IPv6 & Kubernetes, Public Cloud

Pieter Lewyllie

Systems Engineer @ Cisco

Belgian IPv6 Council 2019

IPv6 & containers...

APNIC

IPv6 and containers – a horror story

By **Matt Palmer** on 22 Mar 2018

THENEWSTACK **Ebooks** **Podcasts** **Events** **Newsletter**

| Architecture | Development | Operations |

KUBERNETES

Kubernetes Warms Up to IPv6

25 Feb 2019 11:55am, by Mary Branscombe

Disclaimer

This is my interpretation of current state of things.

I do not have a stake in the public cloud providers presented here.

Kubernetes

- Container orchestrator
- Runs and manages containers
- Supports multiple cloud and bare-metal environments
- Inspired and informed by Google's experiences and internal systems
- 100% Open source, written in Go
- Manage applications, not machines
- Rich ecosystem of plug-ins for scheduling, storage, networking

Kubernetes Master Node(s)

Container
Pod
Kubelet   Kube-Proxy
Worker Node(s)
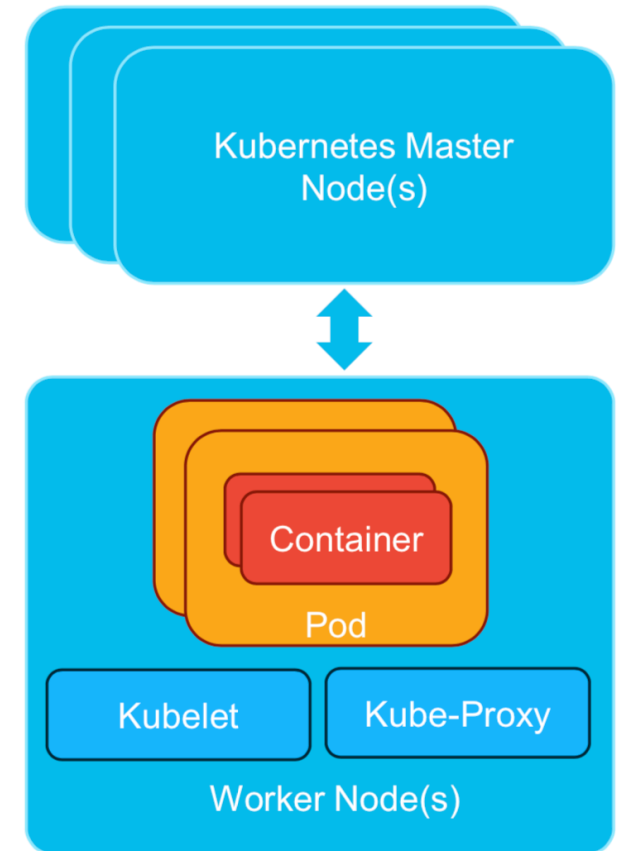
Nodes, Pods, Containers

- Node:
  - A server
- Cluster:
  - Collection of nodes
- Pod:
  - Collection of containers;
  - Nodes can run multiple Pods
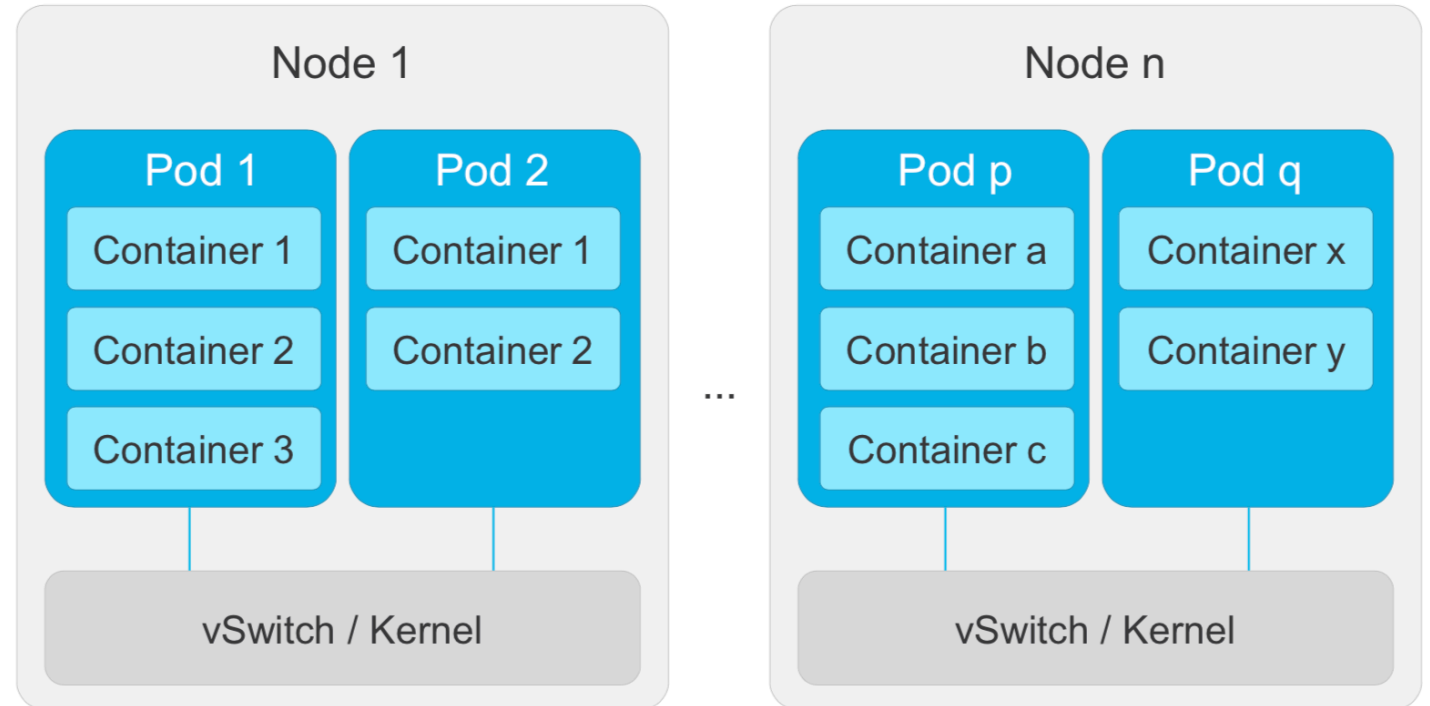
Services overview

- "Pods can come and go, services stay"
- Define a single IP/Port combination that provides access to a pool of pods
- By default a service connects the client to a Pod in a round- robin fashion
- This solves the dilemma of having to keep up with every transient IP address assigned by Docker
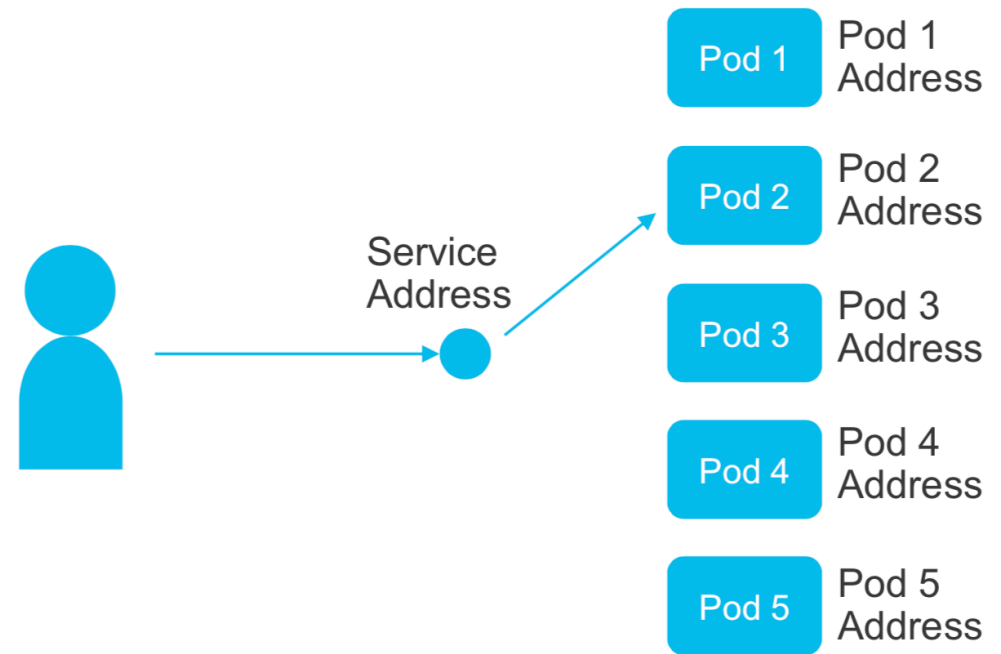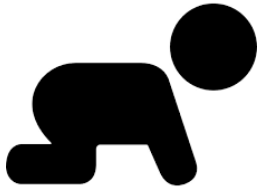
Why IPv6?

- Cleaner
- Easier diagnosis
- We need lots of IPs
- Not easy to find remaining IPv4 space in organization
- Multi cluster
- VNFs: Mobile packet core, 5G…
- IoT

IPv6 in Kubernetes

- IPv4 Parity, no API Changes
- CNI 0.6.0
  - Bridge & Host-Local IPAM
- ip6tables & ipvs
- kubeadm

Rel 1.9 (Alpha)

- Moving to CoreDNS

Rel 1.13

- Phase 1 of dual-stack KEP
- Multiple IPs per pod

Rel 1.15 (targeting)

- Phase 2 of dual-stack KEP
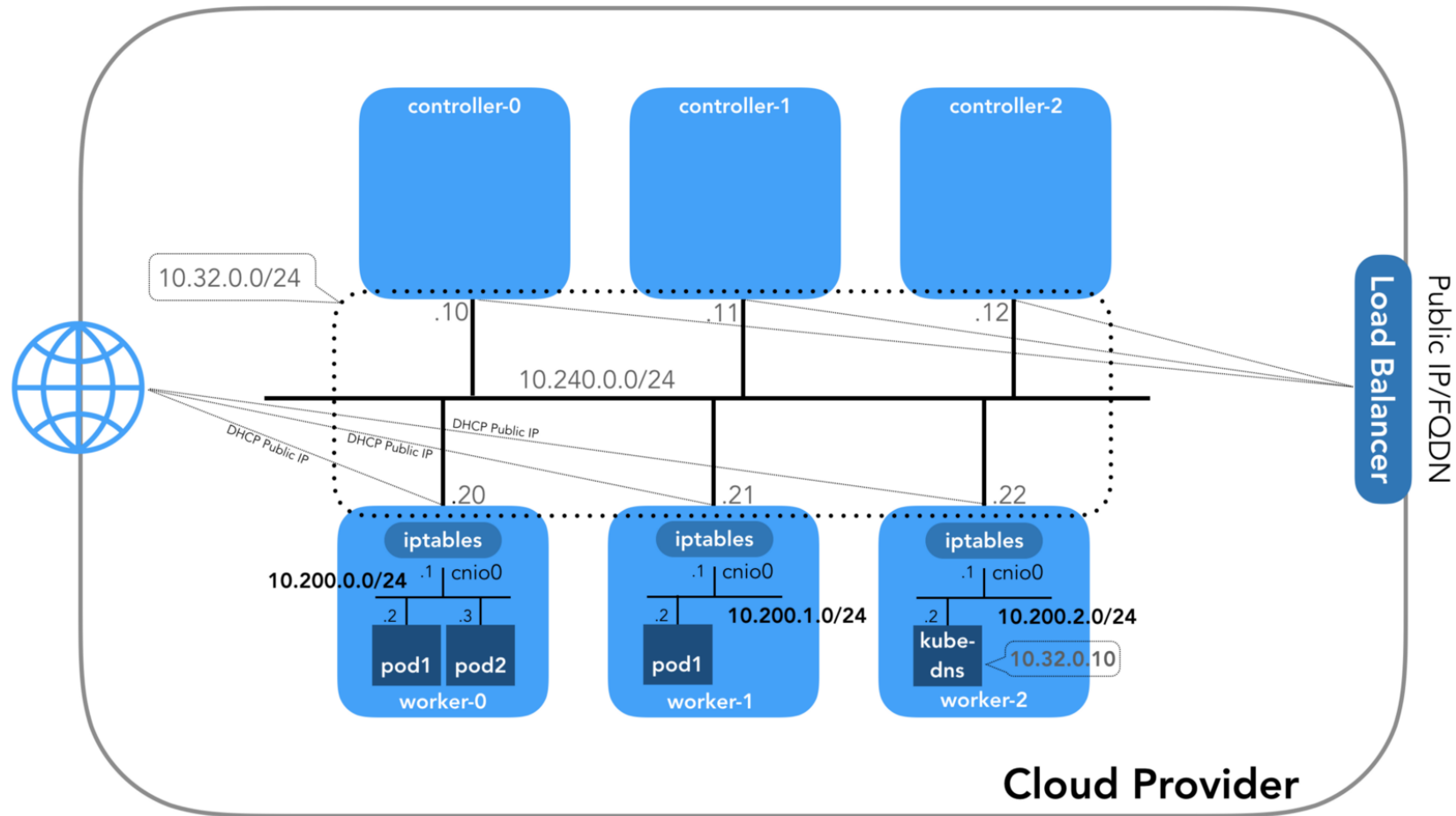- SRv6
- Dual-stack service CIDRs
- Istio IPv6
- …

Planning and Preparing

Original slide source: SRv6LB @ Kubecon https://www.youtube.com/watch?v=RRKUeyFaqEA

Dual stack KEP:
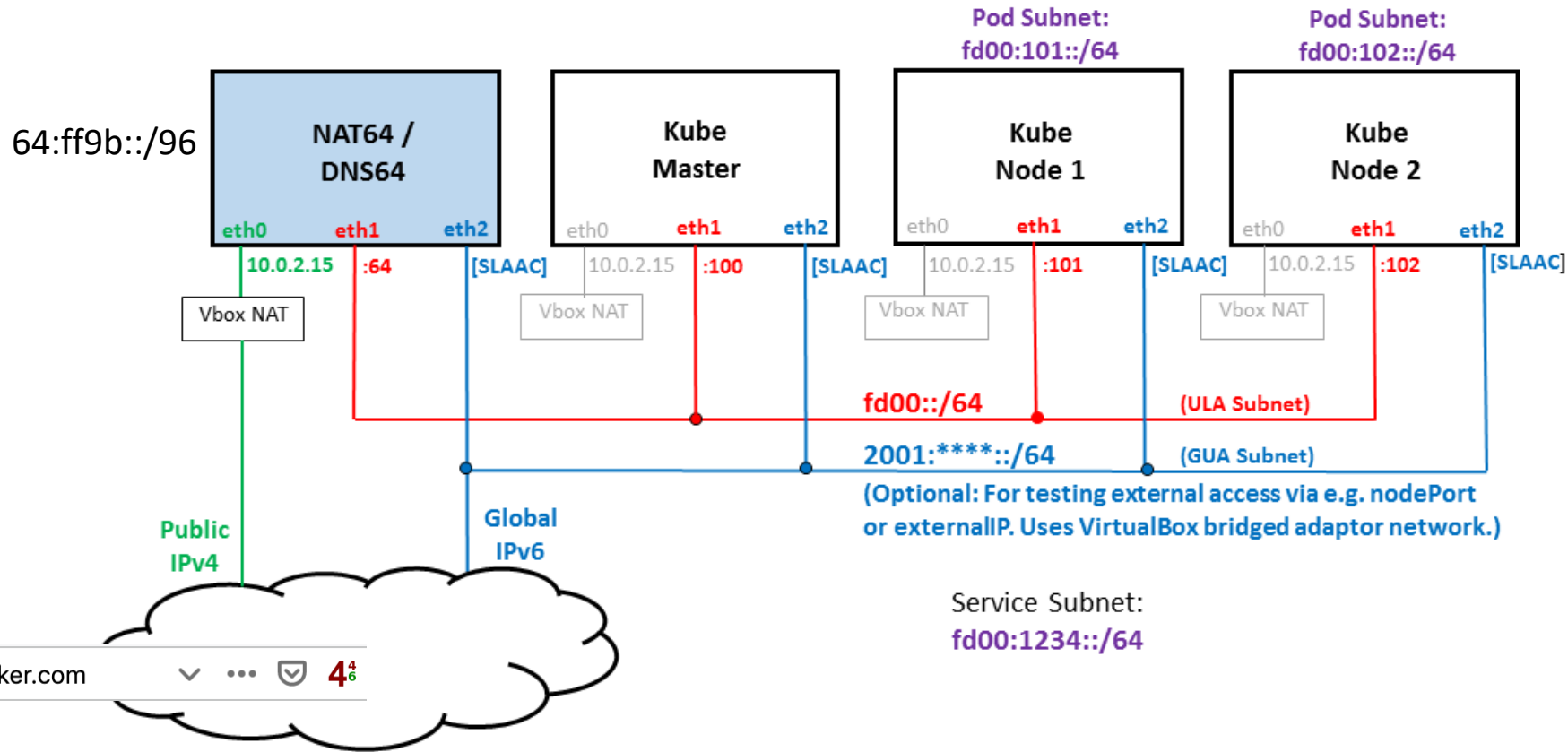https://github.com/kubernetes/enhancements/blob/master/keps/sig-network/20180612-ipv4-ipv6-dual-stack.md#implementation-plan

IPv4 Kubernetes

Multi-node, IPv6-only K8 cluster

Example "IPv6-Only" Topology (Using VirtualBox)

Guide: https://github.com/leblancd/kube-v6
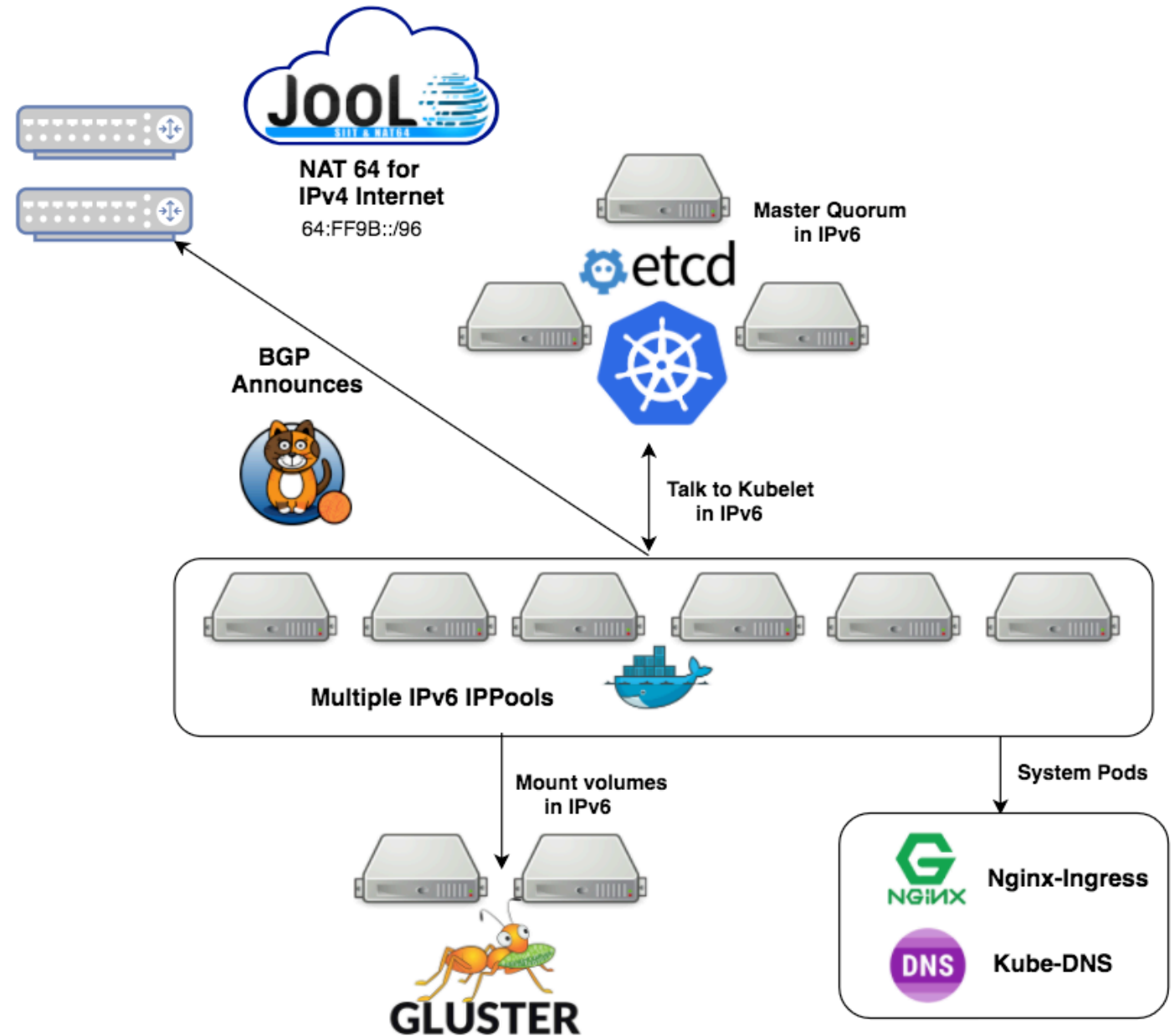
Container Network Interface (CNI)

- Proposed by CoreOS as part of appc specification

- Common interface between container run time and network plugin

- Gives driver freedom to manipulate network namespace

- Network described by JSON config

- Many CNI plugins available:
  - Calico, Flannel, Weave, Contiv...
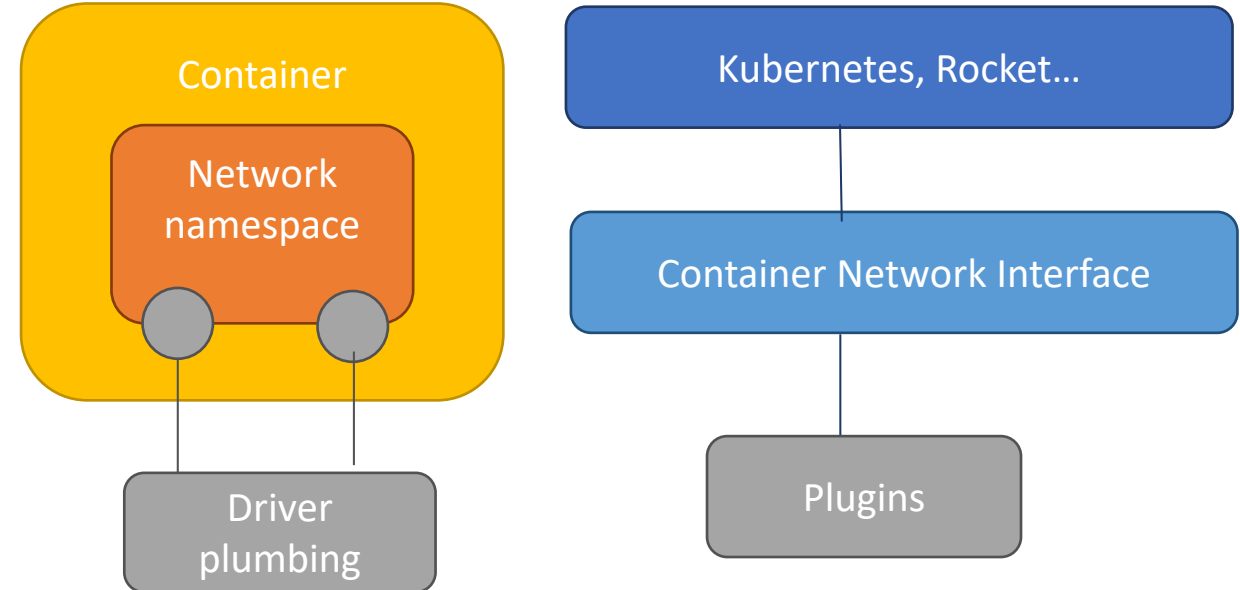
CNI: Calico

- Pure L3 networking with BGP
- IPv6 only clusters
- ULA range by default for PODs
- By default breaks into /122 per node
- **clusterIP: None** on every defined Service
- Nginx-ingress controller

https://opsnotice.xyz/kubernetes-ipv6-only/
https://www.projectcalico.org/enable-ipv6-on-kubernetes-with-project-calico/

IPv6 CNI

- Flannel
  - No IPv6 support
- Contiv-VPP
  - IPv6 only
  - SRv6
  - https://github.com/contiv/vpp
- Cilium
  - IPv6

Multi-cluster IPv4
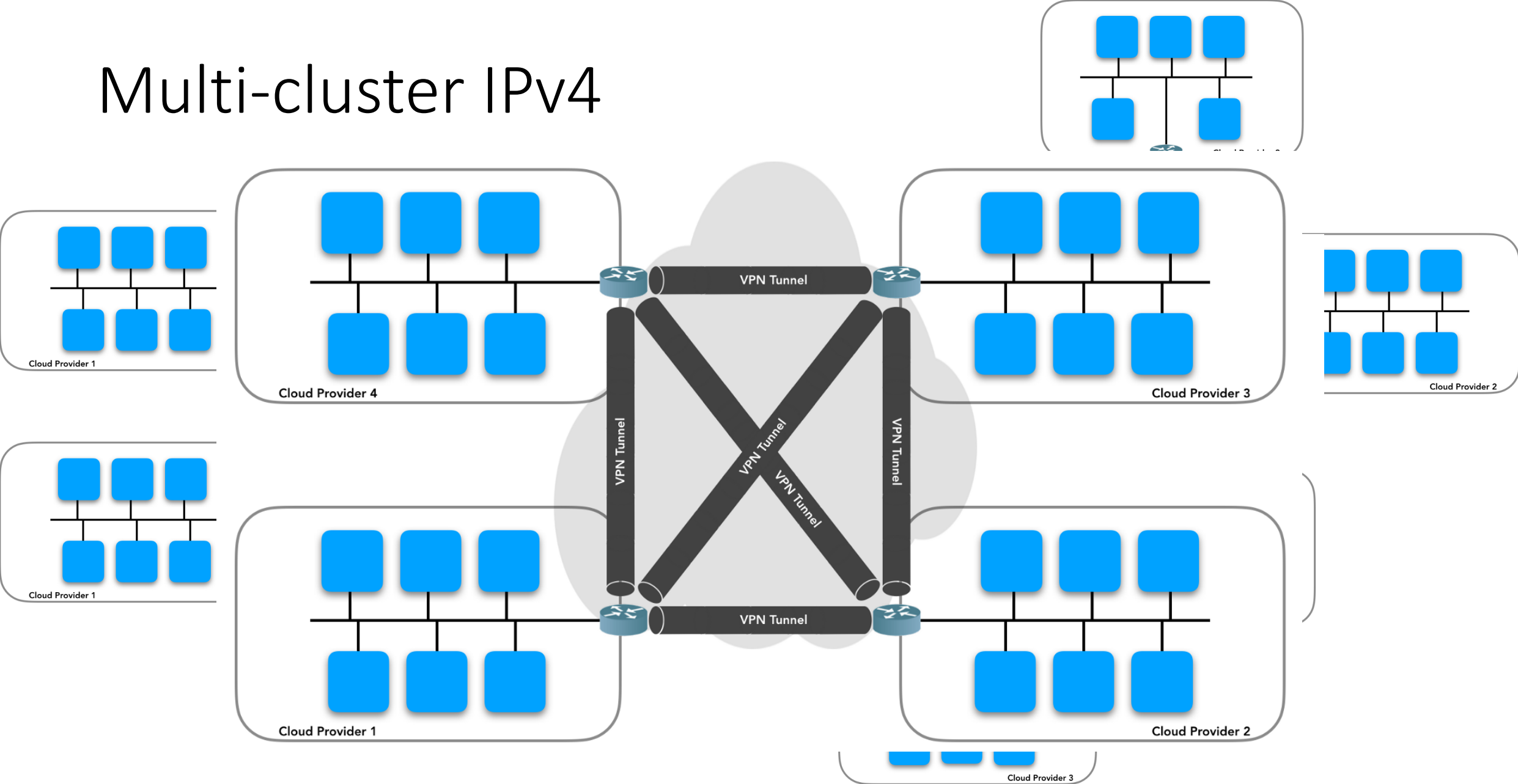
Source: https://itnext.io/kubernetes-multi-cluster-networking-made-simple-c8f26827813 from Nicolas Leiva

Multi-cluster IPv6
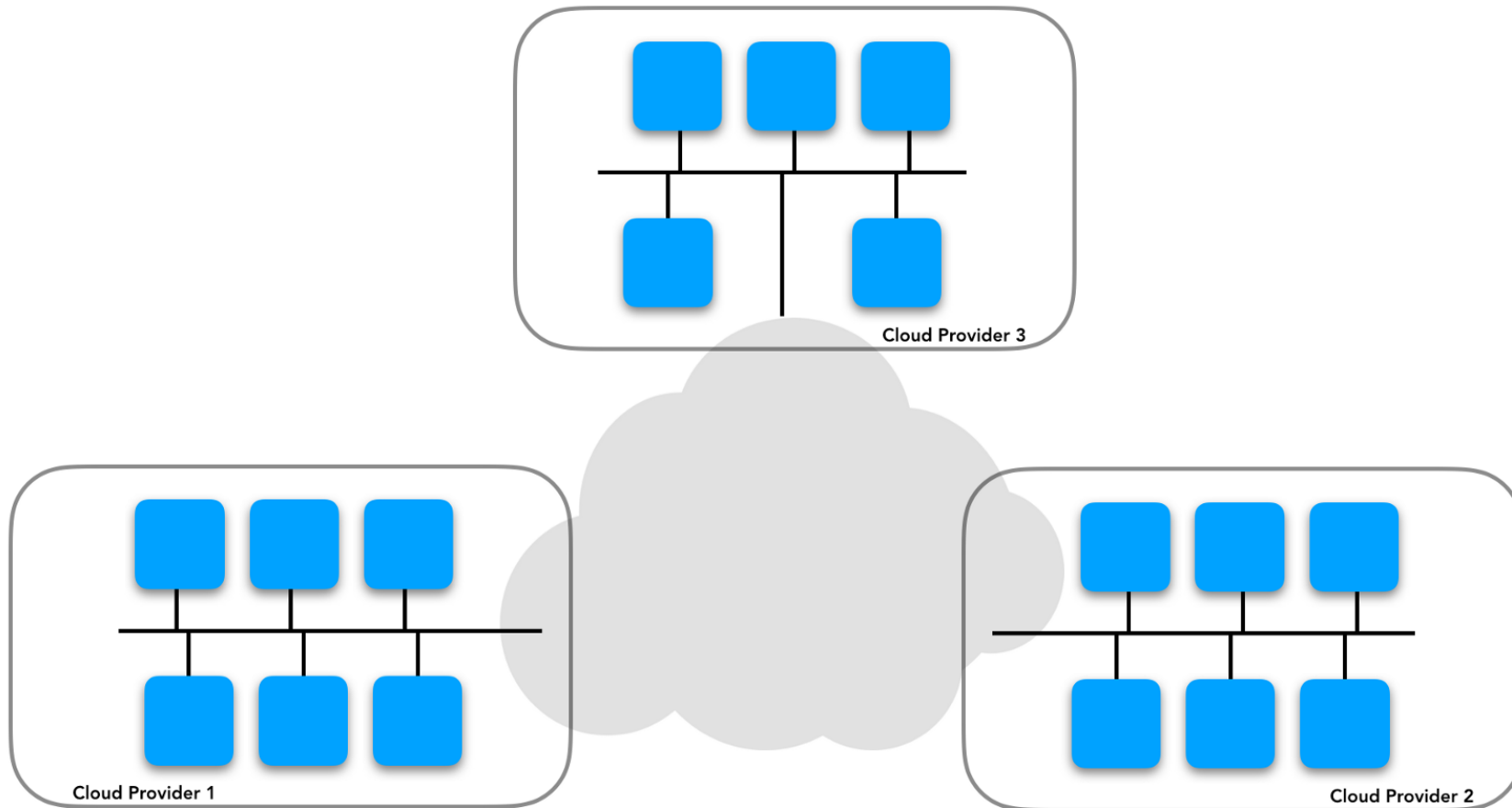
What about the public cloud?

- GCE/GKE does not have IPv6 support
  - VPC networks only support IPv4 unicast traffic. They do not support broadcast, multicast, or IPv6 traffic within the network.
  - Can use IPv6 with load-balancing:
    - https://cloud.google.com/compute/docs/load-balancing/ipv6
- Azure
  - NEW: IPv6 for VNets in public preview https://azure.microsoft.com/en-us/updates/public-preview-microsoft-adds-full-ipv6-support-for-azure-vnets/
  - No IPv6 on AKS
  - IPv6 load-balancer:
    - https://docs.microsoft.com/en-us/azure/load-balancer/load-balancer-ipv6-overview
  - Long list of limitations:
    - A single IPv6 address can be assigned to a single network interface in each VM.
    - The load balancer routes the IPv6 packets to the private IPv6 addresses of the VMs using network address translation (NAT).
    - Azure VMs cannot connect over IPv6 to other VMs, other Azure services, or on-premises devices. They can only communicate with the Azure load balancer over IPv6. However, they can communicate with these other resources using IPv4.
- Amazon
  - No support for IPv6 on EKS
  - Should work with EC2 instances
  - Each VPC is given a unique /56 address prefix from within Amazon's GUA (Global Unicast Address); you can assign a /64 address prefix to each subnet in your VPC
  - Maximum amount of IPv6 addresses per interface: https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/using-eni.html#AvailableIpPerENI

Azure

```
pieter@dsVM1:~$ ip a
1: lo: <LOOPBACK,UP,LOWER_UP> mtu 65536 qdisc noqueue state UNKNOWN group default qlen 1000
    link/loopback 00:00:00:00:00:00 brd 00:00:00:00:00:00
    inet 127.0.0.1/8 scope host lo
       valid_lft forever preferred_lft forever
    inet6 ::1/128 scope host
       valid_lft forever preferred_lft forever
2: eth0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc mq state UP group default qlen 1000
    link/ether 00:0d:3a:1c:92:72 brd ff:ff:ff:ff:ff:ff
    inet 10.0.0.5/24 brd 10.0.0.255 scope global eth0
       valid_lft forever preferred_lft forever
    inet6 ace:cab:deca:deed::5/128 scope global dynamic noprefixroute
       valid_lft 17279311sec preferred_lft 8639311sec
    inet6 fe80::20d:3aff:fe1c:9272/64 scope link
       valid_lft forever preferred_lft forever
```

IPv4 Default Gatew
IPv4 DHCP Server
IPv4 DNS Server
IPv4 WINS Server
NetBIOS over Tcpi
IPv6 Address
Lease Obtained
Lease Expires
Link-local IPv6 Address
IPv6 Default Gateway
IPv6 DNS Server

r the Help Desk

ℹ Your Internet Service Provider (ISP) appears to be Microsoft Corporation

ℹ Since you have IPv6, we are including a tab that shows how well you can reach other IPv6 sites. [more info]

ℹ HTTPS support on this web site is in *beta*. [more info]

✅ Your DNS server (possibly run by your ISP) appears to have IPv6 Internet access.

Your readiness score

10/10   for your IPv6 stability and readiness, when publishers are forced to go IPv6 only

Click to see Test Data

(Updated server side IPv6 readiness stats)

Azure (1)

Public Preview: Microsoft adds full IPv6 support for Azure VNets

Posted on Tuesday, April 23, 2019

IPv6 is used by:

- Government or customers in highly regulated industries

- Internet-of-Things applications that leverage the enormous number of IPv6 addresses available

- Customers who want native (not translated) connectivity from their Azure applications to IPv6 mobile devices

- Customers who deploy service instances in Azure with dedicated per-customer connectivity

- Customers moving to the cloud (or hybrid) struggling with exhaustion of their IPv4 space and inability/expense of acquiring more IPv4.

> ⓘ **Important**
>
> IPv6 for Azure Virtual Network is currently in public preview. This preview is provided without a service level agreement and is not recommended for production workloads. Certain features may not be supported or may have constrained capabilities. See the **Supplemental Terms of Use for Microsoft Azure Previews** for details.

https://docs.microsoft.com/en-us/azure/virtual-network/ipv6-overview

Azure (3)

Create a virtual network

Create a virtual network with az network vnet create. The following example creates a virtual network named *dsVNET* with subnets *dsSubNET_v4* and *dsSubNET_v6*:

```
Azure CLI                                                    Copy

# Create the virtual network
az network vnet create \
--name dsVNET \
--resource-group DsResourceGroup01 \
--location eastus  \
--address-prefixes "10.0.0.0/16" "ace:cab:deca::/48"

# Create a single dual stack subnet

az network vnet subnet create \
--name dsSubNET \
--resource-group DsResourceGroup01 \
--vnet-name dsVNET \
--address-prefixes "10.0.0.0/24" "ace:cab:deca:deed::/64" \
--network-security-group dsNSG1
```

AWS

- Creates a VPC with a /16 IPv4 CIDR block and associates a /56 IPv6 CIDR block with the VPC. For more information, see Your VPC. The size of the IPv6 CIDR block is fixed (/56) and the range of IPv6 addresses is automatically allocated from Amazon's pool of IPv6 addresses (you cannot select the range yourself).
- Attaches an Internet gateway to the VPC. For more information about Internet gateways, see Internet Gateways.
- Creates a subnet with an /24 IPv4 CIDR block and a /64 IPv6 CIDR block in the VPC. The size of the IPv6 CIDR block is fixed (/64).
- Creates a custom route table, and associates it with your subnet, so that traffic can flow between the subnet and the Internet gateway. For more information about route tables, see Route Tables.

Based on DHCPv6 (actual implementation stateless and derived from topology DB)

Instance Type	Maximum Network Interfaces	IPv4 Addresses per Interface	IPv6 Addresses per Interface
a1.medium	2	4	4
a1.large	3	10	10
a1.xlarge	4	15	15
a1.2xlarge	4	15	15
a1.4xlarge	8	30	30

https://docs.aws.amazon.com/vpc/latest/userguide/get-started-ipv6.html
https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/using-eni.html

AWS (1)

Custom route table

Destination	Target
10.0.0.0/16	local
2001:db8:1234:1a00::/56	local
0.0.0.0/0	*igw-id*
::/0	*igw-id*

Subnet
10.0.0.0/24 (IPv4)
2001:db8:1234:1a00::/64 (IPv6)

Availability Zone A

Router Internet Gateway Internet

Availability Zone B

VPC
10.0.0.0/16 (IPv4)
2001:db8:1234:1a00::/56 (IPv6)

Region

Main route table

Destination	Target
10.0.0.0/16	local
2001:db8:1234:1a00::/56	local

AWS (2)

```
[PILEWYLL-M-L2F5:Downloads pilewyll$ ssh -i pieterkeypair.pem ec2-user@2a05:d014:fca:a300:f50:cbeb:2b42:a956
 Last login: Wed Jun 19 17:49:41 2019 from 2001:420:c0c0:1001::2b8


       __|  __|_  )
       _|  (     /   Amazon Linux 2 AMI
      ___|\___|___|

https://aws.amazon.com/amazon-linux-2/
No packages needed for security; 1 packages available
Run "sudo yum update" to apply all updates.
[[ec2-user@ip-10-0-0-92 ~]$ ip a
1: lo: <LOOPBACK,UP,LOWER_UP> mtu 65536 qdisc noqueue state UNKNOWN group default qlen 1000
    link/loopback 00:00:00:00:00:00 brd 00:00:00:00:00:00
    inet 127.0.0.1/8 scope host lo
       valid_lft forever preferred_lft forever
    inet6 ::1/128 scope host
       valid_lft forever preferred_lft forever
2: eth0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 9001 qdisc pfifo_fast state UP group default qlen 1000
    link/ether 0a:e9:14:77:97:42 brd ff:ff:ff:ff:ff:ff
    inet 10.0.0.92/24 brd 10.0.0.255 scope global dynamic eth0
       valid_lft 2469sec preferred_lft 2469sec
    inet6 2a05:d014:fca:a300:f50:cbeb:2b42:a956/64 scope global dynamic
       valid_lft 377sec preferred_lft 77sec
    inet6 fe80::8e9:14ff:fe77:9742/64 scope link
       valid_lft forever preferred_lft forever
```

Where do I track the latest?

- https://github.com/kubernetes/enhancements/issues/508

- https://github.com/kubernetes/enhancements/issues/563

- https://github.com/kubernetes/enhancements/blob/master/keps/sig-network/20180612-ipv4-ipv6-dual-stack.md

- https://discuss.kubernetes.io/t/kubernetes-ipv4-ipv6-dual-stack-support-status/4974

- #k8s-dual-stack channel on Kubernetes.slack.com

- Attending IPv6 Council ☺

Thanks!