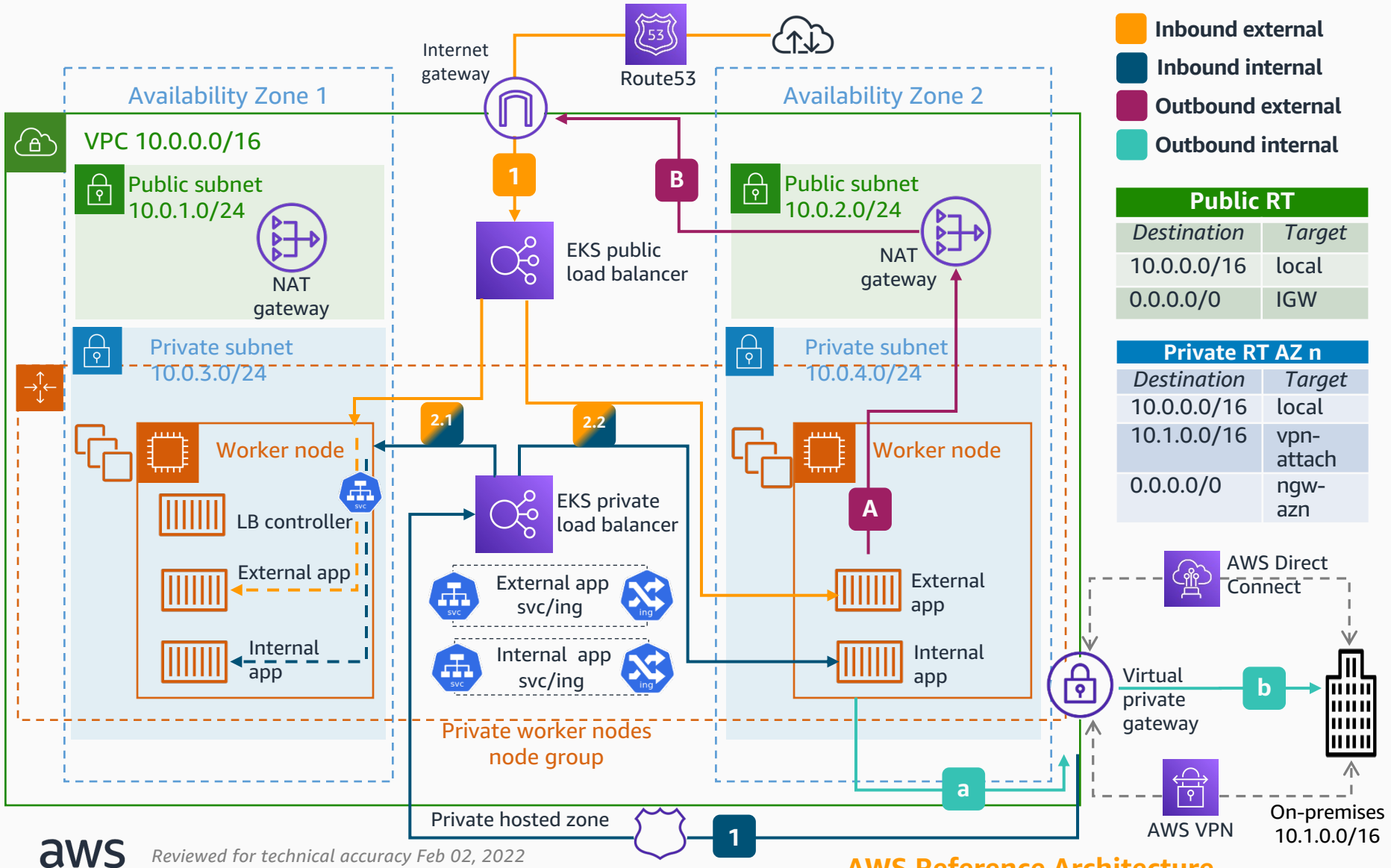
Expose Microservices in a Hybrid Scenario Using Amazon EKS

Expose Amazon Elastic Kubernetes Service (Amazon EKS) microservices hosted in private subnets to the internet and on-premises networks.



Public RT

Destination	*Target*
10.0.0.0/16	local
0.0.0.0/0	IGW

Private RT AZ n

Destination	*Target*
10.0.0.0/16	local
10.1.0.0/16	vpn-attach
0.0.0.0/0	ngw-azn

1 **Amazon Route 53** resolves incoming requests to the public **Elastic Load Balancer (ELB)** deployed by the **AWS Load Balancer Controller.***

1 **Amazon Route 53** resolves incoming requests to the private ELB deployed by the **AWS Load Balancer Controller.***

2 The ELBs forward traffic to applications. You can choose between the two modes**:

2.1 • Instance mode: The traffic is sent to a worker node, then the service redirects traffic to the pod.

2.2 • IP mode: The traffic is directed to the IP of the pod directly.

A When the pod in private subnets initiates an outbound request to the internet, the private route table forwards the traffic to the **NAT gateway** (NGW).***

B The public route table forwards the traffic from the NGW to the **internet gateway** (IGW).

a The pod in private subnets initiates an outbound request to the on-premises network. The private route table forwards the traffic to the **virtual private gateway** (VGW).***

b The traffic is sent to the on-premises network over the **virtual private network** (VPN) or **AWS Direct Connect** connection.

** Recommended way to manage Create, Read, Update, and Delete (CRUD) operations on **EKS**-related ELBs. The AWS LB controller satisfies k8s services with Network Load Balancers (NLBs) and Kubernetes ingresses with Application Load Balancers (ALBs). You can also manage ingresses by implementing other ingress controllers like the NGINX ingress controller.*

*** More information here.*

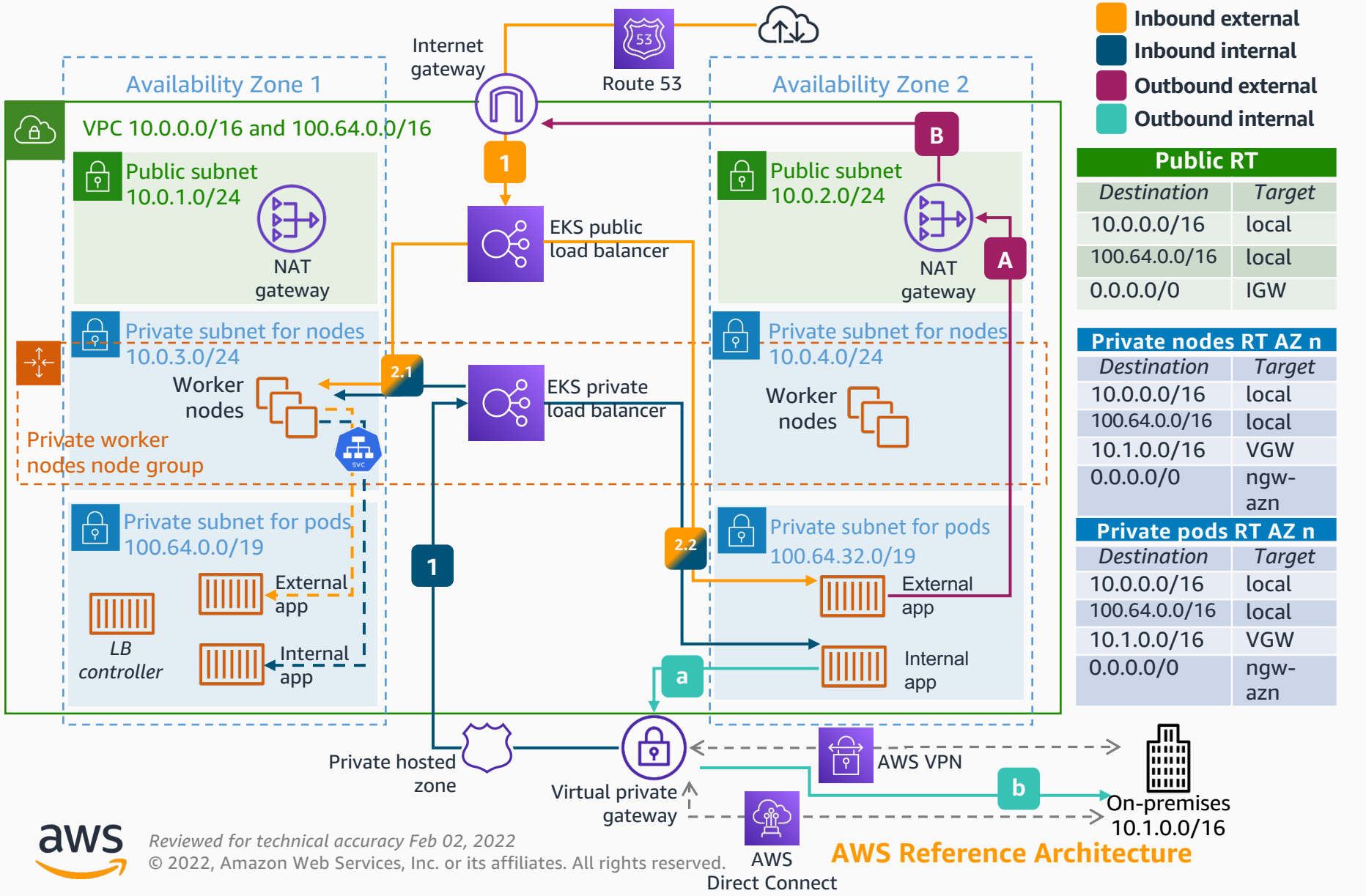
- *You can also enable private access for your **Amazon EKS** cluster's Kubernetes API server endpoint and limit, or completely disable, public access from the internet. More information here.*
- *If you're using AWS Fargate for **Amazon EKS**, you will not have worker nodes but only the pod ENIs in the private subnets. You can only use ELBs with IP mode with **AWS Fargate** pods.*

AWS Reference Architecture

Reviewed for technical accuracy Feb 02, 2022

Deal with Pod IP Exhaustion

Increase the IP addresses available to pods by adding dedicated subnets from the 100.64.0.0/10 and 198.19.0.0/16 ranges.*



Inbound external
Inbound internal
Outbound external
Outbound internal

Public RT

Destination	*Target*
10.0.0.0/16	local
100.64.0.0/16	local
0.0.0.0/0	IGW

Private nodes RT AZ n

Destination	*Target*
10.0.0.0/16	local
100.64.0.0/16	local
10.1.0.0/16	VGW
0.0.0.0/0	ngw-azn

Private pods RT AZ n

Destination	*Target*
10.0.0.0/16	local
100.64.0.0/16	local
10.1.0.0/16	VGW
0.0.0.0/0	ngw-azn

1 **Amazon Route 53** resolves incoming requests to the public ELB deployed by the **AWS Load Balancer Controller**.

1 **Amazon Route 53** resolves incoming requests to the private ELB deployed by the **AWS Load Balancer Controller**.

2 The ELBs forward traffic to applications. You can choose between two modes:

2.1 • Instance mode: The traffic to a worker node and then the service will redirect traffic to the pod.

2.2 • IP mode: The traffic is directed to the IP of the pod directly.

A The pod in private subnets initiates an outbound request to the internet. The private route table forwards the traffic to the **NAT gateway** (NGW).***

B The public route table forwards the traffic from the NGW to the **internet gateway** (IGW).

a The pod in private subnets initiates an outbound request to the on-premises network. The private route table forwards the traffic to the **Virtual Private gateway** (VGW).***

b The traffic is sent on-premises network over the VPN/**Direct Connect** connections.

** By adding secondary CIDR blocks to a VPC from the RFC 6598 address space (in the example 100.64.0.0/16), in conjunction with the CNI Custom Networking feature, it is possible for pods to no longer consume any RFC 1918 IP addresses in a VPC (in the example, pods are in subnets 100.64.0.0/19 and 100.64.32.0/19). Check out this post for a technical how-to.*

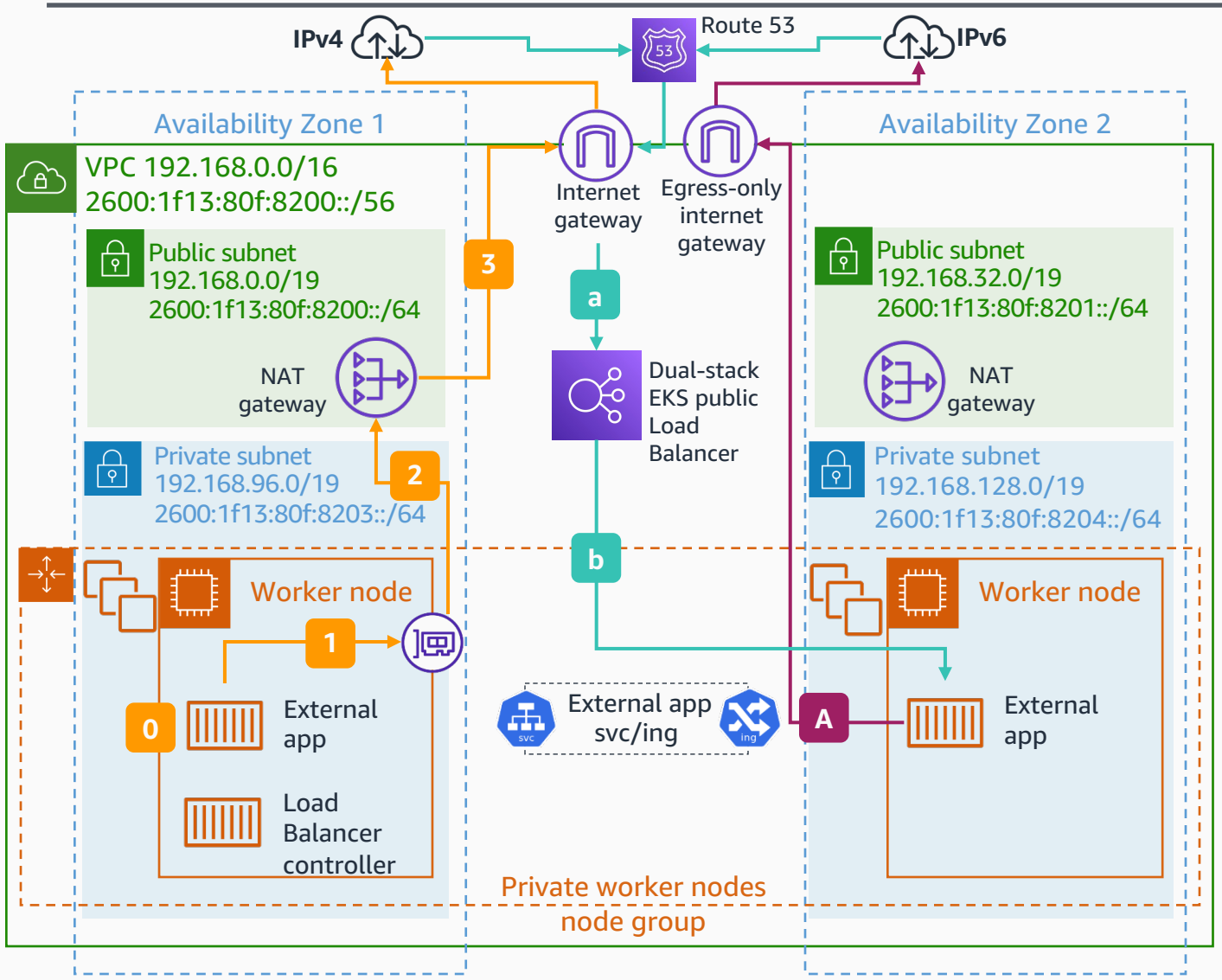
*** More information here.*

- *The default behavior of EKS is to source NAT pod traffic to the primary IP address of the hosting worker node.*
- *Check out this blog for multi-account settings.*
- *AWS Fargate for Amazon EKS supports additional CIDRs.*
- *The ENIConfig custom resource is used to define the subnet in which pods will be scheduled.*

AWS Reference Architecture

Expose Amazon EKS Microservices in IPv6 Clusters

Expose Amazon EKS microservices with IPv6 and connect to both IPv6 and IPv4 endpoints on the internet.



Inbound

Outbound to IPv6

Outbound to IPv4

Public RT

Destination	*Target*
192.168.0.0/16	local
0.0.0.0/0	IGW
2600:1f13:80f:8200::/56	local
::/0	IGW

Private RT AZ n

Destination	*Target*
192.168.0.0/16	local
0.0.0.0/0	ngw-azn
64:ff9b::/96	ngw-azn
2600:1f13:80f:8200::/56	local
::/0	eigw

a **Amazon Route 53** resolves incoming requests to the public ELBs in dual-stack mode* deployed by the **AWS Load Balancer controller.****

b The ELB forwards traffic to the IPv6 pods (the ELB must use the IP mode).

A Any pod communication from within private subnets to IPv6 endpoints outside the cluster will be routed via an **egress-only internet gateway** (EIGW).

0 A pod in private subnet initiates an outbound request to an IPv4 address on the internet and performs a DNS lookup for an endpoint and, upon receiving an IPv4 "A" response, establishes a connection with the IPv4 endpoint using the IPv4 address from the host-local 169.254.172.0/22 IP range***.

1 The pod's node-only unique IPv4 address is translated through NAT to the IPv4 (VPC) address of the primary network interface attached to the node.

2 The private route table forwards the traffic to the NGW, and the private IPv4 address of a node is translated by a NAT gateway to the public IPv4 address of the gateway.

3 The public route table forwards the traffic from the NGW to the IGW.

** At the time of this writing, ALB and NLB support dual-stack for only internet-facing endpoints. More information on the ELB annotation here.*
*** The legacy in-tree service controller does not support IPV6*
****EKS implements a host-local CNI plugin chained along with VPC CNI to allocate and configure an IPv4 address for a pod. The CNI plugin configures a host-specific non-routable IPv4 address for a pod from the 169.254.172.0/22 range.*

- *Moving to IPV6 also solves pod IP exhaustion, because you don't need to work around IPv4 limits.*

AWS Reference Architecture