ANTREA

Antrea Introduction

Vicky Liu, VMware

What’s Antrea

Kubernetes Cluster Networking

Focuses on and optimized for K8s networking

Built with K8s technologies. K8s’ way to operate.

Supports all OSes and compute platforms K8s runs on.

Open vSwitch as Data Plane

Flexible; enables fast feature development

Supports Linux and Windows

Good performance

Rich set of troubleshooting mechanisms

Strong community support

Open Source

Public on Nov 18, 2019
https://github.com/vmware-tanzu/antrea
Apache 2.0 license

Open governance

Leverages existing open source

ANTREA

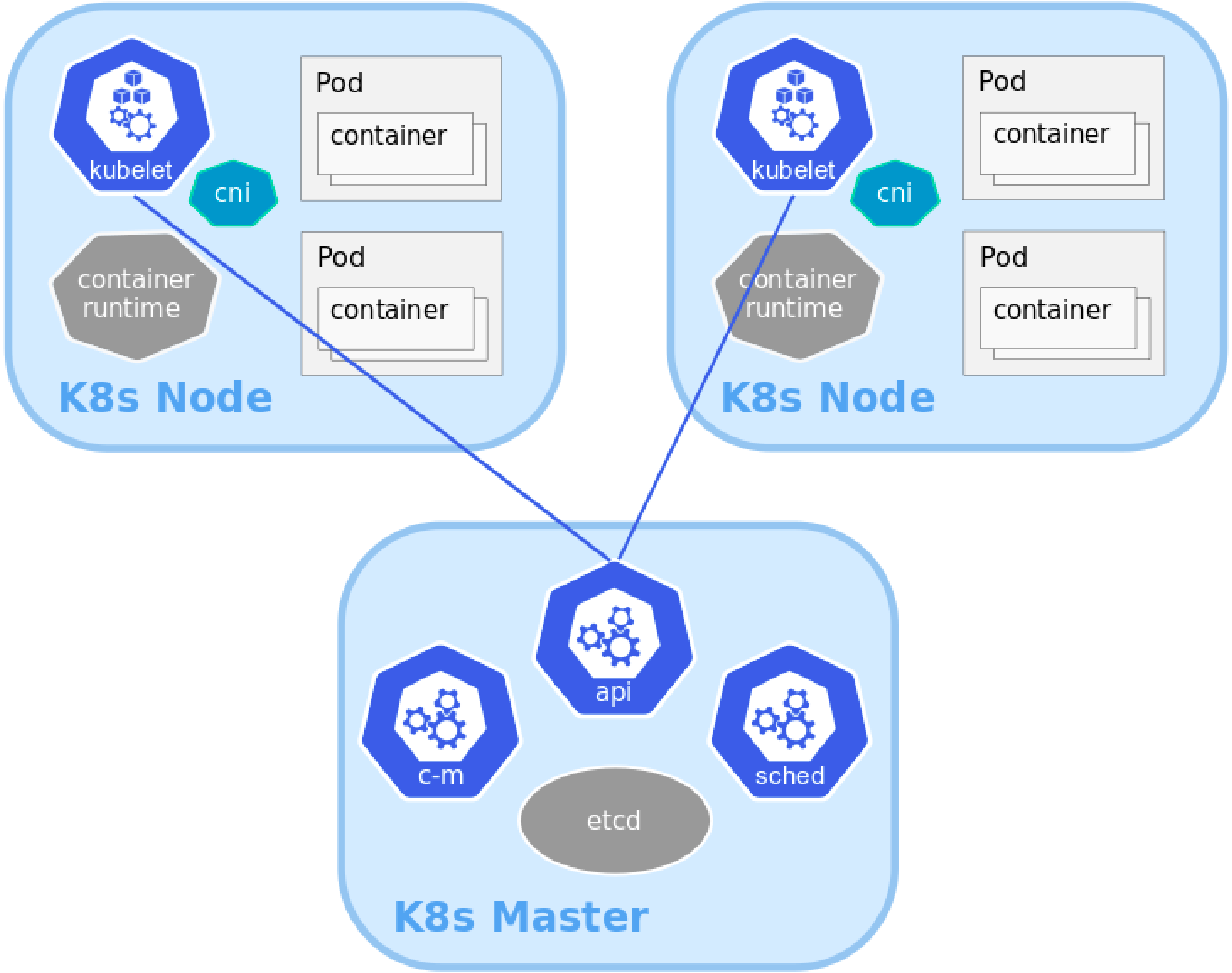
Kubernetes Networking Introduction

Pod

- *"The basic execution unit of a Kubernetes application"*
- Each Pod has its own IP address.
- CNI (Container Network Interface) plugin is responsible for allocating Pod IPs and configuring network interfaces.

Network connectivity assumptions

- *“pods on a node can communicate with all pods on all nodes without NAT”*



Kubernetes Networking Introduction

- **Container Network Interface**
  - ADD

    Add pod to network.

  - DELETE

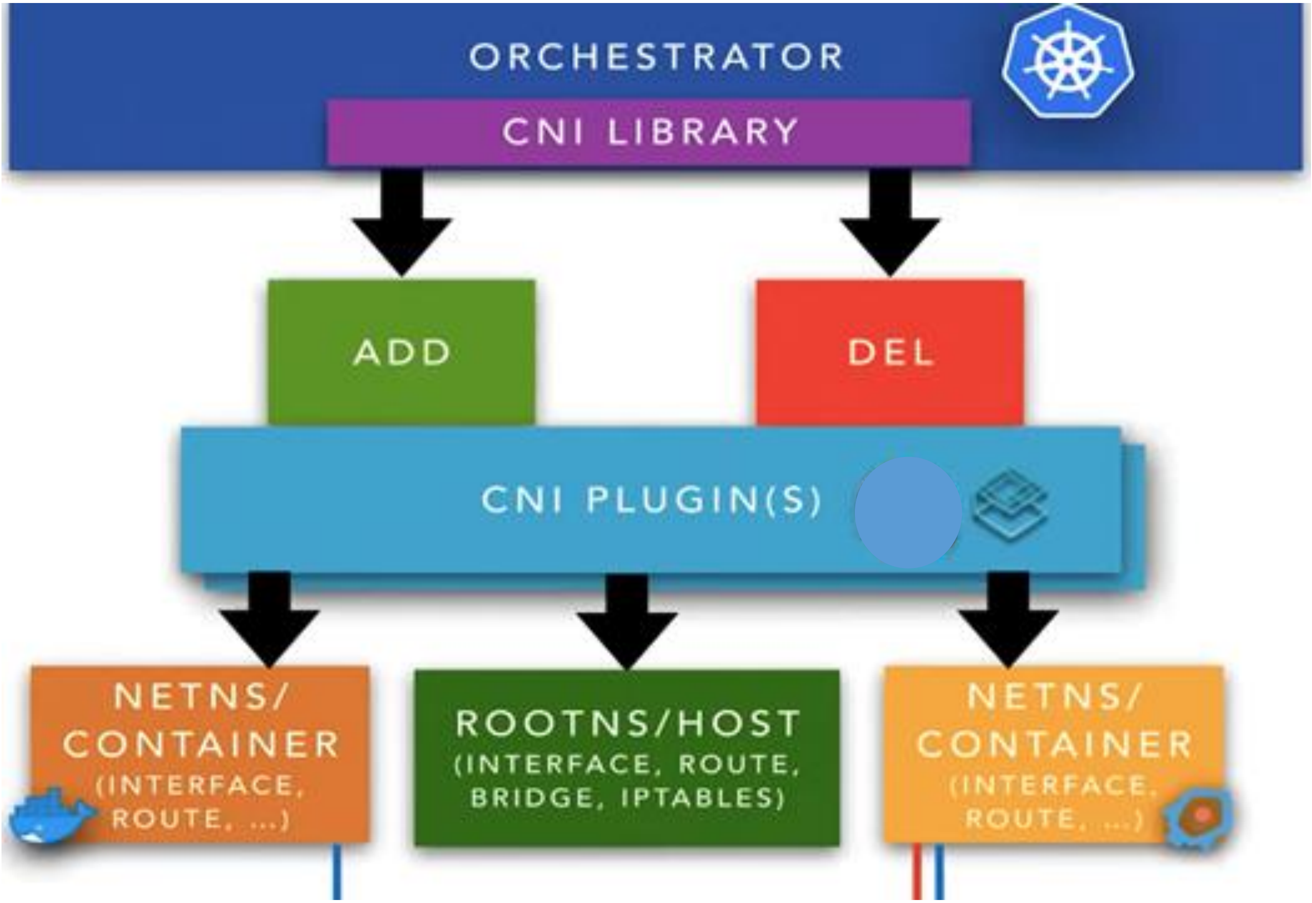
    Remove pod from network.

  - VERSION

    Get the CNI spec versions supported by the plugin

  - CHECK

    Check pod's networking is as expected


ORCHESTRATOR
CNI LIBRARY
ADD
DEL
CNI PLUGIN(S)
NETNS/ CONTAINER (INTERFACE, ROUTE, ...)
ROOTNS/HOST (INTERFACE, ROUTE, BRIDGE, IPTABLES)
NETNS/ CONTAINER (INTERFACE, ROUTE, ...)

Kubernetes Networking Introduction (cont.)

NetworkPolicy

- *"a specification of how groups of pods are allowed to communicate with each other and other network endpoints".*

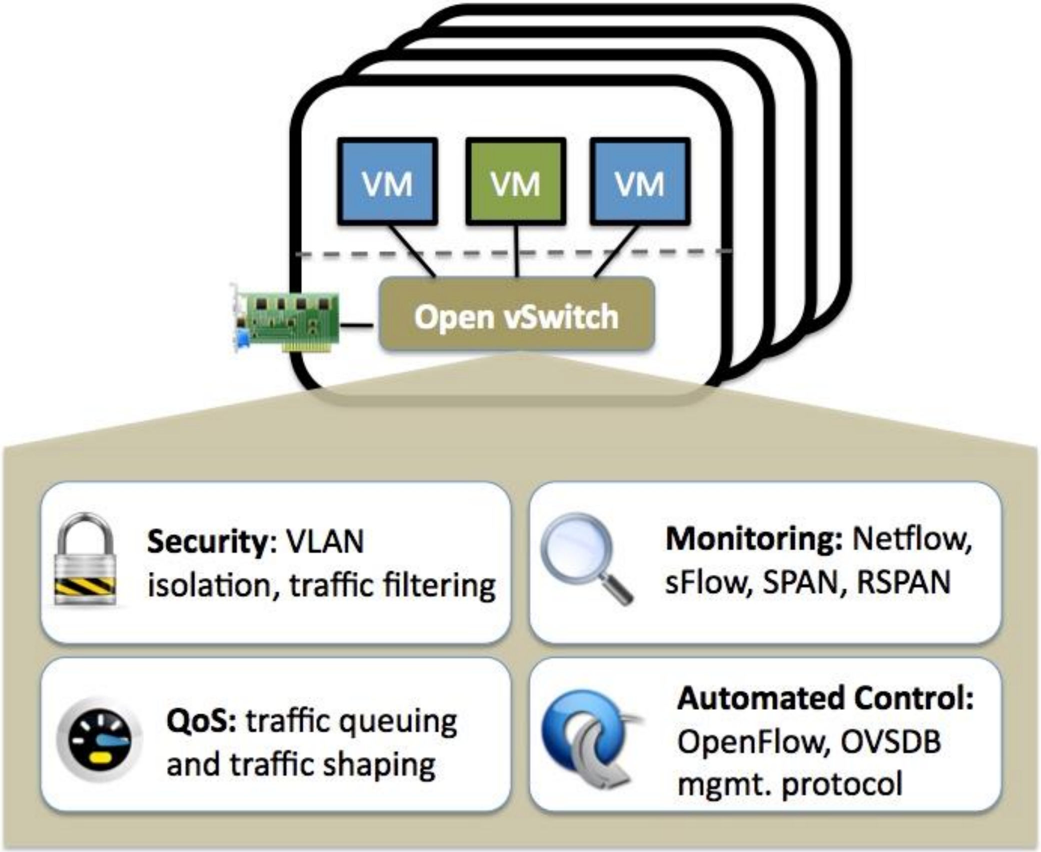
Service

- *"An abstract to expose an application running on a set of Pods"*
- Inside the Cluster exposed via a ClusterIP (VIP allocated by K8s).
  - kube-proxy implements the VIP and distributes the traffic to the Service's backend Pods.
  - kube-proxy has implementations with: iptables, IPVS, and a userspace proxy.

Open vSwitch

Open vSwitch is a production quality, multilayer virtual switch licensed under the open source Apache 2.0 license.

It is designed to enable massive network automation through programmatic extension, while still supporting standard management interfaces and protocols.


VM
VM
VM
Open vSwitch
Security: VLAN isolation, traffic filtering
Monitoring: Netflow, sFlow, SPAN, RSPAN
QoS: traffic queuing and traffic shaping
Automated Control: OpenFlow, OVSDB mgmt. protocol

Antrea technical overview

Supports K8s cluster networking:

- Pod network, NetworkPolicy, Service ClusterIP.

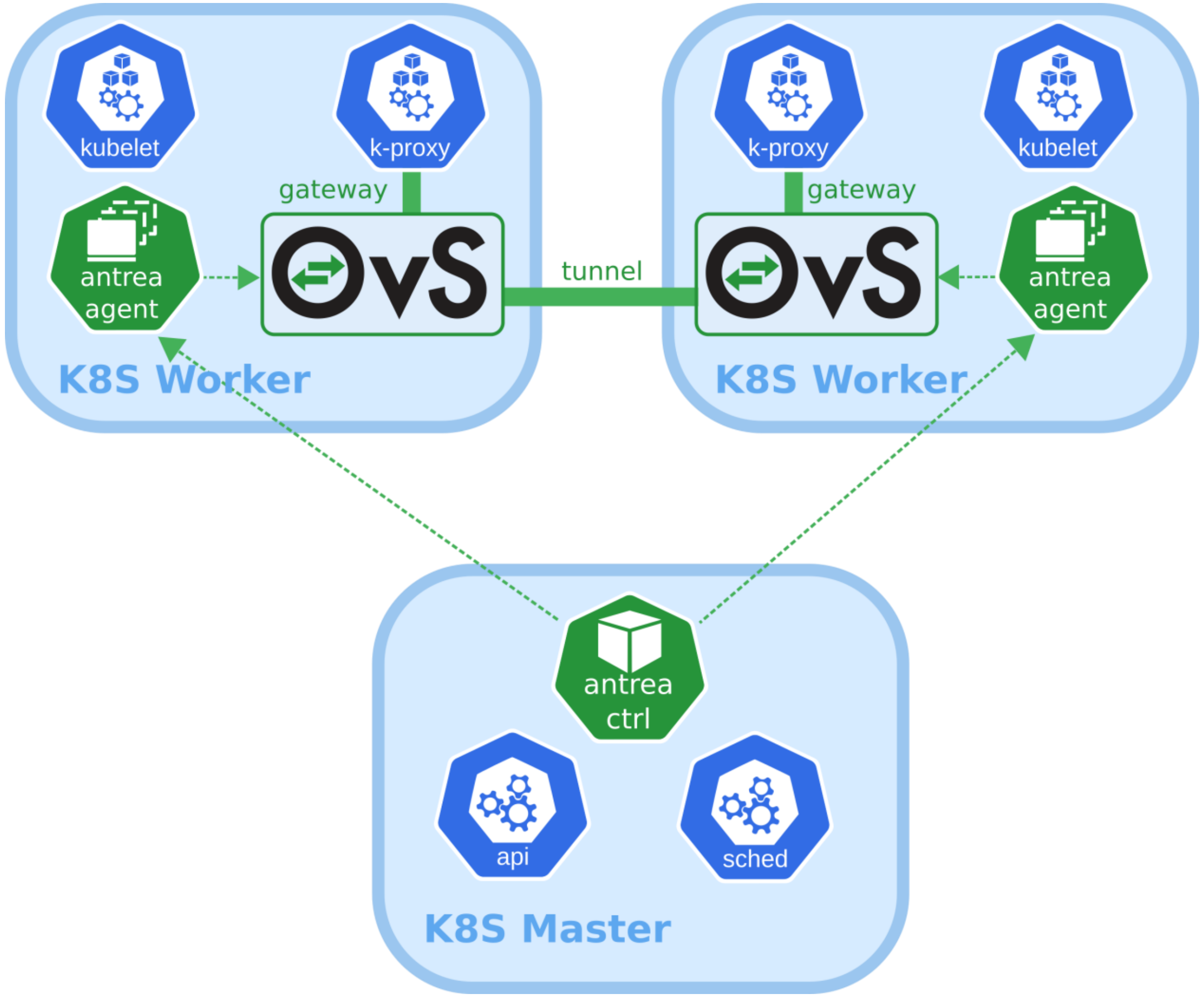
Open vSwitch as data plane

Built with K8s technologies

- Leverages K8s and K8s solutions for API, UI, deployment, control plane, CLI.
- Antrea Controller and Agent are based on K8s controller and apiserver libs.
- All components are deployed using K8s manifests.

Multi-platforms

- VM, bare metal, public clouds
- Windows



Antrea Components

Antrea Controller

- Computes K8s NetworkPolicies, and publishes the results to Antrea Agents.

Antrea Agent

- Manages Pod network interfaces and OVS bridge.
- Creates overlay tunnels across Nodes.
- Implements NetworkPolicies with OVS.

Antrea CNI plugin

- Interface to kubelet. Calls to Antrea Agent for Pod network interface configuration.

Octant UI plugin

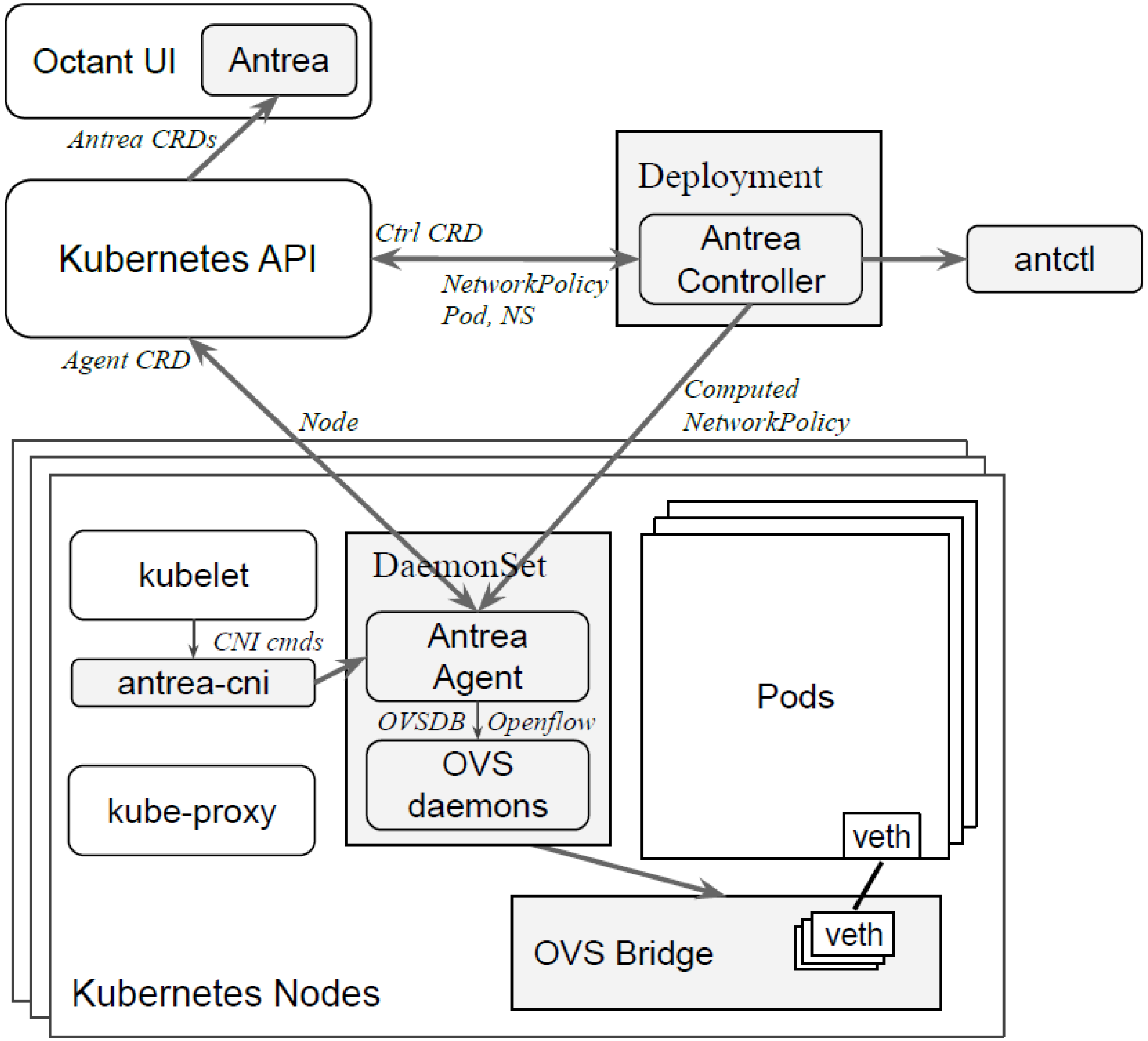
- Shows Antrea runtime information.

antctl – CLI for debugging

- Can be exectuted through kubectl as a plugin.

All bits (inc. OVS daemons) in a Docker image.

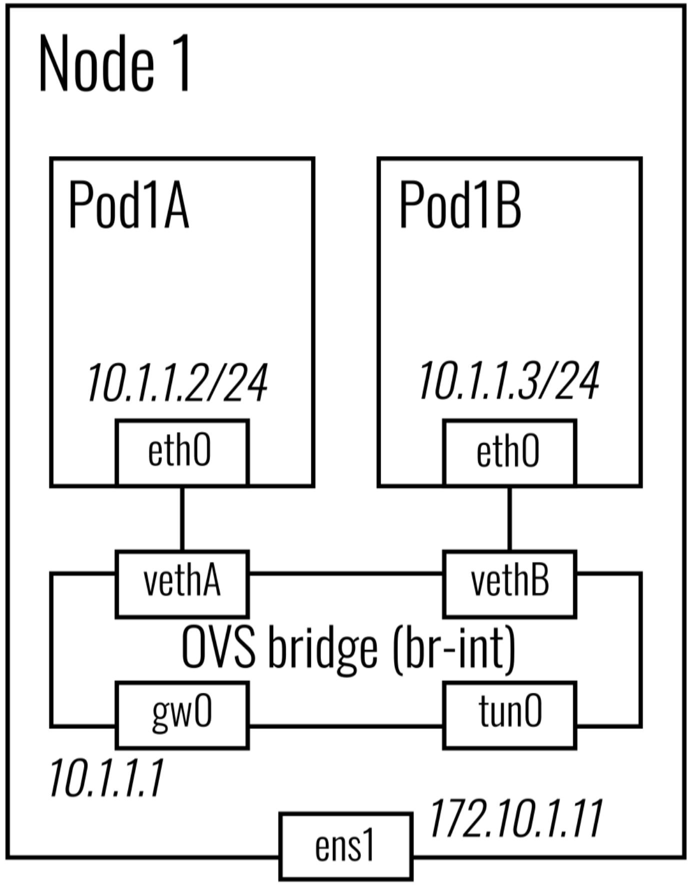
Super simple to deploy: *kubectl apply -f antrea.yml*



Pod Network on a Node

- Each Node has an OVS bridge and is allocated with a single subnet (by K8s NodeIPAM controller).
- The subnet's gateway IP is configured to the "gw0" interface on the bridge.
- A tunnel interface - "tun0" - is created on the bridge too.

- Each Pod's gets an IP from the Node's subnet.
- The Pod network interface is connected to the OVS bridge using the veth devices.



Traffic Walk

Intra-Node traffic

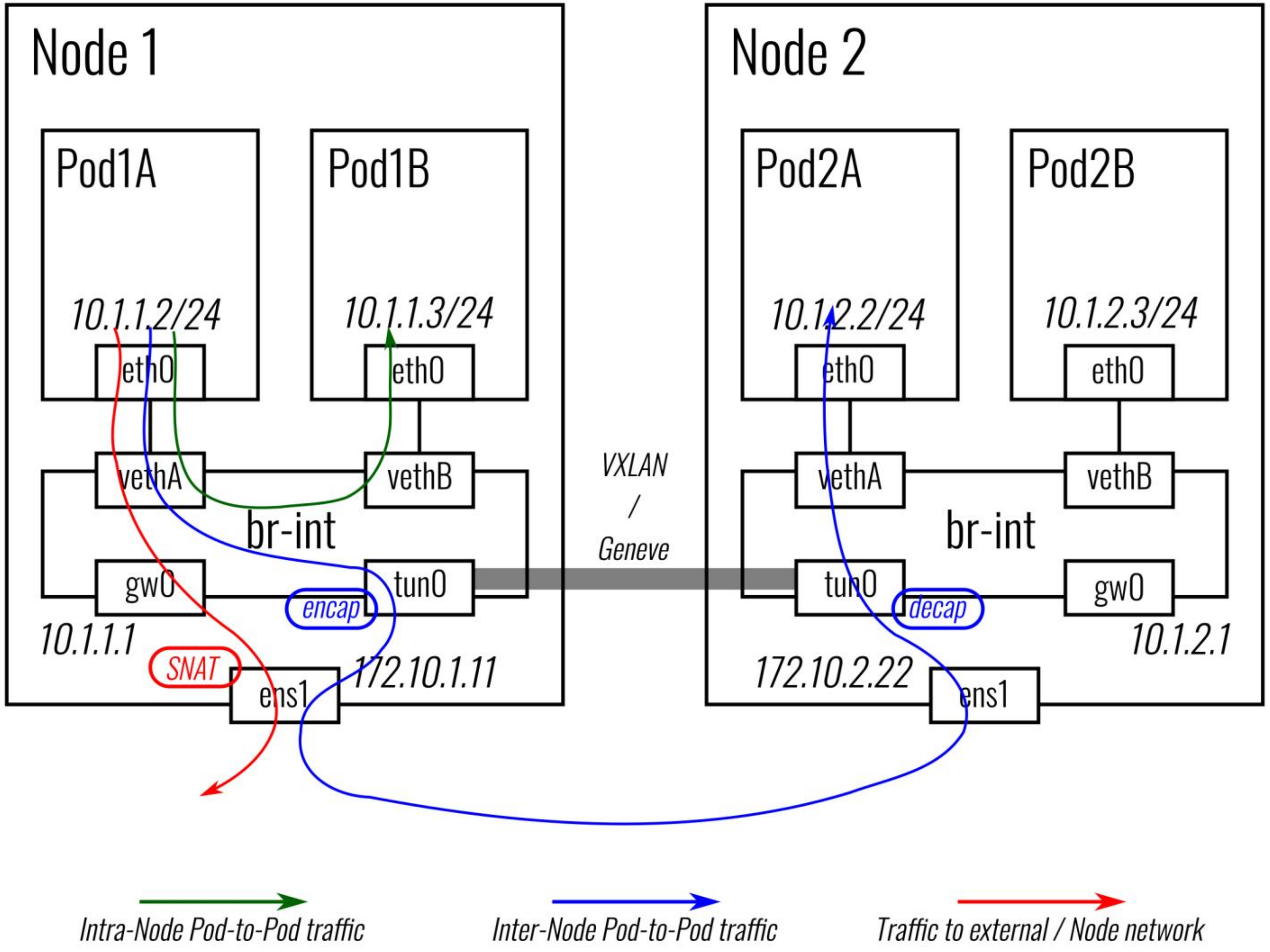
- Does not leave the OVS bridge.

Inter-Node traffic

- Transmitted to the destination Node via overlay tunnels.
- OVS flow based tunneling.

Traffic from a Pod to external network or another Node

- SNAT to the Node IP (by an iptables rule).



ClusterIP Service Traffic

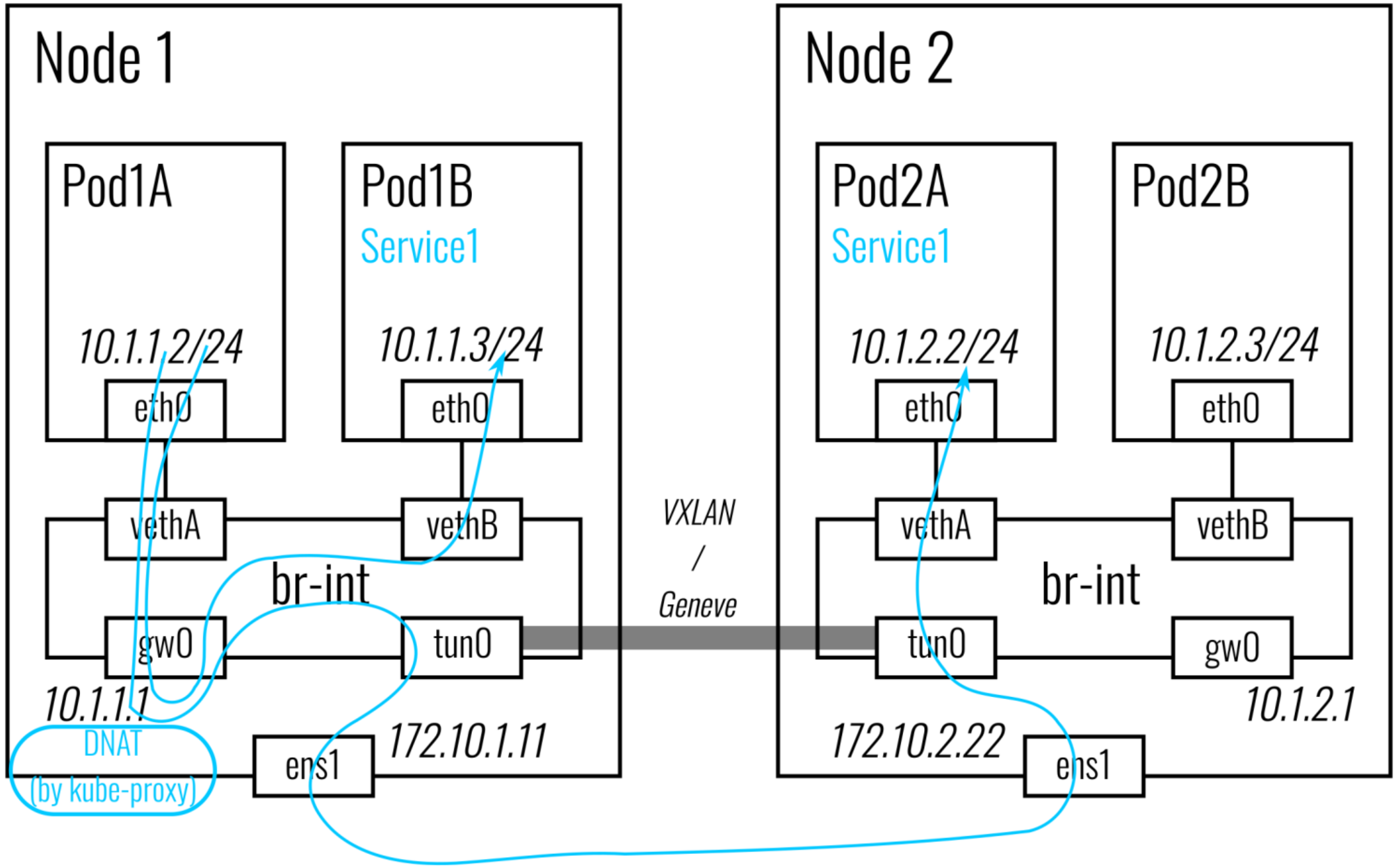
Antrea now leverages *kube-proxy* for Service traffic.

Packets to the Service ClusterIP CIDR will be forwarded to “gw0” (in the default network namespace).

Then kube-proxy is able to intercept the packets.

- kube-proxy can work in any of the iptables, or IPVS, or userspace proxy modes.

Will also support ClusterIP Service with OVS (group and DNAT actions) in future.



Network Policy

- Centralized policy computation

- Antrea Controller watches NetworkPolicy, Pod, and Namespace resources from the Kubernetes API

- Antrea controller calculates SPAN which Nodes need to receive a NetworkPolicy.

```
apiVersion: networking.k8s.io/v1
kind: NetworkPolicy
metadata:
  name: test-network-policy
  namespace: default
spec:
  podSelector:
    matchLabels:
      app: nginx
  policyTypes:
  - Ingress
  - Egress
  ingress:
  - from:
    - podSelector:
        matchLabels:
          app: nginx
    ports:
    - protocol: TCP
      port: 80
  egress:
  ...
```

Antrea Roadmap

Standard K8s Networking

Pod overlay network

- VXLAN or Geneve

K8s Network Policy

Octant UI plugin

CLI for debugging (under development)

Upcoming Features

Windows Node

IPv6 and dual stack

OVS and NetworkPolicy troubleshooting

Data plane extensions

- No encapsulation mode
- IPSec encryption

K8s on public clouds

Open Interfaces to External Policy Manager and Analyzer

Global policy and visibilty

Exporting flow information to a flow/policy analyzer

Next steps

- Run scale tests to detect performance bottlenecks
- Improve performance by using AF_XDP / DPDK?
- kube-proxy with OVS
- Run performance comparisons with iptables-based solutions
- Use OpenFlow bundle
- Metrics and monitoring
  - L2 / L3 / L4
  - Application monitoring (HTTP / DNS)?
  - Integration with Grafana / Prometheus

Community

ANTREA

- Github: https://github.com/vmware-tanzu/antrea
- Slack: #antrea channel in K8s workspace
- Mailing list
  - projectantrea-announce
  - projectantrea
  - projectantrea-dev
- Community meeting: every Wednesday at 9AM PST