

生产环境负载

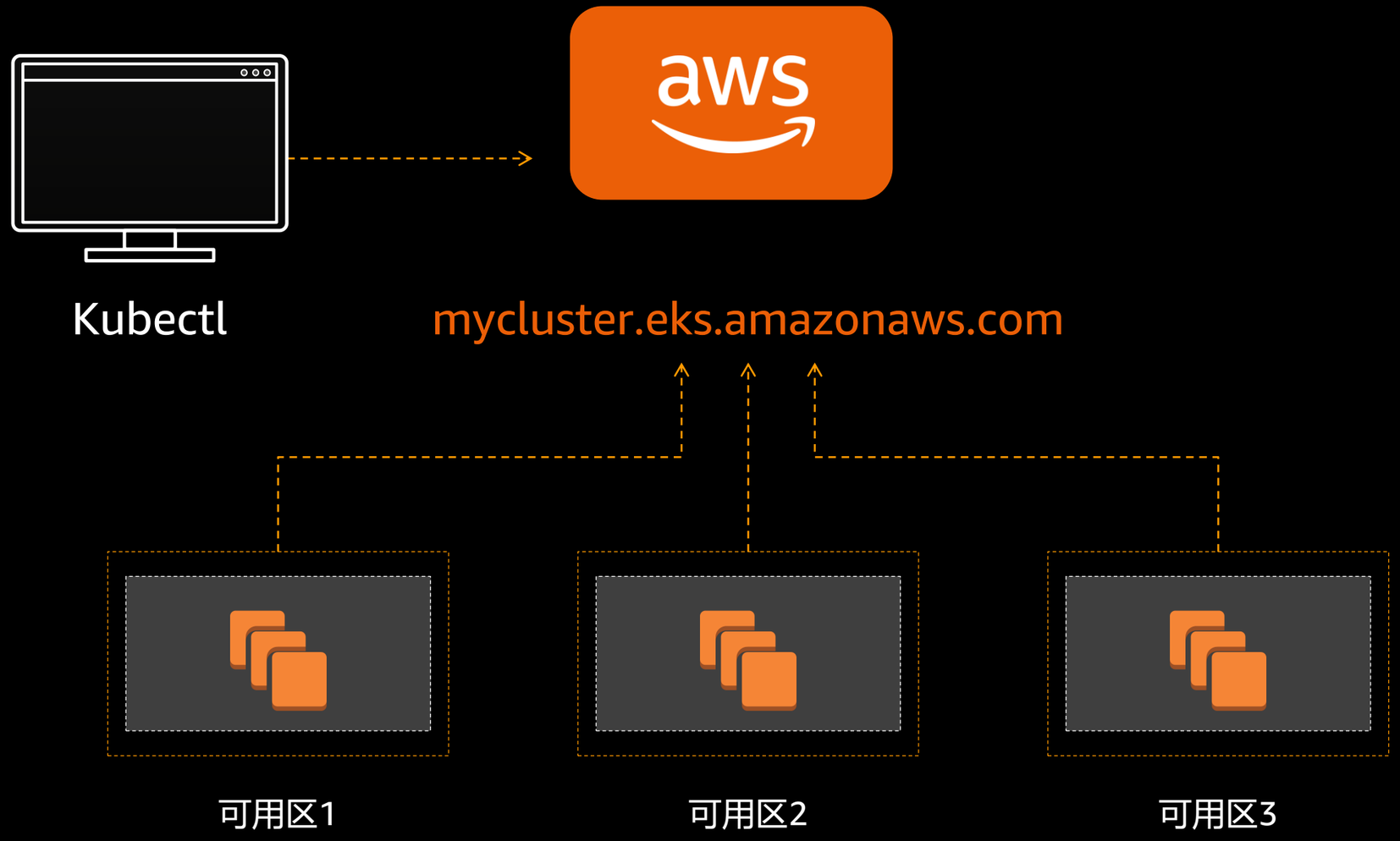
AWS 中国（宁夏）区域由西云数据运营
AWS 中国（北京）区域由光环新网运营

Amazon EKS

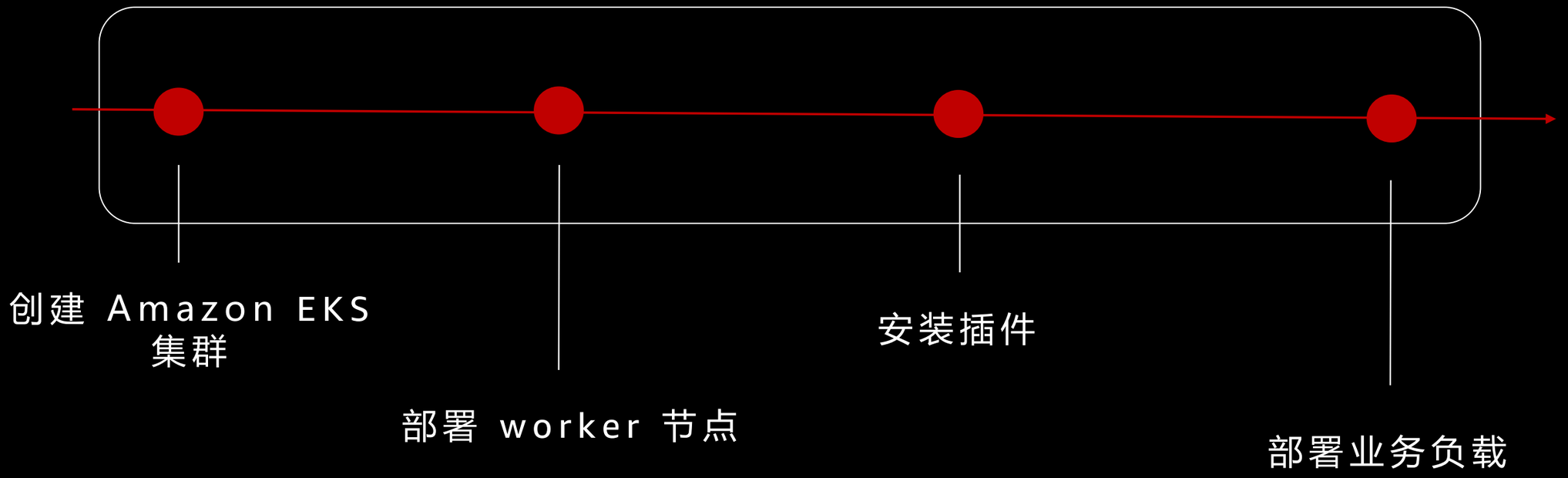
- 完全托管的 masters
- 基于高可用设计
- 软件升级及更新



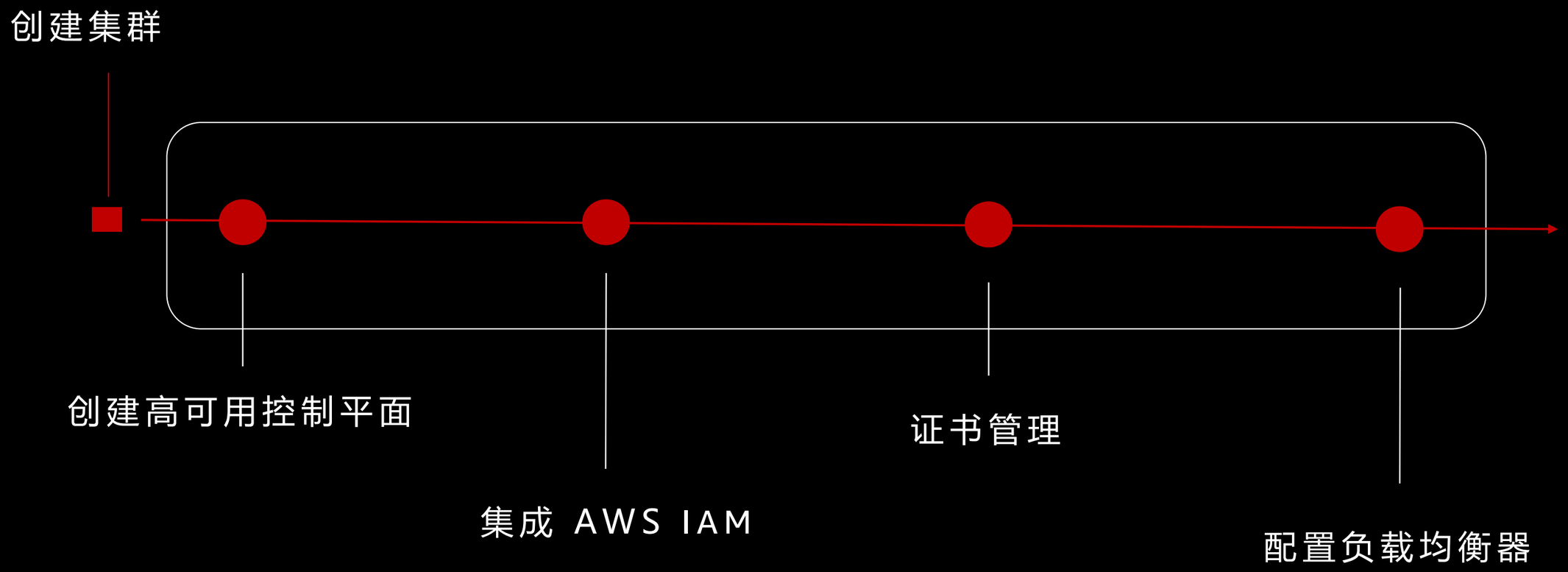
AWS 中国（宁夏）区域由西云数据运营
AWS 中国（北京）区域由光环新网运营



Amazon EKS 使用者

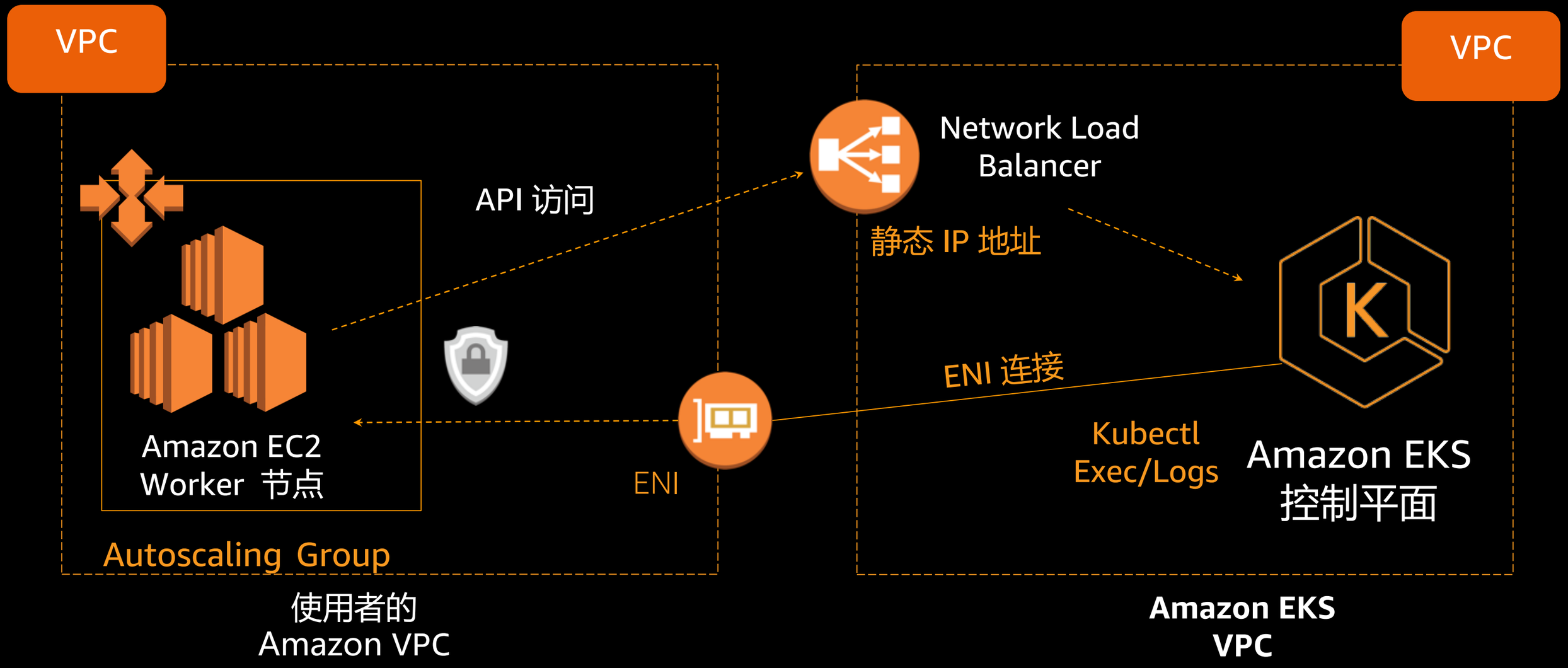


Amazon EKS – Kubernetes 控制平面

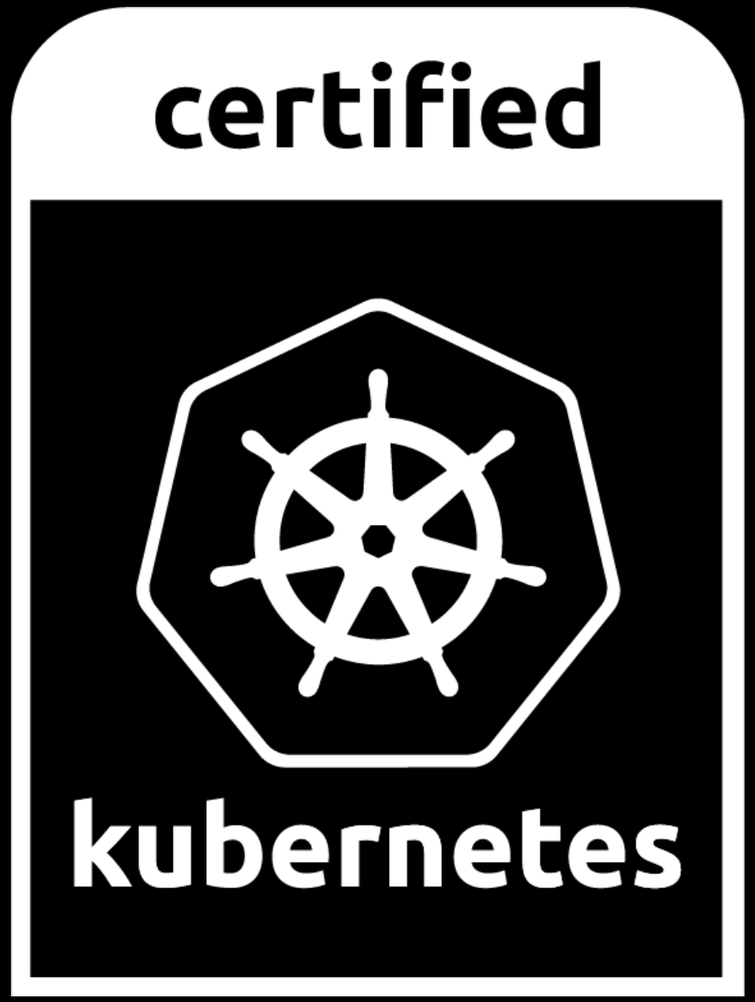


AWS 中国（宁夏）区域由西云数据运营
AWS 中国（北京）区域由光环新网运营

Amazon EKS 架构

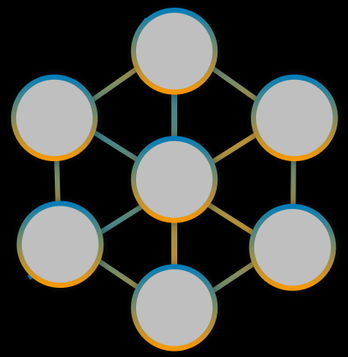


Amazon EKS 通过 Kubernetes 认证



AWS 中国（宁夏）区域由西云数据运营
AWS 中国（北京）区域由光环新网运营

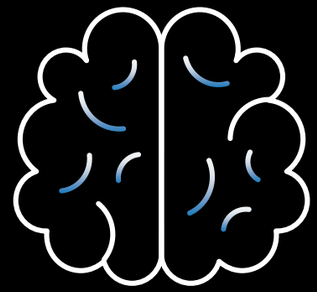
用户正在怎样使用 Amazon EKS?



微服务

平台即服务

企业应用迁移



机器学习

AWS 中国（宁夏）区域由西云数据运营
AWS 中国（北京）区域由光环新网运营

Amazon EKS 网络

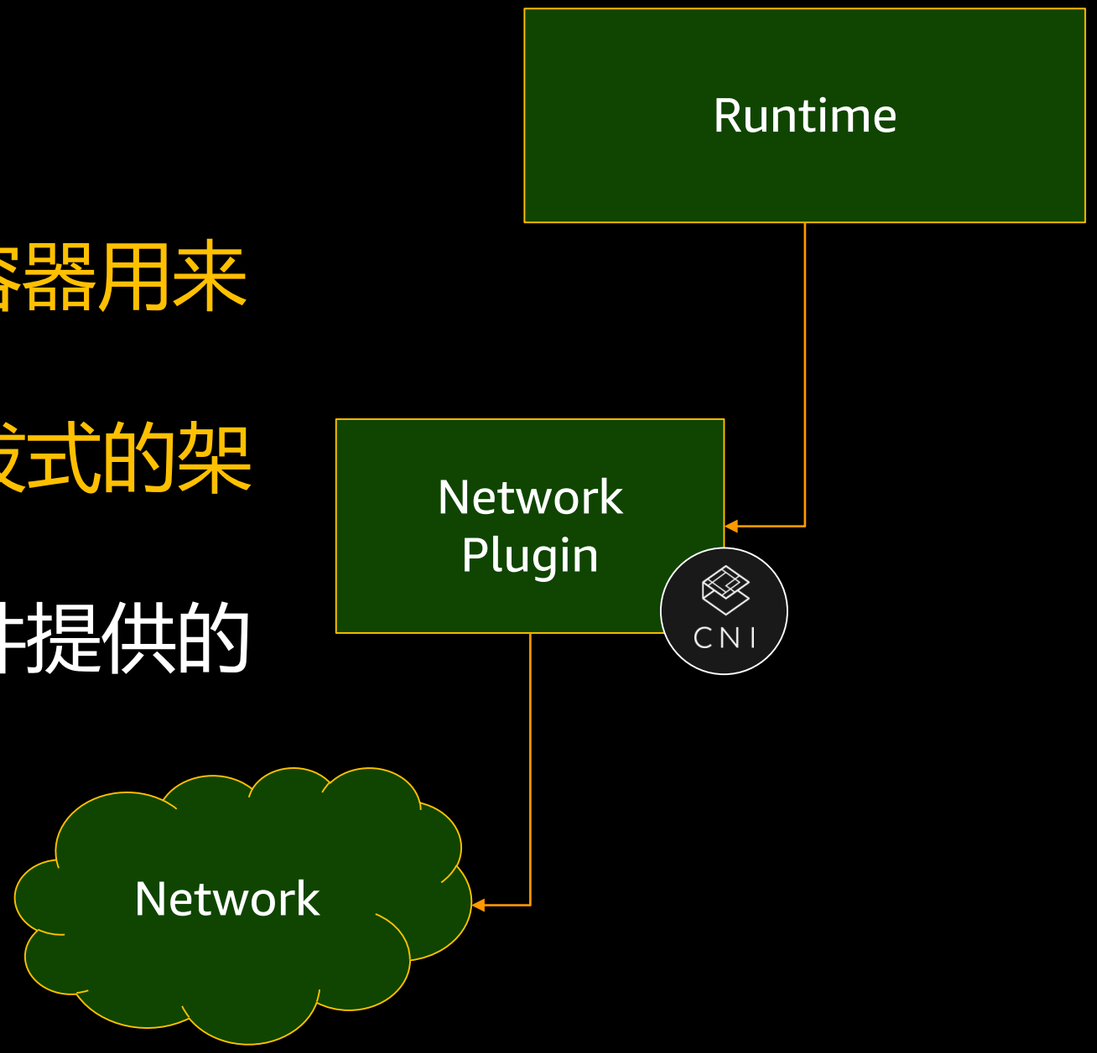
AWS 中国（宁夏）区域由西云数据运营
AWS 中国（北京）区域由光环新网运营

Kubernetes 网络

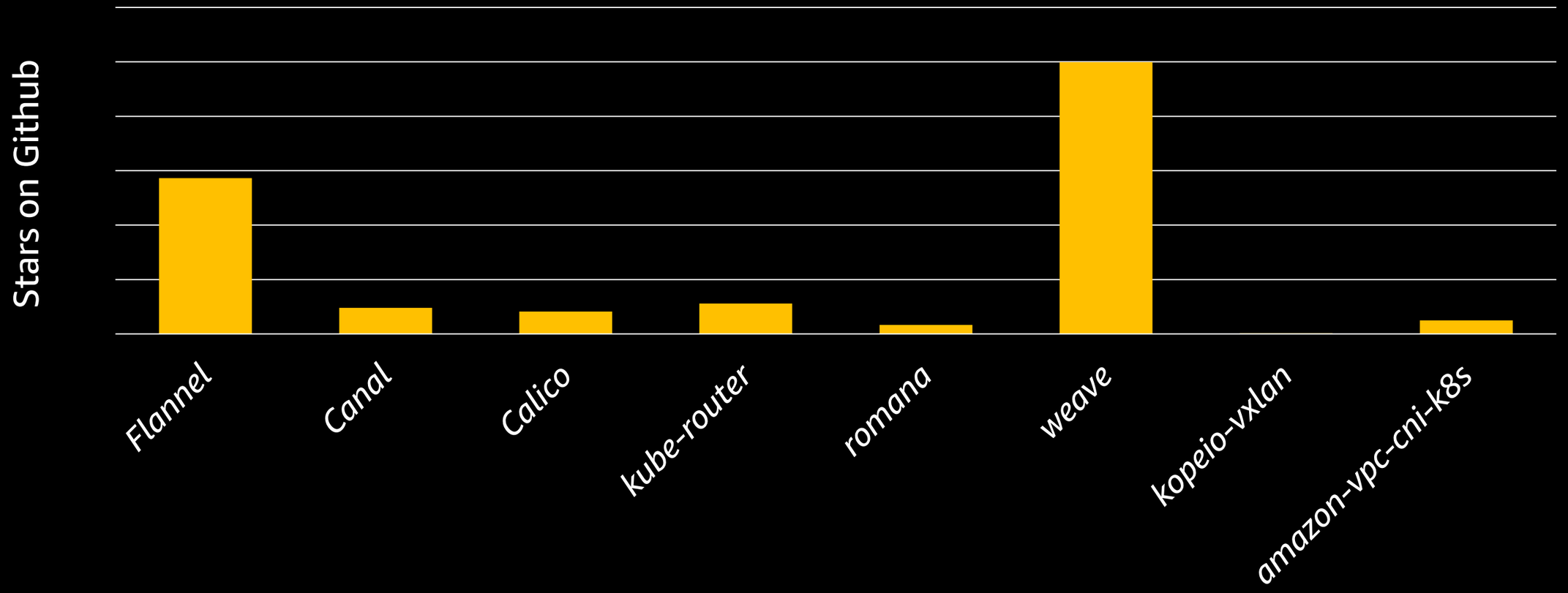
“每个 Pod 都应该拥有自己的 IP 地址，并且所有的 Pod 都应该能够和彼此通信”

什么是 CNI

- Kubernetes 告知底层 SDN 他的容器用来连接网络的方式
- 为容器网络提供一种标准的可插拔式的架构
- 撰写用于配置容器网络接口的插件提供的 API
- CNCF 项目

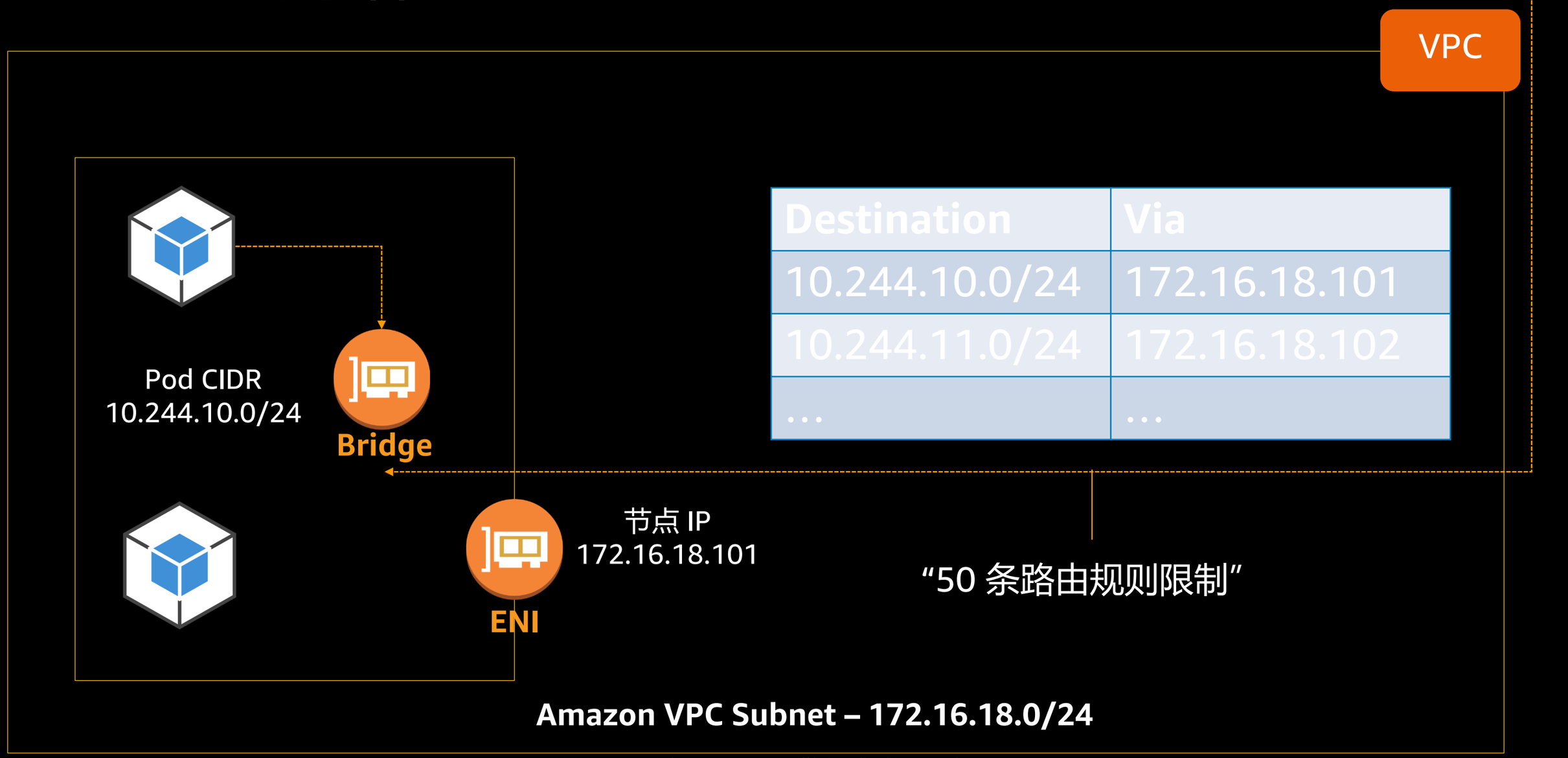


业内主流的 Kubernetes 网络解决方案



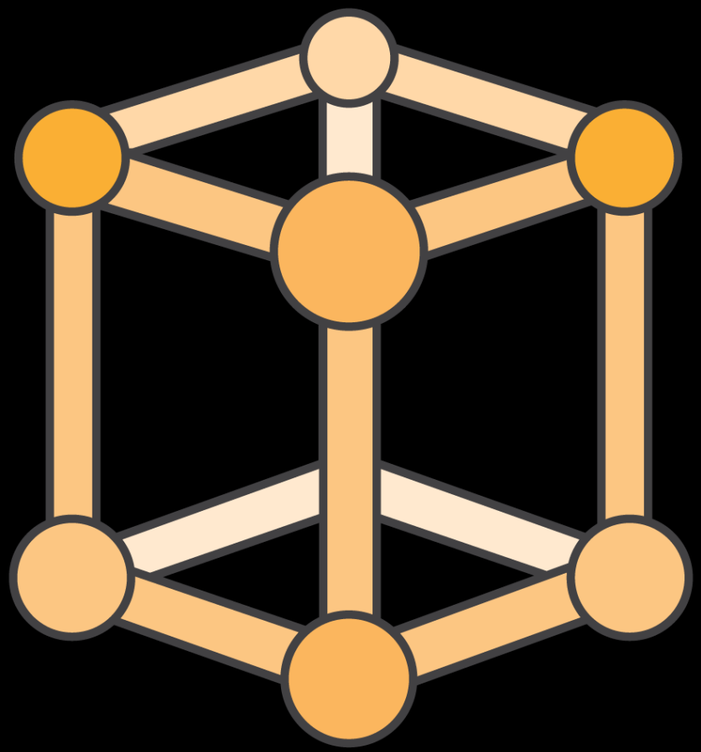
AWS 中国（宁夏）区域由西云数据运营
AWS 中国（北京）区域由光环新网运营

Kubenet 网络

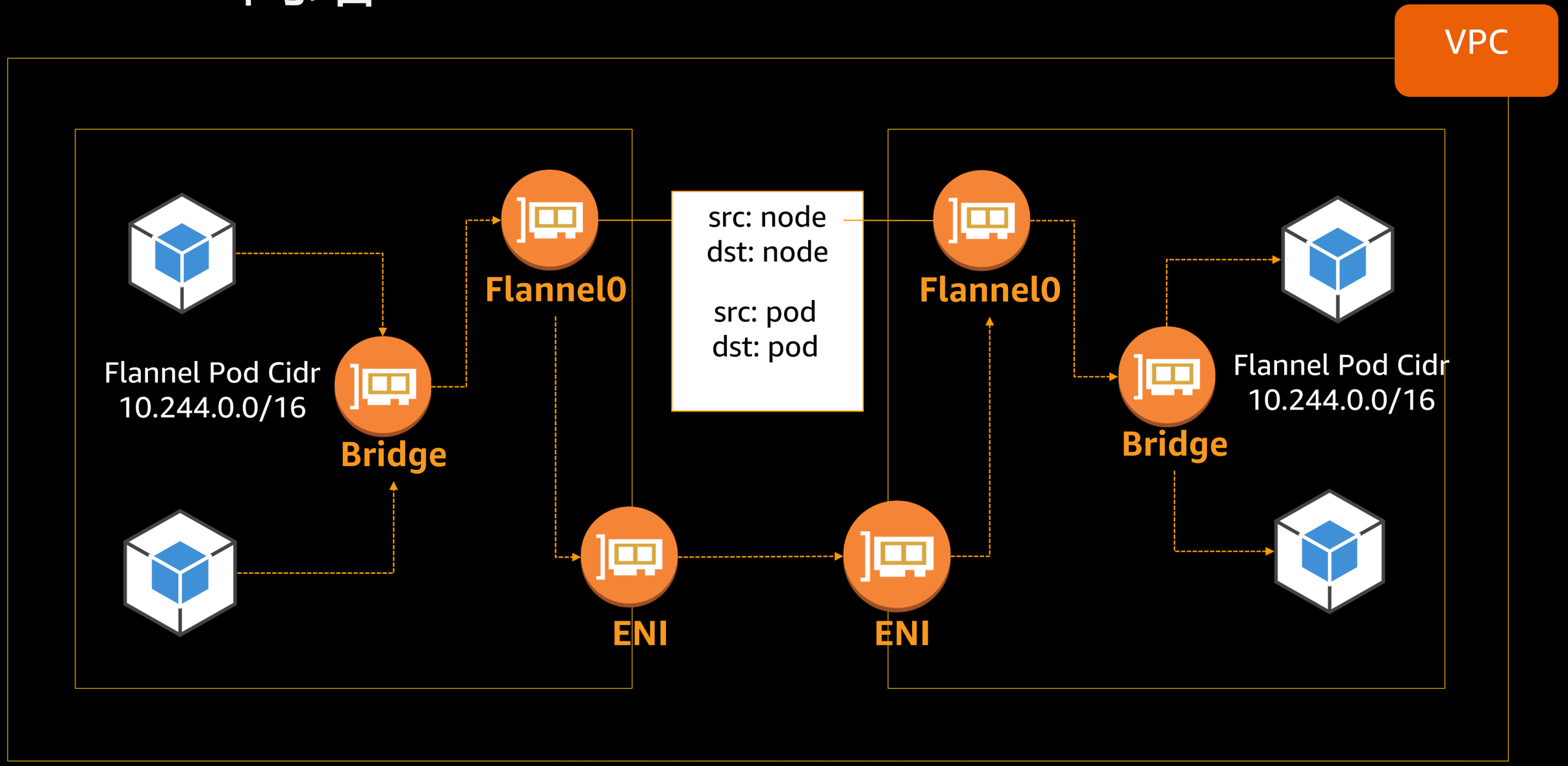


Overlay 网络

- 无法获得足够的 IP 地址空间？（子网 IP 规划不合理）
- 现有网络限制（VPC 路由规则条目限制）无法满足需求
- 需要特定 overlay 网络提供的功能 – *网络规则*



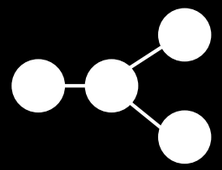
Flannel 网络



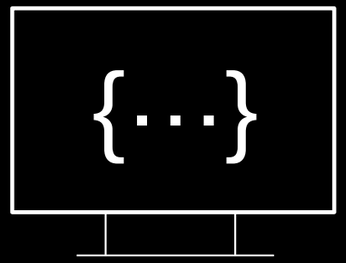
我真的需要一个 overlay 网络么？

CNI (Container Network Interface)

VPC

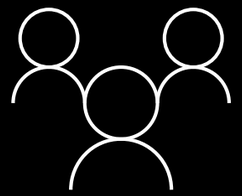


支持原生 VPC 网络的 CNI 插件



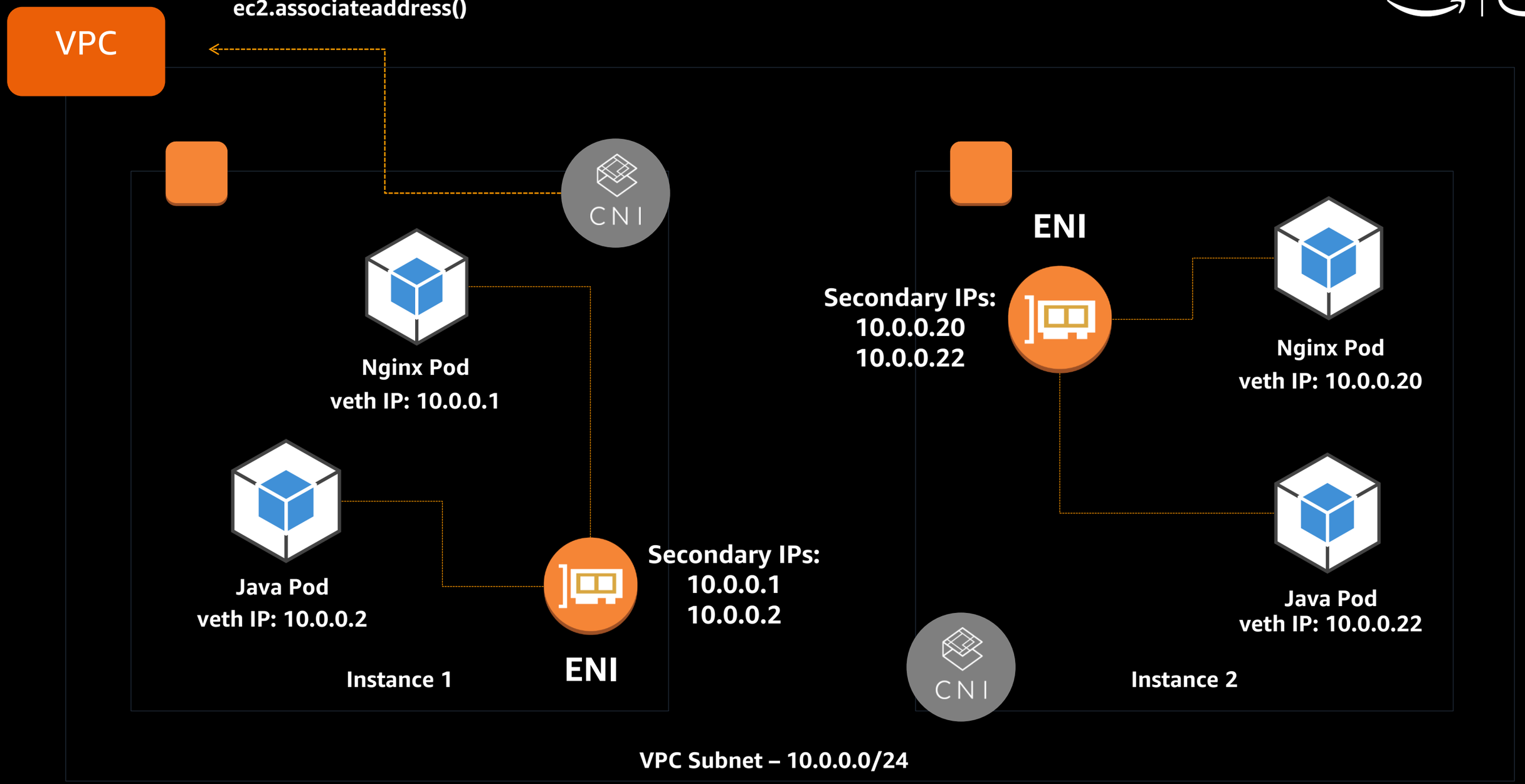
每个 Pod 拥有它所在 VPC 相同的 VPC IP 地址

简单安全的网络

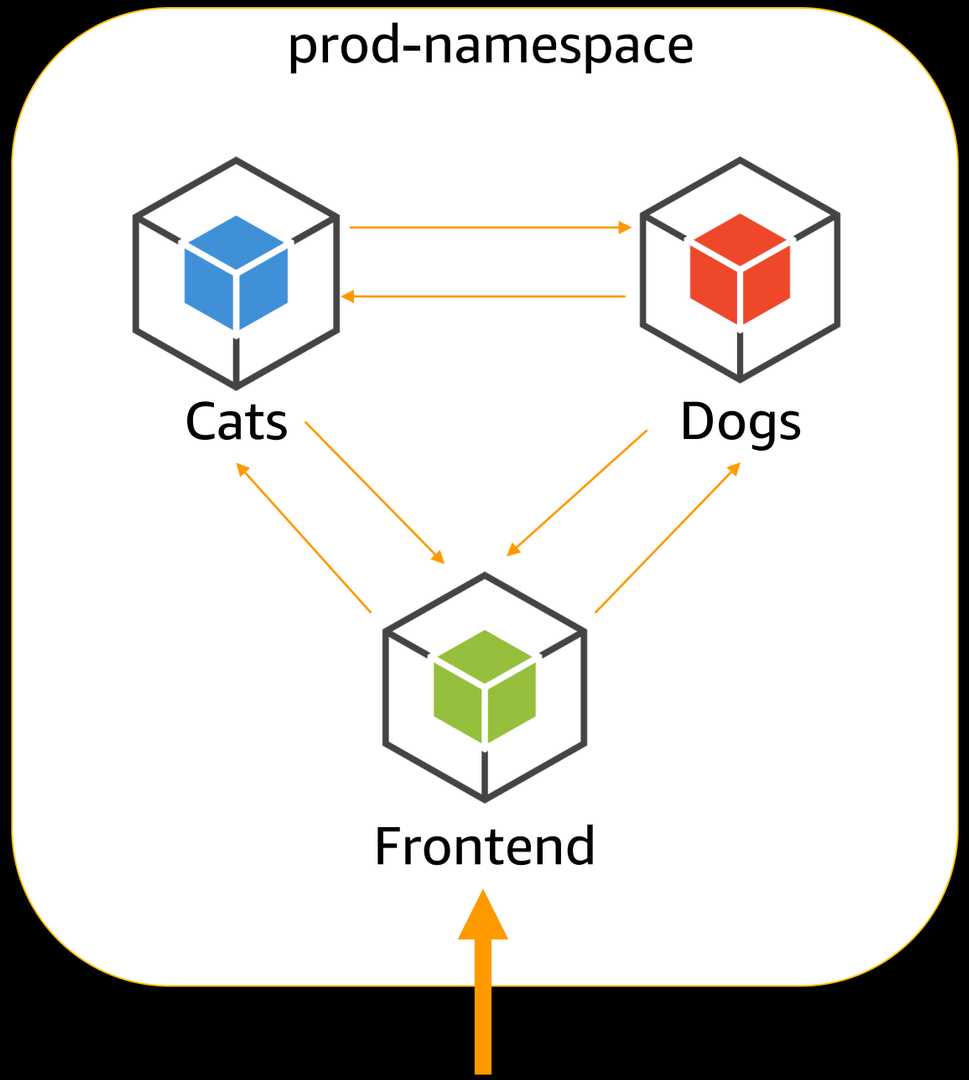


开源项目

https://github.com/aws/amazon-vpc-cni-k8s

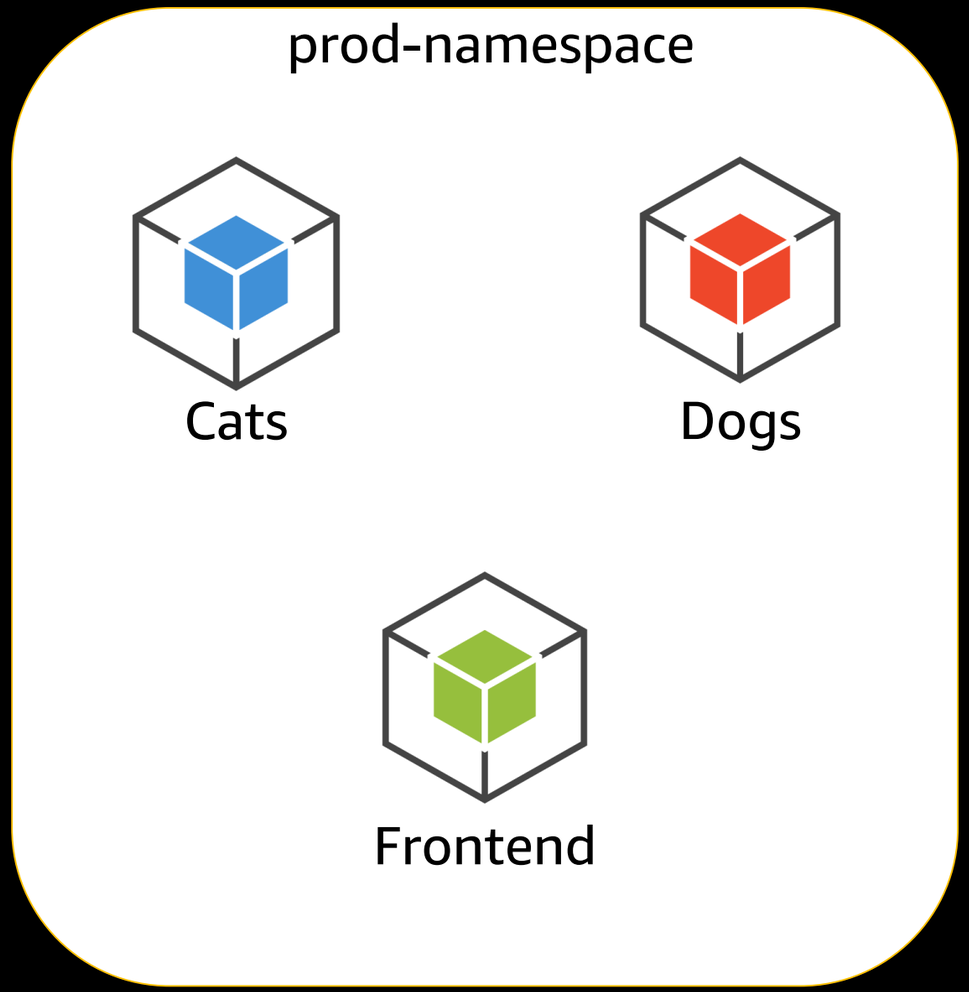


如何做到网络分隔 (Segmentation)?



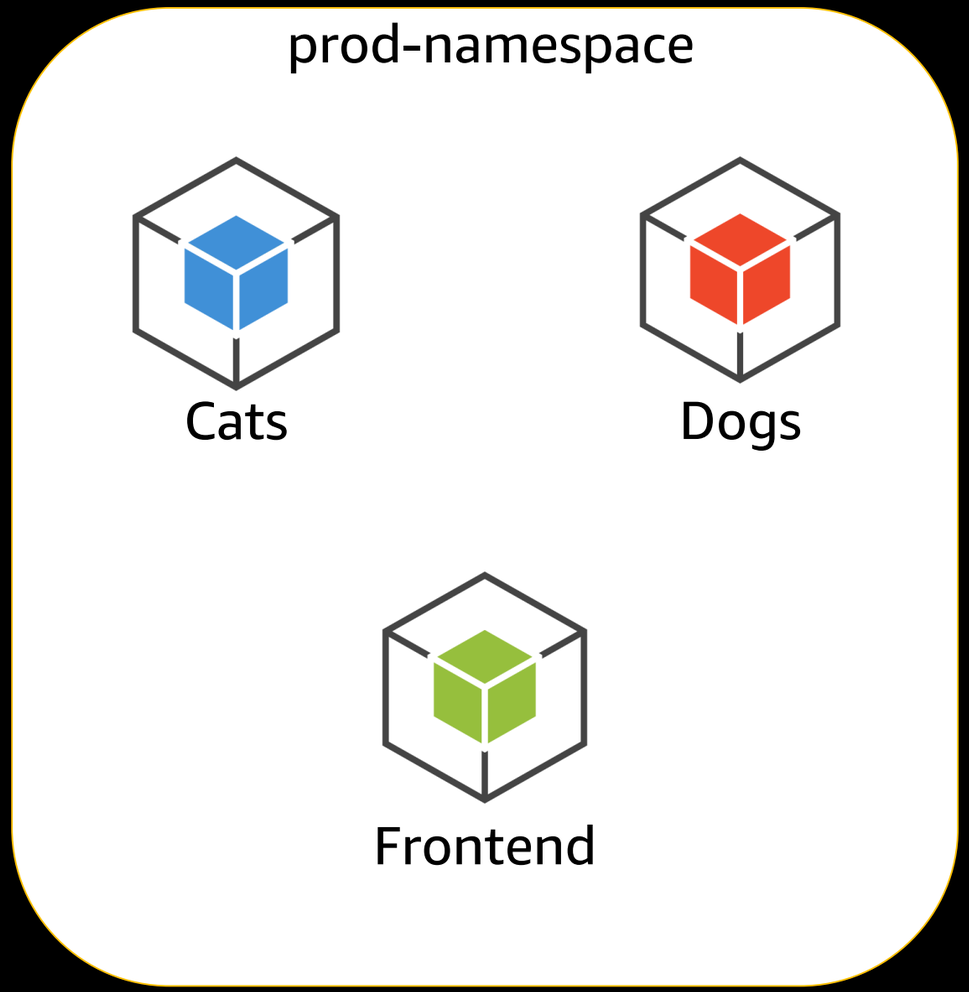

```
kind: NetworkPolicy
apiVersion: extensions/v1beta1
metadata:
 name: default-deny
spec:
 podSelector:
  matchLabels: {}
```

如何做到网络分隔 (Segmentation)?



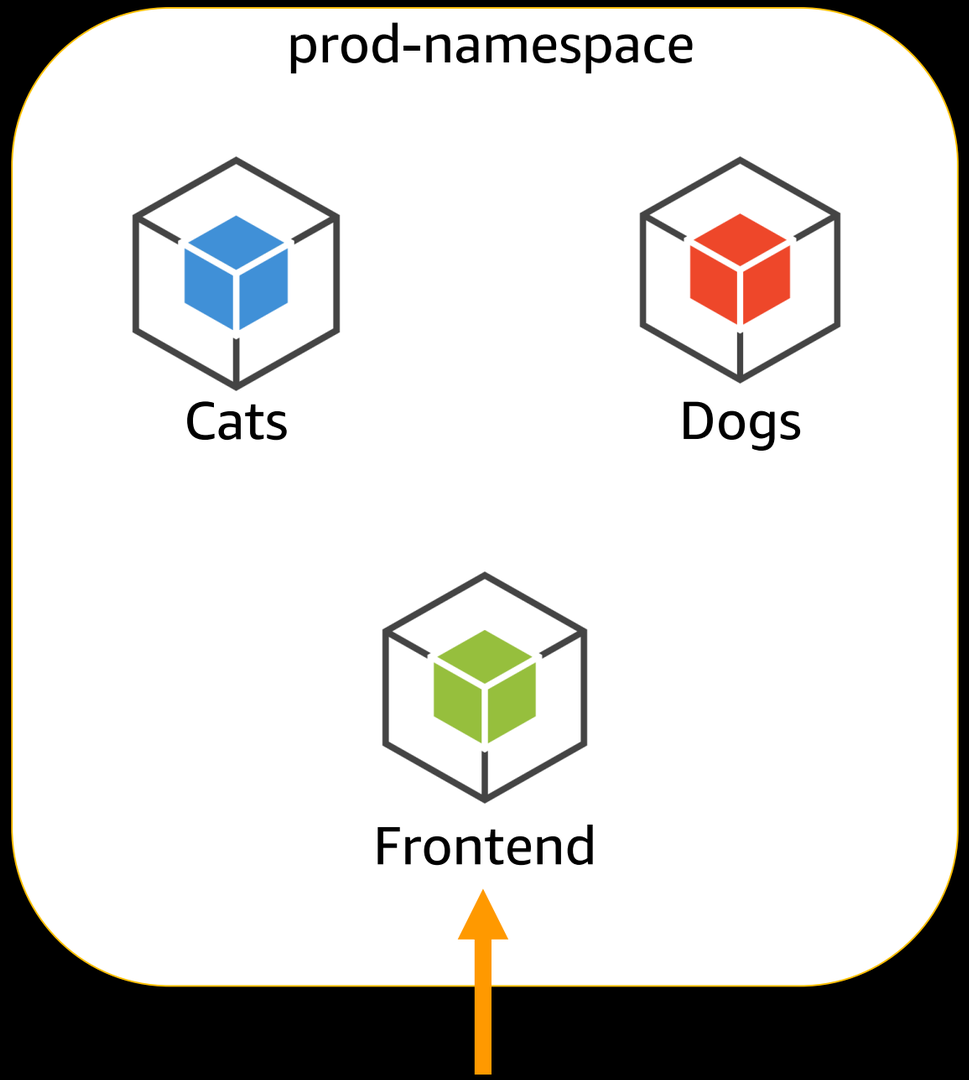

```
kind: NetworkPolicy
apiVersion: extensions/v1beta1
metadata:
 name: default-deny
spec:
 podSelector:
  matchLabels: {}
```

如何做到网络分隔 (Segmentation)?



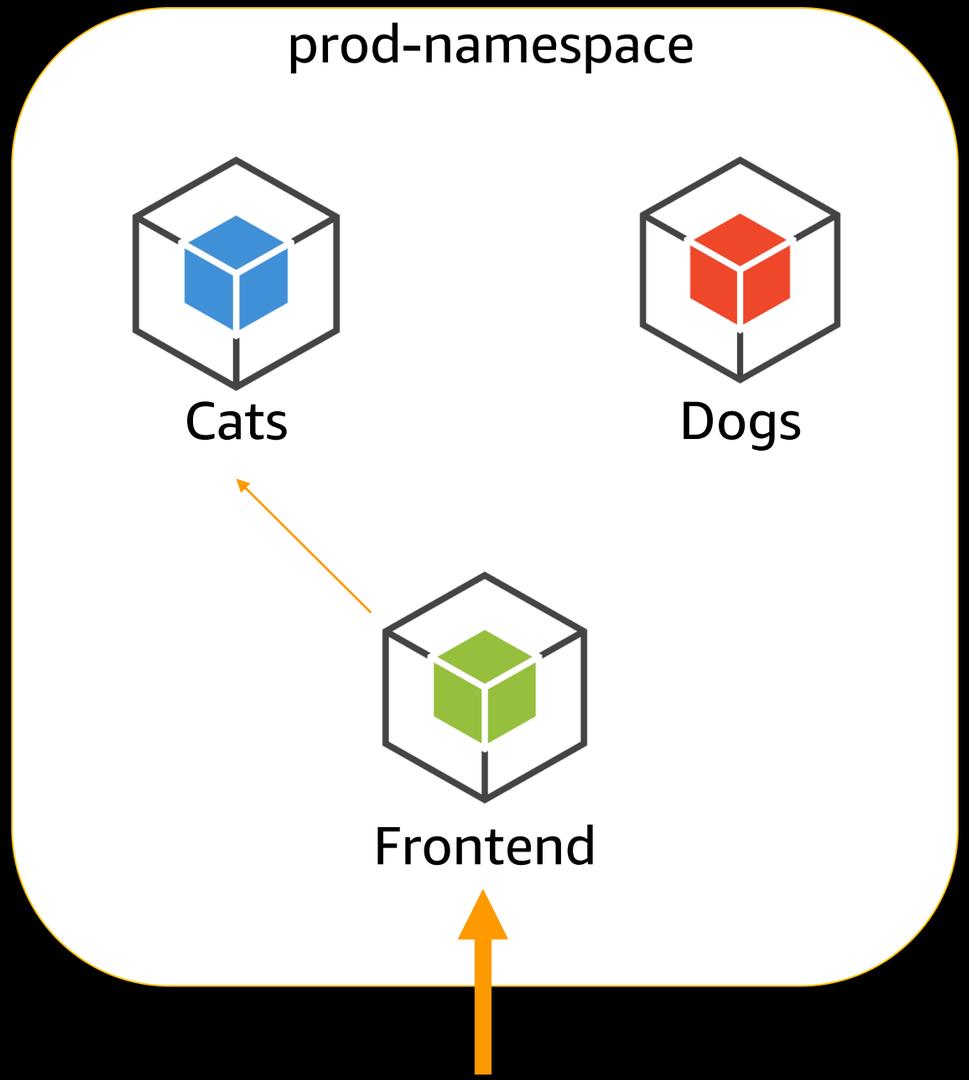

```
apiVersion: extensions/v1beta1
kind: NetworkPolicy
metadata:
 name: public-to-frontend
spec:
 podSelector:
   matchLabels:
     role: frontend
 ingress:
 - from:
   - ipBlock:
       cidr: "0.0.0.0/0"
   ports:
   - protocol: TCP
     port: 80
```

如何做到网络分隔 (Segmentation)?



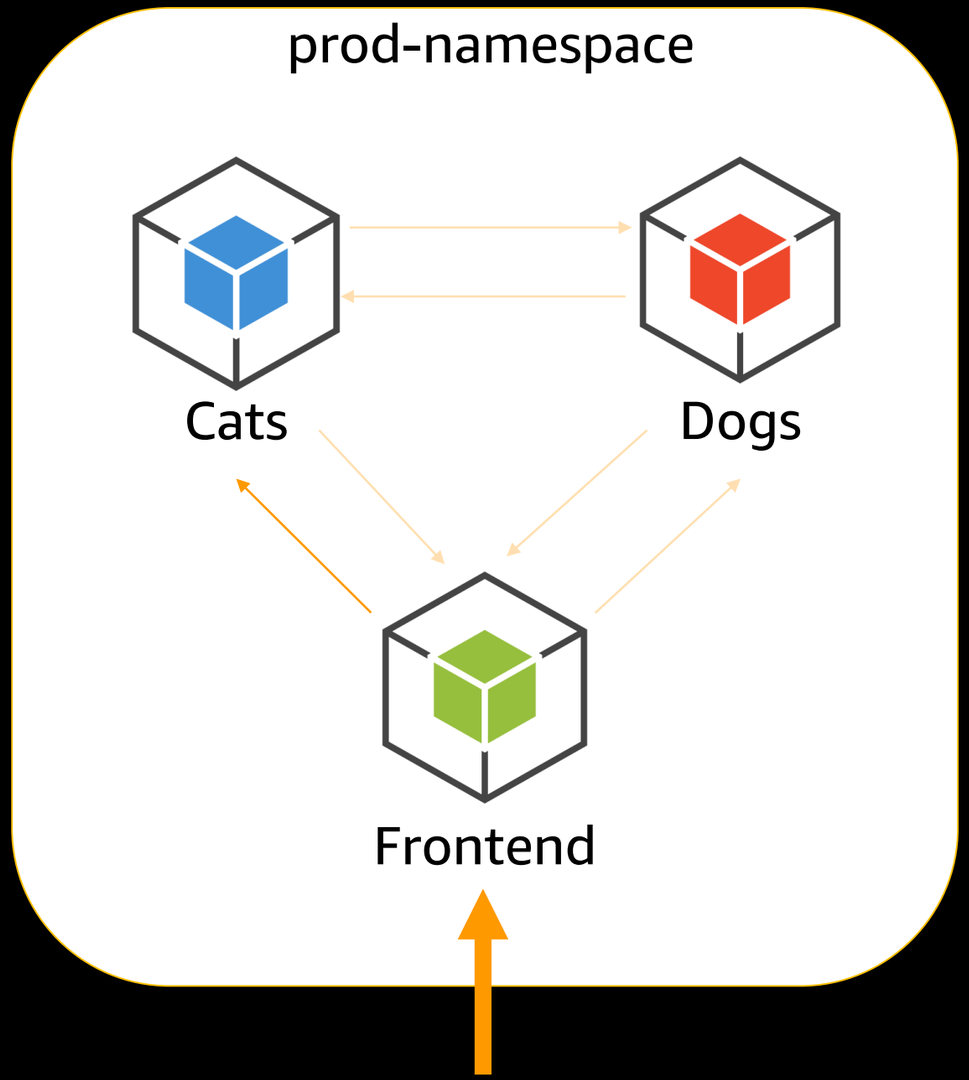

```
apiVersion: extensions/v1beta1
kind: NetworkPolicy
metadata:
 name: public-to-frontend
spec:
 podSelector:
   matchLabels:
     role: frontend
 ingress:
 - from:
   - ipBlock:
       cidr: "0.0.0.0/0"
   ports:
   - protocol: TCP
     port: 80
```

如何做到网络分隔 (Segmentation)?




```
apiVersion: extensions/v1beta1
kind: NetworkPolicy
metadata:
 name: frontend-to-cats
spec:
 podSelector:
  matchLabels:
   role: cats
 ingress:
 - from:
  - podSelector:
     matchLabels:
       role: “frontend”
  ports:
   - protocol: TCP
     port: 80
```

如何做到网络分隔 (Segmentation)?




```
apiVersion: extensions/v1beta1
kind: NetworkPolicy
metadata:
 name: frontend-to-cats
spec:
 podSelector:
  matchLabels:
   role: cats
 ingress:
 - from:
  - podSelector:
     matchLabels:
       role: “frontend”
  ports:
   - protocol: TCP
     port: 80
```

认证

AWS 中国（宁夏）区域由西云数据运营
AWS 中国（北京）区域由光环新网运营

Kubernetes 集成 AWS IAM 认证管理

AWS 中国（宁夏）区域由西云数据运营
AWS 中国（北京）区域由光环新网运营

AWS 认证和访问管理 (AWS IAM)

AWS IAM 提供了安全访问 AWS 服务和资源的机制

为 Amazon EKS 服务提供了两个角色

Service	Access level	Resource
Allow (7 of 109 services) Show remaining 102		
Auto Scaling	**Limited**: List, Write	All resources
EC2	Full access	All resources
EC2 Container Registry	**Full**: List **Limited**: Read	All resources
ELB	Full access	All resources
ELB v2	Full access	All resources
Route 53	**Limited**: List	All resources
S3	**Limited**: List, Read, Write, Permissions management	Multiple

Service	Access level	Resource
Allow (4 of 109 services) Show remaining 105		
EC2	**Full**: List **Limited**: Read, Write	All resources
EC2 Container Registry	**Full**: List **Limited**: Read	All resources
Route 53	**Limited**: List	All resources
S3	**Limited**: List, Read, Write, Permissions management	Multiple

为 Kubectl 和 pods 提供细粒度的控制

Kubernetes + AWS IAM

- AWS 原生的访问控制
- 和开源项目Heptio合作
- Kubectl 和 worker 节点
- 基于 Kubernetes RBAC

AWS IAM 认证 + Kubectl

Kubectl

1) 传递 AWS 认证信息

K8s API

2) 验证 AWS 认证信息

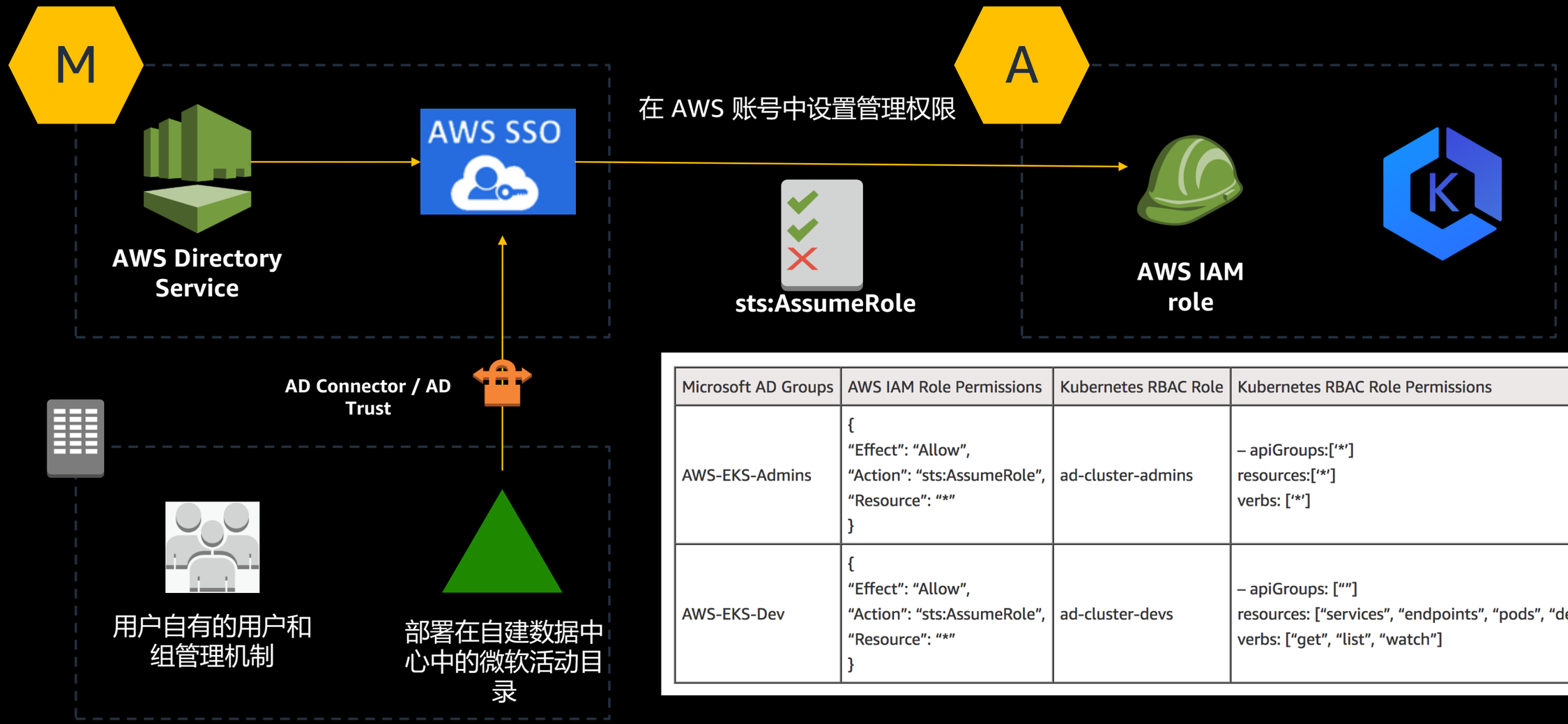
AWS 认证

4) K8s 行为
允许/拒绝

3) 通过 RBAC 授权 AWS 认证信息

heptio

SAML 2.0 – 集成微软活动目录及 SSO



Microsoft AD Groups	AWS IAM Role Permissions	Kubernetes RBAC Role	Kubernetes RBAC Role Permissions

AWS-EKS-Admins	{
	“Effect”: “Allow”,
	“Action”: “sts:AssumeRole”,
	“Resource”: “*”
	}
		ad-cluster-admins	– apiGroups:[‘*’]
			resources:[‘*’]
			verbs: [‘*’]

AWS-EKS-Dev	{
	“Effect”: “Allow”,
	“Action”: “sts:AssumeRole”,
	“Resource”: “*”
	}
		ad-cluster-devs	– apiGroups: [“”]
			resources: [“services”, “endpoints”, “pods”, “deployments”, “ingress”]
			verbs: [“get”, “list”, “watch”]

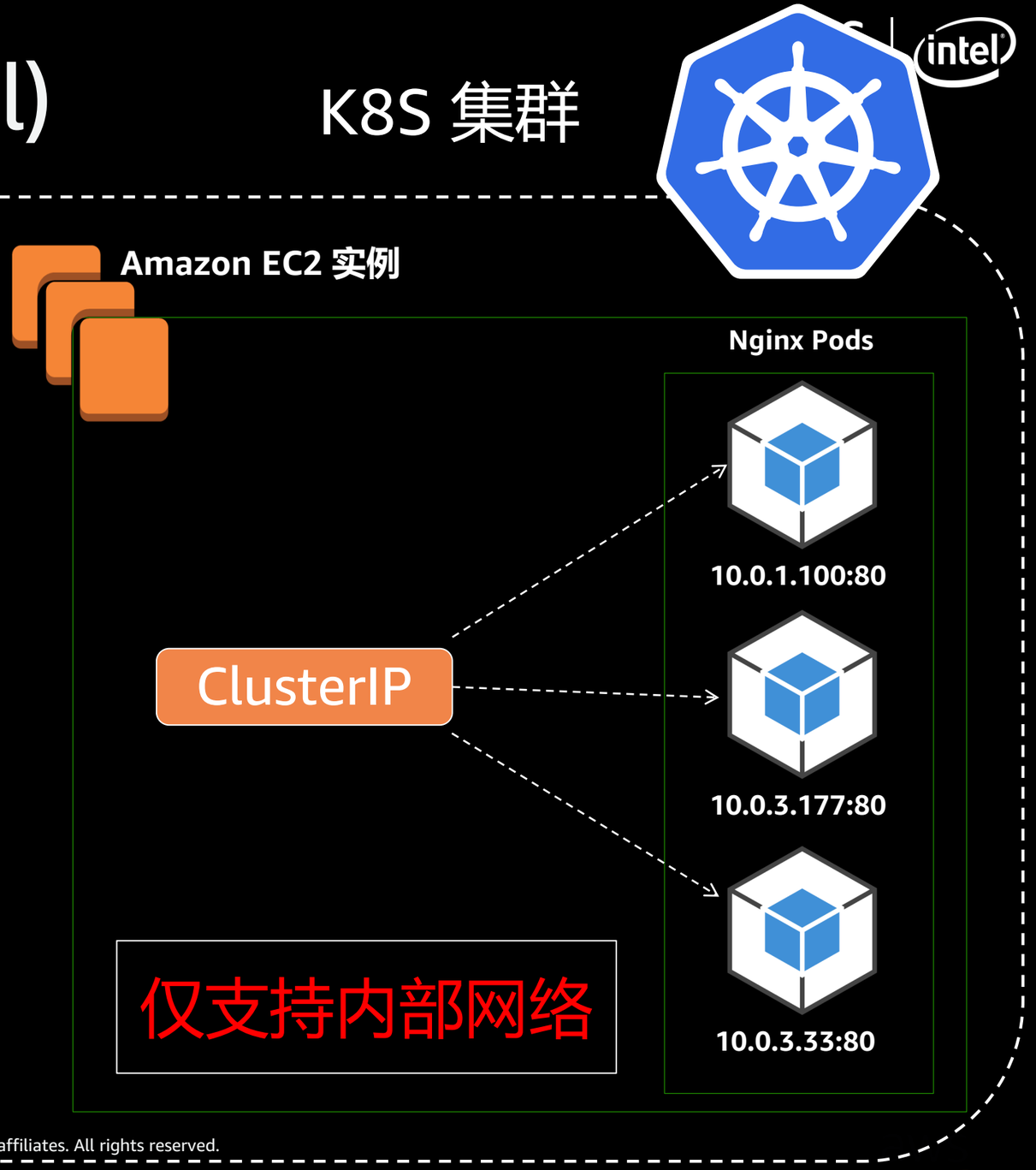
https://aws.amazon.com/tw/blogs/opensource/integrating-ldap-ad-users-kubernetes-rbac-aws-iam-authenticator-project

服务类型与摄取

AWS 中国（宁夏）区域由西云数据运营
AWS 中国（北京）区域由光环新网运营

服务类型 – ClusterIP (virtual)

```
-A KUBE-SEP-JNHR7XFBS7L5NBRR -p tcp -m comment --comment
"default/nginx-service:web" -m tcp -j DNAT --to-destination 10.0.3.33:80
-A KUBE-SEP-MJGBAZNA2WGVIMWN -s 10.0.3.177/32 -m comment --
comment "default/nginx-service:web" -j KUBE-MARK-MASQ
-A KUBE-SEP-MJGBAZNA2WGVIMWN -p tcp -m comment --comment
"default/nginx-service:web" -m tcp -j DNAT --to-destination 10.0.3.177:80
-A KUBE-SEP-XZ3DUOZYSFB3ILKR -s 10.0.1.100/32 -m comment --comment
"default/nginx-service:web" -j KUBE-MARK-MASQ
-A KUBE-SEP-XZ3DUOZYSFB3ILKR -p tcp -m comment --comment
"default/nginx-service:web" -m tcp -j DNAT --to-destination 10.0.1.100:80
-A KUBE-SERVICES -d 172.20.169.0/32 -p tcp -m comment --comment
"default/nginx-service:web cluster IP" -m tcp --dport 80 -j KUBE-SVC-
MCOVNBHDEGIKKKLL
-A KUBE-SVC-MCOVNBHDEGIKKKLL -m comment --comment "default/nginx-
service:web" -m statistic --mode random --probability 0.33332999982 -j
KUBE-SEP-XZ3DUOZYSFB3ILKR
-A KUBE-SVC-MCOVNBHDEGIKKKLL -m comment --comment "default/nginx-
service:web" -m statistic --mode random --probability 0.50000000000 -j
KUBE-SEP-MJGBAZNA2WGVIMWN
-A KUBE-SVC-MCOVNBHDEGIKKKLL -m comment --comment "default/nginx-
service:web" -j KUBE-SEP-JNHR7XFBS7L5NBRR
```



服务类型 – NodePort

K8S 集群

Internet

1) 请求 NGINX Pod
{node}:32001

Amazon EC2 实例

kube-proxy
:32001

2) 代理转发到
k8s 服务

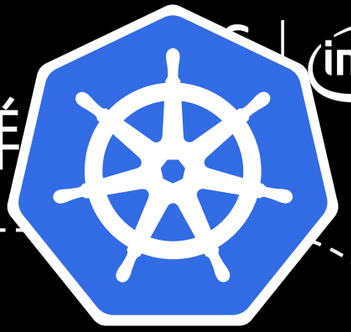
Nginx Pods

10001:8080

10002:8080

10003:8080

服务类型 – LoadBalancer (Amazon ELB)



K8S 集群

Amazon EC2 实例

Nginx Pods

1) 请求 NGINX Pod
{elb}:443

2) ELB 转发到节点
{node:32001}

Internet

Amazon ELB

kube-proxy
:32001

3) 代理转发到
k8s 服务

10001:8080

10002:8080

10003:8080

服务类型– LoadBalancer (NLB)

K8S 集群

Amazon EC2 实例

Nginx Pods

10001:8080

10002:8080

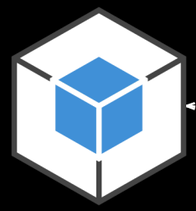
10003:8080

1) 请求 NGINX Pod
{elb}:443

Internet

2) NLB 转发到节点
{node:32001}

NLB

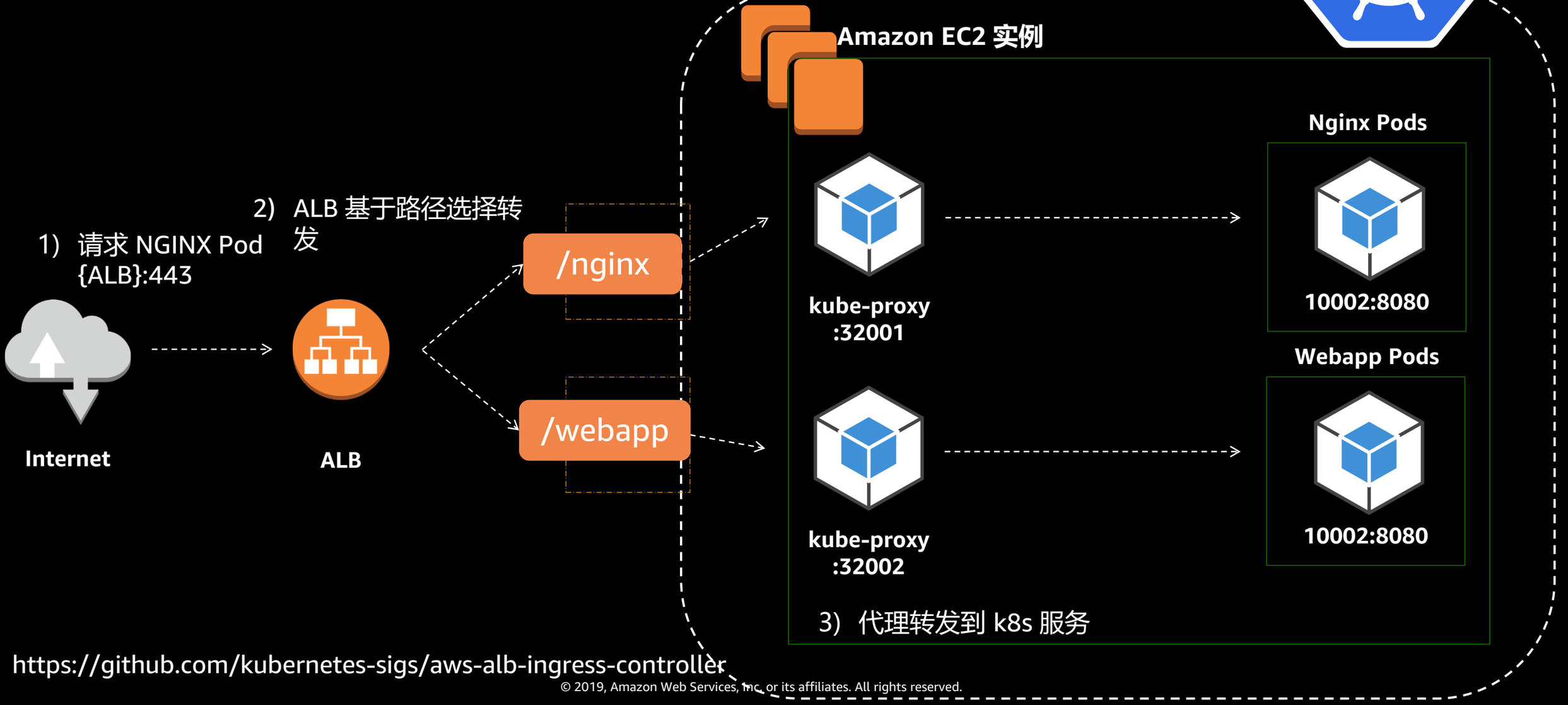


kube-proxy
:32001

3) 代理转发到
k8s 服务

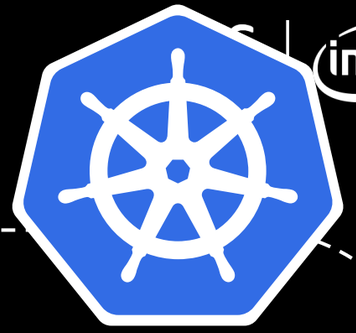
摄取类型 – ALB Ingress

K8S 集群



https://github.com/kubernetes-sigs/aws-alb-ingress-controller

摄取类型– Traefik Ingress



K8S 集群

Amazon EC2 实例

/nginx

Nginx Pods

kube-proxy :32001

traefik-service

10002:8080

1) 请求 ALB:443 2) ALB 基于路径选择转发

/

Internet

ALB

/webapp

Webapp Pods

kube-proxy :32001

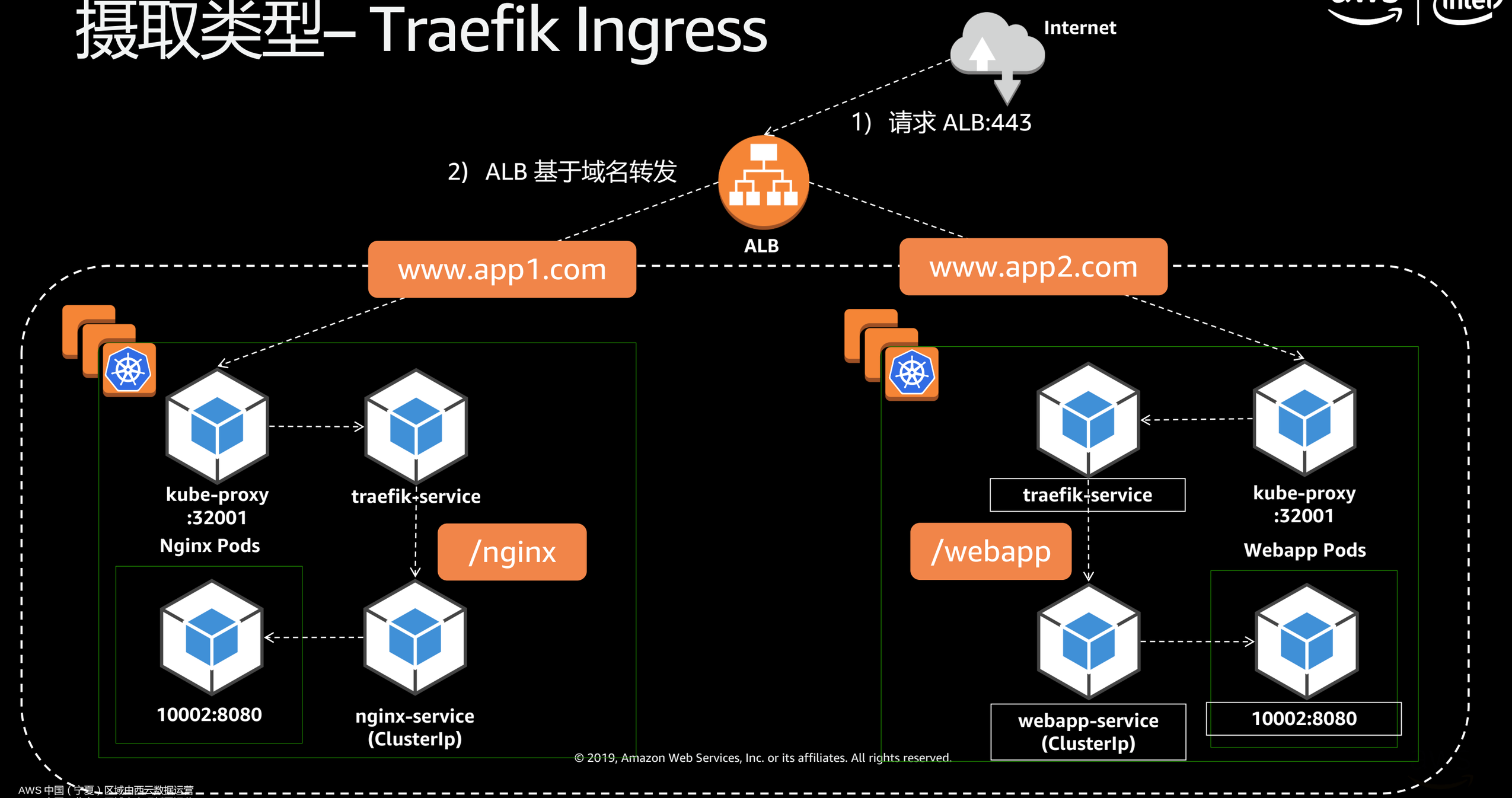
traefik-service

10002:8080

3) 代理转发到 traefik 服务 4) 负载均衡到指定服务

https://docs.traefik.io/user-guide/kubernetes/

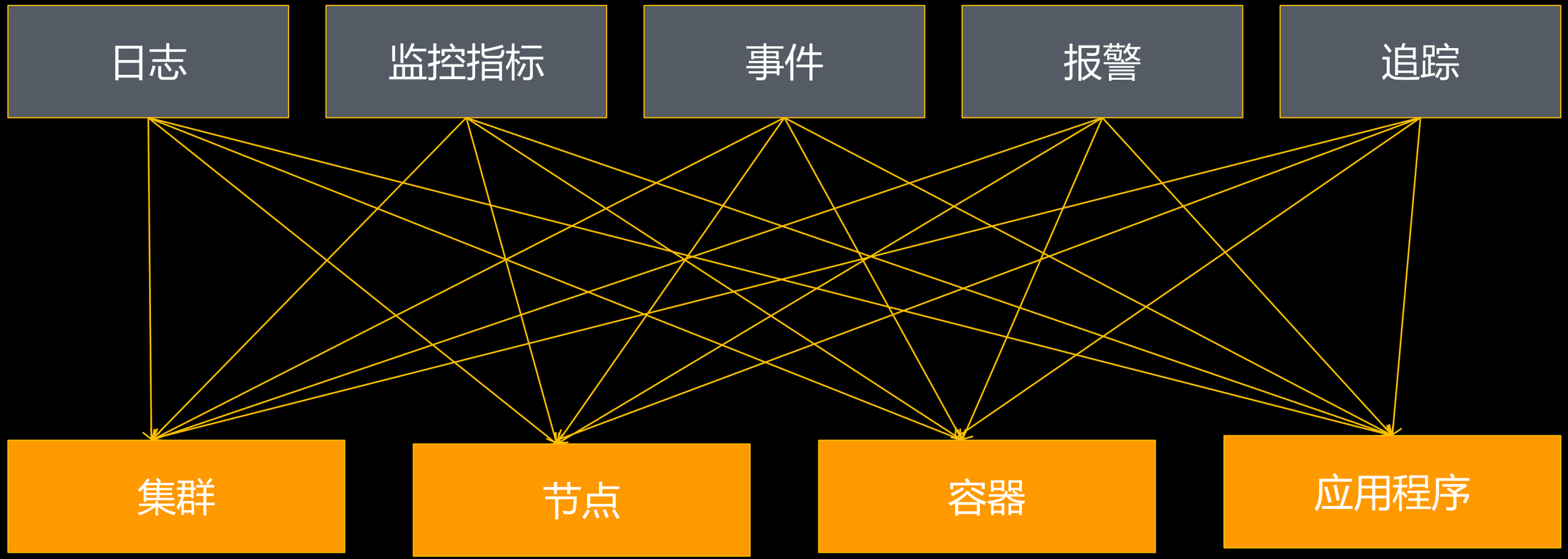
摄取类型– Traefik Ingress



监控

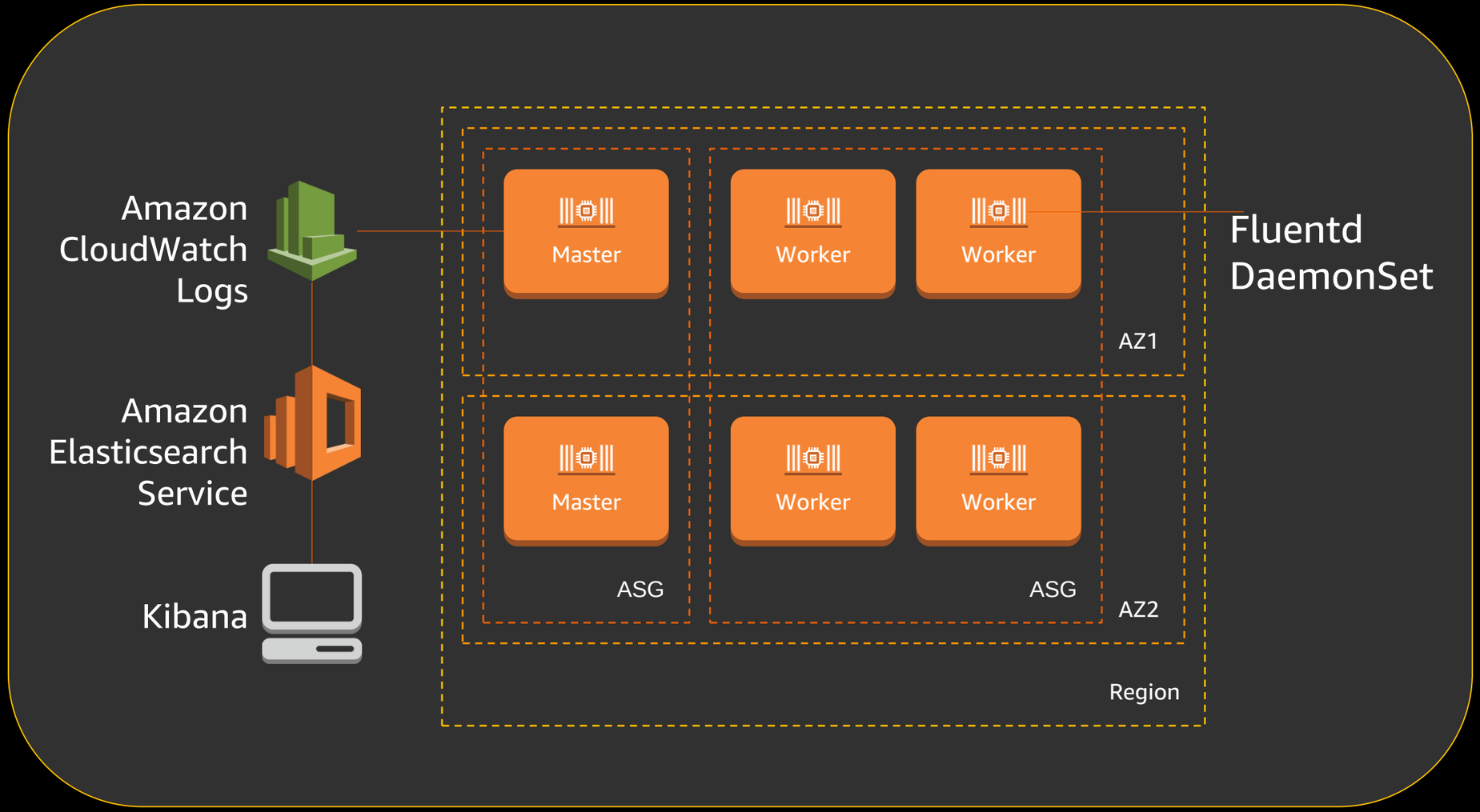
AWS 中国（宁夏）区域由西云数据运营
AWS 中国（北京）区域由光环新网运营

Kubernetes 集群能见度

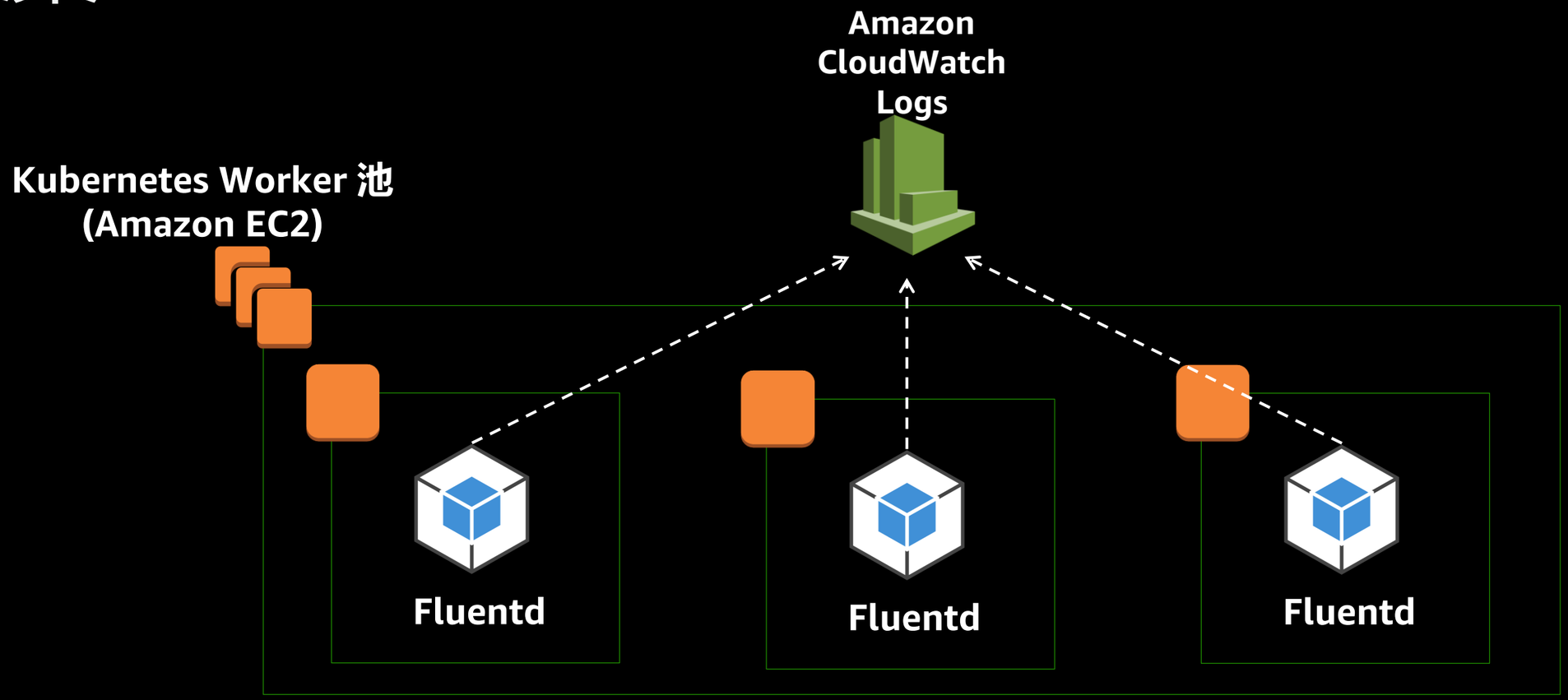


日志

Elasticsearch (索引),
Fluentd (收集), 以及
Kibana (可视化)

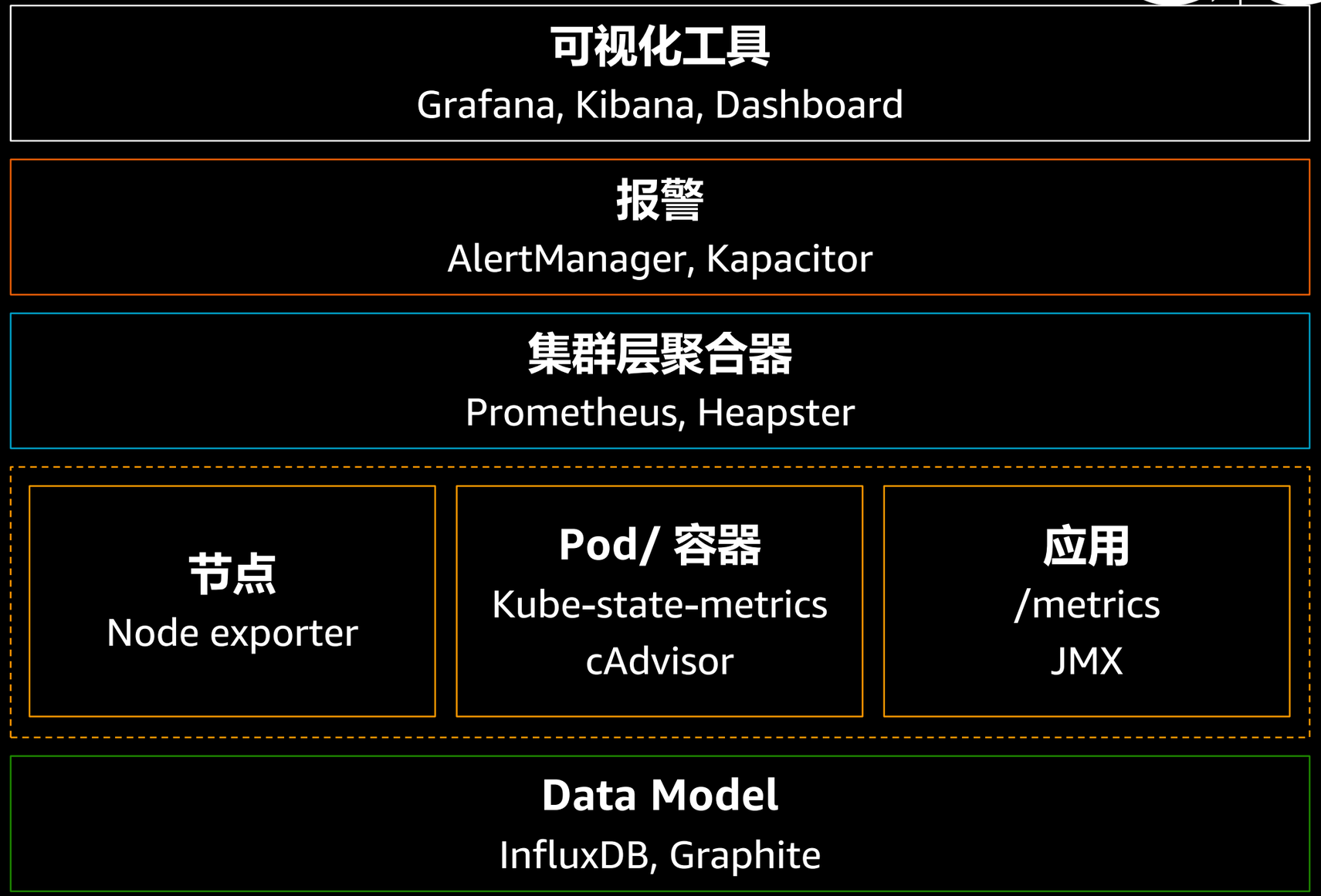


使用 Amazon CloudWatch Logs 功能实现日志聚集



Fluentd DaemonSet
确保每个 worker 节点上都运行一个 Fluentd 容器并且挂接了 /var/lib/docker/containers 路径，这样 fluentd 就可以打包并且发送容器日志到 Amazon CloudWatch Log 中了

监控指标



构建, 发布, 运行...

AWS 中国（宁夏）区域由西云数据运营
AWS 中国（北京）区域由光环新网运营

Kubernetes上的 CI/CD 选择

Jenkins

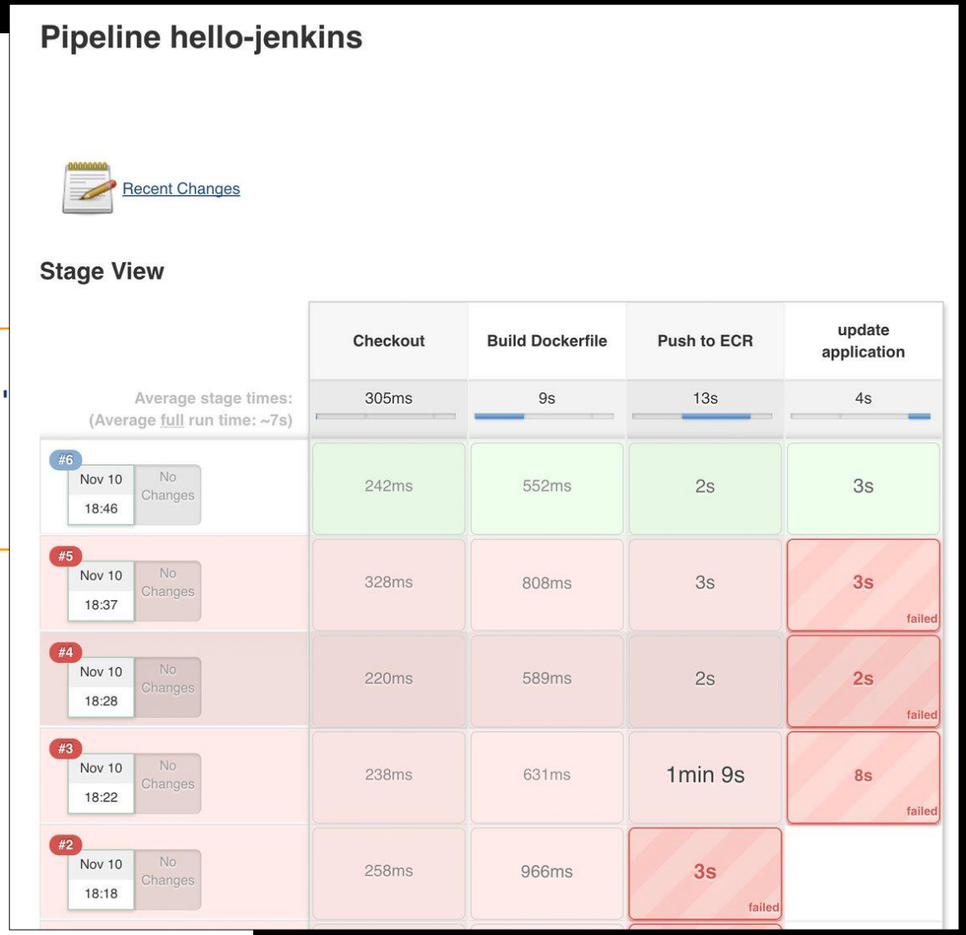
AWS CodePipeline, AWS CodeCommit, AWS CodeBuild

AWS 的 APN 合作伙伴

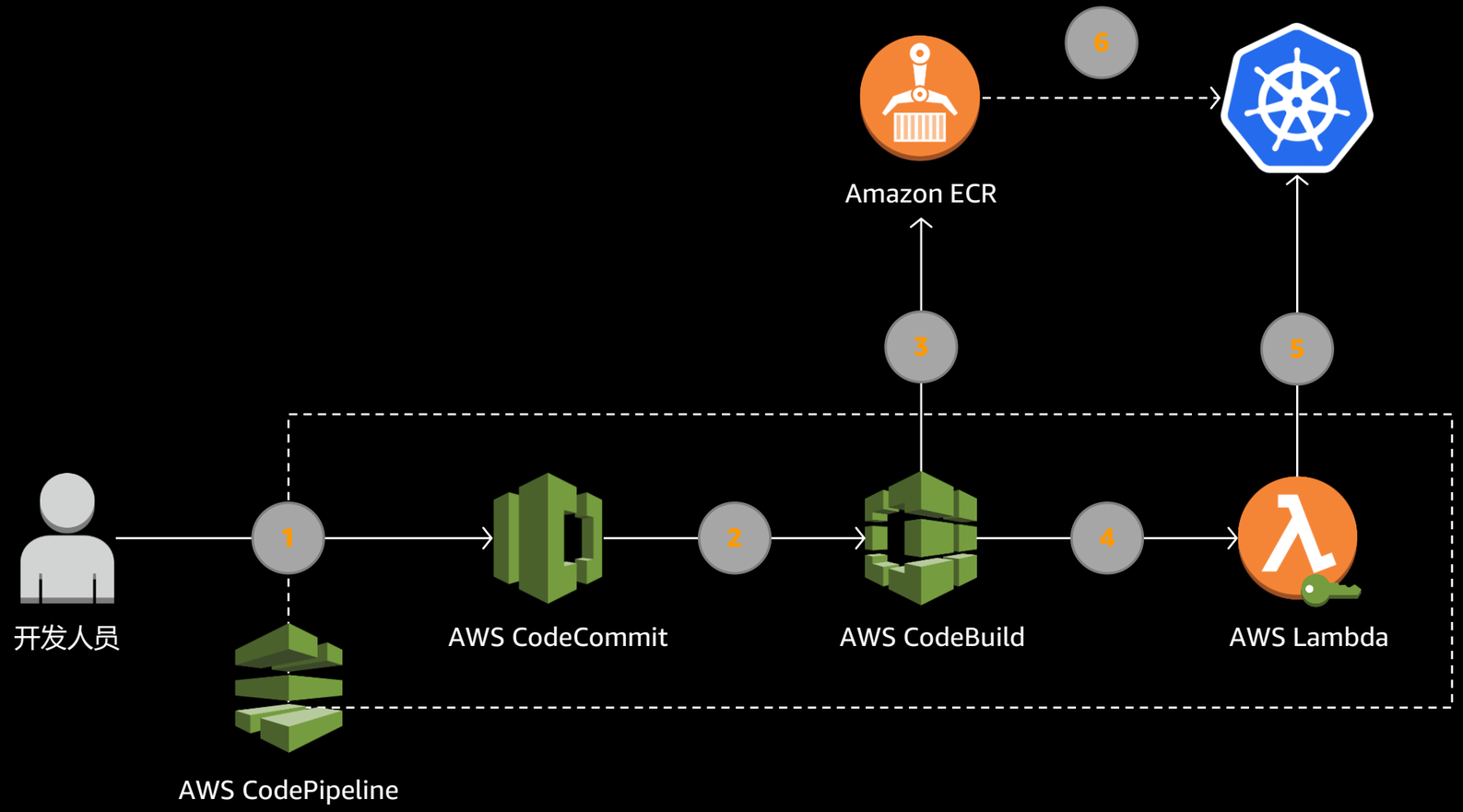
- GitLab
- Shippable
- CircleCI
- Codeship

Jenkins

```
node {

    stage 'Checkout'
    git 'https://github.com/omarlari/aws-container-sample-app.git'

    stage 'Build Dockerfile'
    docker.build('hello')

    stage 'Push to ECR'
    sh ("eval \$(docker run awscli aws ecr get-login --region ${REGION} --no-include-email | sed '
    docker.withRegistry('https://${ECR_REPO}') {
        docker.image('hello').push('${BUILD_NUMBER}')
    }

    stage 'update application'

    kubernetes: { node {
    docker.image('kubectl').inside("--volume=/home/ec2-user/.kube:/config/.kube"){
        sh 'kubectl describe deployment ${APP}'
        sh 'kubectl set image deployment/${APP} hello=${ECR_REPO}/hello:${BUILD_NUMBER}'
        sh 'kubectl describe deployment ${APP}'
        }
    }}
}
```



Kubernetes 持续发布



1. 开发人员在代码仓库中发布更新
2. AWS CodePipline 发现更新后，触发 AWSCodeBuild 服务基于新发布构建新的容器镜像
3. 新的容器镜像被发布到 Amazon ECR 中
4. 触发 AWS Lambda 函数调用应用部署过程
5. 基于 kubernetes 客户端 SDK 更新部署配置
6. 从 Amazon ECR 中获取新的镜像并且发起滚动更新过程

演示 eksctl 工具创建集群及 eks 基本功能

aws INNOVATE
ONLINE CONFERENCE

AWS 中国（宁夏）区域由西云数据运营
AWS 中国（北京）区域由光环新网运营

下一步

Amazon EKS 起步

aws.amazon.com/eks

Amazon EKS 训练营

https://eksworkshop.com

Kubernetes 训练营 (kops)

https://github.com/aws-samples/aws-workshop-for-kubernetes

感谢参加 AWS INNOVATE 2019 在线技术大会

我们希望您在这里找到感兴趣的内容！

也请帮助我们完成**投票打分**和**反馈问卷**。

欲获取关于 AWS 的更多信息和技术内容，可以通过以下方式找到我们：


微信公众号：AWSChina


新浪微博：https://www.weibo.com/amazonaws/


领英：https://www.linkedin.com/company/aws-china/


知乎：https://www.zhihu.com/org/aws-54/activities/


视频中心：http://aws.amazon.bokecc.com/


更多线上活动：https://aws.amazon.com/cn/about-aws/events/webinar/

aws INNOVATE
ONLINE CONFERENCE

AWS 中国（宁夏）区域由西云数据运营
AWS 中国（北京）区域由光环新网运营