2025 WorldSkills Korea Training Camp

DAY 3

WorldSkills Training Camp
제 3과제 풀이 세션

최 영 서

(she/her)
Cloud Support Engineer Intern
Amazon Web Service

과제 풀이 가이드 문서

https://colorful-eggnog-f76.notion.site/3-EKS-TroubleShooting-1fff85b2eedb8034b84df96085f337ed

목차

1. VPC CNI Setting Problems
2. Pod IAM Problem
3. Node Shutdown Problem

1. VPC CNI Setting Problems

1. VPC CNI Setting Problems

(1) IP 주소 부족 문제

Problem

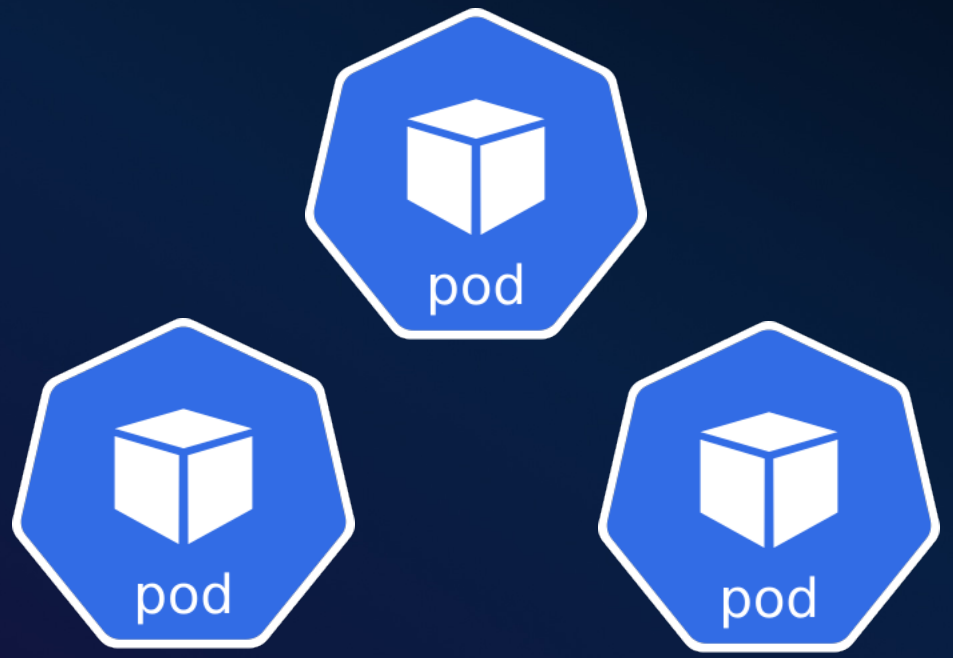
노드가 할당할 수 있는 IP 주소가 부족하여
Pod가 Pending상태에 머무는 현상 발생

Goal

노드가 할당할 수 있는 IP 주소의 개수를 늘리자!

1. VPC CNI Setting Problems

(1) IP 주소 부족 문제

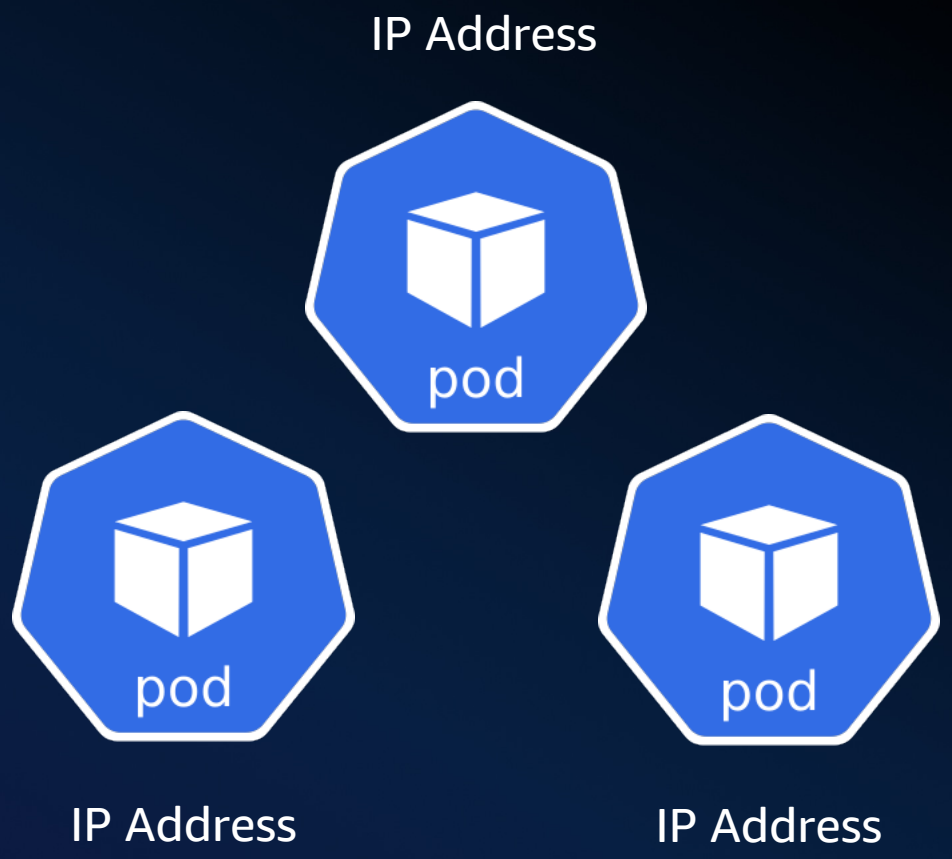


1. VPC CNI Setting Problems

(1) IP 주소 부족 문제

AWS VPC CNI

노드의 ENI
(Elastic Network Interface)



1. VPC CNI Setting Problems

(1) IP 주소 부족 문제

https://docs.aws.amazon.com/ko_kr/ec2/latest/instancetypes/co.html#co_network

Instance type	Baseline / Burst bandwidth (Gbps)	EFA	ENA	ENA Express	Network cards	Max. network interfaces	IP addresses per interface	IPv6
c5.xlarge [1]	1.25 / 10.0	✗	✓	✗	1	4	15	✓

C5.xlarge

최대 ENI 개수 4개 * 각 인터페이스에 붙을 수 있는 IP 개수 15개
=> 4 * (15-1) + 2 = 58

-1 : ENI 자체를 식별하기 위한 IP
+2 : kube-proxy와 aws-node

1. VPC CNI Setting Problems

(1) IP 주소 부족 문제

최대 ENI 개수 4개 * 각 인터페이스에 붙을 수 있는 IP 개수 15개 => 4 * (15-1) + 2 = 58

```
allocatable:
  cpu: 3920m
  ephemeral-storage: "18181869946"
  hugepages-1Gi: "0"
  hugepages-2Mi: "0"
  memory: 6895760Ki
  pods: "58"
capacity:
```

58 * # of Instance (3) = 174

Core Pods : 17

- coredns : 2
- ebs-csi-controller : 2
- ebs-csi-node : 3
- aws-node (vpc-cni) : 3
- kube-proxy : 2
- pod-identity-agent-node : 3
- metrics-server : 2

App Pods : 157

- wsi-deployment : 157 (Running)

남은 43개의 Pod는 Pending 상태

1. VPC CNI Setting Problems

(1) IP 주소 부족 문제

58 * # of Instance (3) = 174

최대 ENI 개수 4개 * 각 인터페이스에 붙을 수 있는
IP 개수 15개 => 4 * (15-1) + 2 = 58

```
allocatable:
  cpu: 3920m
  ephemeral-storage: "18181869945"
  hugepages-1Gi: "0"
  hugepages-2Mi: "0"
  memory: 6895760Ki
  pods: "58"
capacity:
```

IP Prefix Mode

Core Pods : 17

- coredns : 2
- ebs-csi-controller : 2
- ebs-csi-node : 3
- aws-node (vpc-cni) : 3
- kube-proxy : 2
- pod-identity-agent-node : 3
- metrics-server : 2

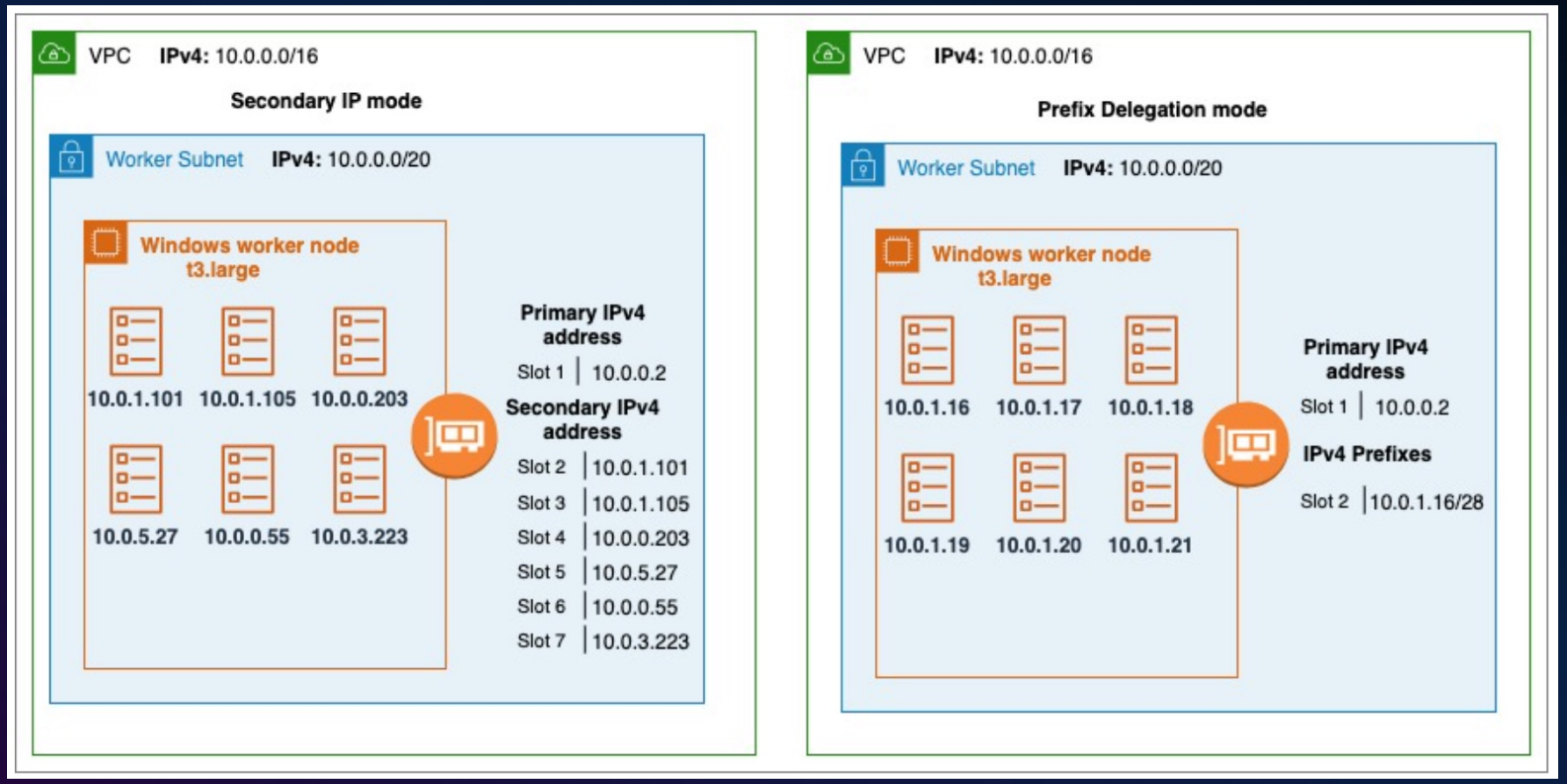
App Pods : 157

- wsi-deployment : 157 (Running)

남은 43개의 Pod는
Pending 상태

1. VPC CNI Setting Problems

(1) IP 주소 부족 문제



1. VPC CNI Setting Problems

(1) IP 주소 부족 문제

ENABLE_PREFIX_DELEGATION
= true

노드가 할당할 수 있는 최대 IP 개수 : 110개

```
allocatable:
  cpu: 3920m
  ephemeral-storage: "18181869946"
  hugepages-1Gi: "0"
  hugepages-2Mi: "0"
  memory: 6895760Ki
  pods: "110"
capacity:
```

1. VPC CNI Setting Problems

(2) SecurityGroup for Pods 문제

Situation

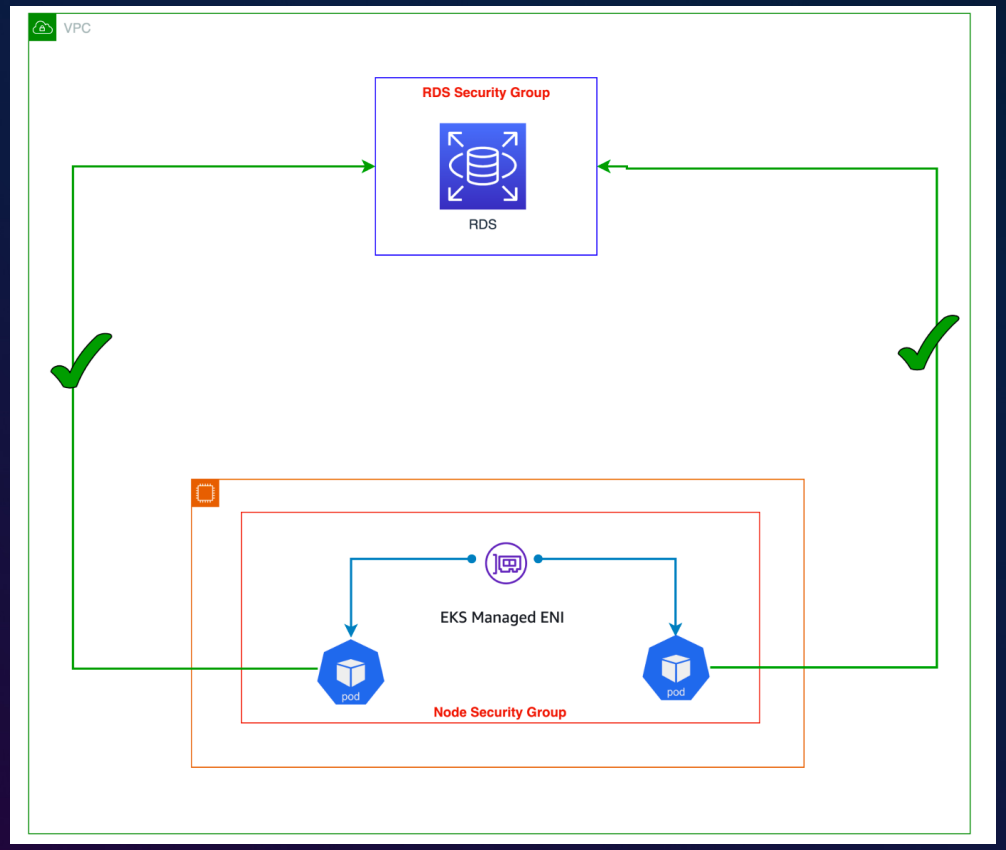
특정 라벨을 갖는 Pod만 VPC 외부에 있는
특정 리소스에 접근하도록 하고 싶음

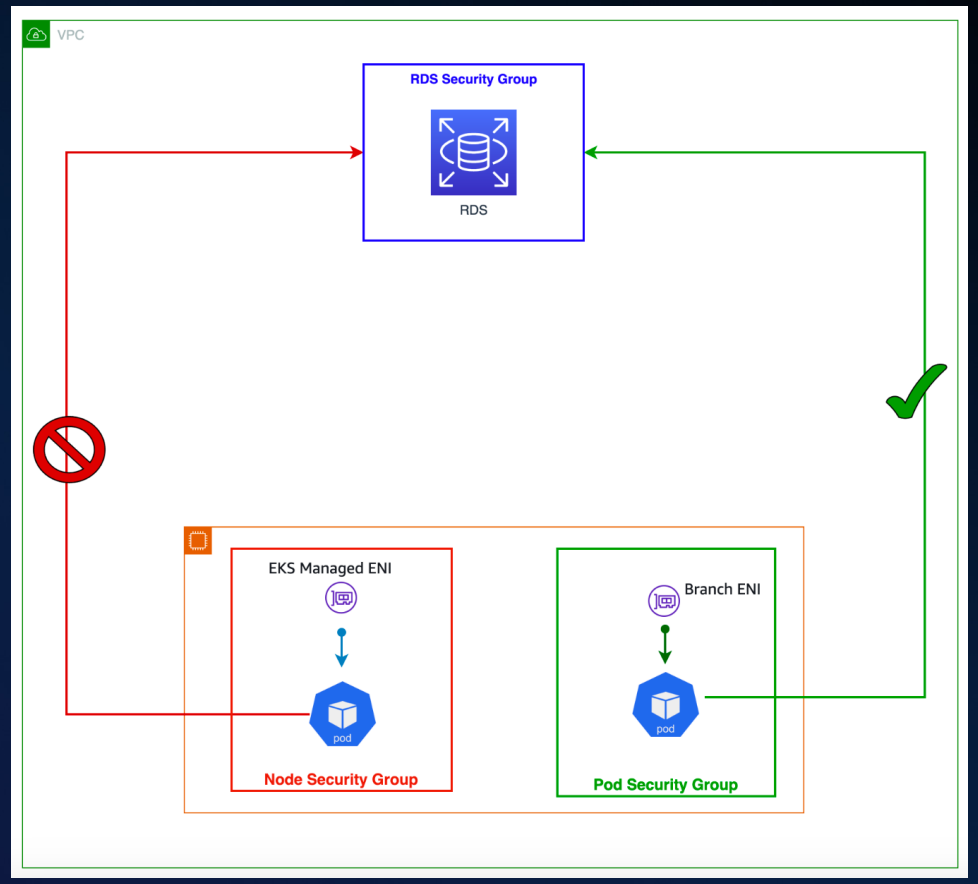
Solution

**SecurityGroupPolicy 리소스를 이용하여
Pod에 보안그룹을 붙이자!**

1. VPC CNI Setting Problems

(2) SecurityGroup for Pods 문제



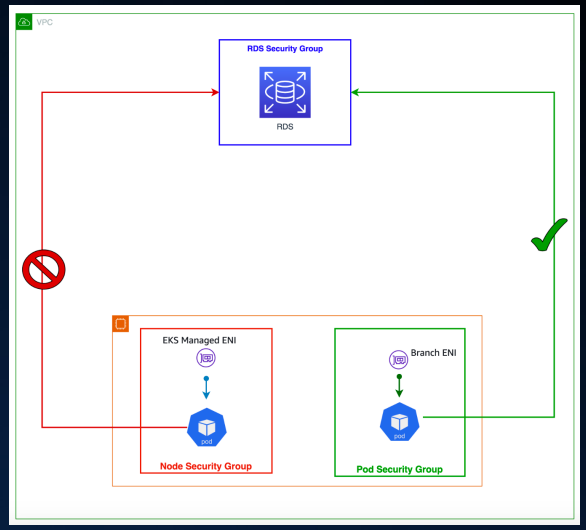


1. VPC CNI Setting Problems

(2) SecurityGroup for Pods 문제

ENABLE_POD_ENI= true

SecurityGroup for Pods



Primary ENI : 보통 노드가 가지고 있는 메인 네트워크 인터페이스

Trunk ENI : VPC Resource Controller에 의해 노드마다 하나씩 생성되는 인터페이스

Branch ENI : Pod에 따로 만들어주는 별도의 네트워크 인터페이스 (여러 Branch ENI가 하나의 Trunk ENI에 연결됨)

VPC Resource Controller : Trunk ENI 및 Branch ENI를 자동으로 생성 / 관리해주는 AWS 서비스

SecurityGroupPolicy: Amazon EKS에서 Pod 단위로 Security Group을 할당하기 위해 사용하는 Kubernetes 리소스

2. Pod IAM Problems

2. Pod IAM Problems

(1) Pod Identity 설정 문제

Problem

EBS Volume이 프로비저닝 되지 않고,
Pod에 PVC가 정상적으로 바인딩 되지 않는 문제 발생

Goal

**Pod Identity 기반 EBS CSI Driver Add on을
설치하여 PVC 바인딩 문제를 해결하자!**

2. Pod IAM Problems

(1) Pod Identity 설정 문제

IMDS
(Instance Metadata Service)

노드에 IAM Role을 부여
➔ Pod가 노드의 IAM Role을 그대로 상속

- 노드의 IAM Role을 그대로 가져가므로 모든 Pod에게 과한 권한이 부여됨
- 보안에 취약함

IRSA
(IAM Role for ServiceAccount)

Pod의 ServiceAccount에 IAM Role을 직접 연결 (OIDC Provider사용)

- 장점 : Pod 단위로 최소 권한 부여 가능
- 단점 : 설정이 복잡하고, serviceAccount에 annotation 붙이는 작업도 필요함

Pod Identity

Pod의 ServiceAccount에 IAM Role을 직접 연결 (OIDC Provider 대신 Pod Identity Association 사용)

- IRSA보다 훨씬 더 간단하고 명확하게 구성 가능
- Pod가 ServiceAccount를 통해 자동으로 IAM Role과 연결됨

3. Non-Graceful Node Shutdown

3. Non-Graceful Node Shutdown

(1) Node Termination Handler 문제

Problem

Spot Instance 종료 시 Non-Graceful 하게
노드가 강제 종료되어 Pod 재스케줄링에 시간이 오래 걸림

Goal

**Node Termination Handler를 사용하여 노드가
자동으로 cordon / drain 되도록 하자!**

cordon : 노드에 새 Pod가 못 뜨도록 “잠금”
drain : 노드에 있는 Pod 전부 “청소”

3. Non-Graceful Node Shutdown

(1) Node Termination Handler 문제

Graceful Node Shutdown

: 노드가 정상적인 절차로 종료

EBS/NFS/PVC 등 볼륨을 정상적으로 detach 후 종료

Graceful Shutdown via Node Termination Handler

1. AWS EC2 Spot instance -> stop interruption
2. NTH DaemonSet 이 알림 감지
3. kubectl cordon + kubectl drain 자동 수행
4. EBS Detach 정상 완료
5. 새로운 노드에서 정상적으로 Pod 재스케줄링

Non-Graceful Node Shutdown

: 노드가 비정상으로 종료

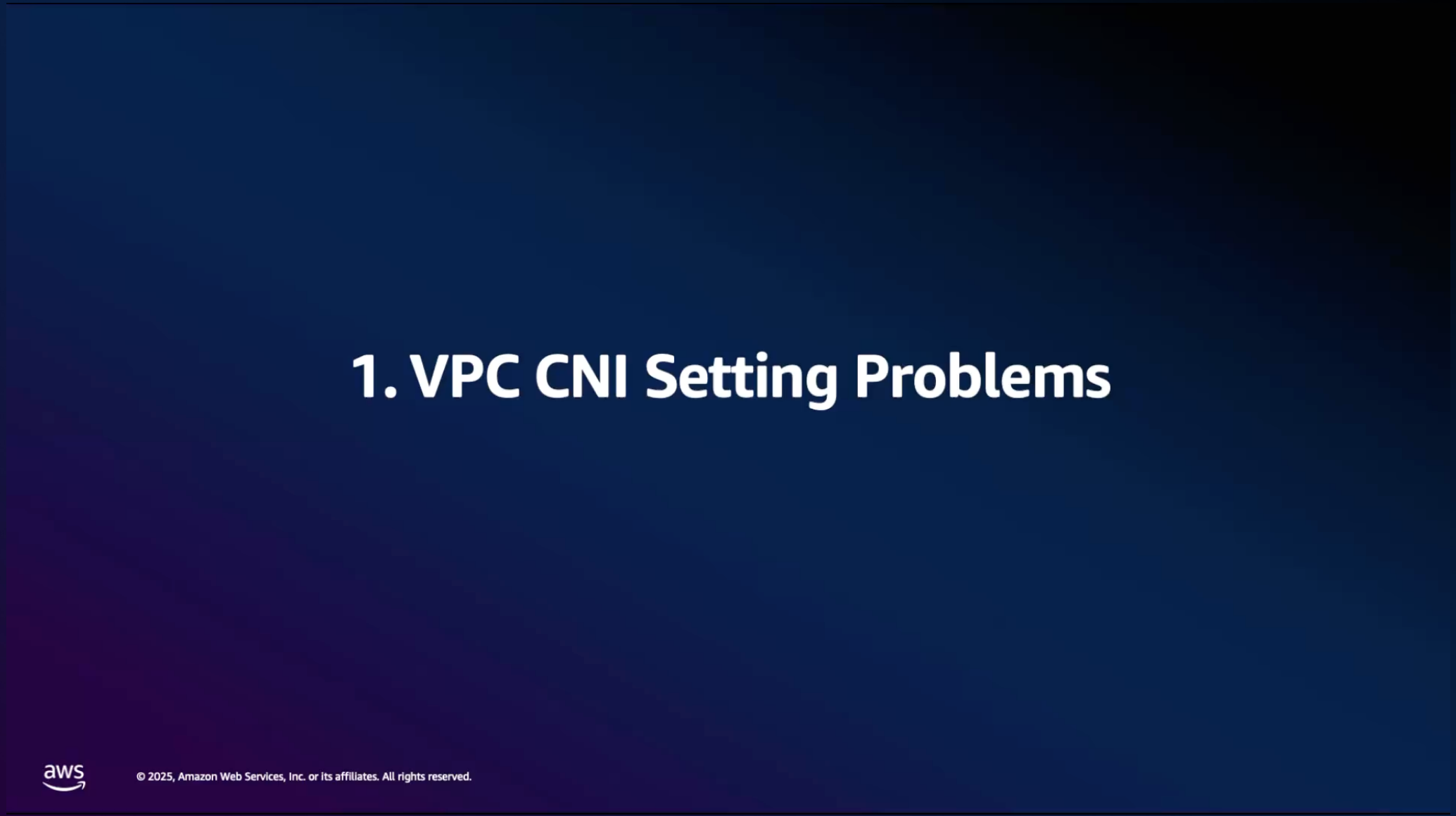
볼륨 detach가 안 된 상태로 강제 종료되어 새로운 노드에 볼륨 attach 시 에러 발생

Non-Graceful Spot Termination

1. AWS EC2 Spot instance 강제 종료 (알림 감지 X)
2. Pod가 날아감
3. 노드에 붙어있던 EBS Volume을 Detach 못 함
4. 새 노드에서 같은 Pod가 재시작 되려고 하지만 Volume Detach를 못 했기 때문에 AttachVolume failed 에러 발생

풀이 영상

https://youtu.be/nZi-cMwQqR8?si=BhGlhGYleXYpe-wr



Q&A

Thank you!

Youngseo Choi

(she/her)
Cloud Support Engineer Intern
Amazon Web Service