

Kube-OVN Document

v1.14.4

Kube-OVN Team

2025 Kube-OVN Team

Table of contents

1.	Kube-OVN	5
1.1	What is Kube-OVN?	5
1.2	Why Kube-OVN?	5
1.3	CNI	6
1.4	OVN/ovn-kubernetes/Kube-OVN	6
1.5		7
2.		8
2.1		8
2.2		10
2.3	Underlay	12
2.4	Talos	17
2.5		18
2.6		20
3.		21
3.1		21
3.2		28
3.3	IP	31
3.4		32
3.5	EIP SNAT	33
3.6	QoS	36
3.7		38
3.8	Webhook	39
3.9		41
3.10	NetworkPolicy	47
3.11	LoadBalancer Service	49
3.12		54
4.	KubeVirt	58
4.1	VM IP	58
4.2		61
4.3		63
4.4		65
4.5		70
4.6	DHCP	73
5.	VPC	74
5.1	VPC	74

5.2 VPC Egress Gateway	83
5.3 VPC QoS	94
5.4 VPC	97
5.5 VPC DNS	101
5.6 SecurityGroup	104
5.7 OVN EIP FIP SNAT DNAT	107
5.8 OVN SNAT ECMP BFD L3 HA	117
5.9 VPC	121
6.	123
6.1 kubectl	123
6.2	133
6.3 ovn-central	135
6.4 OVN	139
6.5 CIDR	141
6.6 Join CIDR	142
6.7	143
6.8	144
7.	148
7.1	148
7.2	155
7.3 FastPath	160
7.4 eBPF TCP	161
7.5 OVN-IC	165
7.6 Submariner	171
7.7 Overlay	173
7.8 BGP	175
7.9 MetalLB Kube-OVN Underlay	179
7.10 Cilium	183
7.11 Cilium NetworkPolicy	186
7.12 Cilium	190
7.13	195
7.14 VIP IP	196
7.15 Mellanox Offload	200
7.16 Offload	209
7.17 Offload	213
7.18 Offload	216
7.19 DPDK	219
7.20 OpenStack	223

7.21	IPsec	227
7.22	OVN	228
7.23	DNS Kube-OVN	231
7.24	VPC NAT	233
8.		234
8.1		234
8.2	What's Next	237
8.3		239
8.4		243
8.5	Underlay	245
8.6	Iptables	252
8.7		257
8.8	OVS/OVN	259
8.9		261
8.10	Kube-OVN	263
8.11	Kube-OVN	269
8.12	Annotation	290
8.13		294
9.		297
9.1		297

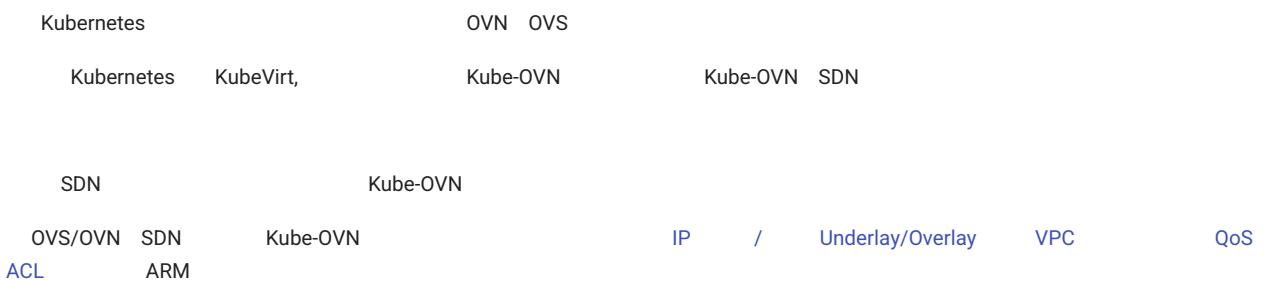
1. Kube-OVN



1.1 What is Kube-OVN?



1.2 Why Kube-OVN?



1.3 CNI



1.3.1 eBPF

[Cilium](#) [Calico eBPF](#)
 Kube-OVN [Open vSwitch](#)

1.3.2 CNI, Ingress, Service Mesh Observability All in One

[Cilium](#)
 Kube-OVN [CNI](#)

1.3.3 OpenShift

[ovn-kubernetes](#)
 OpenShift [CNI](#) [Cluster Network Operator](#) [Kube-OVN](#) [RedHat](#) [Kubernetes](#)

1.3.4 Kubernetes EKS/AKS/GKE

Kubernetes [CNI](#)

1.3.5 AI

[Hostnetwork](#) [host-device](#)
 AI

1.4 OVN/ovn-kubernetes/Kube-OVN

1.4.1 OVN

[OVN](#) [Open vSwitch](#)
 Kubernetes [OpenStack, Kubernetes](#) [CMS](#) [ovn-kubernetes](#) [Kube-OVN](#) [OVN](#)

1.4.2 ovn-kubernetes

[ovn-kubernetes](#) [OVN](#) [OVN](#) [Kubernetes](#) [CNI](#) [OpenShift](#) [OpenShift](#) [UDN, Multihoming, Hardware Acceleration](#)

1.4.3 Kube-OVN

[Kube-OVN](#) [IP](#) [Namespace](#)
[Underlay](#) [VPC](#) [KubeVirt](#) [OVN](#) [ovn-kubernetes](#) [annotation](#) [Pod](#) [join](#)



⌚2025 9 10

⌚2022 5 20



1.5

2.

2.1

Kube-OVN CNI Kubernetes

2.1.1

- Kubernetes >= 1.29
- Docker >= 1.12.6, Containerd >= 1.3.4
- : CentOS 7/8, Ubuntu 16.04/18.04/20.04
- Linux geneve, openvswitch, ip_tables, iptable_nat Kube-OVN

1.	3.10.0-862	netfilter	bug	Kube-OVN	CentOS	bug	Floating IPs broken after kernel upgrade to Centos/RHEL 7.5 - DNAT not working
2.	Rocky Linux 8.6	4.18.0-372.9.1.el8.x86_64	TCP	TCP connection failed in Rocky Linux 8.6		4.18.0-372.13.1.el8_6.x86_64	
3.	4.4	openvswitch		openvswitch			
4.	Geneve	IPv6	cat /proc/cmdline		bug	Geneve tunnels don't work when ipv6 is disabled	

2.1.2

- IPv6 ipv6.disable=1 0
- kube-proxy Kube-OVN Service ClusterIP kube-apiserver
- kubelet CNI , kubelet --network-plugin=cni --cni-bin-dir=/opt/cni/bin --cni-conf-dir=/etc/cni/net.d /etc/cni/net.d/

2.1.3

ovn-central	6641/tcp	ovn nb db server
ovn-central	6642/tcp	ovn sb db server
ovn-central	6643/tcp	ovn northd server
ovn-central	6644/tcp	ovn raft server
ovn-ic	6645/tcp	ovn ic nb db server
ovn-ic	6646/tcp	ovn ic sb db server
ovs-ovn	Geneve 6081/udp, STT 7471/tcp, Vxlan 4789/udp	
kube-ovn-controller	10660/tcp	
kube-ovn-daemon	10665/tcp	
kube-ovn-monitor	10661/tcp	

firewalld Packet Forwarding Masquerade

```
# Packet Forwarding
firewall-cmd --add-forward --permanent
```

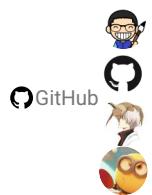
```
#   IPv4 Masquerade
firewall-cmd --add-masquerade --permanent
#   Kube-OVN IPv6/           Masquerade
firewall-cmd --permanent --add-rich-rule 'rule family="ipv6" source address="fd00:10:16::/112" masquerade'

firewall-cmd --reload
```

[!\[\]\(7e49c700e4adaed94ad5398cf2e7059e_img.jpg\) PDF](#)[!\[\]\(5ebcf382a6ee952d6c5b8b948415801e_img.jpg\) Slack](#)[!\[\]\(71ceb62b681518c82e95d615e7265d66_img.jpg\) Support](#)

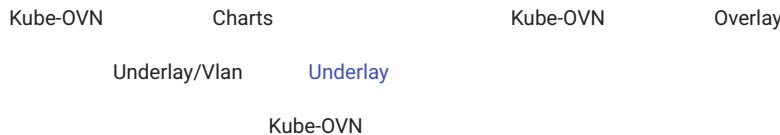
⌚2025 11 20

⌚2022 5 20



2.1.4

2.2



2.2.1

release

```
wget https://raw.githubusercontent.com/kubeovn/kube-ovn/refs/tags/v1.14.23/dist/images/install.sh
```

master

```
wget https://raw.githubusercontent.com/kubeovn/kube-ovn/master/dist/images/install.sh
```

```

REGISTRY="kubeovn"
VERSION="v1.14.23"
POD_CIDR="10.16.0.0/16"
SVC_CIDR="10.96.0.0/12"
JOIN_CIDR="100.64.0.0/16"
LABEL="node-role.kubernetes.io/master"
IFACE="" # OVN DB
TUNNEL_TYPE="geneve" # geneve, vxlan stt stt ovs
  
```

IFACE=enp6s0f0, eth.*

root

```
bash install.sh
```

Kube-OVN

1. [Step 4/6] Pod
2. Kube-OVN

2.2.2 Helm Chart

Kube-OVN Helm Kube-OVN

IP

```
# kubectl get node -o wide
NAME           STATUS    ROLES      AGE   VERSION   INTERNAL-IP   EXTERNAL-IP   OS-IMAGE        KERNEL-VERSION   CONTAINER-RUNTIME
kube-ovn-control-plane  NotReady  control-plane  20h  v1.26.0  172.18.0.3  <none>       Ubuntu 22.04.1 LTS  5.10.104-linuxkit  containerd://1.6.9
kube-ovn-worker     NotReady  <none>      20h  v1.26.0  172.18.0.2  <none>       Ubuntu 22.04.1 LTS  5.10.104-linuxkit  containerd://1.6.9
```

label

```
# kubectl label node -lbeta.kubernetes.io/os=linux kubernetes.io/os=linux --overwrite
node/kube-ovn-control-plane not labeled
node/kube-ovn-worker not labeled

# kubectl label node -lnode-role.kubernetes.io/control-plane kube-ovn/role=master --overwrite
node/kube-ovn-control-plane labeled

#   label    dpdk      dpdk
# kubectl label node -lvn.kubernetes.io/ovs_dp_type!=userspace ovn.kubernetes.io/ovs_dp_type=kernel --overwrite
node/kube-ovn-control-plane labeled
node/kube-ovn-worker labeled
```

Helm Repo

```
# helm repo add kubeovn https://kubeovn.github.io/kube-ovn/
" kubeovn " has been added to your repositories

# helm repo list
NAME          URL
kubeovn      https://kubeovn.github.io/kube-ovn/

# helm repo update kubeovn
Hang tight while we grab the latest from your chart repositories...
...Successfully got an update from the " kubeovn " chart repository
Update Complete. *Happy Helm-ing!*

# helm search repo kubeovn
NAME          CHART VERSION   APP VERSION   DESCRIPTION
kubeovn/kube-ovn   v1.14.23     v1.14.23     Helm chart for Kube-OVN
```

helm install Kube-OVN

Chart values.yaml

```
# helm install kube-ovn kubeovn/kube-ovn --wait -n kube-system --version v1.14.23
NAME: kube-ovn
LAST DEPLOYED: Thu Apr 24 08:30:13 2025
NAMESPACE: kube-system
STATUS: deployed
REVISION: 1
TEST SUITE: None
```

Helm values.yaml

```
helm upgrade -f values.yaml kube-ovn kubeovn/kube-ovn --wait -n kube-system --version v1.14.23
```

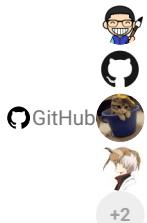
 PDF

 Slack

 Support

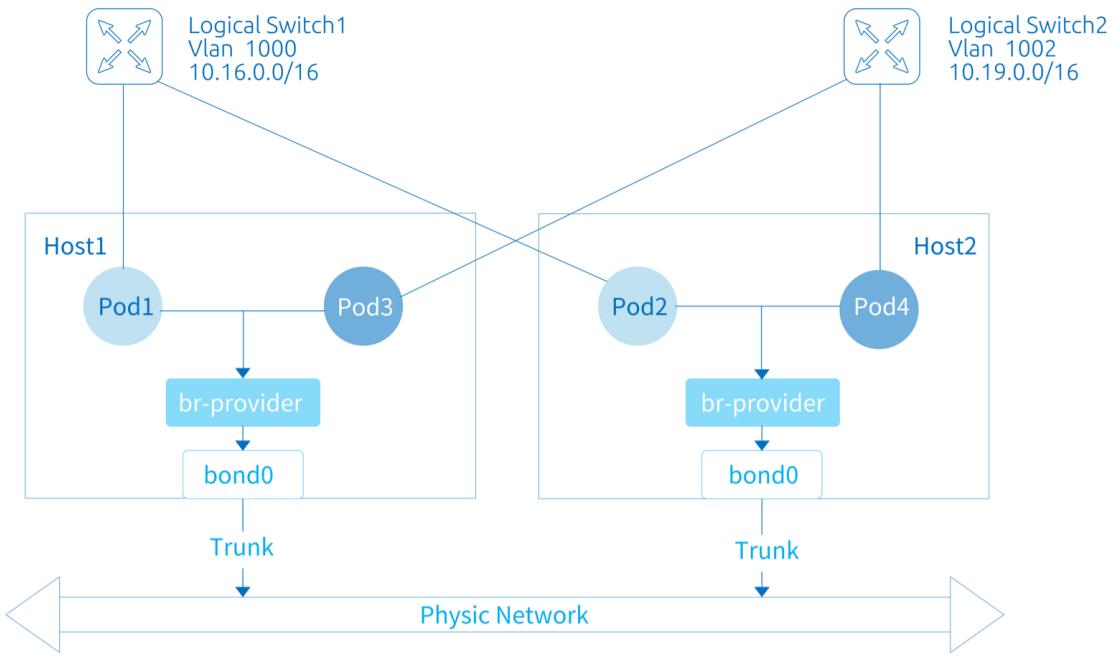
⌚2025 9 17

⌚2022 5 20

**2.2.3**

2.3 Underlay

Kube-OVN Geneve Overlay
 Kube-OVN Underlay



2.3.1

Overlay SNAT/EIP / L3 VPC Underlay

2.3.2 Macvlan

Kube-OVN Underlay Macvlan

1. Macvlan OVS Macvlan
2. Kube-OVN arp-proxy arp
3. Macvlan netfilter Service NetworkPolicy Kube-OVN OVS Service NetworkPolicy
4. Kube-OVN Underlay Macvlan IP QoS

2.3.3

Underlay	OVS	OVS	L2/L3	Vlan
1. OpenStack VM		PortSecurity		
2. VMware vSwitch		MAC Address Changes, Forged Transmits	Promiscuous Mode Operation	allow
3. VMware NSX-T	Underlay			
4. Hyper-V		MAC Address Spoofing		
5. AWS GCE	Mac	Underlay	Underlay	VPC-CNI
6. Linux Bridge				
	Kube-OVN Underlay		Underlay	
	Underlay			
	Kube-OVN	Mac	IP	MTU
PROVIDER_NAME	ProviderNetwork		provider	
1.	OVS Bridge	NetworkManager		DHCP
2.	OVS Bridge	OVS	Kube-OVN	OVS
	Underlay			

2.3.4

Underlay Pod Underlay

```
wget https://raw.githubusercontent.com/kubeovn/kube-ovn/release-1.14/dist/images/install.sh
```

```
ENABLE_ARP_DETECT_IP_CONFLICT #      vlan    arp
NETWORK_TYPE #      vlan
VLAN_INTERFACE_NAME #      eth1
VLAN_ID #      VLAN Tag    0    VLAN
POD_CIDR #      CIDR    192.168.1.0/24
POD_GATEWAY #      192.168.1.1
EXCLUDE_IPS #
ENABLE_LB #      Underlay   Service  true
EXCHANGE_LINK_NAME #
LS_DNAT_MOD_DL_DST #  DNAT     MAC      Service  true
```

```
bash install.sh
```

2.3.5 CRD Underlay

Underlay Pod ProviderNetwork Vlan Subnet

ProviderNetwork

ProviderNetwork Underlay

ProviderNetwork :

```
apiVersion: kubeovn.io/v1
kind: ProviderNetwork
metadata:
  name: net1
```

```
spec:
  defaultInterface: eth1
  customInterfaces:
    - interface: eth2
      nodes:
        - node1
  excludeNodes:
    - node2
```

ProviderNetwork 12

- | | | | | | | |
|---------------------|------|------|--|--------------|---------|---------|
| • defaultInterface: | Bond | Vlan | ProviderNetwork | excludeNodes | br-net1 | br-NAME |
| OVS | | | | | | |
| • customInterfaces: | | | | | | |
| • excludeNodes: | | | net1.provider-network.ovn.kubernetes.io/exclude=true | | | |

Key	Value	
net1.provider-network.ovn.kubernetes.io/ready	true	ProviderNetwork
net1.provider-network.ovn.kubernetes.io/interface	eth1	
net1.provider-network.ovn.kubernetes.io/mtu	1500	MTU

IP IP OVS

VLAN

Vlan Vlan Tag ProviderNetwork

VLAN

```
apiVersion: kubeovn.io/v1
kind: Vlan
metadata:
  name: vlan1
spec:
  id: 0
  provider: net1
```

- | | | | | | | | |
|-------------|-------------|-----------------|------|------|---|-----------------|----------|
| • id: | VLAN ID/Tag | Kube-OVN | Vlan | Vlan | 0 | vlan | localnet |
| • provider: | | ProviderNetwork | | VLAN | | ProviderNetwork | |

Subnet

Vlan

```
apiVersion: kubeovn.io/v1
kind: Subnet
metadata:
  name: subnet1
spec:
  protocol: IPv4
  cidrBlock: 172.17.0.0/16
  gateway: 172.17.0.1
  vlan: vlan1
  disableGatewayCheck: false
```

- | | | | |
|-----------------------|------|----------|------|
| • vlan | VLAN | Subnet | VLAN |
| • disableGatewayCheck | | Underlay | true |

2.3.6

IP

IP Pod IP Mac

2.3.7

Kube-OVN Underlay spec.logicalGateway true

```
apiVersion: kubeovn.io/v1
kind: Subnet
metadata:
  name: subnet1
spec:
  protocol: IPv4
  cidrBlock: 172.17.0.0/16
  gateway: 172.17.0.1
  vlan: wlan1
  logicalGateway: true
```

Pod Kube-OVN Logical Router

2.3.8 Underlay Overlay

Underlay	Overlay	Overlay	Pod	NAT	Underlay	Pod IP	Underlay	Pod	Overlay
	Overlay		Underlay	Pod	Pod IP	Overlay	Pod		
Underlay	Overlay	u2oInterconnection	true	Kube-OVN		Underlay IP	Underlay	ovn-cluster	
	Kube-OVN	Underlay	Overlay						

IP

subnet IP Underlay Subnet u2oInterconnectionIP

Underlay Subnet **VPC**

Underlay Subnet VPC Overlay Subnet VPC u2oInterconnection true subnet.spec.vpc VPC

2.3.9

IP Netplan Ubuntu Netplan renderer NetworkManager IP DHCP

```
network:
  renderer: NetworkManager
  ethernets:
    eth0:
      dhcp4: no
      addresses:
        - 172.16.143.129/24
  version: 2
```

IP netplan

```
netplan generate
nmcli connection reload netplan-eth0
nmcli device set eth0 managed yes
```

Kube-OVN IP OVS

NetworkManager CentOS

```
nmcli connection reload eth0
nmcli device set eth0 managed yes
nmcli -t -f GENERAL.STATE device show eth0 | grep -qw unmanaged || nmcli device reapply eth0
```

IP MAC

2.3.10

hairpin Pod

hairpin Pod Pod OVS MAC

hairpin

Kube-OVN

Pod

Pod 300 ARP OVS resubmit

```
2022-11-13T08:43:46.782Z|00222|ofproto_dpif_upcall(handler5)|WARN|Flow: arp,in_port=331,vlan_tci=0x0000,d1_src=00:00:00:25:eb:39,d1_dst=ff:ff:ff:ff:ff:ff,arp_spa=10.213.131.240,arp_tpa=10.213.159.254,arp_op=1,arp_sha=00:00:00:25:eb:39,arp_tha=ff:ff:ff:ff:ff:ff
bridge("br-int")
-----
0. No match.
    >>> received packet on unknown port 331 <<<
    drop

Final flow: unchanged
Megaflow: recirc_id=0,eth,arp,in_port=331,d1_src=00:00:00:25:eb:39
Datapath actions: drop
2022-11-13T08:44:34.077Z|00224|ofproto_dpif_xlate(handler5)|WARN|over 4096 resubmit actions on bridge br-int while processing
arp,in_port=13483,vlan_tci=0x0000,d1_src=00:00:00:59:ef:13,d1_dst=ff:ff:ff:ff:ff:ff,arp_spa=10.213.152.3,arp_tpa=10.213.159.254,arp_op=1,arp_sha=00:00:00:59:ef:13,arp_tha=ff:ff:ff:ff:ff:ff
```

OVN NB bcast_arp_req_flood false

kubectl ko nbctl set NB_Global . options:bcast_arp_req_flood=false

[PDF](#)[Slack](#)[Support](#)

⌚2025 12 18

⌚2022 5 20



GitHub

+2

2.3.11

2.4 Talos

Talos Linux Kubernetes Linux

2.4.1 Helm Chart Kube-OVN

Talos Linux Kube-OVN

```
helm install kube-ovn kubeovn/kube-ovn --wait \
-n kube-system \
--version v1.14.23 \
--set OVN_DIR=/var/lib/ovn \
--set OPENVSWITCH_DIR=/var/lib/openvswitch \
--set DISABLE_MODULES_MANAGEMENT=true \
--set cni_conf.MOUNT_LOCAL_BIN_DIR=false
```

Underlay

Helm

Chart

```
helm install kubeovn kubeovn/kube-ovn --wait \
-n kube-system \
--version v1.14.23 \
--set OVN_DIR=/var/lib/ovn \
--set OPENVSWITCH_DIR=/var/lib/openvswitch \
--set DISABLE_MODULES_MANAGEMENT=true \
--set cni_conf.MOUNT_LOCAL_BIN_DIR=false \
--set networking.NETWORK_TYPE=vlan \
--set networking.vlan.VLAN_INTERFACE_NAME=enp0s5f1 \
--set networking.vlan.VLAN_ID=0 \
--set networking.NET_STACK=ipv4 \
--set-json networking.EXCLUDE_IPS='["172.99.99.11..172.99.99.99"]' \
--set-json ipv4.POD_CIDR='["172.99.99.8/24"]' \
--set-json ipv4.POD_GATEWAY='["172.99.99.1"]'
```



Note

VLAN Bond Bridge Underlay Underlay Talos ignore=true

```
machine:
  network:
    interfaces:
      - interface: enp0s5f1
        ignore: true
```

[PDF](#)

[Slack](#)

[Support](#)

⌚2025 7 1

⌚2025 4 16



GitHub



2.4.2

2.5

Kube-OVN

pre-v0.0.1

Kube-OVN

2.5.1

- [Linkerd](#)
- [Elasticsearch](#)
- [EMQX](#)
- [KubeSphere](#)

2.5.2 OpenVswitch/OVN

Kube-OVN

OpenVswitch OVN

- [OVN](#)
- [OpenVswitch](#)

ovn-architecture

2.5.3 Kube-OVN

Kube-OVN

Kube-OVN kubectl bash

Kube-OVN

2.5.4

Kube-OVN

Kube-OVN

Kube-OVN E2E

2.5.5

Kube-OVN

[OpenTelemetry](#)[DeepFlow](#)

2.5.6

Kube-OVN

7*24

Github Issue

Github Issue

AI

AI

[!\[\]\(d544e576856cad8f6b886171f7e2e480_img.jpg\) PDF](#)
[!\[\]\(700de1146ce8d8d31c70e1dbe6152e98_img.jpg\) Slack](#)
[!\[\]\(20bc8a2224e0b0c62d32464117ca7264_img.jpg\) Support](#)

⌚2025 9 22

⌚2025 8 29



2.5.7

2.6

Kube-OVN

Kube-OVN

OVS

issue

Kube-OVN

2.6.1 Kubernetes

[Script Uninstall](#) [Helm Uninstall](#)

```
wget https://raw.githubusercontent.com/kubeovn/kube-ovn/release-1.14/dist/images/cleanup.sh
bash cleanup.sh

helm uninstall kube-ovn -n kube-system
```

2.6.2

ovsdb openvswitch

```
rm -rf /var/run/openvswitch
rm -rf /var/run/ovn
rm -rf /etc/origin/openvswitch/
rm -rf /etc/origin/ovn/
rm -rf /etc/cni/net.d/00-kube-ovn.conflist
rm -rf /etc/cni/net.d/01-kube-ovn.conflist
rm -rf /var/log/openvswitch
rm -rf /var/log/ovn
rm -fr /var/log/kube-ovn
```

2.6.3

iptable/ipset

reboot

[PDF](#)

[Slack](#)

[Support](#)

⌚2025 9 10

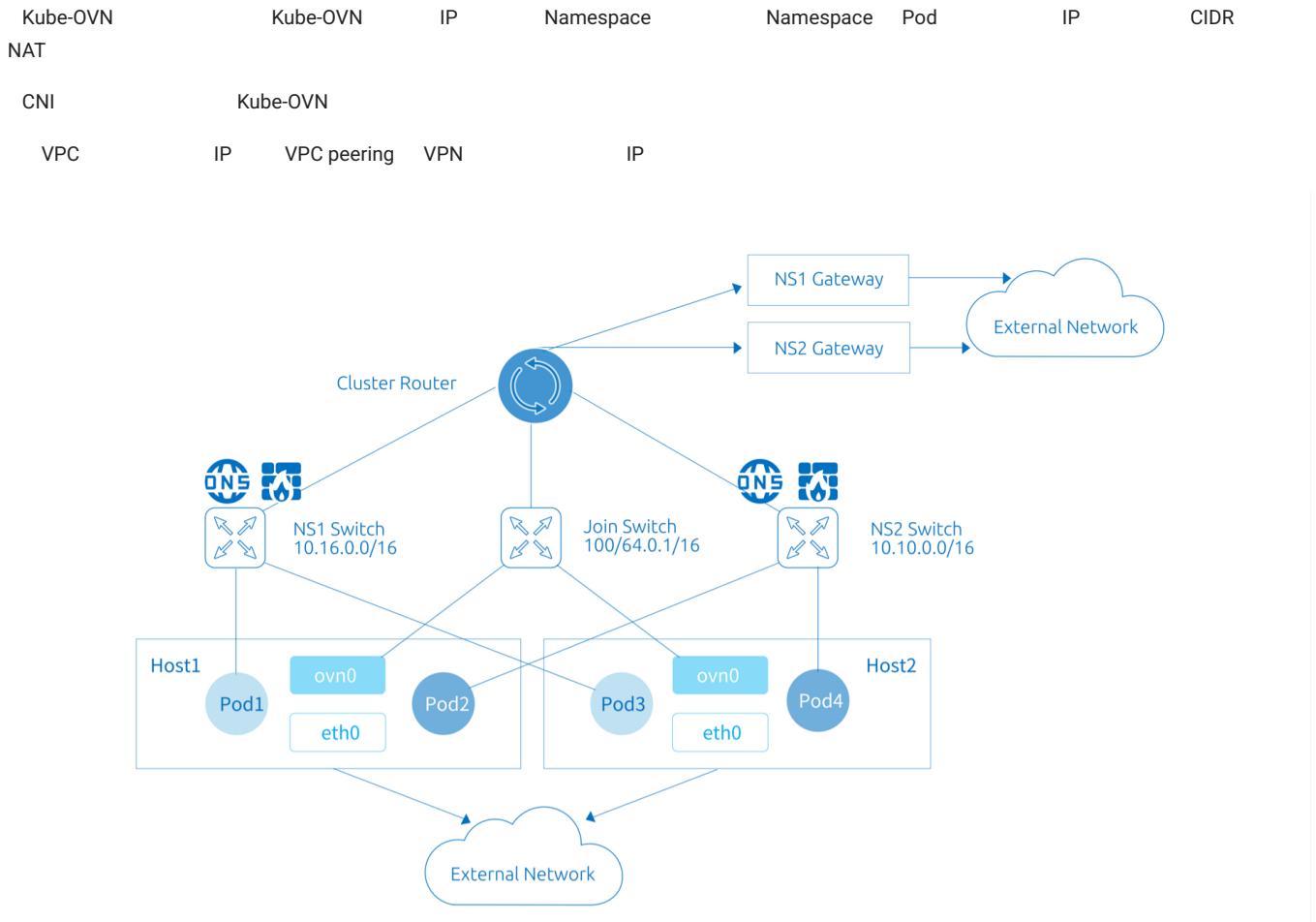
⌚2022 5 24



2.6.4

3.

3.1



Overlay Underlay

3.1.1

Kube-OVN	Namespace	IP	CIDR
Overlay	NAT	Flannel	
Underlay	arping		

spec default true ovn-default

```
# kubectl get subnet ovn-default -o yaml
apiVersion: kubeovn.io/v1
kind: Subnet
metadata:
  creationTimestamp: "2019-08-06T09:33:43Z"
  generation: 1
  name: ovn-default
  resourceVersion: "1571334"
```

```

selfLink: /apis/kubeovn.io/v1/subnets/ovn-default
uid: 7e2451f8-fb44-4f7f-b3e0-cfd27f6fd5d6
spec:
  cidrBlock: 10.16.0.0/16
  default: true
  excludeIps:
  - 10.16.0.1
  gateway: 10.16.0.1
  gatewayType: distributed
  natOutgoing: true
  private: false
  protocol: IPv4

```

3.1.2 Join

Kubernetes Pod	Node	Pod	Overlay	Kube-OVN	join	Node	ovn0	join
Pods	Nodes	ovn0	Node	Pod	ovn0		ovn0	
			join	CIDR		Join		

Join

Pod	hostport	externalTrafficPolicy: Local	NodePort	Service
-----	----------	------------------------------	----------	---------

join	CIDR
------	------

```

# kubectl get subnet join -o yaml
apiVersion: kubeovn.io/v1
kind: Subnet
metadata:
  creationTimestamp: "2019-08-06T09:33:43Z"
  generation: 1
  name: join
  resourceVersion: "1571333"
  selfLink: /apis/kubeovn.io/v1/subnets/join
  uid: 9c744810-c678-4d50-8a7d-b8ec12ef91b8
spec:
  cidrBlock: 100.64.0.0/16
  default: false
  excludeIps:
  - 100.64.0.1
  gateway: 100.64.0.1
  gatewayNode: ""
  gatewayType: ""
  natOutgoing: false
  private: false
  protocol: IPv4

```

Node ovn0

```

# ifconfig ovn0
ovn0: flags=4163<UP,BROADCAST,RUNNING,MULTICAST> mtu 1420
    inet 100.64.0.4 netmask 255.255.0.0 broadcast 100.64.255.255
      inet6 fe80::800:ff:fe40:5 prefixlen 64 scopeid 0x20<link>
        ether 0a:00:00:40:00:05 txqueuelen 1000  (Ethernet)
          RX packets 18 bytes 1428 (1.3 KiB)
          RX errors 0 dropped 0 overruns 0 frame 0
          TX packets 19 bytes 1810 (1.7 KiB)
          TX errors 0 dropped 0 overruns 0 carrier 0 collisions 0

```

3.1.3

Namespace

```

cat <<EOF | kubectl create -f -
apiVersion: kubeovn.io/v1
kind: Subnet
metadata:
  name: subnet1
spec:
  protocol: IPv4
  cidrBlock: 10.66.0.0/16
  excludeIps:

```

```

- 10.66.0.1..10.66.0.10
- 10.66.0.101..10.66.0.151
gateway: 10.66.0.1
gatewayType: distributed
natOutgoing: true
routeTable: ""
namespaces:
- ns1
- ns2
EOF

```

- cidrBlock : CIDR VPC Subnet CIDR
- excludeIps : IP Underlay
- gateway : Overlay Kube-OVN Underlay
- namespaces : Namespace Namespace Pod
- routeTable :

```

# kubectl create ns ns1
namespace/ns1 created

# kubectl run nginx --image=docker.io/library/nginx:alpine -n ns1
deployment.apps/nginx created

# kubectl get pod -n ns1 -o wide
NAME           READY   STATUS    RESTARTS   AGE   IP          NODE   NOMINATED NODE   READINESS GATES
nginx-74d5899f46-n8wtg  1/1     Running   0          10s   10.66.0.11  node1  <none>        <none>

```

Workload

Pod	Namespace	IP	Namespace	Workload	Pod	Annotation ovn.kubernetes.io/logical_switch
-----	-----------	----	-----------	----------	-----	---

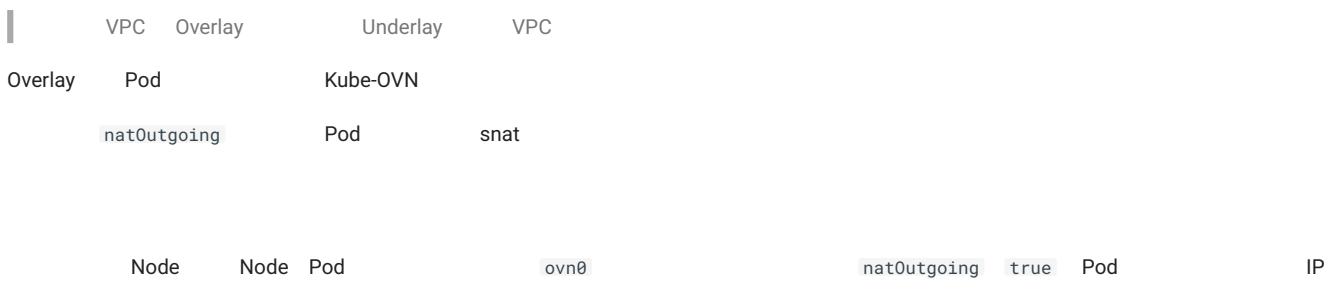
```

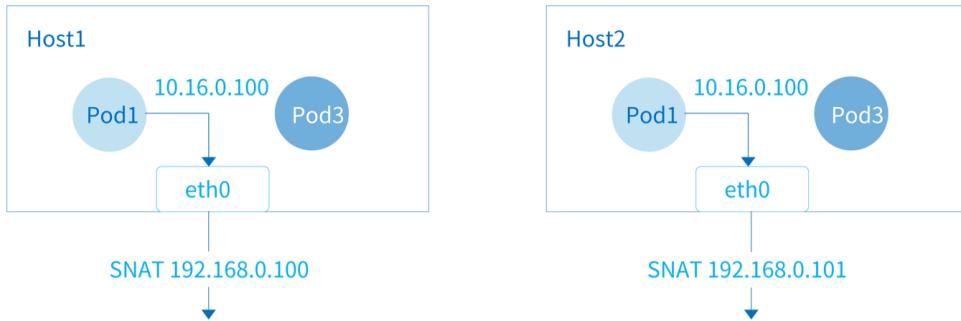
apiVersion: v1
kind: Pod
metadata:
  name: another-subnet
  annotations:
    ovn.kubernetes.io/logical_switch: subnet1
spec:
  containers:
  - name: another-subnet
    image: docker.io/library/nginx:alpine

```

Workload	Deployment	StatefulSet	ovn.kubernetes.io/logical_switch Annotation
spec.template.metadata.annotations			

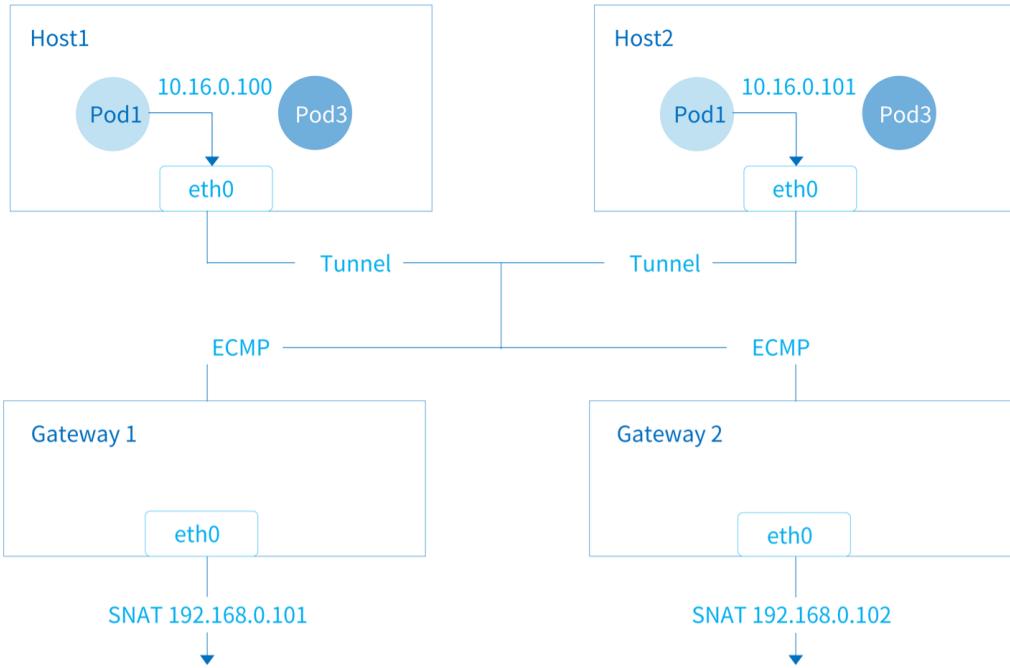
3.1.4 Overlay





```
gatewayType      distributed
```

```
apiVersion: kubeovn.io/v1
kind: Subnet
metadata:
  name: distributed
spec:
  protocol: IPv4
  cidrBlock: 10.166.0.0/16
  default: false
  excludeIps:
  - 10.166.0.1
  gateway: 10.166.0.1
  gatewayType: distributed
  natOutgoing: true
```



Pod	IP	Pod	ovn0	natOutgoing	true
	IP	Kubernetes	NodeName	gatewayNode	
	gatewayType	centralized	gatewayNode		

```

apiVersion: kubeovn.io/v1
kind: Subnet
metadata:
  name: centralized
spec:
  protocol: IPv4
  cidrBlock: 10.166.0.0/16
  default: false
  excludeIps:
  - 10.166.0.1
  gateway: 10.166.0.1
  gatewayType: centralized
  gatewayNode: "node1,node2"
  natOutgoing: true

```

- gatewayNode kube-ovn-worker:172.18.0.2, kube-ovn-control-plane:172.18.0.3
- ECMP ECMP
- Kube-OVN v1.12.0 subnet crd spec enableEcmp ECMP ECMP kube-ovn- controller Deployment enable-ecmp v1.12.0

	ECMP	kube-ovn-controller	ping	5s	5s-10s
Node Ready					

3.1.5 ACL

⚠ Warning

Kube-OVN	NetworkPolicy	Network Policy API	Subnet ACL	Security Group	OVN ACL	NetworkPolicy	NetworkPolicy API
----------	---------------	--------------------	------------	----------------	---------	---------------	-------------------

ACL	Kube-OVN	Subnet	ACL				
Subnet	ACL	OVN	ACL	ovn-nb ACL Table	match	ovn-sb Logical Flow Table	
IP	10.10.0.2	Pod		ACL			

```
apiVersion: kubeovn.io/v1
kind: Subnet
metadata:
  name: acl
spec:
  allowEWTraffic: false
  acls:
    - action: drop
      direction: to-lport
      match: ip4.dst == 10.10.0.2 && ip
      priority: 1002
    - action: allow-related
      direction: from-lport
      match: ip4.src == 10.10.0.2 && ip
      priority: 1002
  cidrBlock: 10.10.0.0/24
```

ACL	allowEWTraffic: true
-----	----------------------

3.1.6

ACL	ACL					
Kube-OVN	Pod					
CRD	private	true		allowSubnets	allowSubnets	

```
apiVersion: kubeovn.io/v1
kind: Subnet
metadata:
  name: private
spec:
  protocol: IPv4
  default: false
  namespaces:
    - ns1
    - ns2
  cidrBlock: 10.69.0.0/16
  private: true
  allowSubnets:
    - 10.16.0.0/16
    - 10.18.0.0/16
```

3.1.7 Underlay

Underlay

vlan:	Underlay	Subnet	Vlan CR	Underlay		
logicalGateway:	Underlay			OVN	Underlay	Overlay
					false	

3.1.8

kube-ovn-cni	Pod	ICMP	ARP	Underlay	ICMP
--------------	-----	------	-----	----------	------

```
apiVersion: kubeovn.io/v1
kind: Subnet
metadata:
  name: disable-gw-check
spec:
  disableGatewayCheck: true
```

3.1.9 Multicast-Snoop

subnet	Pod	OVN	Pod	subnet	multicast snoop	OVN	South Database	Multicast_Group
--------	-----	-----	-----	--------	-----------------	-----	----------------	-----------------

```
apiVersion: kubeovn.io/v1
kind: Subnet
metadata:
  name: sample1
spec:
  enableMulticastSnooop: true
```

3.1.10 Subnet MTU

Subnet	Pod	MTU	Subnet	Pod
--------	-----	-----	--------	-----

```
apiVersion: kubeovn.io/v1
kind: Subnet
metadata:
  name: sample1
spec:
  mtu: 1300
```

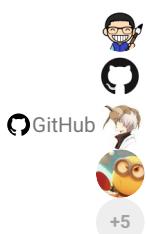
3.1.11

- IP
- VPC NAT
- QoS
-
- DHCP
-
-
- IP

[PDF](#)
[Slack](#)
[Support](#)

⌚2025 11 5

⌚2022 5 24



3.1.12

3.2

Kube-OVN Pod Namespace IP Mac Kube-OVN

- Pod IP/Mac
- Workload IP Pool
- StatefulSet
- KubeVirt VM
- Multus

3.2.1 Pod IP Mac

Pod annotation Pod IP/Mac, kube-ovn-controller

```
apiVersion: v1
kind: Pod
metadata:
  name: static-ip
  annotations:
    ovn.kubernetes.io/ip_address: 10.16.0.15 // 10.16.0.15,fd00:10:16::15
    ovn.kubernetes.io/mac_address: 00:00:00:53:6B:B6
spec:
  containers:
  - name: static-ip
    image: docker.io/library/nginx:alpine
```

annotation Pod IP/Mac

1. IP/Mac IP/Mac
2. IP CIDR
3. IP Mac

3.2.2 Workload IP Pool

Kube-OVN annotation ovn.kubernetes.io/ip_pool Workload Deployment/StatefulSet/DaemonSet/Job/CronJob IP kube-ovn-
controller ovn.kubernetes.io/ip_pool IP
IP Pool Annotation template annotation Kubernetes Workload Workload

Deployment IP

```
apiVersion: apps/v1
kind: Deployment
metadata:
  name: ippool
  labels:
    app: ippool
spec:
  replicas: 2
  selector:
    matchLabels:
      app: ippool
  template:
    metadata:
      labels:
        app: ippool
      annotations:
        ovn.kubernetes.io/ip_pool: 10.16.0.15,10.16.0.16,10.16.0.17 // 10.16.0.15,fd00:10:16::000E;10.16.0.16,fd00:10:16::000F;
        10.16.0.17,fd00:10:16::0010
    spec:
      containers:
      - name: ippool
        image: docker.io/library/nginx:alpine
```

Workload	IP								
1. ovn.kubernetes.io/ip_pool	IP		CIDR						
2. ovn.kubernetes.io/ip_pool	IP		IP						
3. ovn.kubernetes.io/ip_pool	IP	replicas		Pod		Workload		ovn.kubernetes.io/ip_pool	IP

3.2.3 StatefulSet

StatefulSet	IP	Workload	ovn.kubernetes.io/ip_pool	Pod	IP				
StatefulSet			Kube-OVN						
1. Pod	ovn.kubernetes.io/ip_pool	IP	StatefulSet	web	web-0	ovn.kubernetes.io/ip_pool	IP	web-1	IP
2. StatefulSet Pod	OVN	logical_switch_port		Pod	interface	Pod	IP/Mac		StatefulSet Volume
3. 2	ovn.kubernetes.io/ip_pool	StatefulSet Pod			IP/Mac		StatefulSet		

StatefulSet

```
apiVersion: apps/v1
kind: StatefulSet
metadata:
  name: web
spec:
  serviceName: "nginx"
  replicas: 2
  selector:
    matchLabels:
      app: nginx
  template:
    metadata:
      labels:
        app: nginx
    spec:
      containers:
        - name: nginx
          image: docker.io/library/nginx:alpine
          ports:
            - containerPort: 80
              name: web
```

StatefulSet Pod Pod IP

StatefulSet Pod IP

StatefulSet	IP	Pod Name	Statefulset	ovn.kubernetes.io/ip_pool	Annotation	Pod	IP
StatefulSet Pod IP		StatefulSet	scale	0	Annotation	StatefulSet	

3.2.4 KubeVirt VM

KubeVirt	VM	kube-ovn-controller	StatefulSet Pod	IP	VM
VM	IP				

3.2.5 Multus

Multus	Pod	Kube-OVN	annotation	Kube-OVN	CNI	Kube-OVN	IPAM	CNI



⌚2025 11 5

⌚2022 5 20



⌚GitHub 🐱



+2

3.2.6

3.3 IP

IP IPPool Subnet IPAM IP Namespace

3.3.1

```
apiVersion: kubeovn.io/v1
kind: IPPool
metadata:
  name: pool-1
spec:
  subnet: ovn-default
  ips:
    - "10.16.0.201"
    - "10.16.0.210/30"
    - "10.16.0.220..10.16.0.230"
  namespaces:
    - ns-1
```

	subnet			
ips	IP	..	IPv6	
namespaces			Pod	IP IP

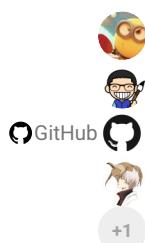
3.3.2

1. Workload IP Pool IP IP
2. IP .spec.ips IP IP .spec.ips CIDR
3. IP IP IP
4. IP .spec.ips
5. IP IP IP IP IP
6. IP IP
7. IP Namespace



⌚2025 11 20

⌚2023 7 10



3.3.3

3.4

Pod Annotations

```
apiVersion: v1
kind: Pod
metadata:
  name: custom-routes
  annotations:
    ovn.kubernetes.io/routes: |
      [{ "dst": "192.168.0.101/24", "gw": "10.16.0.254" },
       { "gw": "10.16.0.254" }]
spec:
  containers:
  - name: nginx
    image: docker.io/library/nginx:alpine
```

dst

Deployment	DaemonSet	StatefulSet	Annotation	.spec.template.metadata.annotations

```
apiVersion: apps/v1
kind: Deployment
metadata:
  name: custom-routes
  labels:
    app: nginx
spec:
  replicas: 2
  selector:
    matchLabels:
      app: nginx
  template:
    metadata:
      labels:
        app: nginx
      annotations:
        ovn.kubernetes.io/routes: |
          [{ "dst": "192.168.0.101/24", "gw": "10.16.0.254" },
           { "gw": "10.16.0.254" }]
    spec:
      containers:
      - name: nginx
        image: docker.io/library/nginx:alpine
```

[PDF](#)
[Slack](#)
[Support](#)

⌚2023 9 26

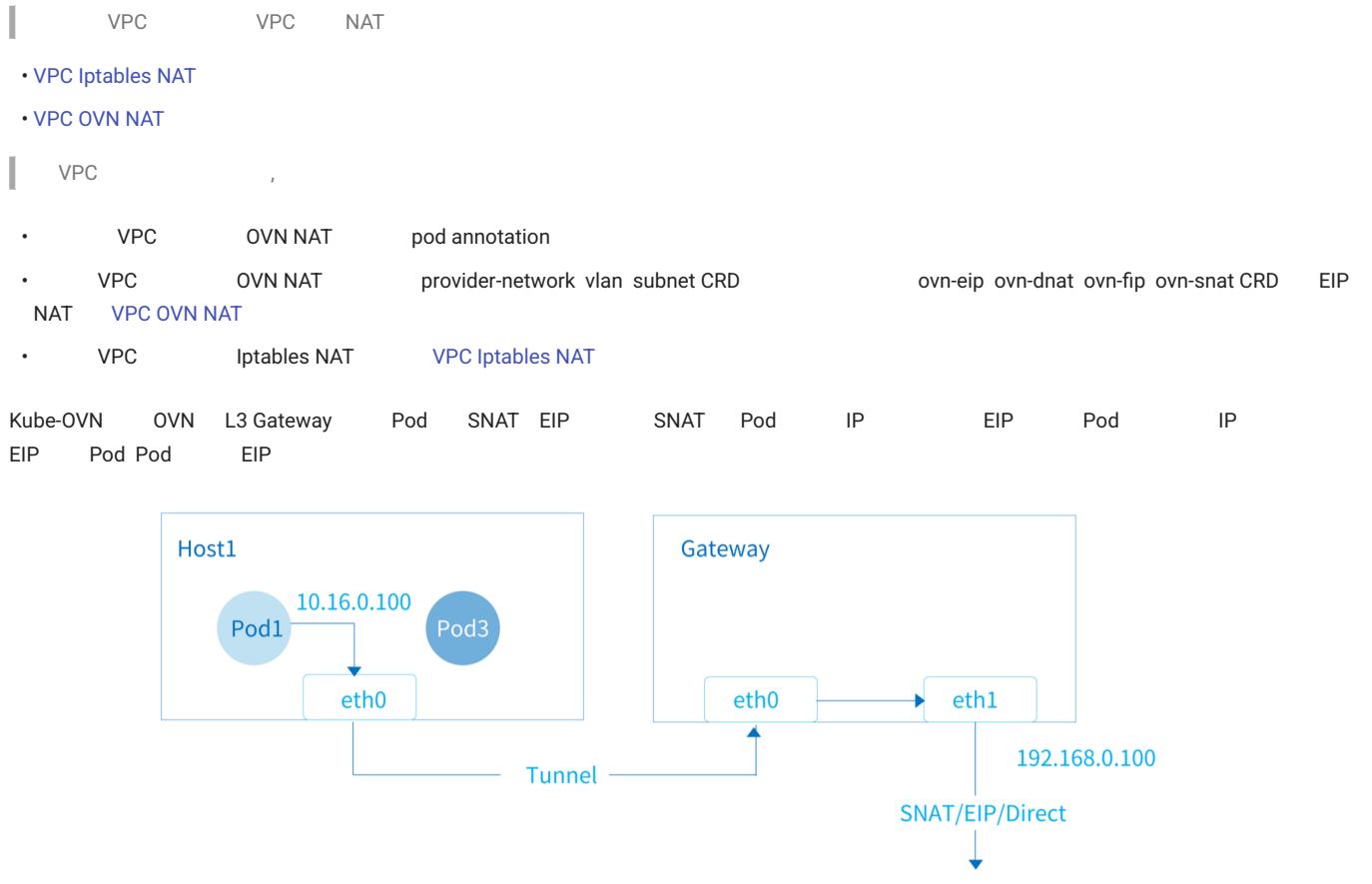
⌚2023 2 16



GitHub 


3.4.1

3.5 EIP SNAT



3.5.1

- OVN L3 Gateway OVS Overlay Underlay
- NAT Underlay
- EIP SNAT

3.5.2

kube-system ConfigMap ovn-external-gw-config

```
apiVersion: v1
kind: ConfigMap
metadata:
  name: ovn-external-gw-config
  namespace: kube-system
data:
  enable-external-gw: "true"
  external-gw-nodes: "kube-ovn-worker"
  external-gw-nic: "eth1"
  external-gw-addr: "172.56.0.1/16"
```

```

nic-ip: "172.56.0.254/16"
nic-mac: "16:52:f3:13:6a:25"

• enable-external-gw: SNAT EIP
• type: centralized distributed centralized distributed
• external-gw-nodes: centralized
• external-gw-nic:
• external-gw-addr: IP
• nic-ip, nic-mac: IP Mac IP Mac

```

3.5.3 OVN OVS

OVN-NB	ovn-external	ovn-cluster-ovn-external	chassis
--------	--------------	--------------------------	---------

```

# kubectl get nbct1 show
switch 3de4cead7-1a71-43f3-8b62-435a57ef16a6 (external)
  port localnet.external
    type: localnet
    addresses: ["unknown"]
  port external-ovn-cluster
    type: router
    router-port: ovn-cluster-external
router e1eb83ad-34be-4ed5-9a02-fcc8b1d357c4 (ovn-cluster)
  port ovn-cluster-external
    mac: "ac:1f:6b:2d:33:f1"
    networks: ["172.56.0.100/16"]
  gateway chassis: [a5682814-2e2c-46dd-9c1c-6803ef0dab66]

```

OVS	br-external
-----	-------------

```

# kubectl get vsctl ${gateway node name} show
e7d81150-7743-4d6e-9e6f-5c688232e130
  Bridge br-external
    Port br-external
      Interface br-external
        type: internal
    Port eth1
      Interface eth1
      Port patch-localnet.external-to-br-int
        Interface patch-localnet.external-to-br-int
          type: patch
          options: {peer=patch-br-int-to-localnet.external}

```

3.5.4 Pod EIP SNAT

Pod	ovn.kubernetes.io/snat	ovn.kubernetes.io/eip annotation	SNAT EIP
-----	------------------------	----------------------------------	----------

```

apiVersion: v1
kind: Pod
metadata:
  name: pod-snat
  annotations:
    ovn.kubernetes.io/snat: 172.56.0.200
spec:
  containers:
    - name: pod-snat
      image: docker.io/library/nginx:alpine
---
apiVersion: v1
kind: Pod
metadata:
  name: pod-eip
  annotations:
    ovn.kubernetes.io/eip: 172.56.0.233
spec:
  containers:
    - name: pod-eip
      image: docker.io/library/nginx:alpine

```

kubectl	Pod	EIP SNAT	ovn.kubernetes.io/routed annotation
---------	-----	----------	-------------------------------------

```

kubectl annotate pod pod-gw ovn.kubernetes.io/eip=172.56.0.221 --overwrite
kubectl annotate pod pod-gw ovn.kubernetes.io/routed-

```

EIP SNAT	ovn.kubernetes.io/routed annotation
----------	-------------------------------------

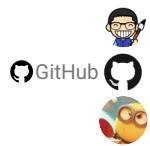
3.5.5

```
kube-ovn-controller      SNAT   EIP  
• --external-gateway-config-ns: Configmap ovn-external-gw-config  Namespace  kube-system  
• --external-gateway-net:           external  
• --external-gateway-vlanid:     Vlan Tag    0    Vlan
```

[!\[\]\(05abdec45d3d9667a7f3c64e46754c68_img.jpg\) PDF](#)[!\[\]\(b246cff486eb7c4dec085261afe05404_img.jpg\) Slack](#)[!\[\]\(89b88375a21fd57206ff9ec1ea24d933_img.jpg\) Support](#)

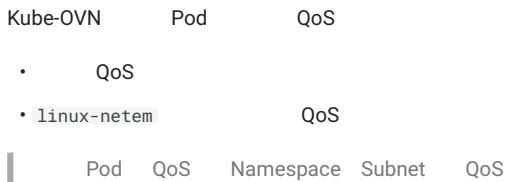
⌚2025 9 10

⌚2022 5 24



3.5.6

3.6 QoS



3.6.1 QoS

QoS Pod annotation Pod Mbit/s

```

apiVersion: v1
kind: Pod
metadata:
  name: qos
  namespace: ls1
  annotations:
    ovn.kubernetes.io/ingress_rate: "3"
    ovn.kubernetes.io/egress_rate: "1"
spec:
  containers:
  - name: qos
    image: docker.io/library/nginx:alpine
  
```

annotation QoS

```
kubectl annotate --overwrite pod nginx-74d5899f46-d7qkn ovn.kubernetes.io/ingress_rate=3
```

QoS

```

kind: DaemonSet
apiVersion: apps/v1
metadata:
  name: perf
  namespace: ls1
  labels:
    app: perf
spec:
  selector:
    matchLabels:
      app: perf
  template:
    metadata:
      labels:
        app: perf
    spec:
      containers:
      - name: nginx
        image: docker.io/kubeovn/perf
  
```

Pod iperf3 server

```
# kubectl exec -it perf-4n4gt -n ls1 sh
# iperf3 -s
-----
Server listening on 5201
-----
```

Pod Pod

```

# kubectl exec -it perf-d4mqc -n ls1 sh
# iperf3 -c 10.66.0.12
Connecting to host 10.66.0.12, port 5201
[  4] local 10.66.0.14 port 51544 connected to 10.66.0.12 port 5201
[ ID] Interval           Transfer     Bandwidth   Retr  Cwnd
[  4]  0.00-1.00  sec   86.4 MBytes   725 Mbits/sec  3   350 KBytes
[  4]  1.00-2.00  sec   89.9 MBytes   754 Mbits/sec 118  473 KBytes
[  4]  2.00-3.00  sec   101 MBytes   848 Mbits/sec 184  586 KBytes
[  4]  3.00-4.00  sec   104 MBytes   875 Mbits/sec 217  671 KBytes
[  4]  4.00-5.00  sec   111 MBytes   935 Mbits/sec 175  772 KBytes
  
```

```
[ 4] 5.00-6.00 sec 100 MBytes 840 Mbits/sec 658 598 KBytes
[ 4] 6.00-7.00 sec 106 MBytes 890 Mbits/sec 742 668 KBytes
[ 4] 7.00-8.00 sec 102 MBytes 857 Mbits/sec 764 724 KBytes
[ 4] 8.00-9.00 sec 97.4 MBytes 817 Mbits/sec 1175 764 KBytes
[ 4] 9.00-10.00 sec 111 MBytes 934 Mbits/sec 1083 838 KBytes
-----
[ ID] Interval Transfer Bandwidth Retr
[ 4] 0.00-10.00 sec 1010 MBytes 848 Mbits/sec 5119 sender
[ 4] 0.00-10.00 sec 1008 MBytes 846 Mbits/sec receiver

iperf Done.
```

Pod QoS

```
kubectl annotate --overwrite pod perf-4n4gt -n ls1 ovn.kubernetes.io/ingress_rate=30
```

Pod Pod

```
# iperf3 -c 10.66.0.12
Connecting to host 10.66.0.12, port 5201
[ 4] local 10.66.0.14 port 52372 connected to 10.66.0.12 port 5201
[ ID] Interval Transfer Bandwidth Retr Cwnd
[ 4] 0.00-1.00 sec 3.66 MBytes 30.7 Mbits/sec 2 76.1 KBytes
[ 4] 1.00-2.00 sec 3.43 MBytes 28.8 Mbits/sec 0 104 KBytes
[ 4] 2.00-3.00 sec 3.50 MBytes 29.4 Mbits/sec 0 126 KBytes
[ 4] 3.00-4.00 sec 3.50 MBytes 29.3 Mbits/sec 0 144 KBytes
[ 4] 4.00-5.00 sec 3.43 MBytes 28.8 Mbits/sec 0 160 KBytes
[ 4] 5.00-6.00 sec 3.43 MBytes 28.8 Mbits/sec 0 175 KBytes
[ 4] 6.00-7.00 sec 3.50 MBytes 29.3 Mbits/sec 0 212 KBytes
[ 4] 7.00-8.00 sec 3.68 MBytes 30.9 Mbits/sec 0 294 KBytes
[ 4] 8.00-9.00 sec 3.74 MBytes 31.4 Mbits/sec 0 398 KBytes
[ 4] 9.00-10.00 sec 3.80 MBytes 31.9 Mbits/sec 0 526 KBytes
-----
[ ID] Interval Transfer Bandwidth Retr
[ 4] 0.00-10.00 sec 35.7 MBytes 29.9 Mbits/sec 2 sender
[ 4] 0.00-10.00 sec 34.5 MBytes 29.0 Mbits/sec receiver

iperf Done.
```

3.6.2 linux-netem QoS

RHEL netem yum install -y kernel-modules-extra && modprobe sch_netem

Pod	annotation	linux-netem	QoS	netem	QoS	Pod	Ingress
• ovn.kubernetes.io/latency	Pod			ms			
• ovn.kubernetes.io/jitter	Pod			ms			
• ovn.kubernetes.io/limit		qdisc			1000		
• ovn.kubernetes.io/loss			float	20	20%		

[PDF](#)

[Slack](#)

[Support](#)

⌚2025 10 10

⌚2022 5 24



3.6.3

3.7



3.7.1

CIDR cidr=<IPv4 CIDR>,<IPv6 CIDR> CIDR IPv4 IPv6

```

apiVersion: kubeovn.io/v1
kind: Subnet
metadata:
  name: ovn-test
spec:
  cidrBlock: 10.16.0.0/16,fd00:10:16::/64
  excludeIps:
  - 10.16.0.1
  - fd00:10:16::1
  gateway: 10.16.0.1,fd00:10:16::1
  
```

```

POD_CIDR="10.16.0.0/16,fd00:10:16::/64"
JOIN_CIDR="100.64.0.0/16,fd00:100:64::/64"
  
```

3.7.2 Pod

Pod IPv4 IPv6 Pod annotation :

```

apiVersion: v1
kind: Pod
metadata:
  annotations:
    ovn.kubernetes.io/allocated: "true"
    ovn.kubernetes.io/cidr: 10.16.0.0/16,fd00:10:16::/64
    ovn.kubernetes.io/gateway: 10.16.0.1,fd00:10:16::1
    ovn.kubernetes.io/ip_address: 10.16.0.9,fd00:10:16::9
    ovn.kubernetes.io/logical_switch: ovn-default
    ovn.kubernetes.io/mac_address: 00:00:00:14:88:09
    ovn.kubernetes.io/network_types: geneve
    ovn.kubernetes.io/routed: "true"
...
podIP: 10.16.0.9
  podIPs:
  - ip: 10.16.0.9
  - ip: fd00:10:16::9
  
```

[PDF](#)

[Slack](#)

[Support](#)

⌚2025 9 10

⌚2022 5 24



3.7.3

3.8 Webhook

```
Webhook   Kube-OVN   CRD           Webhook   IP   Subnet CIDR
Webhook   Subnet   Pod           Kube-OVN   Webhook   Pod
```

3.8.1 Cert-Manager

```
Webhook   cert-manager   Webhook   cert-manager
```

cert-manager:

```
kubectl apply -f https://github.com/cert-manager/cert-manager/releases/download/v1.8.0/cert-manager.yaml
```

```
cert-manager   cert-manager
```

3.8.2 Webhook

```
Webhook   yaml   :
```

```
# kubectl apply -f https://raw.githubusercontent.com/kubeovn/kube-ovn/release-1.14/yamls/webhook.yaml
deployment.apps/kube-ovn-webhook created
service/kube-ovn-webhook created
validatingwebhookconfiguration.admissionregistration.k8s.io/kube-ovn-webhook created
certificate.cert-manager.io/kube-ovn-webhook-serving-cert created
issuer.cert-manager.io/kube-ovn-webhook-selfsigned-issuer created
```

3.8.3 Webhook

Pod Pod IP 10.16.0.15

```
# kubectl get pod -o wide
NAME          READY   STATUS    RESTARTS   AGE     IP          NODE      NOMINATED NODE   READINESS GATES
static-7584848b74-fw9dm   1/1     Running   0        2d13h   10.16.0.15   kube-ovn-worker   <none>
```

yaml IP Pod

```
apiVersion: v1
kind: Pod
metadata:
  annotations:
    ovn.kubernetes.io/ip_address: 10.16.0.15
    ovn.kubernetes.io/mac_address: 00:00:00:53:6B:B6
  labels:
    app: static
  managedFields:
    name: staticip-pod
    namespace: default
spec:
  containers:
  - image: docker.io/library/nginx:alpine
    imagePullPolicy: IfNotPresent
    name: qatest
```

yaml Pod IP

```
# kubectl apply -f pod-static.yaml
Error from server (annotation ip address 10.16.0.15 is conflict with ip crd static-7584848b74-fw9dm.default 10.16.0.15): error when creating "pod-static.yaml": admission webhook "pod-ip-validation.kube-ovn.io" denied the request: annotation ip address 10.16.0.15 is conflict with ip crd static-7584848b74-fw9dm.default 10.16.0.15
```



PDF



Slack



Support

⌚2023 5 9

⌚2022 5 24

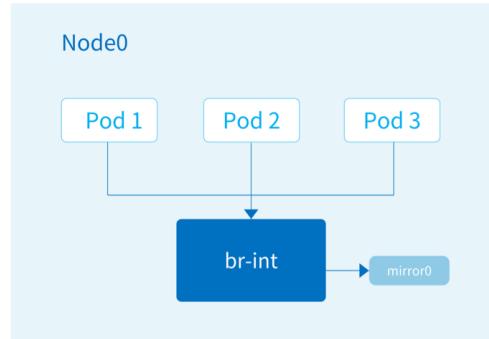
⌚GitHub 🧑‍💻

3.8.4

3.9

NPM

CPU	5%~10%	CPU
-----	--------	-----



3.9.1

kube-ovn-cni DaemonSet

- `--enable-mirror=true`
- `--mirror-iface=mirror0:` br-int Kube-OVN
- `mirror0`
- tcpdump `mirror0`

```
tcpdump -ni mirror0
```

3.9.2 Pod

Pod Pod ovn.kubernetes.io/mirror annotation Pod

```

apiVersion: v1
kind: Pod
metadata:
  name: mirror-pod
  namespace: ls1
  annotations:
    ovn.kubernetes.io/mirror: "true"
spec:
  containers:
  - name: mirror-pod
    image: docker.io/library/nginx:alpine

```

3.9.3

1. Pod to Pod in the same Nodes

Size	TCP Latency	TCP Bandwidth	UDP Latency	UDP Lost Rate	UDP Bandwidth
64	12.7 us	289 Mbits/sec	12.6 us	(1.8%)	77.9 Mbits/sec
128	15.5 us	517 Mbits/sec	12.7 us	(0%)	155 Mbits/sec
512	12.2 us	1.64 Gbits/sec	12.4 us	(0%)	624 Mbits/sec
1k	13 us	2.96 Gbits/sec	11.4 us	(0.53%)	1.22 Gbits/sec
4k	18 us	7.67 Gbits/sec	25.7 us	(0.41%)	1.50 Gbits/sec

Size	TCP Latency	TCP Bandwidth	UDP Latency	UDP Lost Rate	UDP Bandwidth
64	11.9 us	324 Mbits/sec	12.2 us	(0.22%)	102 Mbits/sec
128	10.5 us	582 Mbits/sec	9.5 us	(0.21%)	198 Mbits/sec
512	11.6 us	1.84 Gbits/sec	9.32 us	(0.091%)	827 Mbits/sec
1k	10.5 us	3.44 Gbits/sec	10 us	(1.2%)	1.52 Gbits/sec
4k	16.7 us	8.52 Gbits/sec	18.2 us	(1.3%)	2.42 Gbits/sec

2. Pod to Pod in the different Nodes

Size	TCP Latency	TCP Bandwidth	UDP Latency	UDP Lost Rate	UDP Bandwidth
64	258 us	143 Mbits/sec	237 us	(61%)	28.5 Mbits/sec
128	240 us	252 Mbits/sec	231 us	(64%)	54.9 Mbits/sec
512	236 us	763 Mbits/sec	256 us	(68%)	194 Mbits/sec
1k	242 us	969 Mbits/sec	225 us	(62%)	449 Mbits/sec
4k	352 us	1.12 Gbits/sec	382 us	(0.71%)	21.4 Mbits/sec

Size	TCP Latency	TCP Bandwidth	UDP Latency	UDP Lost Rate	UDP Bandwidth
64	278 us	140 Mbits/sec	227 us	(24%)	59.6 Mbits/sec
128	249 us	265 Mbits/sec	265 us	(23%)	114 Mbits/sec
512	233 us	914 Mbits/sec	235 us	(21%)	468 Mbits/sec
1k	238 us	1.14 Gbits/sec	240 us	(15%)	891 Mbits/sec
4k	370 us	1.25 Gbits/sec	361 us	(0.43%)	7.54 Mbits/sec

3. Node to Node

Size	TCP Latency	TCP Bandwidth	UDP Latency	UDP Lost Rate	UDP Bandwidth
64	205 us	162 Mbits/sec	183 us	(11%)	74.2 Mbits/sec
128	222 us	280 Mbits/sec	206 us	(6.3%)	155 Mbits/sec
512	220 us	1.04 Gbits/sec	177 us	(20%)	503 Mbits/sec
1k	213 us	2.06 Gbits/sec	201 us	(8.6%)	1.14 Gbits/sec
4k	280 us	5.01 Gbits/sec	315 us	(37%)	1.20 Gbits/sec

Size	TCP Latency	TCP Bandwidth	UDP Latency	UDP Lost Rate	UDP Bandwidth
64	204 us	157 Mbits/sec	204 us	(8.8%)	81.9 Mbits/sec
128	213 us	262 Mbits/sec	225 us	(19%)	136 Mbits/sec
512	220 us	1.02 Gbits/sec	227 us	(21%)	486 Mbits/sec
1k	217 us	1.79 Gbits/sec	218 us	(29%)	845 Mbits/sec
4k	275 us	5.27 Gbits/sec	336 us	(34%)	1.21 Gbits/sec

4. Pod to the Node where the Pod is located

Size	TCP Latency	TCP Bandwidth	UDP Latency	UDP Lost Rate	UDP Bandwidth
64	12.2 us	295 Mbits/sec	12.7 us	(0.27%)	74.1 Mbits/sec
128	14.1 us	549 Mbits/sec	10.6 us	(0.41%)	153 Mbits/sec
512	13.5 us	1.83 Gbits/sec	12.7 us	(0.23%)	586 Mbits/sec
1k	12 us	2.69 Gbits/sec	13 us	(1%)	1.16 Gbits/sec
4k	18.9 us	4.51 Gbits/sec	21.8 us	(0.42%)	1.81 Gbits/sec

Size	TCP Latency	TCP Bandwidth	UDP Latency	UDP Lost Rate	UDP Bandwidth
64	10.4 us	335 Mbits/sec	12.2 us	(0.75%)	95.4 Mbits/sec
128	12.1 us	561 Mbits/sec	11.3 us	(0.25%)	194 Mbits/sec
512	11.6 us	1.87 Gbits/sec	10.7 us	(0.66%)	745 Mbits/sec
1k	12.7 us	3.12 Gbits/sec	10.9 us	(1.2%)	1.46 Gbits/sec
4k	16.5 us	8.23 Gbits/sec	17.9 us	(1.5%)	2.51 Gbits/sec

5. Pod to the Node where the Pod is not located

Size	TCP Latency	TCP Bandwidth	UDP Latency	UDP Lost Rate	UDP Bandwidth
64	234 us	153 Mbits/sec	232 us	(63%)	29.4 Mbits/sec
128	237 us	261 Mbits/sec	238 us	(49%)	76.1 Mbits/sec
512	231 us	701 Mbits/sec	238 us	(57%)	279 Mbits/sec
1k	256 us	1.05 Gbits/sec	228 us	(56%)	524 Mbits/sec
4k	330 us	1.08 Gbits/sec	359 us	(1.5%)	35.7 Mbits/sec

Size	TCP Latency	TCP Bandwidth	UDP Latency	UDP Lost Rate	UDP Bandwidth
64	283 us	141 Mbits/sec	230 us	(26%)	55.8 Mbits/sec
128	234 us	255 Mbits/sec	234 us	(25%)	113 Mbits/sec
512	246 us	760 Mbits/sec	234 us	(22%)	458 Mbits/sec
1k	268 us	1.23 Gbits/sec	242 us	(20%)	879 Mbits/sec
4k	326 us	1.20 Gbits/sec	369 us	(0.5%)	7.87 Mbits/sec

6. Pod to the cluster ip service

Size	TCP Latency	TCP Bandwidth	UDP Latency	UDP Lost Rate	UDP Bandwidth
64	237 us	133 Mbits/sec	213 us	(65%)	25.5 Mbits/sec
128	232 us	271 Mbits/sec	222 us	(62%)	54.8 Mbits/sec
512	266 us	800 Mbits/sec	234 us	(60%)	232 Mbits/sec
1k	248 us	986 Mbits/sec	239 us	(50%)	511 Mbits/sec
4k	314 us	1.03 Gbits/sec	367 us	(0.6%)	13.2 Mbits/sec

TCP-Conn-Number	QPS	Avg-Resp-Time	Stdev-Resp-Time	Max-Resp-Time
10	14305.17	0.87ms	1.48ms	24.46ms
100	29082.07	3.87ms	4.35ms	102.85ms

Size	TCP Latency	TCP Bandwidth	UDP Latency	UDP Lost Rate	UDP Bandwidth
64	241 us	145 Mbits/sec	225 us	(19%)	60.2 Mbits/sec
128	245 us	261 Mbits/sec	212 us	(15%)	123 Mbits/sec
512	252 us	821 Mbits/sec	219 us	(14%)	499 Mbits/sec
1k	253 us	1.08 Gbits/sec	242 us	(16%)	852 Mbits/sec
4k	320 us	1.32 Gbits/sec	360 us	(0.47%)	6.70 Mbits/sec

TCP-Conn-Number	QPS	Avg-Resp-Time	Stdev-Resp-Time	Max-Resp-Time
10	13634.07	0.96ms	1.72ms	30.07ms
100	30215.23	3.59ms	3.20ms	77.56ms

7. Host to the Node port service where the Pod is not located on the target Node

TCP-Conn-Number	QPS	Avg-Resp-Time	Stdev-Resp-Time	Max-Resp-Time
10	14802.73	0.88ms	1.66ms	31.49ms
100	29809.58	3.78ms	4.12ms	105.34ms

TCP-Conn-Number	QPS	Avg-Resp-Time	Stdev-Resp-Time	Max-Resp-Time
10	14273.33	0.90ms	1.60ms	37.16ms
100	30757.81	3.62ms	3.41ms	59.78ms

8. Host to the Node port service where the Pod is located on the target Node

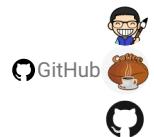
TCP-Conn-Number	QPS	Avg-Resp-Time	Stdev-Resp-Time	Max-Resp-Time
10	15402.39	802.50us	1.42ms	30.91ms
100	29424.66	4.05ms	4.31ms	90.60ms

TCP-Conn-Number	QPS	Avg-Resp-Time	Stdev-Resp-Time	Max-Resp-Time
10	14649.21	0.91ms	1.72ms	43.92ms
100	32143.61	3.66ms	3.76ms	67.02ms



⌚2025 11 20

⌚2022 5 24



3.9.4

3.10 NetworkPolicy

NetworkPolicy	Kubernetes	Kube-OVN	OVN	ACL	NetworkPolicy	NetworkPolicy
Kube-OVN	NetworkPolicy			NetworkPolicy Drop		

⚠ Warning

NetworkPolicy	Drop	CPU	OVN	ACL Log Meter
ACL Log				

3.10.1 NetworkPolicy

NetworkPolicy annotation ovn.kubernetes.io/enable_log

```
apiVersion: networking.k8s.io/v1
kind: NetworkPolicy
metadata:
  name: default-deny-ingress
  namespace: kube-system
  annotations:
    ovn.kubernetes.io/enable_log: "true"
spec:
  podSelector: {}
  policyTypes:
  - Ingress
```

Pod /var/log/ovn/ovn-controller.log

```
# tail -f /var/log/ovn/ovn-controller.log
2022-07-20T05:55:03.229Z|00394|acl_log(ovn_pinctrl0)|INFO|name="np/default-deny-ingress.kube-system/IPv4/0", verdict=drop, severity=warning, direction=to-lport: udp,vlan_tci=0x0000,d1_src=00:00:00:21:b7:d1,d1_dst=00:00:00:8d:0b:86,nw_src=10.16.0.10,nw_dst=10.
16.0.7,nw_tos=0,nw_ecn=0,nw_ttl=63,tp_src=54343,tp_dst=53
2022-07-20T05:55:06.229Z|00395|acl_log(ovn_pinctrl0)|INFO|name="np/default-deny-ingress.kube-system/IPv4/0", verdict=drop, severity=warning, direction=to-lport: udp,vlan_tci=0x0000,d1_src=00:00:00:21:b7:d1,d1_dst=00:00:00:8d:0b:86,nw_src=10.16.0.9,nw_dst=10.
16.0.7,nw_tos=0,nw_ecn=0,nw_ttl=63,tp_src=44187,tp_dst=53
2022-07-20T05:55:08.230Z|00396|acl_log(ovn_pinctrl0)|INFO|name="np/default-deny-ingress.kube-system/IPv4/0", verdict=drop, severity=warning, direction=to-lport: udp,vlan_tci=0x0000,d1_src=00:00:00:21:b7:d1,d1_dst=00:00:00:8d:0b:86,nw_src=10.16.0.10,nw_dst=10.
16.0.7,nw_tos=0,nw_ecn=0,nw_ttl=63,tp_src=54274,tp_dst=53
2022-07-20T05:55:11.231Z|00397|acl_log(ovn_pinctrl0)|INFO|name="np/default-deny-ingress.kube-system/IPv4/0", verdict=drop, severity=warning, direction=to-lport: udp,vlan_tci=0x0000,d1_src=00:00:00:21:b7:d1,d1_dst=00:00:00:8d:0b:86,nw_src=10.16.0.9,nw_dst=10.
16.0.7,nw_tos=0,nw_ecn=0,nw_ttl=63,tp_src=32778,tp_dst=53
2022-07-20T05:55:11.231Z|00398|acl_log(ovn_pinctrl0)|INFO|name="np/default-deny-ingress.kube-system/IPv4/0", verdict=drop, severity=warning, direction=to-lport: udp,vlan_tci=0x0000,d1_src=00:00:00:21:b7:d1,d1_dst=00:00:00:8d:0b:86,nw_src=10.16.0.9,nw_dst=10.
16.0.7,nw_tos=0,nw_ecn=0,nw_ttl=63,tp_src=34188,tp_dst=53
```

3.10.2 NetworkPolicy

ovn.kubernetes.io/enable_log annotation	Drop ACL	ACL		
Kube-OVN v1.13.0	annotation ovn.kubernetes.io/log_acl_actions	ACL	annotation	allow
NetworkPolicy	annotation ovn.kubernetes.io/log_acl_actions			

```
apiVersion: networking.k8s.io/v1
kind: NetworkPolicy
metadata:
  name: default-deny-ingress
  namespace: kube-system
  annotations:
    ovn.kubernetes.io/enable_log: "true"
    ovn.kubernetes.io/log_acl_actions: "allow"
spec:
  podSelector: {}
  policyTypes:
  - Ingress
```

Pod Pod /var/log/ovn/ovn-controller.log Allow ACL

```
2024-08-14T09:27:49.590Z|00004|acl_log(ovn_pinctrl0)|INFO|name="np/test.default/ingress/IPv4/0", verdict=allow, severity=info, direction=to-lport: icmp,vlan_tci=0x0000,d1_src=96:7b:0:2f:a0:1a,d1_dst=a6:e5:1b:c2:1b:f8,nw_src=10.16.0.7,nw_dst=10.
16.0.12,nw_tos=0,nw_ecn=0,nw_ttl=64,nw_frag=no,icmp_type=8,icmp_code=0
```

3.10.3 NetworkPolicy

```
NetworkPolicy annotation ovn.kubernetes.io/enable_log false NetworkPolicy
```

```
kubectl annotate networkpolicy -n kube-system default-deny-ingress ovn.kubernetes.io/enable_log=false --overwrite
```

3.10.4 AdminNetworkPolicy BaselineAdminNetworkPolicy

v1.13.0	Kube-OVN	AdminNetworkPolicy	BaselineAdminNetworkPolicy	AdminNetworkPolicy	BaselineAdminNetworkPolicy
		Network Policy API			

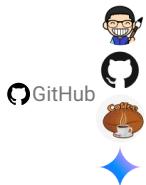
Network Policy API	ovn.kubernetes.io/log_acl_actions annotation	action ACL	annotation	allow, drop, pass
--------------------	--	------------	------------	-------------------

ovn.kubernetes.io/enable_log annotation	annotation	ovn.kubernetes.io/log_acl_actions
annotation		

[!\[\]\(974977f0ae7a9bd8eb2b52f26d2700fe_img.jpg\) PDF](#)
[!\[\]\(646869a5c9c88531ae41e50ad8af3b92_img.jpg\) Slack](#)
[!\[\]\(9500e4e0b06f494c3391c456117a0e37_img.jpg\) Support](#)

 2025 9 17

 2022 7 20



3.10.5

3.11 LoadBalancer Service

```
Kube-OVN VPC VPC VPC
VPC VPC VPC LoadBalancer Service LoadBalancerIP VPC Service
```

1. multus-cni macvlan cni
2. LoadBalancer Service VPC vpc-nat-gw macvlan
3. VPC VPC LoadBalancer VPC VPC

3.11.1 VPC LoadBalancer Service

```
kube-system namespace deployment kube-ovn-controller args --enable-lb-svc=true false
```

```
containers:
- args:
  - /kube-ovn/start-controller.sh
  - --default-cidr=10.16.0.0/16
  - --default-gateway=10.16.0.1
  - --default-gateway-check=true
  - --enable-lb-svc=true
    // true
```

NetworkAttachmentDefinition CRD

```
yaml net-attach-def
```

```
apiVersion: "k8s.cni.cncf.io/v1"
kind: NetworkAttachmentDefinition
metadata:
  name: lb-svc-attachment
  namespace: kube-system
spec:
  config: '{
    "cniVersion": "0.3.0",
    "type": "macvlan",
    "master": "eth0",
    "mode": "bridge"
  }'
```

```
eth0 master
```

Subnet

```
Subnet LoadBalancer Service LoadBalancerIP Underlay Subnet
```

```
yaml
```

```
apiVersion: kubeovn.io/v1
kind: Subnet
metadata:
  name: attach-subnet
spec:
  protocol: IPv4
  provider: lb-svc-attachment.kube-system # provider
  cidrBlock: 172.18.0.0/16
  gateway: 172.18.0.1
  excludeIps:
  - 172.18.0.0..172.18.0.10
```

```
Subnet provider ovn .ovn Kube-OVN logical switch
provider ovn .ovn Kube-OVN IPAM IP
```

LoadBalancer Service

yaml LoadBalancer Service

```
apiVersion: v1
kind: Service
metadata:
  annotations:
    lb-svc-attachment.kube-system.kubernetes.io/logical_switch: attach-subnet  #
    ovn.kubernetes.io/attachmentprovider: lb-svc-attachment.kube-system          #
  labels:
    app: dynamic
    name: test-service
    namespace: default
spec:
  loadBalancerIP: 172.18.0.18
  ports:
    - name: test
      protocol: TCP
      port: 80
      targetPort: 80
  selector:
    app: dynamic
    sessionAffinity: None
  type: LoadBalancer
```

yaml annotation ovn.kubernetes.io/attachmentprovider net-attach-def Name.Namespace annotation Pod
net-attach-def

annotation	annotation key	net-attach-def	Name.Namespace.kubernetes.io/logical_switch
LoadBalancerIP	LoadBalancerIP		
LoadBalancerIP	spec.loadBalancerIP		

yaml Service Service Namespace Pod

```
# kubectl get pod
NAME                               READY   STATUS    RESTARTS   AGE
lb-svc-test-service-6869d98dd8-cjvll   1/1     Running   0          107m
# kubectl get svc
NAME            TYPE        CLUSTER-IP      EXTERNAL-IP      PORT(S)        AGE
test-service    LoadBalancer  10.109.201.193  172.18.0.18   80:30056/TCP   107m
```

service.spec.loadBalancerIP service external-ip

Pod yaml

```
# kubectl get pod -o yaml lb-svc-test-service-6869d98dd8-cjvll
apiVersion: v1
kind: Pod
metadata:
  annotations:
    k8s.v1.cni.cncf.io/network-status: |-
      [ {
        "name": "kube-ovn",
        "ips": [
          "10.16.0.2"
        ],
        "default": true,
        "dns": {}
      }, {
        "name": "default/test-service",
        "interface": "net1",
        "mac": "ba:85:f7:02:9f:42",
        "dns": {}
      } ]
    k8s.v1.cni.cncf.io/networks: default/test-service
    k8s.v1.cni.cncf.io/networks-status: |-
      [ {
        "name": "kube-ovn",
        "ips": [
          "10.16.0.2"
        ],
        "default": true,
        "dns": {}
      }, {
        "name": "default/test-service",
        "interface": "net1",
        "mac": "ba:85:f7:02:9f:42",
        "dns": {}
      } ]
  ovn.kubernetes.io/allocated: "true"
  ovn.kubernetes.io/cidr: 10.16.0.0/16
  ovn.kubernetes.io/gateway: 10.16.0.1
```

```

ovn.kubernetes.io/ip_address: 10.16.0.2
ovn.kubernetes.io/logical_router: ovn-cluster
ovn.kubernetes.io/logical_switch: ovn-default
ovn.kubernetes.io/mac_address: 00:00:00:45:F4:29
ovn.kubernetes.io/pod_nic_type: veth-pair
ovn.kubernetes.io/routed: "true"
test-service.default.kubernetes.io/allocated: "true"
test-service.default.kubernetes.io/cidr: 172.18.0.0/16
test-service.default.kubernetes.io/gateway: 172.18.0.1
test-service.default.kubernetes.io/ip_address: 172.18.0.18
test-service.default.kubernetes.io/logical_switch: attach-subnet
test-service.default.kubernetes.io/mac_address: 00:00:00:AF:AA:BF
test-service.default.kubernetes.io/pod_nic_type: veth-pair

```

Service

```

# kubectl get svc -o yaml test-service
apiVersion: v1
kind: Service
metadata:
  annotations:
    kubectl.kubernetes.io/last-applied-configuration: |
      {"apiVersion":"v1","kind":"Service","metadata":{"annotations":{"test-service.default.kubernetes.io/logical_switch":"attach-subnet"},"labels":{"app":"dynamic","name":"test-service","namespace":"default"},"spec":{"ports":[{"name":"test","port":80,"protocol":"TCP","targetPort":80}],"selector":{"app":"dynamic","sessionAffinity":"None","type":"LoadBalancer"}}}
    ovn.kubernetes.io/vpc: ovn-cluster
    test-service.default.kubernetes.io/logical_switch: attach-subnet
  creationTimestamp: "2022-06-15T09:01:58Z"
  labels:
    app: dynamic
    name: test-service
    namespace: default
  resourceVersion: "38485"
  uid: 161edee1-7f6e-40f5-9e09-5a52c44267d0
spec:
  allocateLoadBalancerNodePorts: true
  clusterIP: 10.109.201.193
  clusterIPs:
  - 10.109.201.193
  externalTrafficPolicy: Cluster
  internalTrafficPolicy: Cluster
  ipFamilies:
  - IPv4
  ipFamilyPolicy: SingleStack
  ports:
  - name: test
    nodePort: 30056
    port: 80
    protocol: TCP
    targetPort: 80
  selector:
    app: dynamic
    sessionAffinity: None
    type: LoadBalancer
  status:
    loadBalancer:
      ingress:
      - ip: 172.18.0.18

```

3.11.2 LoadBalancerIP

yaml, Pod Service Endpoints

```

apiVersion: apps/v1
kind: Deployment
metadata:
  labels:
    app: dynamic
    name: dynamic
    namespace: default
spec:
  replicas: 2
  selector:
    matchLabels:
      app: dynamic
  strategy:
    rollingUpdate:
      maxSurge: 25%
      maxUnavailable: 25%
    type: RollingUpdate
  template:
    metadata:
      creationTimestamp: null
      labels:
        app: dynamic
    spec:
      containers:
      - image: docker.io/library/nginx:alpine

```

```
imagePullPolicy: IfNotPresent
name: nginx
dnsPolicy: ClusterFirst
restartPolicy: Always
```

Service LoadBalancerIP:Port

```
# curl 172.18.0.18:80
<html>
<head>
<title>Hello World!</title>
<link href='//fonts.googleapis.com/css?family=Open+Sans:400,700' rel='stylesheet' type='text/css'>
<style>
body {
    background-color: white;
    text-align: center;
    padding: 50px;
    font-family: "Open Sans", "Helvetica Neue", Helvetica, Arial, sans-serif;
}
#logo {
    margin-bottom: 40px;
}
</style>
</head>
<body>
    <h1>Hello World!</h1>
        <h3>Links found</h3>
    <h3>I am on dynamic-7d8d7874f5-hsgc4</h3>
    <h3>Cookie =/>
        <b>KUBERNETES</b> listening in 443 available at tcp://10.96.0.1:443<br />
            <h3>my name is hanhouchao!</h3>
        <h3> RequestURI='/'</h3>
</body>
</html>
```

Service Pod

```
# ip a
4: net1@if62: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc noqueue state UP group default
    link/ether ba:85:f7:02:9f:42 brd ff:ff:ff:ff:ff:ff link-netnsid 0
    inet 172.18.0.18/16 scope global net1
        valid_lft forever preferred_lft forever
    inet6 fe80::ba85:f7ff:fe02:9f42/64 scope link
        valid_lft forever preferred_lft forever
36: eth0@if37: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1400 qdisc noqueue state UP group default
    link/ether 00:00:00:45:f4:29 brd ff:ff:ff:ff:ff:ff link-netnsid 0
    inet 10.16.0.2/16 brd 10.16.255.255 scope global eth0
        valid_lft forever preferred_lft forever
    inet6 fe80::200:ff:fe45:f429/64 scope link
        valid_lft forever preferred_lft forever

# ip rule
0: from all lookup local
32764: from all iif eth0 lookup 100
32765: from all iif net1 lookup 100
32766: from all lookup main
32767: from all lookup default

# ip route show table 100
default via 172.18.0.1 dev net1
10.109.201.193 via 10.16.0.1 dev eth0
172.18.0.0/16 dev net1 scope link

# iptables -t nat -L -n -v
Chain PREROUTING (policy ACCEPT 0 packets, 0 bytes)
  pkts bytes target     prot opt in     out     source               destination
    0     0 DNAT       tcp   --  *      *       0.0.0.0/0          172.18.0.18           tcp  dpt:80  to:10.109.201.193:80

Chain INPUT (policy ACCEPT 0 packets, 0 bytes)
  pkts bytes target     prot opt in     out     source               destination

Chain OUTPUT (policy ACCEPT 0 packets, 0 bytes)
  pkts bytes target     prot opt in     out     source               destination

Chain POSTROUTING (policy ACCEPT 0 packets, 0 bytes)
  pkts bytes target     prot opt in     out     source               destination
    0     0 MASQUERADE  all   --  *      *       0.0.0.0/0          10.109.201.193
```

Ib service Pod nodeSelector

```
ovn-vpc-nat-config ConfigMap nodeSelector LoadBalancer service Pod
```

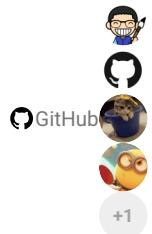
```
apiVersion: v1
data:
  image: docker.io/kubeovn/vpc-nat-gateway:v1.14.0
  nodeSelector: |
    kubernetes.io/hostname: kube-ovn-control-plane
```

```
kubernetes.io/os: linux
kind: ConfigMap
metadata:
  name: ovn-vpc-nat-config
  namespace: kube-system
```

[!\[\]\(8b6ba4e01d044f020536a6874a89c057_img.jpg\) PDF](#)[!\[\]\(39fe6a447a07f214586d259b16ffba42_img.jpg\) Slack](#)[!\[\]\(2a4282dc455b24a8719bbd3b8683d6a8_img.jpg\) Support](#)

⌚2025 9 10

⌚2022 7 19



3.11.3

3.12



3.12.1 Prometheus Monitor

Kube-OVN Prometheus Monitor CRD

```

#
kubectl apply -f https://raw.githubusercontent.com/kubeovn/kube-ovn/master/dist/monitoring/pinger-monitor.yaml
# kube-ovn-controller
kubectl apply -f https://raw.githubusercontent.com/kubeovn/kube-ovn/master/dist/monitoring/controller-monitor.yaml
# kube-ovn-cni
kubectl apply -f https://raw.githubusercontent.com/kubeovn/kube-ovn/master/dist/monitoring/cni-monitor.yaml
# ovn
kubectl apply -f https://raw.githubusercontent.com/kubeovn/kube-ovn/master/dist/monitoring/ovn-monitor.yaml
  
```

Prometheus 15s yaml interval

3.12.2 Grafana

Kube-OVN Grafana Dashboard

Dashboard

```

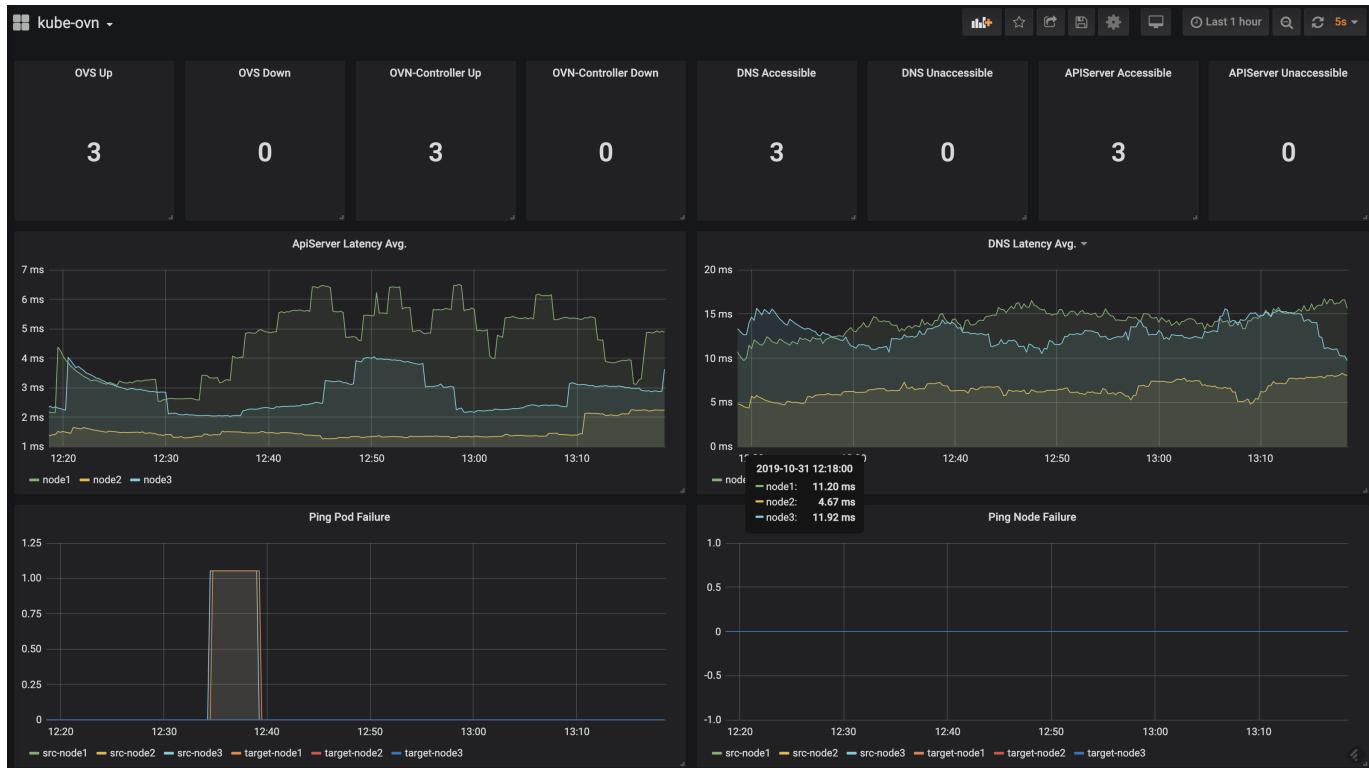
#
wget https://raw.githubusercontent.com/kubeovn/kube-ovn/master/dist/monitoring/pinger-grafana.json
# kube-ovn-controller
wget https://raw.githubusercontent.com/kubeovn/kube-ovn/master/dist/monitoring/controller-grafana.json
# kube-ovn-cni
wget https://raw.githubusercontent.com/kubeovn/kube-ovn/master/dist/monitoring/cni-grafana.json
# ovn
wget https://raw.githubusercontent.com/kubeovn/kube-ovn/master/dist/monitoring/ovn-grafana.json
# ovs
wget https://raw.githubusercontent.com/kubeovn/kube-ovn/master/dist/monitoring/ovs-grafana.json
  
```

Grafana Prometheus Dashboard

kube-ovn-controller



kube-ovn-pinger

**kube-ovn-cni**[PDF](#)[Slack](#)[Support](#)

⌚2025 9 17

⌚2022 5 23



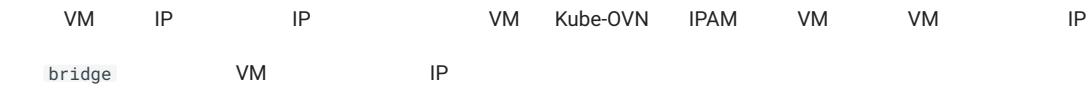
3.12.3

4. KubeVirt

4.1 VM IP



4.1.1 IP VM



1. VM

```

apiVersion: kubevirt.io/v1
kind: VirtualMachine
metadata:
  name: testvm
spec:
  runStrategy: Always
  template:
    metadata:
      labels:
        kubevirt.io/size: small
        kubevirt.io/domain: testvm
      annotations:
        kubevirt.io/allow-pod-bridge-network-live-migration: "true"
    spec:
      domain:
        devices:
          disks:
            - name: containerdisk
              disk:
                bus: virtio
            - name: cloudinitdisk
              disk:
                bus: virtio
          interfaces:
            - name: default
              bridge: {}
        resources:
          requests:
            memory: 64M
      networks:
        - name: default
          pod: {}
      volumes:
        - name: containerdisk
          containerDisk:
            image: quay.io/kubevirt/cirros-container-disk-demo
        - name: cloudinitdisk

```

```
cloudInitNoCloud:
  userDataBase64: SGkuXG4=
```

1. VM

```
kubectl get vmi testvm
```

1. VM

```
virtctl restart testvm
```

1. VM

```
virtctl migrate testvm
```

bridge VM IP

4.1.2 IP/Mac

VM	IP/Mac	VM	annotation	VM	IP	KubeVirt
----	--------	----	------------	----	----	----------

```
apiVersion: kubevirt.io/v1
kind: VirtualMachine
metadata:
  name: testvm
spec:
  runStrategy: Always
  template:
    metadata:
      labels:
        kubevirt.io/size: small
        kubevirt.io/domain: testvm
    annotations:
      ovn.kubernetes.io/ip_address: 10.16.0.15 #(1)
      ovn.kubernetes.io/mac_address: 00:00:00:53:6B:B6 #(2)
      kubevirt.io/allow-pod-bridge-network-live-migration: "true"
    spec:
      domain:
        devices:
          disks:
            - name: containerdisk
              disk:
                bus: virtio
            - name: cloudinitdisk
              disk:
                bus: virtio
          interfaces:
            - name: default
              bridge: {}
        resources:
          requests:
            memory: 64M
      networks:
        - name: default
          pod: {}
      volumes:
        - name: containerdisk
          containerDisk:
            image: quay.io/kubevirt/cirros-container-disk-demo
        - name: cloudinitdisk
          cloudInitNoCloud:
            userDataBase64: SGkuXG4=
```

1. 📱 IP

2. 📱 Mac

⚠ Warning

KubeVirt VM API	spec.template.spec.domain.devices.interfaces.macAddress	Mac	Mac	Kube-OVN
annotation	Mac	Mac		

4.1.3 VM IP

Kube-OVN	VM IP	IP	VM
----------	-------	----	----

VM IP

1. VM Annotation IP
2. `virtctl restart <vm name>` VM IP

4.1.4

VM IP VM ovn.kubernetes.io/logical_switch

```
apiVersion: kubevirt.io/v1
kind: VirtualMachine
metadata:
  name: testvm
spec:
  runStrategy: Always
  template:
    metadata:
      labels:
        kubevirt.io/size: small
        kubevirt.io/domain: testvm
    annotations:
      ovn.kubernetes.io/logical_switch: subnet1 #(1)
      kubevirt.io/allow-pod-bridge-network-live-migration: "true"
  spec:
    domain:
      devices:
        disks:
          - name: containerdisk
            disk:
              bus: virtio
          - name: cloudinitdisk
            disk:
              bus: virtio
        interfaces:
          - name: default
            bridge: {}
    resources:
      requests:
        memory: 64M
  networks:
    - name: default
      pod: {}
  volumes:
    - name: containerdisk
      containerDisk:
        image: quay.io/kubevirt/cirros-container-disk-demo
    - name: cloudinitdisk
      cloudInitNoCloud:
        userDataBase64: SGkuXG4=
```

1. 🤖 VM

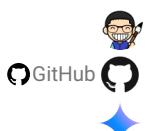
[PDF](#)

[Slack](#)

[Support](#)

⌚ 2025 11 27

⌚ 2025 3 3



4.1.5

4.2

KubeVirt	Clone	Label	Annotation	Kube-OVN	Annotation	IP	MAC
----------	-------	-------	------------	----------	------------	----	-----

4.2.1 Annotation

VirtualMachineClone	Kube-OVN	Annotation
---------------------	----------	------------

```

kind: VirtualMachineClone
apiVersion: "clone.kubevirt.io/v1beta1"
metadata:
  name: testclone
spec:
  source:
    apiGroup: kubevirt.io
    kind: VirtualMachine
    name: vm-source
  target:
    apiGroup: kubevirt.io
    kind: VirtualMachine
    name: vm-target
  template:
    annotationFilters:
      - "ovn.kubernetes.io/*"
  
```

4.2.2



Note

patches KubeVirt 1.6

IP

```

kind: VirtualMachineClone
apiVersion: "clone.kubevirt.io/v1beta1"
metadata:
  name: testclone
spec:
  source:
    apiGroup: kubevirt.io
    kind: VirtualMachine
    name: vm-source
  target:
    apiGroup: kubevirt.io
    kind: VirtualMachine
    name: vm-target
  patches:
    - {"op": "replace", "path": "/spec/template/metadata/annotations/ovn.kubernetes.io~1ip_address", "value": "10.16.0.15"}
  
```

```

kind: VirtualMachineClone
apiVersion: "clone.kubevirt.io/v1beta1"
metadata:
  name: testclone
spec:
  source:
    apiGroup: kubevirt.io
    kind: VirtualMachine
    name: vm-source
  target:
    apiGroup: kubevirt.io
    kind: VirtualMachine
    name: vm-target
  patches:
    - {"op": "remove", "path": "/spec/template/metadata/annotations/ovn.kubernetes.io~1ip_address"}
  
```

Annotation

KubeVirt Clone API



PDF



Slack



Support

⌚2025 9 10

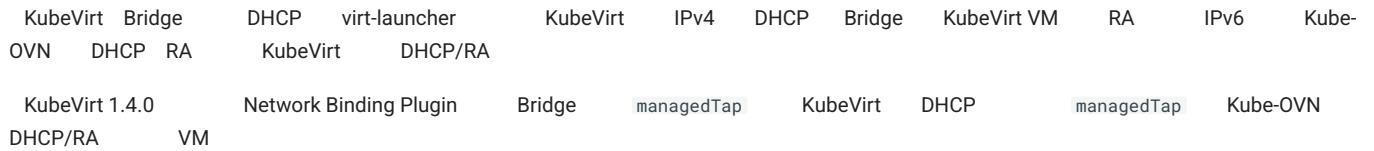
⌚2025 8 13



⌚GitHub⌚

4.2.3

4.3



4.3.1 DHCP

Kube-OVN Subnet DHCP IPv6 RA YAML

```

apiVersion: kubeovn.io/v1
kind: Subnet
metadata:
  name: dual-stack-subnet
spec:
  cidrBlock: "10.244.0.0/16,fd00:10:244::/64"
  enableDHCPS: true
  enableIPv6RA: true
  
```

4.3.2 managedTap

KubeVirt managedTap Network Binding Plugin:

```

# kubectl patch kubevirt -n kubevirt kubevirt --type=json -p=
'[{"op": "add", "path": "/spec/configuration/network", "value": {
  "binding": {
    "managedtap": {
      "domainAttachmentType": "managedTap"
    }
  }
}]'
  
```

4.3.3 managedTap

```

apiVersion: kubevirt.io/v1
kind: VirtualMachine
metadata:
  name: dual-stack-vm
  namespace: default
spec:
  running: false
  template:
    spec:
      domain:
        devices:
          interfaces:
            - name: default
              binding:
                name: managedtap
      networks:
        - name: default
      pod: {}
  
```

VM DHCP IPv6 RA IPv4/IPv6

[PDF](#)

[Slack](#)

[Support](#)

⌚2025 9 10

⌚2025 3 3



4.3.4

4.4

KubeVirt

- KubeVirt Bridge
- KubeVirt
- IP
-

Kube-OVN 0.5 TCP

4.4.1

VM Spec `kubevirt.io/allow-pod-bridge-network-live-migration: "true"` annotation Kube-OVN

1. VM

```
kubectl apply -f - <<EOF
apiVersion: kubevirt.io/v1
kind: VirtualMachine
metadata:
  name: testvm
spec:
  runStrategy: Always
  template:
    metadata:
      labels:
        kubevirt.io/size: small
        kubevirt.io/domain: testvm
    annotations:
      kubevirt.io/allow-pod-bridge-network-live-migration: "true"
  spec:
    domain:
      devices:
        disks:
          - name: containerdisk
            disk:
              bus: virtio
          - name: cloudinitdisk
            disk:
              bus: virtio
        interfaces:
          - name: default
            bridge: {}
    resources:
      requests:
        memory: 64M
    networks:
      - name: default
        pod: {}
    volumes:
      - name: containerdisk
        containerDisk:
          image: quay.io/kubevirt/cirros-container-disk-demo
      - name: cloudinitdisk
        cloudInitNoCloud:
          userDataBase64: SGkuXG4=
EOF
```

1. SSH

```
# password: gocubsgo
virtctl ssh cirros@testvm
ping 8.8.8.8
```

1.

```
virtctl migrate testvm
```

VM	SSH	ping
----	-----	------

4.4.2

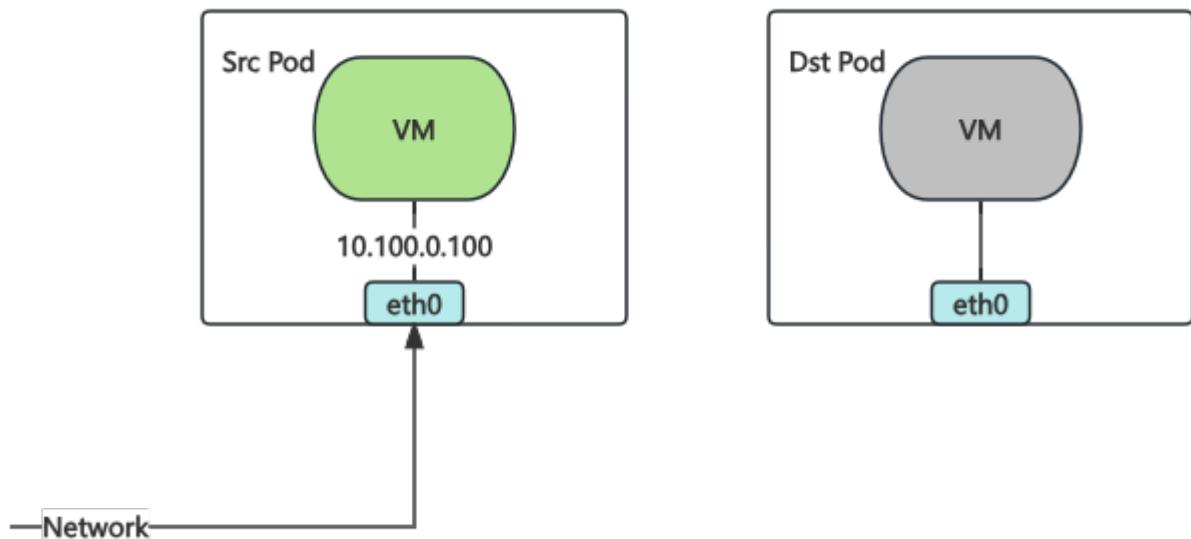
Kube-OVN

Live migration - Reducing downtime with multichassis port bindings

1. KubeVirt

IP

Pod

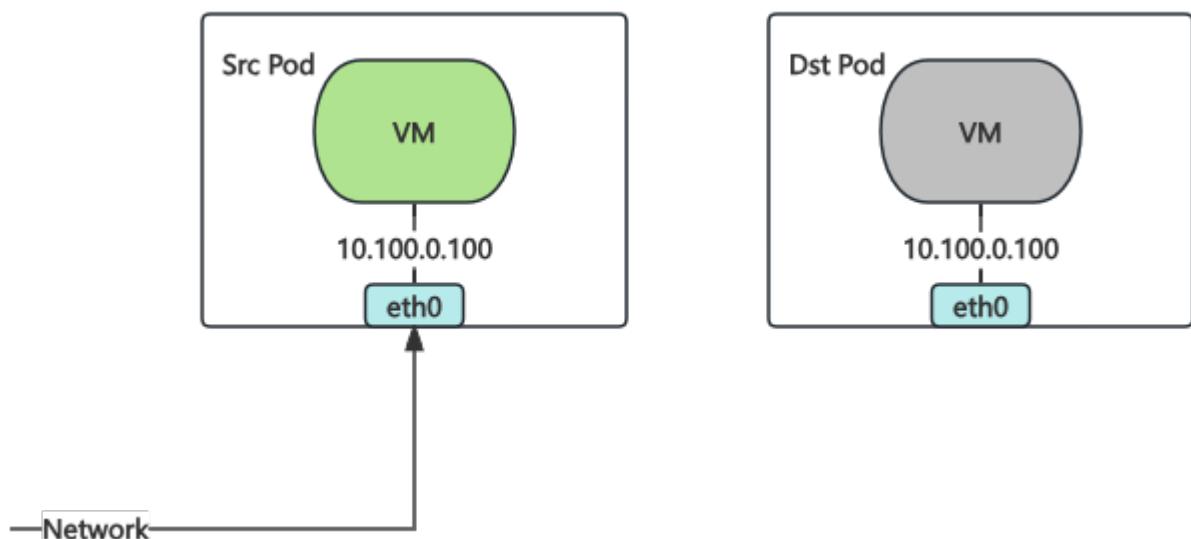


1. Kube-OVN

Pod

Pod

Pod



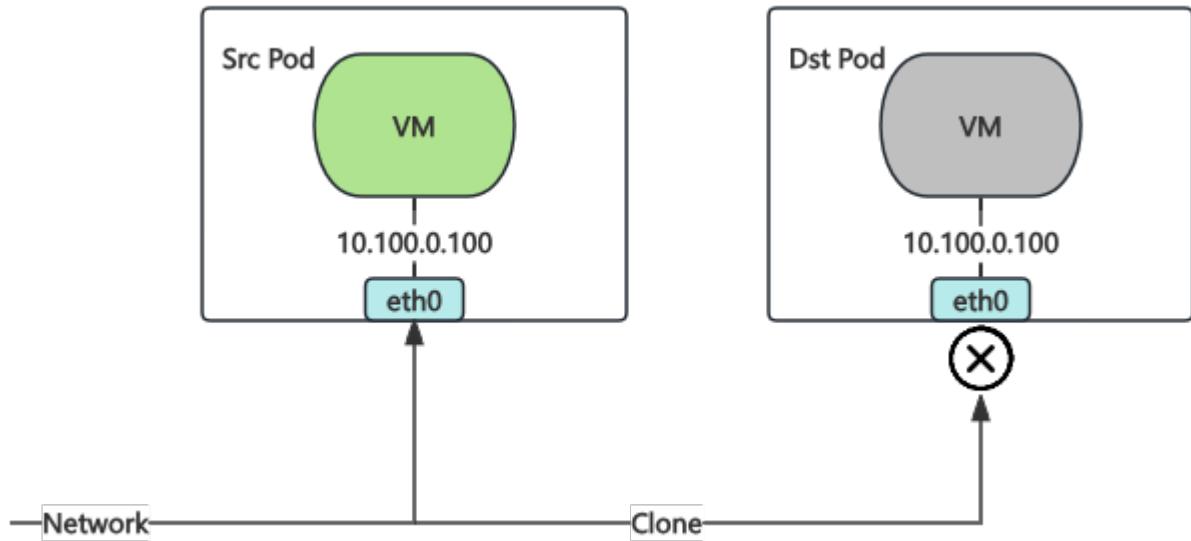
1. Kube-OVN

Pod

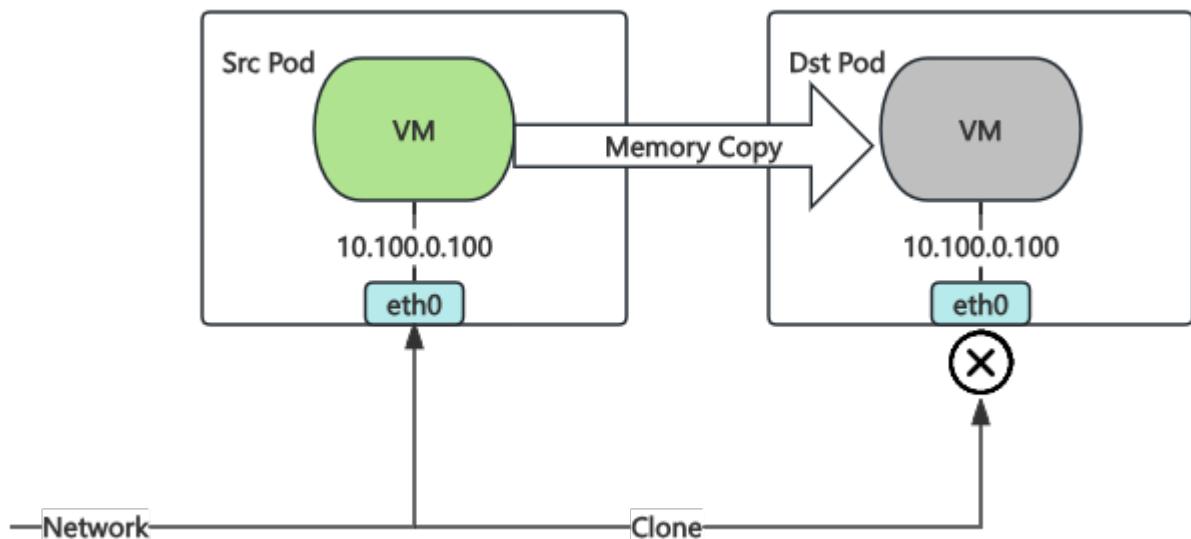
Pod

Pod

Pod

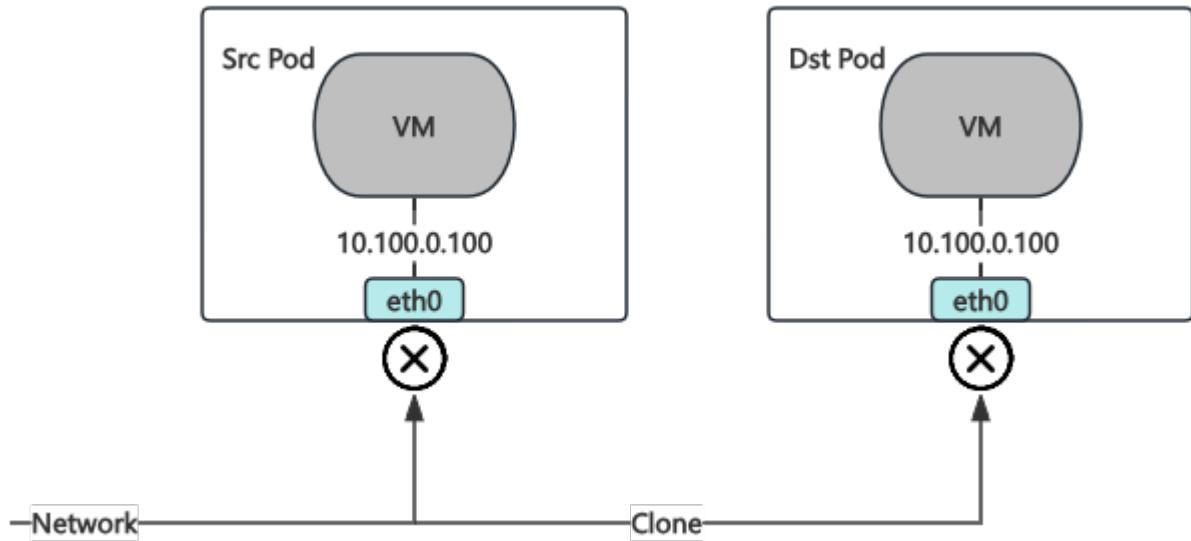


1. KubeVirt VM

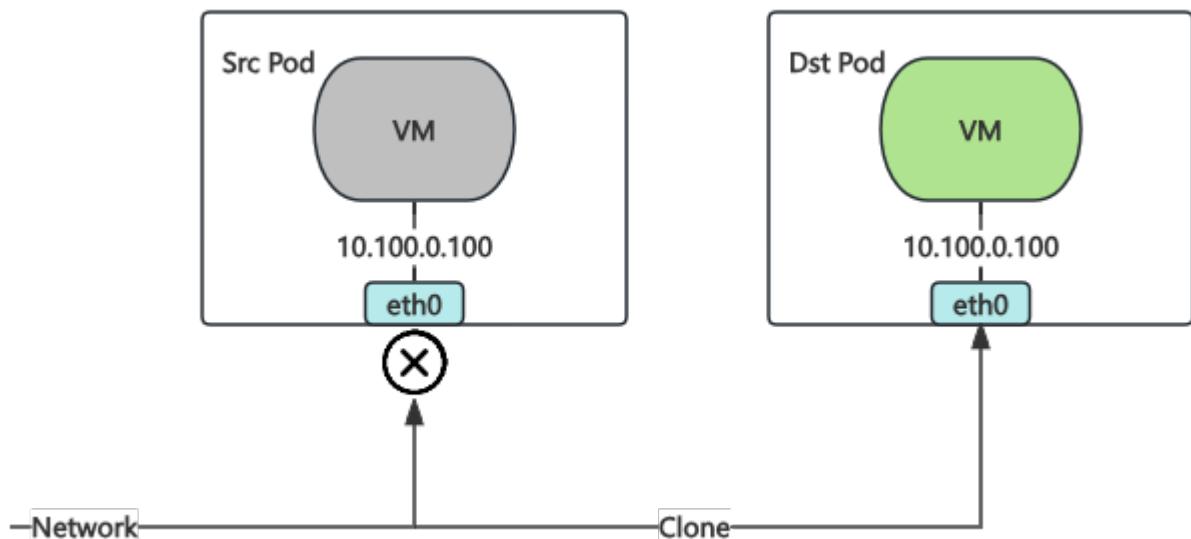


1. KubeVirt

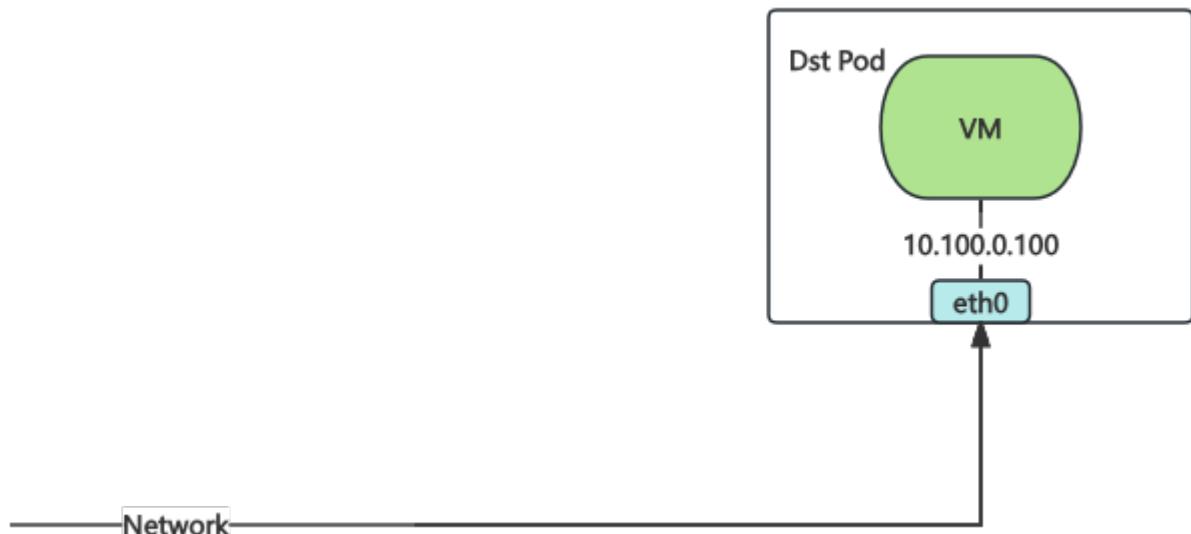
Pod Pod



1. KubeVirt Pod libvirt RARP Pod Pod



1. KubeVirt Pod Kube-OVN Watch Migration CR



5 6

libvirt RARP

0.5 TCP



⌚2025 9 10

⌚2025 3 3



4.4.3

4.5

Kube-OVN Multus Dynamic Networks Controller KubeVirt v1.4.0 VM

4.5.1

Thick Multus

```
kubectl apply -f https://raw.githubusercontent.com/k8snetworkplumbingwg/multus-cni/refs/heads/master/deployments/multus-daemonset-thick.yml
```

Multus Dynamic Networks Controller

```
kubectl apply -f https://raw.githubusercontent.com/k8snetworkplumbingwg/multus-dynamic-networks-controller/refs/heads/main/manifests/dynamic-networks-controller.yaml
```

4.5.2

NetworkAttachmentDefinition

provider ovn

```
apiVersion: "k8s.cni.cncf.io/v1"
kind: NetworkAttachmentDefinition
metadata:
  name: attachnet
  namespace: default
spec:
  config: '{
    "cniVersion": "0.3.0",
    "type": "kube-ovn",
    "server_socket": "/run/openvswitch/kube-ovn-daemon.sock",
    "provider": "attachnet.default.ovn"
  }'
```

- spec.config.type: kube-ovn CNI Kube-OVN
- server_socket: Kube-OVN socket /run/openvswitch/kube-ovn-daemon.sock
- provider: NetworkAttachmentDefinition <name>.<namespace>.ovn , Kube-OVN Subnet ovn

Kube-OVN Subnet

Kube-OVN provider NetworkAttachmentDefinition <name>.<namespace>.ovn ovn Kube-OVN Subnet

```
apiVersion: kubeovn.io/v1
kind: Subnet
metadata:
  name: attachnet
spec:
  protocol: IPv4
  provider: attachnet.default.ovn
  cidrBlock: 172.17.0.0/16
  gateway: 172.17.0.1
  excludeIps:
    - 172.17.0.0..172.17.0.10
```

4.5.3 VM

yaml VM

```
apiVersion: kubevirt.io/v1
kind: VirtualMachine
metadata:
  name: vm-fedora
spec:
  runStrategy: Always
  template:
    spec:
      domain:
```

```

devices:
  disks:
    - disk:
        bus: virtio
        name: containerdisk
  interfaces:
    - masquerade: {}
      name: defaultnetwork
  resources:
    requests:
      memory: 1024M
  networks:
    - name: defaultnetwork
      pod: {}
  terminationGracePeriodSeconds: 0
  volumes:
    - containerDisk:
        image: quay.io/kubevirt/fedora-with-test-tooling-container-disk:devel
        name: containerdisk

```

VM Spec

```

apiVersion: kubevirt.io/v1
kind: VirtualMachine
metadata:
  name: vm-fedora
template:
  spec:
    domain:
      devices:
        interfaces:
          - name: defaultnetwork
            masquerade: {}
            # new interface
          - name: dyniface1
            bridge: {}
    networks:
      - name: defaultnetwork
        pod: {}
        # new network
      - name: dyniface1
        multus:
          networkName: attachnet

```

interface state absent

```

apiVersion: kubevirt.io/v1
kind: VirtualMachine
metadata:
  name: vm-fedora
template:
  spec:
    domain:
      devices:
        interfaces:
          - name: defaultnetwork
            masquerade: {}
            # set the interface state to absent
          - name: dyniface1
            state: absent
            bridge: {}
    networks:
      - name: defaultnetwork
        pod: {}
      - name: dyniface1
        multus:
          networkName: attachnet

```



PDF



Slack



Support

⌚2025 11 18

⌚2025 11 18



4.5.4

4.6 DHCP

managedTap	SR-IOV	DPDK	KubeVirt	DHCP	Kube-OVN	OVN	DHCP	DHCP	KubeVirt
DHCP	IP		Kube-OVN	DHCP	DHCPv6, IPv6RA, DNS	TFTP	DHCP	DHCP	

⚠ Warning

1. bridge KubeVirt DHCP Kube-OVN DHCP Kube-OVN DHCP Kube-OVN DHCP managedTap
bridge managedTap managedTap
2. DHCP Pod DHCP

DHCP

```
apiVersion: kubeovn.io/v1
kind: Subnet
metadata:
  name: sn-dual
spec:
  cidrBlock: "10.0.0.0/24,240e::a00/120"
  default: false
  disableGatewayCheck: true
  disableInterConnection: false
  excludeIps:
    - 10.0.0.1
    - 240e::a01
  gateway: 10.0.0.1,240e::a01
  gatewayNode: ''
  gatewayType: distributed
  natOutgoing: false
  private: false
  protocol: Dual
  provider: ovn
  vpc: vpc-test
  enableDHCP: true
  dhcpV4Options: "lease_time=3600,router=10.0.0.1,server_id=169.254.0.254,server_mac=00:00:00:2E:2F:B8"
  dhcpV6Options: "server_id=00:00:00:2E:2F:C5"
  enableIPv6RA: true
  ipv6RAConfigs: "address_mode=dhcpv6_stateful,max_interval=30,min_interval=5,send_periodic=true"
```

- enableDHCP : DHCP
- dhcpV4Options , dhcpV6Options : ovn-nb DHCP Options "lease_time=3600, router=\$ipv4_gateway, server_id=169.254.0.254, server_mac=\$random_mac" server_id=\$random_mac
- enableIPv6RA : DHCPv6
- ipv6RAConfigs : ovn-nb Logical_Router_Port Logical Router Port address_mode=dhcpv6_stateful, max_interval=30, min_interval=5, send_periodic=true

[PDF](#)
[Slack](#)
[Support](#)
 2025 9 10

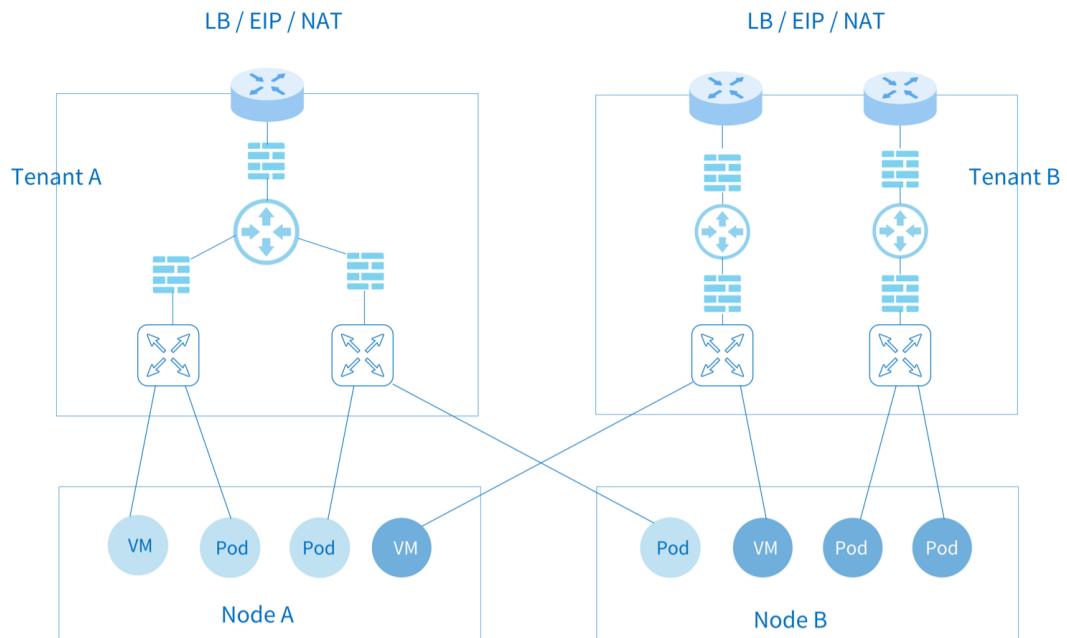
 2022 5 24


4.6.1

5. VPC

5.1 VPC

Kube-OVN	VPC	VPC	Subnet	EIP				
VPC	Kube-OVN	VPC	Kubernetes	Pod	NodePort	DNS	Kubernetes	Kubernetes
VPC	VPC	VPC	Subnet	VPC	VPC	NAT	VPC	ACL



5.1.1

Kube-OVN	VPC	OVN	IP	VPC	IP	IP	VPC	OVN	Datapath ID
Datapath ID									

5.1.2 VPC

VPC

```

kind: Vpc
apiVersion: kubeovn.io/v1
metadata:
  name: test-vpc-1
spec:
  namespaces:
  - ns1
---
kind: Vpc
apiVersion: kubeovn.io/v1
metadata:
  name: test-vpc-2
spec:
  namespaces:
  - ns2
  
```

- namespaces

Namespace

VPC

VPC CIDR:

```

kind: Subnet
apiVersion: kubeovn.io/v1
metadata:
  name: net1
spec:
  vpc: test-vpc-1
  cidrBlock: 10.0.1.0/24
  protocol: IPv4
  namespaces:
    - ns1
---
kind: Subnet
apiVersion: kubeovn.io/v1
metadata:
  name: net2
spec:
  vpc: test-vpc-2
  cidrBlock: 10.0.1.0/24
  protocol: IPv4
  namespaces:
    - ns2
  
```

Namespace Pod:

```

apiVersion: v1
kind: Pod
metadata:
  namespace: ns1
  name: vpc1-pod
spec:
  containers:
    - name: vpc1-pod
      image: docker.io/library/nginx:alpine
---
apiVersion: v1
kind: Pod
metadata:
  namespace: ns2
  name: vpc2-pod
spec:
  containers:
    - name: vpc2-pod
      image: docker.io/library/nginx:alpine
  
```

Pod CIDR VPC Pod

5.1.3 VPC

VPC	VPC	VPC	IP SNAT	DNAT
VPC	Multus-CNI		multus-cni	



Note

VPC	VPC								VPC
VPC	VPC NAT	OVN	Egress Gateway	VPC NAT	Kube-OVN	VPC NAT	Pod		VPC
Macvlan	Pod	iptables							
OVN	OVN	NAT		OVN	BFD	OVN	OVN		
Egress Gateway		VPC NAT							

```

apiVersion: kubeovn.io/v1
kind: Subnet
metadata:
  name: ovn-vpc-external-network
spec:
  protocol: IPv4
  provider: ovn-vpc-external-network.kube-system
  cidrBlock: 192.168.0.0/24
  gateway: 192.168.0.1 # IP address of the physical gateway
  excludeIps:
    
```

```

- 192.168.0.1..192.168.0.10
---
apiVersion: "k8s.cni.cncf.io/v1"
kind: NetworkAttachmentDefinition
metadata:
  name: ovn-vpc-external-network
  namespace: kube-system
spec:
  config: '{
    "cniVersion": "0.3.0",
    "type": "macvlan",
    "master": "eth1",
    "mode": "bridge",
    "ipam": {
      "type": "kube-ovn",
      "server_socket": "/run/openvswitch/kube-ovn-daemon.sock",
      "provider": "ovn-vpc-external-network.kube-system"
    }
}'

```

• Subnet	Macvlan	VPC	IP
• VPC	Macvlan	NetworkAttachmentDefinition	master
• name			
Macvlan	L2/L3	Vlan	
1. OpenStack VM PortSecurity			
2. VMware vSwitch	MAC Address Changes, Forged Transmits	Promiscuous Mode Operation	allow
3. Hyper-V	MAC Address Spoofing		
4. AWS GCE	Mac	Macvlan	
5. Macvlan	Macvlan	VpcNATGateway Pod	Pod
6.	Trunk	Macvlan	

VPC

VPC kube-system ovn-vpc-nat-gw-config nodeSelector

```

---
kind: ConfigMap
apiVersion: v1
metadata:
  name: ovn-vpc-nat-config
  namespace: kube-system
data:
  image: 'docker.io/kubeovn/vpc-nat-gateway:v1.14.23'
  nodeSelector: |
    kubernetes.io/hostname: kube-ovn-control-plane
---
kind: ConfigMap
apiVersion: v1
metadata:
  name: ovn-vpc-nat-gw-config
  namespace: kube-system
data:
  enable-vpc-nat-gw: 'true'

```

- image: Pod
- enable-vpc-nat-gw: VPC

VPC

```

kind: VpcNatGateway
apiVersion: kubeovn.io/v1
metadata:
  name: gw1
spec:
  vpc: test-vpc-1
  subnet: net1
  lanIp: 10.0.1.254
  selector:
    - "kubernetes.io/hostname: kube-ovn-worker"
    - "kubernetes.io/os: linux"

```

```

externalSubnets:
  - ovn-vpc-external-network

• vpc  VpcNatGateway  VPC
• subnet  VPC  Subnet  VPC  Pod      lanIp
• lanIp  subnet       IP  VPC  Pod      IP  VPC      VpcNatGateway  nextHopIP      lanIp
• selector  VpcNatGateway Pod      Kubernetes  NodeSelector
• externalSubnets  VPC      ovn-vpc-external-network

• tolerations  VPC
• affinity  VPC  Pod

VPC-NAT-GW

```

1. nat gw pod net1 arp ping eip arp ping

EIP

EIP IP VPC DNAT SNAT IP

EIP

```

kind: IptablesEIP
apiVersion: kubeovn.io/v1
metadata:
  name: eip-random
spec:
  natGwDp: gw1

```

EIP

```

kind: IptablesEIP
apiVersion: kubeovn.io/v1
metadata:
  name: eip-static
spec:
  natGwDp: gw1
  v4ip: 192.168.0.100

```

EIP

```

kind: IptablesEIP
apiVersion: kubeovn.io/v1
metadata:
  name: eip-random
spec:
  natGwDp: gw1
  externalSubnet: ovn-vpc-external-network

```

• externalSubnet EIP ovn-vpc-external-network VPC externalSubnets

DNAT

DNAT EIP VPC IP

```

kind: IptablesEIP
apiVersion: kubeovn.io/v1
metadata:
  name: eipd01
spec:
  natGwDp: gw1

---
kind: IptablesDnatRule
apiVersion: kubeovn.io/v1
metadata:
  name: dnat01
spec:
  eip: eipd01
  externalPort: '8888'

```

```
internalIp: 10.0.1.10  
internalPort: '80'  
protocol: tcp
```

SNAT

```
kind: IptablesEIP
apiVersion: kubeovn.io/v1
metadata:
  name: eips01
spec:
  natGwDp: gw1
---
kind: IptablesSnatRule
apiVersion: kubeovn.io/v1
metadata:
  name: snat01
spec:
  eip: eips01
  internalCIDR: 10.0.1.0/24
```

IP

IP VPC IP EIP EIP VPC IP VPC IP SNAT EIF

```
---
```

```
kind: IptablesEIP
apiVersion: kubeovn.io/v1
metadata:
  name: eipf01
spec:
  natGwDp: gw1
```



```
---
```

```
kind: IptablesFIPRule
apiVersion: kubeovn.io/v1
metadata:
  name: fipf01
spec:
  eip: eipf01
  internalIp: 10.0.1.5
```

5.1.4

```
kind: Vpc
apiVersion: kubeovn.io/v1
metadata:
  name: test-vpc-1
spec:
  staticRoutes:
    - cidr: 0.0.0.0/0
      nextHopIP: 10.0.1.254
      policy: policyDst
    - cidr: 172.31.0.0/24
      nextHopIP: 10.0.1.253
      policy: policySrc
      routeTable: "rtb1"
```

- policy: policyDst policySrc
 - CIDR
 - routeTable:

OVN

Logical Router Policy

```

kind: Vpc
apiVersion: kubeovn.io/v1
metadata:
  name: test-vpc-1
spec:
  policyRoutes:
    - action: drop
      match: ip4.src==10.0.1.0/24 && ip4.dst==10.0.1.250
      priority: 11
    - action: reroute
      match: ip4.src==10.0.1.0/24
      nextHopIP: 10.0.1.252
      priority: 10

```

5.1.5

Kubernetes	Service	Kubernetes	Service	IP	VPC	VPC
Kubernetes	Service					

Kube-OVN SwitchLBRule

SwitchLBRule

```

apiVersion: kubeovn.io/v1
kind: SwitchLBRule
metadata:
  name: cjh-slr-nginx
spec:
  vip: 1.1.1.1
  sessionAffinity: ClientIP
  namespace: default
  selector:
    - app: nginx
  ports:
    - name: dns
      port: 8888
      targetPort: 80
      protocol: TCP

```

- vip
- namespace Pod Namespace
- sessionAffinity Service sessionAffinity
- selector Service selector
- ports Service port

```

# kubectl get slr
NAME      VIP        PORT(S)           SERVICE          AGE
vpc-dns-test-cjh2  10.96.0.3  53/UDP,53/TCP   kube-system/slr-vpc-dns-test-cjh2  88m

```

5.1.6 vpc-dns

VPC	VPC	VPC	Pod	coredns	VPC	CoreDNS	Service	Kube-OVN	vpc-dns
-----	-----	-----	-----	---------	-----	---------	---------	----------	---------

```

apiVersion: "k8s.cni.cncf.io/v1"
kind: NetworkAttachmentDefinition
metadata:
  name: ovn-nad
  namespace: default
spec:
  config: '{
    "cniVersion": "0.3.0",
    "type": "kube-ovn",
    "server_socket": "/run/openvswitch/kube-ovn-daemon.sock",
    "provider": "ovn-nad.default.ovn"
  }'

```

ovn-default provider

ovn-default provider nad provider ovn-nad.default.ovn

```
apiVersion: kubeovn.io/v1
kind: Subnet
metadata:
  name: ovn-default
spec:
  cidrBlock: 10.16.0.0/16
  default: true
  disableGatewayCheck: false
  disableInterConnection: false
  enableDHCP: false
  enableIPv6RA: false
  excludeIps:
  - 10.16.0.1
  gateway: 10.16.0.1
  gatewayType: distributed
  logicalGateway: false
  natOutgoing: true
  private: false
  protocol: IPv4
  provider: ovn-nad.default.ovn
  vpc: ovn-cluster
```

vpc-dns ConfigMap

kube-system configmap vpc-dns vpc-dns

```
apiVersion: v1
kind: ConfigMap
metadata:
  name: vpc-dns-config
  namespace: kube-system
data:
  coredns-vip: 10.96.0.3
  enable-vpc-dns: "true"
  nad-name: ovn-nad
  nad-provider: ovn-nad.default.ovn
```

- enable-vpc-dns true false true
- coredns-image dns coredns
- coredns-template dns URL yamls/coredns-template.yaml
- coredns-vip coredns lb vip
- nad-name network-attachment-definitions
- nad-provider provider
- k8s-service-host coredns k8s apiserver ip
- k8s-service-port coredns k8s apiserver port

vpc-dns

```
apiVersion: rbac.authorization.k8s.io/v1
kind: ClusterRole
metadata:
  labels:
    kubernetes.io/bootstrapping: rbac-defaults
  name: system:vpc-dns
rules:
- apiGroups:
  - ""
  resources:
  - endpoints
  - services
  - pods
  - namespaces
  verbs:
  - list
  - watch
- apiGroups:
  - discovery.k8s.io
  resources:
  - endpointslices
  verbs:
  - list
  - watch
```

```

---
apiVersion: rbac.authorization.k8s.io/v1
kind: ClusterRoleBinding
metadata:
  annotations:
    rbac.authorization.kubernetes.io/autoupdate: "true"
  labels:
    kubernetes.io/bootstrapping: rbac-defaults
  name: vpc-dns
roleRef:
  apiGroup: rbac.authorization.k8s.io
  kind: ClusterRole
  name: system:vpc-dns
subjects:
- kind: ServiceAccount
  name: vpc-dns
  namespace: kube-system
---
apiVersion: v1
kind: ServiceAccount
metadata:
  name: vpc-dns
  namespace: kube-system
---
apiVersion: v1
kind: ConfigMap
metadata:
  name: vpc-dns-corefile
  namespace: kube-system
data:
  Corefile: |
    .:53 {
      errors
      health {
        lameduck 5s
      }
      ready
      kubernetes cluster.local in-addr.arpa ip6.arpa {
        pods insecure
        fallthrough in-addr.arpa ip6.arpa
      }
      prometheus :9153
      forward . /etc/resolv.conf {
        prefer_udp
      }
      cache 30
      loop
      reload
      loadbalance
    }
}

```

vpc-dns

```

kind: VpcDns
apiVersion: kubeovn.io/v1
metadata:
  name: test-cjh1
spec:
  vpc: cjh-vpc-1
  subnet: cjh-subnet-1

```

- vpc dns vpc
- subnet dns

```

[root@hci-dev-mst-1 kubeovn]# kubectl get vpc-dns
NAME      ACTIVE   VPC      SUBNET
test-cjh1  false    cjh-vpc-1  cjh-subnet-1
test-cjh2  true     cjh-vpc-1  cjh-subnet-2

```

- ACTIVE: true dns false
- VPC DNS
- VPC vpc-dns VPC subnet vpc-dns true false
- true vpc-dns false vpc-dns

5.1.7

VPC

VPC

```
kind: Vpc
apiVersion: kubeovn.io/v1
metadata:
  name: test-vpc-1
spec:
  namespaces:
    - ns1
  defaultSubnet: test
```

- defaultSubnet VPC

Namespace ovn.kubernetes.io/logical_switch ovn.kubernetes.io/logical_switch Pod

VPC Pod livenessProbe readinessProbe

VPC Pod kubelet VPC Pod Kube-OVN TProxy kubelet VPC Pod

DaemonSet kube-ovn-cni --enable-tproxy=true

```
spec:
  template:
    spec:
      containers:
        - args:
            - --enable-tproxy=true
```

1. VPC Pod IP

2. tcpSocket httpGet

 PDF

 Slack

 Support

 2025 9 30

 2022 5 24



5.1.8

5.2 VPC Egress Gateway

Note									
VPC	VPC								
VPC	VPC NAT	OVN	Egress Gateway	VPC NAT	Kube-OVN	VPC NAT	Pod	VPC	
Macvlan	Pod	iptables							
OVN	OVN	NAT		OVN	BFD	OVN	OVN		
Egress Gateway	VPC NAT								

VPC Egress Gateway VPC VPC Pod

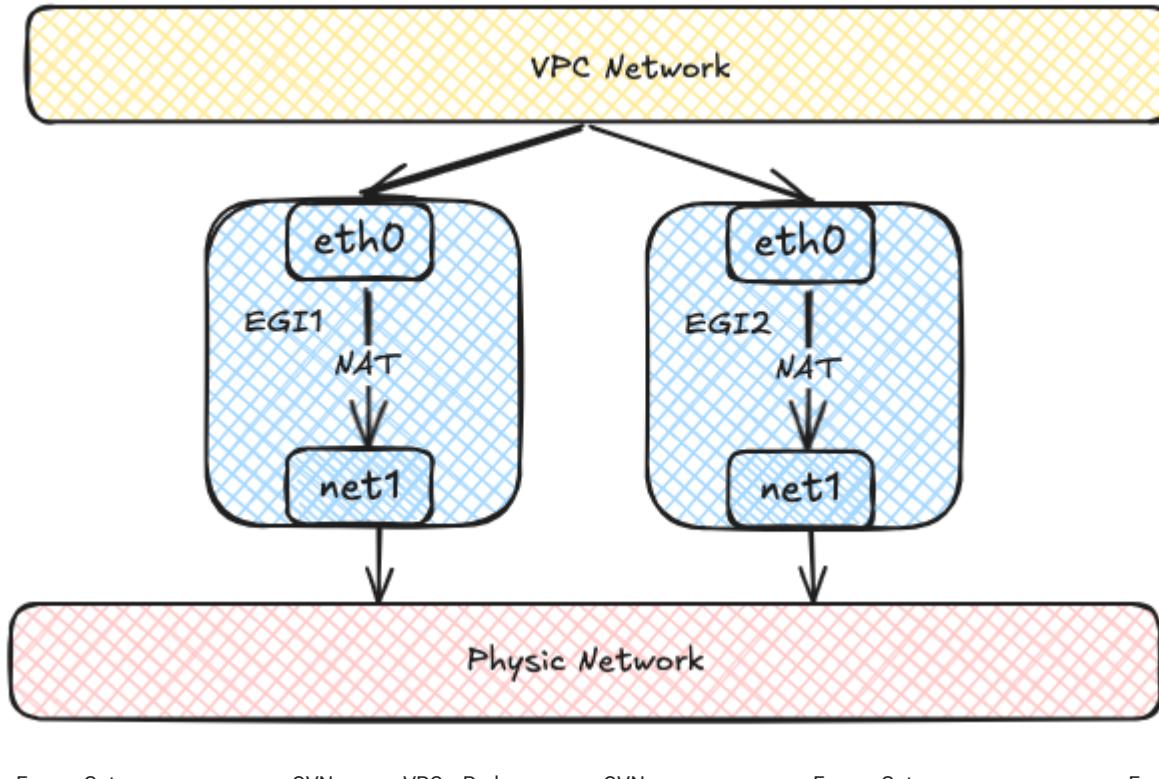
- ECMP Active-Active
- BFD <1s
- IPv6
- Namespace Pod
- Node Egress Gateway

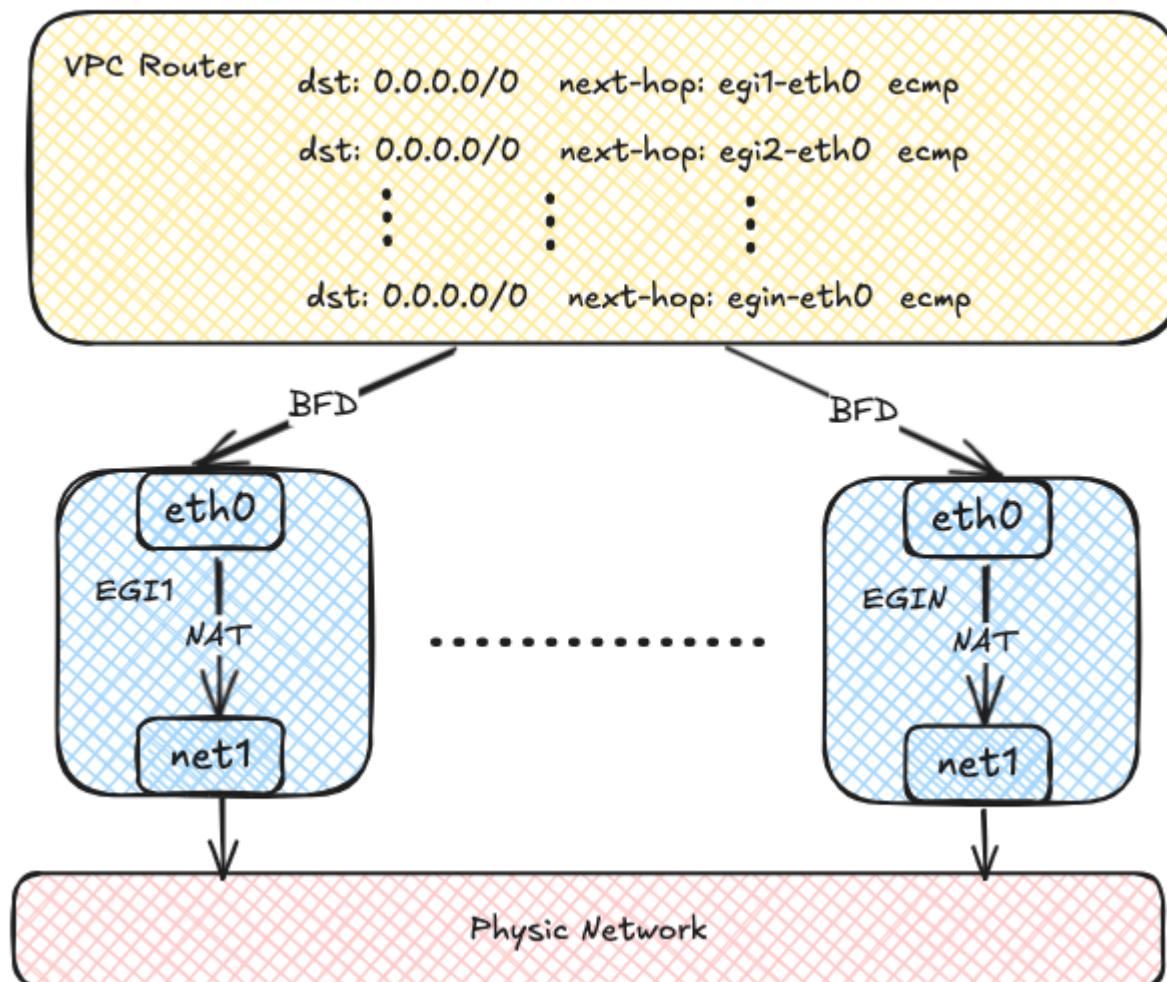
VPC Egress Gateway

- Macvlan [Underlay](#)
- Gateway Egress IP
- SNAT EIP DNAT
-

5.2.1

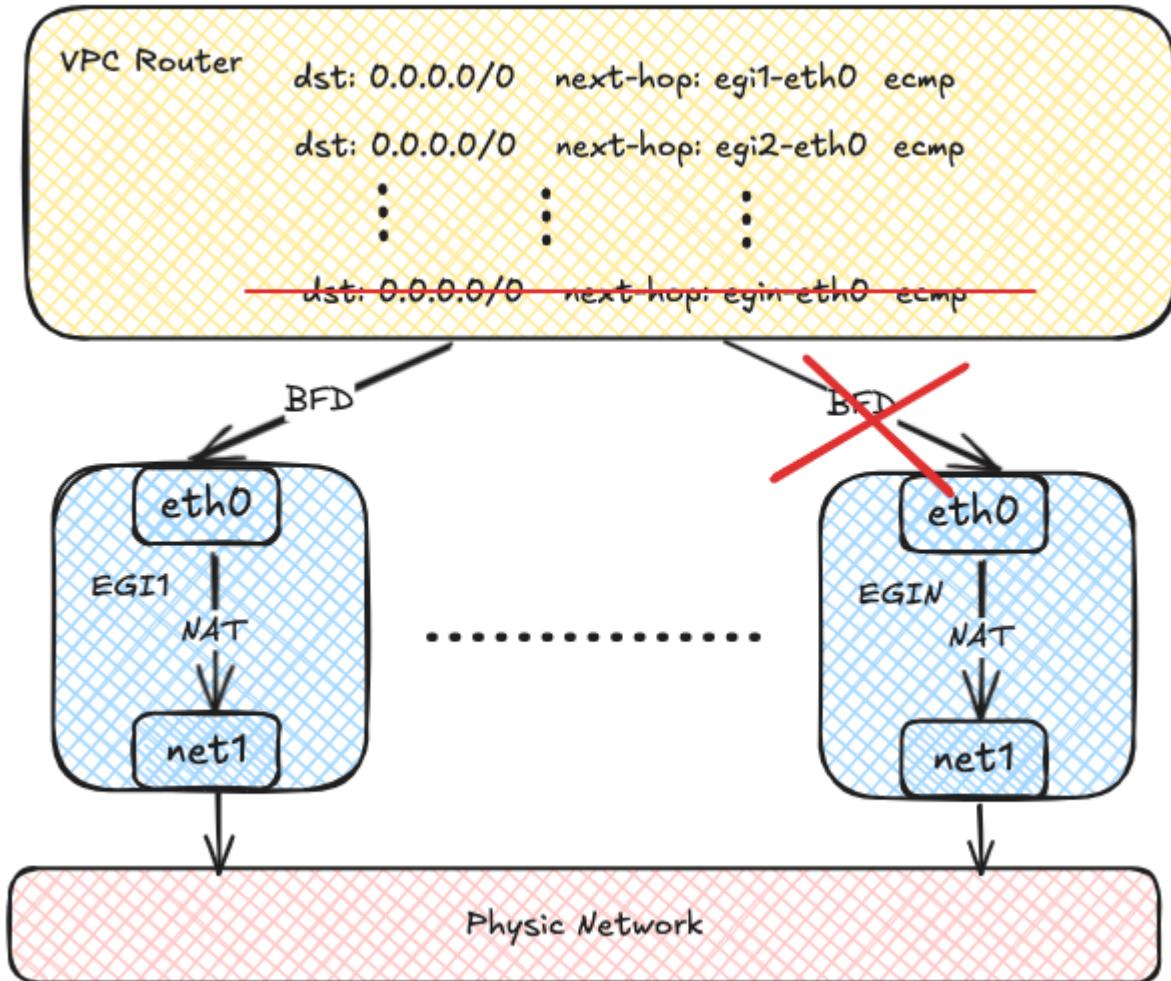
Egress Gateway	Pod	Pod	VPC	Macvlan	Egress Gateway	NAT
----------------	-----	-----	-----	---------	----------------	-----





OVN BFD Egress Gateway

Egress Gateway OVN



5.2.2

VPC Egress Gateway	VPC NAT Gateway	Multus-CNI
VPC Egress Gateway	ConfigMap	

5.2.3

NetworkAttachmentDefinition

VPC Egress Gateway	VPC	NetworkAttachmentDefinition	macvlan	Kube-OVN	IPAM
--------------------	-----	-----------------------------	---------	----------	------

```

apiVersion: k8s.cni.cncf.io/v1
kind: NetworkAttachmentDefinition
metadata:
  name: eth1
  namespace: default
spec:
  config: '{
    "cniVersion": "0.3.0",
    "type": "macvlan",
    "master": "eth1",
    "mode": "bridge",
    "ipam": {
      "type": "kube-ovn",
      "server_socket": "/run/openvswitch/kube-ovn-daemon.sock",
      "provider": "eth1.default"
    }
  }'
---
apiVersion: kubeovn.io/v1
kind: Subnet
metadata:
  name: macvlan1

```

```
spec:
  protocol: IPv4
  provider: eth1.default
  cidrBlock: 172.17.0.0/16
  gateway: 172.17.0.1
  excludeIps:
    - 172.17.0.0..172.17.0.10
```

CNI NetworkAttachmentDefinition VPC Egress Gateway

VPC Egress Gateway

VPC Egress Gateway

```
apiVersion: kubeovn.io/v1
kind: VpcEgressGateway
metadata:
  name: gateway1
  namespace: default
spec:
  vpc: ovn-cluster
  replicas: 1
  externalSubnet: macvlan1
  policies:
    - snat: true
      subnets:
        - ovn-default
```

default	VPC	ovn-cluster	gateway1	VPC Egress Gateway	ovn-cluster	ovn-default	10.16.0.0/16	Pod
macvlan1								SNAT

VPC Egress Gateway

```
$ kubectl get veg gateway1
NAME      VPC      REPLICAS   BFD ENABLED   EXTERNAL SUBNET   PHASE   READY   AGE
gateway1  ovn-cluster  1          false        macvlan1       Completed  true    13s
```

```
kubectl get veg gateway1 -o wide
NAME      VPC      REPLICAS   BFD ENABLED   EXTERNAL SUBNET   PHASE   READY   INTERNAL IPS   EXTERNAL IPS   WORKING NODES   AGE
gateway1  ovn-cluster  1          false        macvlan1       Completed  true    ["10.16.0.12"]  ["172.17.0.11"]  ["kube-ovn-worker"]  82s
```

```
$ kubectl get deployment -l ovn.kubernetes.io/vpc-egress-gateway=gateway1
NAME      READY   UP-TO-DATE   AVAILABLE   AGE
gateway1  1/1     1           1           4m40s
```

```
$ kubectl get pod -l ovn.kubernetes.io/vpc-egress-gateway=gateway1 -o wide
NAME          READY   STATUS   RESTARTS   AGE   IP           NODE   NOMINATED NODE   READINESS GATES
gateway1-b9f8b4448-761hm  1/1     Running   0        4m48s  10.16.0.12  kube-ovn-worker  <none>  <none>
```

Pod IP iptables

```
$ kubectl exec gateway1-b9f8b4448-761hm -c gateway -- ip address show
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: net1@if13: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc noqueue state UP group default qlen 1000
  link/ether 62:d8:71:90:7b:86 brd ff:ff:ff:ff:ff:ff link-netnsid 0
  inet 172.17.0.11/16 brd 172.17.255.255 scope global net1
    valid_lft forever preferred_lft forever
  inet6 fe80::60d8:71ff:fe90:7b86/64 scope link
    valid_lft forever preferred_lft forever
17: eth0@if18: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1400 qdisc noqueue state UP group default
  link/ether 36:7c:6b:c7:82:6b brd ff:ff:ff:ff:ff:ff link-netnsid 0
  inet 10.16.0.12/16 brd 10.16.255.255 scope global eth0
    valid_lft forever preferred_lft forever
  inet6 fe80::347c:6bff:fec7:826b/64 scope link
    valid_lft forever preferred_lft forever
```

```
$ kubectl exec gateway1-b9f8b4448-761hm -c gateway -- ip route show
default via 172.17.0.1 dev net1
10.16.0.0/16 dev eth0 proto kernel scope link src 10.16.0.12
```

```
172.17.0.0/16 dev net1 proto kernel scope link src 172.17.0.11

$ kubectl exec gateway1-b9f8b4448-76lhm -c gateway -- iptables -t nat -S
-P PREROUTING ACCEPT
-P INPUT ACCEPT
-P OUTPUT ACCEPT
-P POSTROUTING ACCEPT
-A POSTROUTING -s 10.16.0.0/16 -j MASQUERADE --random-fu
```

Gateway Pod

```
$ kubectl exec -ti gateway1-b9f8b4448-76lhm -c gateway -- bash
nobody@gateway1-b9f8b4448-76lhm:/kube-ovn$ tcpdump -i any -nnve icmp and host 172.17.0.1
tcpdump: data link type LINUX_SLL2
tcpdump: listening on any, link-type LINUX_SLL2 (Linux cooked v2), snapshot length 262144 bytes
06:50:58.936528 eth0 In ifindex 17 92:26:b8:9e:f2:1c ethertype IPv4 (0x0800), length 104: (tos 0x0, ttl 63, id 30481, offset 0, flags [DF], proto ICMP (1), length 84)
    10.16.0.9 > 172.17.0.1: ICMP echo request, id 37989, seq 0, length 64
06:50:58.936574 net1 Out ifindex 2 62:d8:71:90:7b:86 ethertype IPv4 (0x0800), length 104: (tos 0x0, ttl 62, id 30481, offset 0, flags [DF], proto ICMP (1), length 84)
    172.17.0.11 > 172.17.0.1: ICMP echo request, id 39449, seq 0, length 64
06:50:58.936613 net1 In ifindex 2 02:42:39:79:7f:08 ethertype IPv4 (0x0800), length 104: (tos 0x0, ttl 64, id 26701, offset 0, flags [none], proto ICMP (1), length 84)
    172.17.0.1 > 172.17.0.11: ICMP echo reply, id 39449, seq 0, length 64
06:50:58.936621 eth0 Out ifindex 17 36:7c:6b:c7:82:6b ethertype IPv4 (0x0800), length 104: (tos 0x0, ttl 63, id 26701, offset 0, flags [none], proto ICMP (1), length 84)
    172.17.0.1 > 10.16.0.9: ICMP echo reply, id 37989, seq 0, length 64
```

OVN Logical Router

VPC

```
$ kubectl ko nbctl lr-policy-list ovn-cluster
Routing Policies
 31000          ip4.dst == 10.16.0.0/16      allow
 31000          ip4.dst == 100.64.0.0/16     allow
 30000          ip4.dst == 172.18.0.2       reroute   100.64.0.3
 30000          ip4.dst == 172.18.0.3       reroute   100.64.0.2
 30000          ip4.dst == 172.18.0.4       reroute   100.64.0.4
 29100          ip4.src == 10.16.0.0/16      reroute   10.16.0.12
 29000 ip4.src == $ovn.default.kube.ovn.control_plane_ip4  reroute   100.64.0.2
 29000          ip4.src == $ovn.default.kube.ovn.worker2_ip4  reroute   100.64.0.4
 29000          ip4.src == $ovn.default.kube.ovn.worker_ip4   reroute   100.64.0.3
```

.spec.replicas

```
$ kubectl scale veg gateway1 --replicas=2
vpceregressgateway.kubeovn.io/gateway1 scaled

$ kubectl get veg gateway1
NAME      VPC      REPLICAS      BFD ENABLED      EXTERNAL SUBNET      PHASE      READY      AGE
gateway1   ovn-cluster   2           false          macvlan          Completed   true      39m

$ kubectl get pod -l ovn.kubernetes.io/vpc-egress-gateway=gateway1 -o wide
NAME                  READY   STATUS   RESTARTS   AGE      IP          NODE      NOMINATED NODE   READINESS GATES
gateway1-b9f8b4448-76lhm   1/1    Running   0          40m   10.16.0.12   kube-ovn-worker   <none>        <none>
gateway1-b9f8b4448-zd4dl   1/1    Running   0          64s   10.16.0.13   kube-ovn-worker2  <none>        <none>

$ kubectl ko nbctl lr-policy-list ovn-cluster
Routing Policies
 31000          ip4.dst == 10.16.0.0/16      allow
 31000          ip4.dst == 100.64.0.0/16     allow
 30000          ip4.dst == 172.18.0.2       reroute   100.64.0.3
 30000          ip4.dst == 172.18.0.3       reroute   100.64.0.2
 30000          ip4.dst == 172.18.0.4       reroute   100.64.0.4
 29100          ip4.src == 10.16.0.0/16      reroute   10.16.0.12, 10.16.0.13
 29000 ip4.src == $ovn.default.kube.ovn.control_plane_ip4  reroute   100.64.0.2
 29000          ip4.src == $ovn.default.kube.ovn.worker2_ip4  reroute   100.64.0.4
 29000          ip4.src == $ovn.default.kube.ovn.worker_ip4   reroute   100.64.0.3
```

Egress Gateway IP

externalIPs	nodeSelector	Egress Gateway Pods	Egress IP

```
apiVersion: kubeovn.io/v1
kind: VpcEgressGateway
metadata:
  name: gateway1
  namespace: default
spec:
  vpc: ovn-cluster
  replicas: 2
  externalSubnet: macvlan1
  policies:
    - snat: true
      subnets:
        - ovn-default
```

```
externalIPs:
  - 172.17.0.10
  - 172.17.0.11
nodeSelector:
  - matchLabels:
      kubernetes.io/hostname: kube-ovn-worker
```

BFD

BFD VPC BFD LRP VPC BFD Port

```
apiVersion: kubeovn.io/v1
kind: Vpc
metadata:
  name: vpc1
spec:
  bfdPort:
    enabled: true
    ip: 10.255.255.255
  ...
  ...
apiVersion: kubeovn.io/v1
kind: Subnet
metadata:
  name: subnet1
spec:
  vpc: vpc1
  protocol: IPv4
  cidrBlock: 192.168.0/0
```

BFD Port OVN LR BFD LRP

```
$ kubectl get nbctl show vpc1
router 0c1d1e8f-4c86-4d96-88b2-c4171c7ff824 (vpc1)
  port bfd@vpc1
    mac: "Be:51:4b:16:3c:90"
    networks: ["10.255.255.255"]
  port vpc1-subnet1
    mac: "de:c9:5c:38:7a:61"
    networks: ["192.168.0.1/24"]
```

VPC Egress Gateway .spec.bfd.enabled true

```
apiVersion: kubeovn.io/v1
kind: VpcEgressGateway
metadata:
  name: gateway2
  namespace: default
spec:
  vpc: vpc1
  replicas: 2
  internalSubnet: subnet1
  externalSubnet: macvlan
  bfd:
    enabled: true
  policies:
    - snat: true
      ipBlocks:
        - 192.168.0.0/24
```

VPC Egress Gateway

```

$ kubectl get veg gateway2 -o wide
NAME      VPC      REPLICAS   BFD  ENABLED    EXTERNAL SUBNET    PHASE      READY     INTERNAL  IPS                                     EXTERNAL IPs          WORKING
NODES          AGE
gateway2   vpc1    2           true   macvlan        Completed   true   [ "192.168.0.2", "192.168.0.3" ]   [ "172.17.0.13", "172.17.0.14" ]   [ "kube-ovn-
worker", "kube-ovn-worker2" ]   58s

$ kubectl get pod -l ovn.kubernetes.io/vpc-egress-gateway=gateway2 -o wide
NAME            READY   STATUS    RESTARTS   AGE   IP           NODE   NOMINATED NODE   READINESS GATES
gateway2-fcc6b8b87-8lgvx  1/1    Running   0          2m18s  192.168.0.3  kube-ovn-worker2  <none>  <none>
gateway2-fcc6b8b87-wmmw6  1/1    Running   0          2m18s  192.168.0.2  kube-ovn-worker   <none>  <none>

$ kubectl ko nbctl lr-route-list vpc1
IPv4 Routes
Route Table <main>:
  192.168.0.0/24          192.168.0.2 src-ip ecmp ecmp-symmetric-reply bfd
  192.168.0.0/24          192.168.0.3 src-ip ecmp ecmp-symmetric-reply bfd

$ kubectl ko nbctl list bfd
_uuid          : 223ede10-9169-4c7d-9524-a546e24bfab5
detect_mult    : 3
dst_ip         : "192.168.0.2"
external_ids   : {af="4", vendor=kube-ovn, vpc-egress-gateway="default/gateway2"}
logical_port   : "bfd@vpc1"
min_rx         : 1000

```

```

min_tx      : 1000
options     : {}
status      : up

_uuid       : b050c75e-2462-470b-b89c-7bd38889b758
detect_mult : 3
dst_ip      : "192.168.0.3"
external_ids: {pf="4", vendor=kube-ovn, vpc-egress-gateway="default/gateway2"}
logical_port: "bfd@vpc1"
min_rx      : 1000
min_tx      : 1000
options     : {}
status      : up

```

Pod BFD

```

$ kubectl exec gateway2-fcc6b8b87-8lgvx -c bfdd -- bfdd-control status
There are 1 sessions:
Session 1
id=1 local=192.168.0.3 (p) remote=10.255.255.255 state=Up

$ kubectl exec gateway2-fcc6b8b87-wmww6 -c bfdd -- bfdd-control status
There are 1 sessions:
Session 1
id=1 local=192.168.0.2 (p) remote=10.255.255.255 state=Up

```

VPC BFD PORT

			false	BFD Port	true
enabled	boolean				
ip	string	-	BFD Port	IP	169.255.255.252
				IPv6	169.255.255.251
nodeSelector	object	-	BFD Port	BFD Port	-
				OVN HA Chassis Group Active/ Backup Active nodeSelector Kube-OVN	
			kubectl ko nbctl list ha_chassis_group OVN HA Chassis Group		
nodeSelector.matchLabels	dict/map	-			-
nodeSelector.matchExpressions	object array	-			-

VPC EGRESS GATEWAY

Spec

vpc	string	VPC	ovn-cluster	VPC	vpc1
replicas	integer/int32	1			2
prefix	string	-		Deployment	veg-
image	string	-		Deployment	docker.io/kubeovn/kube-ovn
internalSubnet	string	VPC		VPC	subnet1
externalSubnet	string	-			ext1
internalIPs	string array	-	VPC IPv6	IP IP	10.16.0.101 / fd00::11 / 1
			<replicas> + 1 Pod		
externalIPs	string array	-	IPv6	IP IP	10.16.0.101 / fd00::11 / 1
			<replicas> + 1 Pod		
bfd	object	-		BFD	-
policies	object array	-		Egress selectors	-
selectors	object array	-		Namespace Selector Pod Selector Egress Pod SNAT/ MASQUERADE policies	-
nodeSelector	object array	-		Deployment/Pod	-
trafficPolicy	string	Cluster		Cluster / Local BFD Local Egress VPC	Local
				Egress Gateway VPC Egress Gateway Egress	

BFD

<code>enabled</code>	<code>boolean</code>	<code>false</code>	<code>BFD</code>	<code>true</code>
<code>minRX</code>	<code>integer/int32</code>	<code>1000</code>	<code>BFD minRX</code>	<code>500</code>
<code>minTX</code>	<code>integer/int32</code>	<code>1000</code>	<code>BFD minTX</code>	<code>500</code>
<code>multiplier</code>	<code>integer/int32</code>	<code>3</code>	<code>BFD multiplier</code>	<code>1</code>

Egress

<code>snat</code>	<code>boolean</code>	<code>false</code>	<code>SNAT/MASQUERADE</code>	<code>true</code>
<code>ipBlocks</code>	<code>string array</code>	<code>-</code>	<code>Gateway IP</code>	<code>192.168.0.1 / 192.168.0.0/24</code>
<code>subnets</code>	<code>string array</code>	<code>-</code>	<code>Gateway VPC</code>	<code>subnet1 IPv6</code>

Selectors

<code>namespaceSelector</code>	<code>object</code>	<code>-</code>	<code>Namespace</code>	<code>-</code>
<code>podSelector</code>	<code>object</code>	<code>-</code>	<code>Pod</code>	<code>-</code>
<code>namespaceSelector.matchLabels</code>	<code>dict/map</code>	<code>-</code>	<code>Namespace</code>	<code>-</code>
<code>namespaceSelector.matchExpressions</code>	<code>object array</code>	<code>-</code>	<code>-</code>	<code>-</code>
<code>podSelector.matchLabels</code>	<code>dict/map</code>	<code>-</code>	<code>Pod</code>	<code>-</code>
<code>podSelector.matchExpressions</code>	<code>object array</code>	<code>-</code>	<code>Pod</code>	<code>-</code>

<code>matchLabels</code>	<code>dict/map</code>	<code>-</code>	<code>-</code>
<code>matchExpressions</code>	<code>object array</code>	<code>-</code>	<code>-</code>
<code>matchFields</code>	<code>object array</code>	<code>-</code>	<code>-</code>

Status

ready	boolean	Gateway	true
phase	string	Gateway	Pending / Processing / Completed
internalIPs	string array	VPC	IP
externalIPs	string array	IP	-
workload	object		-
workload.apiVersion	string	API	apps/v1
workload.kind	string		Deployment
workload.name	string		gateway1
workload.nodes	string array		-
conditions	object array	-	-

[PDF](#)[Slack](#)[Support](#)

⌚2025 11 20

⌚2024 12 12



GitHub



5.2.4

5.3 VPC QoS

Kube-OVN QoS Policy CRD VPC

5.3.1 EIP QoS

EIP	1Mbps	1	shared=false	QoS Policy	EIP	QoS Policy	QoS
-----	-------	---	--------------	------------	-----	------------	-----

QoS Policy

```
apiVersion: kubeovn.io/v1
kind: QoS Policy
metadata:
  name: qos-eip-example
spec:
  shared: false
  bindingType: EIP
  bandwidthLimitRules:
    - name: eip-ingress
      rateMax: "1" # Mbps
      burstMax: "1" # Mbps
      priority: 1
      direction: ingress
    - name: eip-egress
      rateMax: "1" # Mbps
      burstMax: "1" # Mbps
      priority: 1
      direction: egress
```

Iptables EIP

```
kind: IptablesEIP
apiVersion: kubeovn.io/v1
metadata:
  name: eip-1
spec:
  natGwDp: gw1
  qosPolicy: qos-eip-example
```

.spec.qosPolicy

5.3.2 QoS EIP

label	qos	eip
-------	-----	-----

```
# kubectl get eip -l ovn.kubernetes.io/qos=qos-eip-example
NAME     IP          MAC        NAT      NATGWDP   READY
eip-1   172.18.11.24  00:00:00:34:41:0B  fip    gw1      true
```

5.3.3 VPC NATGW net1 QoS

VPC NATGW	net1	10Mbps	3	shared=true	QoS Policy	QoS Policy
-----------	------	--------	---	-------------	------------	------------

QoS Policy

```
apiVersion: kubeovn.io/v1
kind: QoS Policy
metadata:
  name: qos-natgw-example
spec:
  shared: true
  bindingType: NATGW
  bandwidthLimitRules:
    - name: net1-ingress
      interface: net1
      rateMax: "10" # Mbps
      burstMax: "10" # Mbps
      priority: 3
      direction: ingress
    - name: net1-egress
      interface: net1
      rateMax: "10" # Mbps
      burstMax: "10" # Mbps
```

```
priority: 3
direction: egress
```

VpcNatGateway

```
kind: VpcNatGateway
apiVersion: kubeovn.io/v1
metadata:
  name: gw1
spec:
  vpc: test-vpc-1
  subnet: net1
  lanIp: 10.0.1.254
  qosPolicy: qos-natgw-example
  selector:
    - "kubernetes.io/hostname: kube-ovn-worker"
    - "kubernetes.io/os: linux"
```

```
.spec.qosPolicy
```

5.3.4 net1 QoS

net1	5Mbps	2	shared=true	QoS Policy	QoS Policy
------	-------	---	-------------	------------	------------

QoS Policy

```
apiVersion: kubeovn.io/v1
kind: QoS Policy
metadata:
  name: qos-natgw-example
spec:
  shared: true
  bindingType: NATGW
  bandwidthLimitRules:
    - name: net1-extip-ingress
      interface: net1
      rateMax: "5" # Mbps
      burstMax: "5" # Mbps
      priority: 2
      direction: ingress
      matchType: ip
      matchValue: src 172.18.11.22/32
    - name: net1-extip-egress
      interface: net1
      rateMax: "5" # Mbps
      burstMax: "5" # Mbps
      priority: 2
      direction: egress
      matchType: ip
      matchValue: dst 172.18.11.23/32
```

VpcNatGateway

```
kind: VpcNatGateway
apiVersion: kubeovn.io/v1
metadata:
  name: gw1
spec:
  vpc: test-vpc-1
  subnet: net1
  lanIp: 10.0.1.254
  qosPolicy: qos-natgw-example
  selector:
    - "kubernetes.io/hostname: kube-ovn-worker"
    - "kubernetes.io/os: linux"
```

5.3.5 QoS NATGW

label	qos	eip
-------	-----	-----

```
# kubectl get vpc-nat-gw -l ovn.kubernetes.io/qos=qos-natgw-example
NAME    VPC      SUBNET   LANIP
gw1    test-vpc-1  net1    10.0.1.254
```

5.3.6 qos

```
# kubectl get qos -A
NAME          SHARED   BINDINGTYPE
```

qos-eip-example	false	EIP
qos-natgw-example	true	NATGW

5.3.7



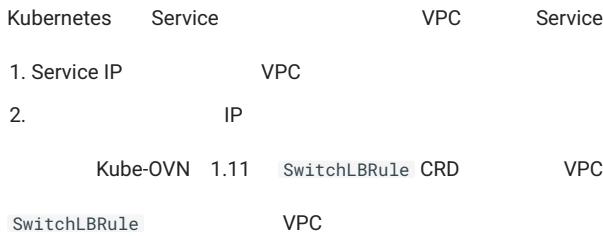
⌚2025 3 3

⌚2023 5 9



5.3.8

5.4 VPC



5.4.1 Selector

```

selector      label      pod
SwitchLBRule
  
```

```

apiVersion: kubeovn.io/v1
kind: SwitchLBRule
metadata:
  name: cjh-slr-nginx
spec:
  vip: 1.1.1.1
  sessionAffinity: ClientIP
  namespace: default
  selector:
    - app:nginx
  ports:
    - name: dns
      port: 8888
      targetPort: 80
      protocol: TCP
  
```

- selector, sessionAffinity, port Kubernetes Service
- vip IP
- namespace, selector Pod

```

Kube-OVN   SwitchLBRule   Pod   Pod   VPC   L2 LB
  
```

5.4.2 Endpoints

```

endpoints      selector      kubevirt      vm
SwitchLBRule
  
```

```

apiVersion: kubeovn.io/v1
kind: SwitchLBRule
metadata:
  name: cjh-slr-nginx
spec:
  vip: 1.1.1.1
  sessionAffinity: ClientIP
  namespace: default
  endpoints:
    - 192.168.0.101
    - 192.168.0.102
    - 192.168.0.103
  ports:
    - name: dns
      port: 8888
      targetPort: 80
      protocol: TCP
  
```

- sessionAffinity, port Kubernetes Service
- vip IP
- namespace, selector Pod
- endpoints IP

```
    selector  endpoints ,      selector
```

5.4.3

OVN IPv4

[Health Checks](<https://www.ovn.org/support/dist-docs/ovn-nb.5.html>)

ovn	SwitchLBRule	SwitchLBRule	VPC	subnet	vip	ip_port_mappings
load_balancer_health_check						
•	vip	subnet	subnet		SwitchLBRule	
•	Selector					

```
root@server:~# kubectl get po -o wide -n vulpecula
NAME           READY   STATUS    RESTARTS   AGE     IP          NODE   NOMINATED NODE   READINESS GATES
nginx-78d9578975-f4qn4   1/1    Running   3        4d16h  10.16.0.4  worker <none>        <none>
nginx-78d9578975-t8tm5   1/1    Running   3        4d16h  10.16.0.6  worker <none>        <none>

#   slr
root@server:~# cat << END > slr.yaml
apiVersion: kubeovn.io/v1
kind: SwitchLBRule
metadata:
  name: nginx
  namespace: vulpecula
spec:
  vip: 1.1.1.1
  sessionAffinity: ClientIP
  namespace: default
  selector:
    - app:nginx
  ports:
    - name: dns
      port: 8888
      targetPort: 80
      protocol: TCP
END
root@server:~# kubectl apply -f slr.yaml
root@server:~# kubectl get slr
NAME           VIP       PORT(S)      SERVICE           AGE
vulpecula-nginx  1.1.1.1  8888/TCP   default/slr-vulpecula-nginx  3d21h
```

subnet vip

```
#   vip

root@server:~# kubectl get vip
NAME   NS     V4IP      MAC          V6IP      PMAC      SUBNET      READY   TYPE
vulpecula-subnet  10.16.0.2  00:00:00:39:95:C1  <nil>      vulpecula-subnet  true
```

Load_Balancer_Health_Check Service_Monitor

```
root@server:~# kubectl ko nbctl list Load_Balancer
_uuid          : 3ccb6d43-44aa-4028-962f-30d2dba9f0b8
external_ids   : {}
health_check   : [5bee3f12-6b54-411c-9cc8-c9def8f67356]
ip_port_mappings : {"10.16.0.4":"nginx-78d9578975-f4qn4.default:10.16.0.2", "10.16.0.6":"nginx-78d9578975-t8tm5.default:10.16.0.2"}
name           : cluster-tcp-session-loadbalancer
options         : {affinity_timeout="10800"}
protocol        : tcp
selection_fields : [ip_src]
vips            : {"1.1.1.1:8888":"10.16.0.4:80,10.16.0.6:80"}

root@server:~# kubectl ko nbctl list Load_Balancer_Health_Check
_uuid          : 5bee3f12-6b54-411c-9cc8-c9def8f67356
external_ids   : {switch_lb_subnet=vulpecula-subnet}
options         : {failure_count="3", interval="5", success_count="3", timeout="20"}
vip             : "1.1.1.1:8888"

root@server:~# kubectl ko sbctl list Service_Monitor
_uuid          : 1bddc541-cc49-44ea-9935-a4208f627a91
external_ids   : {}
ip              : "10.16.0.4"
logical_port   : nginx-78d9578975-f4qn4.default
options         : {failure_count="3", interval="5", success_count="3", timeout="20"}
port            : 80
protocol        : tcp
```

```

src_ip : "10.16.0.2"
src_mac : "c6:d4:b8:08:54:e7"
status : online

_uuid : 84dd24c5-e1b4-4e97-9daa-13687ed59785
external_ids : {}
ip : "10.16.0.6"
logical_port : nginx-78d9578975-t8tm5.default
options : {"failure_count="3", interval="5", success_count="3", timeout="20"}
port : 80
protocol : tcp
src_ip : "10.16.0.2"
src_mac : "c6:d4:b8:08:54:e7"
status : online

```

vip

```

root@server:~# kubectl exec -it -n vulpecula nginx-78d9578975-t8tm5 -- curl 1.1.1.1:8888
<!DOCTYPE html>
<html>
<head>
<title>Welcome to nginx!</title>
<p><em>Thank you for using nginx.</em></p>
</body>
</html>

```

pod

```

kubectl delete po nginx-78d9578975-f4qn4
kubectl get po -o wide -n vulpecula
NAME           READY   STATUS    RESTARTS   AGE     IP          NODE   NOMINATED NODE   READINESS GATES
nginx-78d9578975-lxmvh  1/1    Running   0          31s    10.16.0.8   worker <none>        <none>
nginx-78d9578975-t8tm5  1/1    Running   3          4d16h   10.16.0.6   worker <none>        <none>

```

Load_Balancer_Health_Check Service_Monitor

```

root@server:~# kubectl ko nbctl list Load_Balancer
_uuid : 3ccb6d43-44aa-4028-962f-30d2dba9f0b8
external_ids : {}
health_check : [5bee3f12-6b54-411c-9cc8-c9def8f67356]
ip_port_mappings : {"10.16.0.4":"nginx-78d9578975-f4qn4.default:10.16.0.2", "10.16.0.6":"nginx-78d9578975-t8tm5.default:10.16.0.2", "10.16.0.8":"nginx-78d9578975-lxmvh.default:10.16.0.2"}
name : cluster-tcp-session-loadbalancer
options : {"affinity_timeout="10800"}
protocol : tcp
selection_fields : [ip_src]
vips : {"1.1.1.1:8888"="10.16.0.6:80,10.16.0.8:80"}

root@server:~# kubectl ko nbctl list Load_Balancer_Health_Check
_uuid : 5bee3f12-6b54-411c-9cc8-c9def8f67356
external_ids : {switch_lb_subnet=vulpecula-subnet}
options : {"failure_count="3", interval="5", success_count="3", timeout="20"}
vip : "1.1.1.1:8888"

root@server:~# kubectl ko sbctl list Service_Monitor
_uuid : 84dd24c5-e1b4-4e97-9daa-13687ed59785
external_ids : {}
ip : "10.16.0.6"
logical_port : nginx-78d9578975-t8tm5.default
options : {"failure_count="3", interval="5", success_count="3", timeout="20"}
port : 80
protocol : tcp
src_ip : "10.16.0.2"
src_mac : "c6:d4:b8:08:54:e7"
status : online

_uuid : 5917b7b7-a999-49f2-a42d-da81f1eeb28f
external_ids : {}
ip : "10.16.0.8"
logical_port : nginx-78d9578975-lxmvh.default
options : {"failure_count="3", interval="5", success_count="3", timeout="20"}
port : 80
protocol : tcp
src_ip : "10.16.0.2"
src_mac : "c6:d4:b8:08:54:e7"
status : online

```

SwitchLBRule Load_Balancer_Health_Check Service_Monitor vip

```

root@server:~# kubectl delete -f slr.yaml
switchlrule.kubeovn.io "vulpecula-nginx" deleted
root@server:~# kubectl get vip
No resources found

```

```
root@server:~# kubectl ko sbctl list Service_Monitor
root@server:~#
root@server:~# kubectl ko nbctl list Load_Balancer_Health_Check
root@server:~#
```

[PDF](#)[Slack](#)[Support](#)

⌚2025 9 10

⌚2022 12 15



5.4.4

5.5 VPC DNS



5.5.1 vpc-dns

```

apiVersion: rbac.authorization.k8s.io/v1
kind: ClusterRole
metadata:
  labels:
    kubernetes.io/bootstrapping: rbac-defaults
  name: system:vpc-dns
rules:
- apiGroups:
  - ""
  resources:
  - endpoints
  - services
  - pods
  - namespaces
  verbs:
  - list
  - watch
- apiGroups:
  - discovery.k8s.io
  resources:
  - endpointslices
  verbs:
  - list
  - watch
---
apiVersion: rbac.authorization.k8s.io/v1
kind: ClusterRoleBinding
metadata:
  annotations:
    rbac.authorization.kubernetes.io/autoupdate: "true"
  labels:
    kubernetes.io/bootstrapping: rbac-defaults
  name: vpc-dns
roleRef:
  apiGroup: rbac.authorization.k8s.io
  kind: ClusterRole
  name: system:vpc-dns
subjects:
- kind: ServiceAccount
  name: vpc-dns
  namespace: kube-system
---
apiVersion: v1
kind: ServiceAccount
metadata:
  name: vpc-dns
  namespace: kube-system
---
apiVersion: v1
kind: ConfigMap
metadata:
  name: vpc-dns-corefile
  namespace: kube-system
data:
  Corefile: |
    .:53 {
      errors
      health {
        lameduck 5s
      }
      ready
      kubernetes cluster.local in-addr.arpa ip6.arpa {
        pods insecure
        fallthrough in-addr.arpa ip6.arpa
      }
      prometheus :9153
      forward . /etc/resolv.conf {
        prefer_udp
      }
    }
    cache 30
  
```

```

    loop
    reload
    loadbalance
}

```

nat-gw-pod

5.5.2

```

apiVersion: "k8s.cni.cncf.io/v1"
kind: NetworkAttachmentDefinition
metadata:
  name: ovn-nad
  namespace: default
spec:
  config: '{
    "cniVersion": "0.3.0",
    "type": "kube-ovn",
    "server_socket": "/run/openvswitch/kube-ovn-daemon.sock",
    "provider": "ovn-nad.default.ovn"
}'

```

5.5.3 vpc-dns Configmap

kube-system configmap vpc-dns vpc-dns

```

apiVersion: v1
kind: ConfigMap
metadata:
  name: vpc-dns-config
  namespace: kube-system
data:
  coredns-vip: 10.96.0.3
  enable-vpc-dns: "true"
  nad-name: ovn-nad
  nad-provider: ovn-nad.default.ovn

```

- enable-vpc-dns true
- coredns-image dns coredns
- coredns-vip coredns lb vip
- coredns-template coredns URL ovn coredns-template.yaml <https://raw.githubusercontent.com/kubeovn/kube-ovn/> /yaml/coredns-template.yaml
- nad-name network-attachment-definitions
- nad-provider provider
- k8s-service-host coredns k8s apiserver ip apiserver
- k8s-service-port coredns k8s apiserver port apiserver

5.5.4 vpc-dns

vpc-dns yaml

```

kind: VpcDns
apiVersion: kubeovn.io/v1
metadata:
  name: test-cjh1
spec:
  vpc: cjh-vpc-1
  subnet: cjh-subnet-1
  replicas: 2

```

- vpc dns vpc
- subnet dns
- replicas: vpc dns deployment replicas

```
# kubectl get vpc-dns
NAME      ACTIVE   VPC      SUBNET
test-cjh1  false    cjh-vpc-1  cjh-subnet-1
test-cjh2  true     cjh-vpc-1  cjh-subnet-2
```

```
ACTIVE : true      dns  false
          VPC      DNS
•  VPC      vpc-dns      VPC      subnet      vpc-dns      true      false
•  true   vpc-dns      false  vpc-dns
```

5.5.5

vpc-dns Pod label app=vpc-dns vpc-dns Pod

```
# kubectl -n kube-system get pods -l app=vpc-dns
NAME                  READY   STATUS    RESTARTS   AGE
vpc-dns-test-cjh1-7b878d96b4-g5979  1/1    Running   0          28s
vpc-dns-test-cjh1-7b878d96b4-ltmf9   1/1    Running   0          28s
```

slr

```
# kubectl -n kube-system get slr
NAME      VIP      PORT(S)      SERVICE      AGE
vpc-dns-test-cjh1  10.96.0.3  53/UDP, 53/TCP  kube-system/vpc-dns-test-cjh1  113s
```

VPC Pod dns :

```
nslookup kubernetes.default.svc.cluster.local 10.96.0.3
```

VPC switch lb rule VPC Pod

[PDF](#)

[Slack](#)

[Support](#)

⌚2025 9 10

⌚2022 12 26



5.5.6

5.6 SecurityGroup

Kube-OVN	Pod	Pod	IPAM	L2/L3	MAC	IP	
⚠ Warning							
Kube-OVN	NetworkPolicy	Network Policy API	Subnet ACL	Security Group	OVN ACL	NetworkPolicy	NetworkPolicy API

5.6.1

```
apiVersion: kubeovn.io/v1
kind: SecurityGroup
metadata:
  name: sg-example
spec:
  allowSameGroupTraffic: true
  egressRules:
  - ipVersion: ipv4
    policy: allow
    priority: 1
    protocol: all
    remoteAddress: 10.16.0.13 # 10.16.0.0/16
    remoteType: address
  ingressRules:
  - ipVersion: ipv4
    policy: deny
    priority: 1
    protocol: icmp
    remoteAddress: 10.16.0.14
    remoteType: address
```

Kube-OVN

Pod	annotation	annotation	Kube-OVN	IPAM	L2/L3	Pod	L2/L3	Pod
• port_security:			Kube-OVN	IPAM	L2/L3	Pod	L2/L3	Pod
• security_groups:		ACL						
annotation								
ovn.kubernetes.io/port_security: "true" ovn.kubernetes.io/security_groups: sg-example								

5.6.2

•	ACL	OVN	ACL	ACL				
•	priority	1-200	ACL	ACL	ACL	= 2300 -		ACL
•		Kube-OVN	CNI	Pod	Pod	Pod	ContainerCreating	Running

5.6.3

YAML Pod annotation

```
apiVersion: v1
kind: Pod
metadata:
  labels:
    app: static
  annotations:
    ovn.kubernetes.io/port_security: 'true'
    ovn.kubernetes.io/security_groups: 'sg-example'
  name: sg-test-pod
  namespace: default
spec:
  nodeName: kube-ovn-worker
  containers:
```

```
- image: docker.io/library/nginx:alpine
  imagePullPolicy: IfNotPresent
  name: qatest
```

```
# kubectl get pod -o wide
NAME           READY   STATUS      RESTARTS   AGE     IP          NODE
sg-test-pod    0/1     ContainerCreating  0          5h32m  <none>      kube-ovn-worker
test-99fffff86-52h9r  1/1     Running     0          5h41m  10.16.0.14  kube-ovn-control-plane
test-99fffff86-qcgjw  1/1     Running     0          5h43m  10.16.0.13  kube-ovn-worker
```

kubectl describe pod Pod

```
# kubectl describe pod sg-test-pod
Name:         sg-test-pod
Namespace:    default
Priority:    0
Node:        kube-ovn-worker/172.18.0.2
Start Time:  Tue, 28 Feb 2023 10:29:36 +0800
Labels:      app=static
Annotations: ovn.kubernetes.io/allocated: true
             ovn.kubernetes.io/cidr: 10.16.0.0/16
             ovn.kubernetes.io/gateway: 10.16.0.1
             ovn.kubernetes.io/ip_address: 10.16.0.15
             ovn.kubernetes.io/logical_router: ovn-cluster
             ovn.kubernetes.io/logical_switch: ovn-default
             ovn.kubernetes.io/mac_address: 00:00:00:FA:17:97
             ovn.kubernetes.io/pod_nic_type: veth-pair
             ovn.kubernetes.io/port_security: true
             ovn.kubernetes.io/routed: true
             ovn.kubernetes.io/security_groups: sg-allow-reject
Status:      Pending
IP:
IPs:        <none>
.
.
Events:
  Type  Reason     Age     From      Message
  ----  ----     --  ----      -----
  Warning FailedCreatePodSandBox 5m3s (x70 over 4h59m) kubelet (combined from similar events): Failed to create pod sandbox: rpc error: code = Unknown desc = failed to setup network for sandbox "40636e0c7f1ade5500fa958486163d74f2e2300051a71522a9af7ba0538afb6": plugin type="kube-ovn" failed (add): RPC failed; request ip return 500 configure nic failed 10.16.0.15 network not ready after 200 ping 10.16.0.1
```

```
apiVersion: kubeovn.io/v1
kind: SecurityGroup
metadata:
  name: sg-gw-both
spec:
  allowSameGroupTraffic: true
  egressRules:
  - ipVersion: ipv4
    policy: allow
    priority: 2
    protocol: all
    remoteAddress: 10.16.0.13
    remoteType: address
  - ipVersion: ipv4
    policy: allow
    priority: 1
    protocol: all
    remoteAddress: 10.16.0.1
    remoteType: address
  ingressRules:
  - ipVersion: ipv4
    policy: deny
    priority: 2
    protocol: icmp
    remoteAddress: 10.16.0.14
    remoteType: address
  - ipVersion: ipv4
    policy: allow
    priority: 1
    protocol: icmp
    remoteAddress: 10.16.0.1
    remoteType: address
```

yaml Pod Pod

```
apiVersion: v1
kind: Pod
```

```

metadata:
  labels:
    app: static
  annotations:
    ovn.kubernetes.io/port_security: 'true'
    ovn.kubernetes.io/security_groups: 'sg-gw-both'
  name: sg-gw-both
  namespace: default
spec:
  nodeName: kube-ovn-worker
  containers:
  - image: docker.io/library/nginx:alpine
    imagePullPolicy: IfNotPresent
  name: qatest

```

Pod

# kubectl get pod -o wide									
NAME	READY	STATUS	RESTARTS	AGE	IP	NODE	NOMINATED NODE	NODE	READINESS GATES
sg-test-pod	0/1	ContainerCreating	0	5h41m	<none>	kube-ovn-worker	<none>	<none>	<none>
sg-gw-both	1/1	Running	0	5h37m	10.16.0.19	kube-ovn-worker	<none>	<none>	<none>

[PDF](#)
[Slack](#)
[Support](#)

⌚2025 9 10

⌚2023 2 28



5.6.4

5.7 OVN EIP FIP SNAT DNAT

Note									
VPC	VPC								
VPC	VPC NAT	OVN	Egress Gateway	VPC NAT	Kube-OVN	VPC NAT	Pod	VPC	
Macvlan	Pod	iptables							
OVN	OVN	NAT		OVN	BFD	OVN	OVN		
Egress Gateway	VPC NAT								

VPC OVN NAT provider-network vlan (external) subnet VPC EIP/SNAT

5.7.1

- kube-ovn-controller kube-ovn-cni ovn-external-gw-config VPC spec enableExternal
- CRD provider-network vlan subnet VPC spec extraExternalSubnets ovn-eip ovn-dnat ovn-fip ovn-snats CRD

```
graph LR
pod-->subnet-->vpc-->lsp--bind-->gw-chassis-->snat-->lsp-->external-subnet
lsp--peer-->lsp
```

Pod SNAT Pod Fip

```
graph LR
pod-->subnet-->vpc-->lsp--bind-->local-chassis-->snat-->lsp-->external-subnet
lsp--peer-->lsp
```

Pod FIP (dnat_and_snat)

- CRD iptables nat gw
- ovn eip: ip underlay provider network vlan subnet
 - ovn fip dnat snat VPC ip vip
 - ovn snat VPC ip snat
 - ovn dnat router lb , ip + VPC endpoints

5.7.2 1.

OpenStack Neutron ovn provider network VPC EIP/SNAT

vlan vlan 0 vlan id

```
#
# 1. kube-ovn-controller
- --external-gateway-vlanid=204
- --external-gateway-switch=external204

# 2. kube-ovn-cni :
- --external-gateway-switch=external204
```

```
###          vlan id          underlay
•      provider network vlan subnet
•      VPC enable_eip_snat      vlan subnet      ip ipam
•      VPC enable_eip_snat ,    pod annotation fip snat
•      VPC enable_eip_snat      vlan subnet      VPC eip snat
```

1.1 underlay

```
# provider-network vlan subnet
# cat 01-provider-network.yaml

apiVersion: kubeovn.io/v1
kind: ProviderNetwork
metadata:
  name: external204
spec:
  defaultInterface: vlan

# cat 02-vlan.yaml

apiVersion: kubeovn.io/v1
kind: Vlan
metadata:
  name: vlan204
spec:
  id: 204
  provider: external204

# cat 03-vlan-subnet.yaml

apiVersion: kubeovn.io/v1
kind: Subnet
metadata:
  name: external204
spec:
  protocol: IPv4
  cidrBlock: 10.5.204.0/24
  gateway: 10.5.204.254
  vlan: vlan204
  excludeIps:
  - 10.5.204.1..10.5.204.100
```

1.2 VPC eip_snat

```
# VPC underlay provider subnet
# cat 00-centralized-external-gw-no-ip.yaml

apiVersion: v1
kind: ConfigMap
metadata:
  name: ovn-external-gw-config
  namespace: kube-system
data:
  enable-external-gw: "true"
  external-gw-nodes: "pc-node-1,pc-node-2,pc-node-3"
  type: "centralized"
  external-gw-nic: "vlan" # ovs
  external-gw-addr: "10.5.204.254/24" # underlay ip
```

logical router port (lrp)	ip	mac	underlay	lrp	ovn	eip
ip	lrp	ovn-eip		lrp	ovn	eip

1.3 VPC eip snat fip

node

```
# external-gw-nodes
kubectl label nodes pc-node-1 pc-node-2 pc-node-3 ovn.kubernetes.io/external-gw=true

# cat 00-ns.yaml

apiVersion: v1
```

```

kind: Namespace
metadata:
  name: vpc1

# cat 01-vpc-ecmp-enable-external-bfd.yml

kind: Vpc
apiVersion: kubeovn.io/v1
metadata:
  name: vpc1
spec:
  namespaces:
    - vpc1
  enableExternal: true
  staticRoutes:
    - cidr: 0.0.0.0/0
      nextHopIP: 10.5.204.254
      policy: policyDst

# VPC      enableExternal      lrp

# cat 02-subnet.yml

apiVersion: kubeovn.io/v1
kind: Subnet
metadata:
  name: vpc1-subnet1
spec:
  cidrBlock: 192.168.0.0/24
  default: false
  disableGatewayCheck: false
  disableInterConnection: true
  enableEcmp: true
  gatewayNode: ""
  gatewayType: distributed
  #gatewayType: centralized
  natOutgoing: false
  private: false
  protocol: IPv4
  provider: ovn
  vpc: vpc1
  namespaces:
    - vpc1

#
#          subnet

```

```

# kubectl ko nbctl show vpc1

router 87ad06fd-71d5-4ff8-a1f0-54fa3bba1a7f (vpc1)
  port vpc1-vpc1-subnet1
    mac: "00:00:00:ED:8E:C7"
    networks: ["192.168.0.1/24"]
  port vpc1-external204
    mac: "00:00:00:EF:05:C7"
    networks: ["10.5.204.105/24"]
    gateway chassis: [7cedd14f-265b-42e5-ac17-e03e7a1f2342 276bacb-fe9c-4476-b41d-05872a94976d fd9f140c-c45d-43db-a6c0-0d4f8ea298dd]
  nat 21d853b0-f7b4-40bd-9a53-31d2e2745739
    external ip: "10.5.204.115"
    logical ip: "192.168.0.0/24"
    type: "snat"

```

```

# kubectl ko nbctl lr-route-list vpc1

IPv4 Routes
Route Table <main>:
  0.0.0.0/0           10.5.204.254 dst-ip
#   VPC CRD

```

enableExternal

VPC CRD

1.4

1.4.1 UNDERLAY

eip snat fip 1 eip snat fip

```

#   provider-network  vlan  subnet
# cat 01-extra-provider-network.yaml
apiVersion: kubeovn.io/v1
kind: ProviderNetwork
metadata:
  name: extra
spec:
  defaultInterface: vlan
# cat 02-extra-vlan.yaml

```

```

apiVersion: kubeovn.io/v1
kind: Vlan
metadata:
  name: wlan0
spec:
  id: 0
  provider: extra
# cat 03-extra-vlan-subnet.yaml
apiVersion: kubeovn.io/v1
kind: Subnet
metadata:
  name: extra
spec:
  protocol: IPv4
  cidrBlock: 10.10.204.0/24
  gateway: 10.10.204.254
  vlan: wlan0
  excludeIps:
    - 10.10.204.1..10.10.204.100

```

1.4.2 VPC

```

apiVersion: kubeovn.io/v1
kind: Vpc
metadata:
  name: vpc1
spec:
  namespaces:
    - vpc1
  enableExternal: true #   enableExternal   VPC           external   ls
  extraExternalSubnets: #   extraExternalSubnets
    - extra

```

```

# kubectl get nbctl show vpc1
router 87ad06fd-71d5-4ff8-a1f0-54fa3bba1a7f (vpc1)
  port vpc1-vpc1-subnet1
    mac: "00:00:00:ED:8E:C7"
    networks: ["192.168.0.1/24"]
  port vpc1-external1204
    mac: "00:00:00:EF:05:C7"
    networks: ["10.5.204.105/24"]
    gateway chassis: [7cedd14f-265b-42e5-ac17-e03e7a1f2342 276baccc-f9c-4476-b41d-05872a94976d fd9f140c-c45d-43db-a6c0-0d4f8ea298dd]
  port vpc1-extra
    mac: "00:00:00:EF:6A:C7"
    networks: ["10.10.204.105/24"]
    gateway chassis: [7cedd14f-265b-42e5-ac17-e03e7a1f2342 276baccc-f9c-4476-b41d-05872a94976d fd9f140c-c45d-43db-a6c0-0d4f8ea298dd]

```

5.7.3 2. ovn-eip

iptables-eip	ovn-eip	type
• nat:	ovn dnat fip, snat	nat
• lrp:	underlay	lrp ip dnat snat
• lsp:	ovn bfd ecmp	ovs internal port ecmp

```

---
kind: OvnEip
apiVersion: kubeovn.io/v1
metadata:
  name: eip-static
spec:
  externalSubnet: external1204
  type: nat
#       eip          fip

```

externalSubnet	external204
externalSubnet	extra

2.1 ovn-fip pod fip

```

# kubectl get po -o wide -n vpc1 vpc-1-busybox01
NAME        READY   STATUS    RESTARTS   AGE      IP          NODE
vpc-1-busybox01  1/1     Running   0          3d15h   192.168.0.2   pc-node-2

# kubectl get ip vpc-1-busybox01.vpc1

```

```

NAME      V4IP      V6IP      MAC      NODE      SUBNET
vpc-1-busybox01.vpc1  192.168.0.2          00:00:00:0A:DD:27  pc-node-2  vpc1-subnet1

---
kind: OvnEip
apiVersion: kubeovn.io/v1
metadata:
  name: eip-static
spec:
  externalSubnet: external204
  type: nat

---
kind: OvnFip
apiVersion: kubeovn.io/v1
metadata:
  name: eip-static
spec:
  ovnEip: eip-static
  ipName: vpc-1-busybox01.vpc1 #      ip crd
  type: "centralized"           # centralized     distributed

--#
#      VPC      ip

kind: OvnFip
apiVersion: kubeovn.io/v1
metadata:
  name: eip-static
spec:
  ovnEip: eip-static
  vpc: vpc1
  v4Ip: 192.168.0.2
  type: "centralized"           # centralized     distributed

```

```

# kubectl get ofip
NAME      VPC      V4EIP      V4IP      READY    IPTYPE    IPNAME
eip-for-vip  vpc1  10.5.204.106  192.168.0.3  true    vip      test-fip-vip
eip-static   vpc1  10.5.204.101  192.168.0.2  true    vpc1-busybox01.vpc1
# kubectl get ofip eip-static
NAME      VPC      V4EIP      V4IP      READY    IPTYPE    IPNAME
eip-static   vpc1  10.5.204.101  192.168.0.2  true    vpc1-busybox01.vpc1

[root@pc-node-1 03-cust-vpc]# ping 10.5.204.101
PING 10.5.204.101 (10.5.204.101) 56(84) bytes of data.
64 bytes from 10.5.204.101: icmp_seq=2 ttl=62 time=1.21 ms
64 bytes from 10.5.204.101: icmp_seq=3 ttl=62 time=0.624 ms
64 bytes from 10.5.204.101: icmp_seq=4 ttl=62 time=0.368 ms
^C
--- 10.5.204.101 ping statistics ---
4 packets transmitted, 3 received, 25% packet loss, time 3049ms
rtt min/avg/max/mdev = 0.368/0.734/1.210/0.352 ms
[root@pc-node-1 03-cust-vpc]#

```

```

#      node ping      VPC      pod      ip

#      ip
# kubectl ko nbctl show vpc1
router 87ad6fd-71d5-4ff8-a1f0-54fa3bba1a7f (vpc1)
  port vpc1-vpc1-subnet1
    mac: "00:00:00:ED:8E:C7"
    networks: ["192.168.0.1/24"]
  port vpc1-external204
    mac: "00:00:00:EF:05:C7"
    networks: ["10.5.204.105/24"]
  gateway chassis: [7cedd14f-265b-42e5-ac17-e03e7a1f2342 276baccb-fe9c-4476-b41d-05872a94976d fd9f140c-c45d-43db-a6c0-0d4f8ea298dd]
  nat 813523e7-c68c-408f-bd8c-cba30cb2e4f4
    external ip: "10.5.204.101"
    logical ip: "192.168.0.2"
    type: "dnat_and_snat"

```

2.2 ovn-fip vip fip

```

vip      kubevirt      vip      keepalived kube-vip

fip      VPC      vip      vip

```

```

#      vip      eip      eip      vip
# cat vip.yaml

apiVersion: kubeovn.io/v1
kind: Vip
metadata:
  name: test-fip-vip
spec:
  subnet: vpc1-subnet1

```

```

# cat 04-fip.yaml

---
kind: OvnEip
apiVersion: kubeovn.io/v1
metadata:
  name: eip-for-vip
spec:
  externalSubnet: external204
  type: nat

---
kind: OvnFip
apiVersion: kubeovn.io/v1
metadata:
  name: eip-for-vip
spec:
  ovnEip: eip-for-vip
  ipType: vip      #      fip      pod ip          vip
  ipName: test-fip-vip

---
#      VPC      ip

kind: OvnFip
apiVersion: kubeovn.io/v1
metadata:
  name: eip-for-vip
spec:
  ovnEip: eip-for-vip
  ipType: vip      #      fip      pod ip          vip
  vpc: vpc1
  v4Ip: 192.168.0.3

# kubectl get ofip
NAME      VPC      V4EIP      V4IP      READY    IPTYPE    IPNAME
eip-for-vip  vpc1  10.5.204.106  192.168.0.3  true    vip      test-fip-vip

[root@pc-node-1 fip-vip]# ping 10.5.204.106
PING 10.5.204.106 (10.5.204.106) 56(84) bytes of data.
64 bytes from 10.5.204.106: icmp_seq=1 ttl=62 time=0.694 ms
64 bytes from 10.5.204.106: icmp_seq=2 ttl=62 time=0.436 ms

# node ping

# pod ip

[root@pc-node-1 fip-vip]# kubectl -n vpc1 exec -it vpc-1-busybox03 -- bash
[root@vpc-1-busybox03 /]#
[root@vpc-1-busybox03 /]# 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
1568: eth0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc noqueue state UP group default
  link/ether 00:00:00:56:40:e5 brd ff:ff:ff:ff:ff:ff link-netnsid 0
  inet 192.168.0.5/24 brd 192.168.0.255 scope global eth0
    valid_lft forever preferred_lft forever
  inet 192.168.0.3/24 scope global secondary eth0 #      vip
    valid_lft forever preferred_lft forever
  inet6 fe80::200:ff:fe56:40e5/64 scope link
    valid_lft forever preferred_lft forever

[root@vpc-1-busybox03 /]# tcpdump -i eth0 host 192.168.0.3 -netvv
tcpdump: listening on eth0, link-type EN10MB (Ethernet), capture size 262144 bytes
00:00:00:ed:8e:c7 > 00:00:00:56:40:e5, ethertype IPv4 (0x0800), length 98: (tos 0x0, ttl 62, id 44830, offset 0, flags [DF], proto ICMP (1), length 84)
  10.5.32.51 > 192.168.0.3: ICMP echo request, id 177, seq 1, length 64
00:00:00:56:40:e5 > 00:00:00:ed:8e:c7, ethertype IPv4 (0x0800), length 98: (tos 0x0, ttl 64, id 43962, offset 0, flags [none], proto ICMP (1), length 84)
  192.168.0.3 > 10.5.32.51: ICMP echo reply, id 177, seq 1, length 64

# pod fip icmp

```

5.7.4 3. ovn-snat

3.1 ovn-snat subnet cidr

iptables-snat

```
# cat 03-subnet-snat.yaml

---
kind: OvnEip
apiVersion: kubeovn.io/v1
metadata:
```

```
        name: snat-for-subnet-in-vpc
spec:
  externalSubnet: external1204
  type: nat

---
kind: OvnSnatRule
apiVersion: kubeovn.io/v1
metadata:
  name: snat-for-subnet-in-vpc
spec:
  ovnEip: snat-for-subnet-in-vpc
  vpcSubnet: vpc1-subnet1 # eip

---
#          VPC      subnet    cidr
kind: OvnSnatRule
apiVersion: kubeovn.io/v1
metadata:
  name: snat-for-subnet-in-vpc
spec:
  ovnEip: snat-for-subnet-in-vpc
  vpc: vpc1
  v4IpCidr: 192.168.0.0/24 #           cidr
  ip
```

externalSubnet extra

3.2 ovn-snat pod ip

iptables-snat

```
# cat 03-pod-snat.yaml

---
kind: OvnEip
apiVersion: kubeovn.io/v1
metadata:
  name: snat-for-pod-vpc-ip
spec:
  externalSubnet: external204
  type: nat

---
kind: OvnSnatRule
apiVersion: kubeovn.io/v1
metadata:
  name: snat01
spec:
  ovnEip: snat-for-pod-vpc-ip
  ipName: vpc-1-busybox02.vpc1 # eip
  pod ip

---
#          VPC      ip

kind: OvnSnatRule
apiVersion: kubeovn.io/v1
metadata:
  name: snat-for-subnet-in-vpc
spec:
  ovnEip: snat-for-subnet-in-vpc
  vpc: vpc1
  v4Ipcidr: 192.168.0.4
```

`externalSubnet` `extra`

snat

```
# kubectl get vpc1
router 87ad06fd-71d5-4ff8-a1f0-54fa3bba1a7f (vpc1)
  port vpc1-vpc1-subnet1
    mac: "00:00:00:ED:8E:C7"
    networks: ["192.168.0.1/24"]
  port vpc1-external204
    mac: "00:00:00:EF:05:C7"
    networks: ["10.5.204.105/24"]
    gateway chassis: [7cedd14f-265b-42e5-ac17-e03e7a1f2342 276baccc-fc9c-4476-b41d-05872a94976d fd9f140c-c45d-43db-a6c0-0d4f8ea298dd]
nat 21d853b0-f7b4-40bd-9a53-31d2e2745739
  external ip: "10.5.204.115"
  logical ip: "192.168.0.0/24"
  type: "snat"
nat da77a11f-c523-439c-b1d1-72c664196a0f
  external ip: "10.5.204.116"
  logical ip: "192.168.0.4"
  type: "snat"
```

```
[root@pc-node-1 03-cust-vpc]# kubectl get po -A -o wide | grep busy
vpc1           vpc-1-busybox01          1/1   Running   0      3d15h  192.168.0.2  pc-node-2  <none>    <none>
vpc1           vpc-1-busybox02          1/1   Running   0      17h   192.168.0.4  pc-node-1  <none>    <none>
vpc1           vpc-1-busybox03          1/1   Running   0      17h   192.168.0.5  pc-node-1  <none>    <none>
vpc1           vpc-1-busybox04          1/1   Running   0      17h   192.168.0.6  pc-node-3  <none>    <none>
vpc1           vpc-1-busybox05          1/1   Running   0      17h   192.168.0.7  pc-node-1  <none>    <none>

# kubectl exec -it -n vpc1           vpc-1-busybox04 bash
kubectl exec [POD] [COMMAND] is DEPRECATED and will be removed in a future version. Use kubectl exec [POD] -- [COMMAND] instead.
[root@vpc-1-busybox04 /]#
[root@vpc-1-busybox04 /]#
[root@vpc-1-busybox04 /]# 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 fe80::1:1/128 scope host
    valid_lft forever preferred_lft forever
17095: eth0@if17096: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc noqueue state UP group default
  link/ether 00:00:00:76:94:55 brd ff:ff:ff:ff:ff:ff link-netnsid 0
    inet 192.168.0.6/24 brd 192.168.0.255 scope global eth0
      valid_lft forever preferred_lft forever
    inet6 fe80::200:ff:fe76:9455/64 scope link
      valid_lft forever preferred_lft forever
[root@vpc-1-busybox04 /]# ping 223.5.5.5
PING 223.5.5.5 (223.5.5.5) 56(84) bytes of data.
64 bytes from 223.5.5.5: icmp_seq=1 ttl=114 time=22.2 ms
64 bytes from 223.5.5.5: icmp_seq=2 ttl=114 time=21.8 ms

[root@pc-node-1 03-cust-vpc]# kubectl exec -it -n vpc1           vpc-1-busybox02 bash
kubectl exec [POD] [COMMAND] is DEPRECATED and will be removed in a future version. Use kubectl exec [POD] -- [COMMAND] instead.
[root@vpc-1-busybox02 /]#
[root@vpc-1-busybox02 /]#
[root@vpc-1-busybox02 /]#
[root@vpc-1-busybox02 /]# 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 fe80::1:1/128 scope host
    valid_lft forever preferred_lft forever
1566: eth0@if1567: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc noqueue state UP group default
  link/ether 00:00:00:0b:e9:d0 brd ff:ff:ff:ff:ff:ff link-netnsid 0
    inet 192.168.0.4/24 brd 192.168.0.255 scope global eth0
      valid_lft forever preferred_lft forever
    inet6 fe80::200:ff:fe0b:e9d0/64 scope link
      valid_lft forever preferred_lft forever
[root@vpc-1-busybox02 /]# ping 223.5.5.5
PING 223.5.5.5 (223.5.5.5) 56(84) bytes of data.
64 bytes from 223.5.5.5: icmp_seq=2 ttl=114 time=22.7 ms
64 bytes from 223.5.5.5: icmp_seq=3 ttl=114 time=22.6 ms
64 bytes from 223.5.5.5: icmp_seq=4 ttl=114 time=22.1 ms
^C
--- 223.5.5 ping statistics ---
4 packets transmitted, 3 received, 25% packet loss, time 3064ms
rtt min/avg/max/mdev = 22.126/22.518/22.741/0.278 ms

#       pod      snat

#
```

5.7.5 4. ovn-dnat

4.1 ovn-dnat pod dnat

```
kind: OvnEip
apiVersion: kubeovn.io/v1
metadata:
  name: eip-dnat
spec:
  externalSubnet: underlay

---
kind: OvnDnatRule
apiVersion: kubeovn.io/v1
metadata:
  name: eip-dnat
spec:
  ovnEip: eip-dnat
  ipName: vpc-1-busybox01.vpc1 #      pod ip crd
  protocol: tcp
  internalPort: "22"
  externalPort: "22"

---
#       VPC      ip

kind: OvnDnatRule
apiVersion: kubeovn.io/v1
metadata:
  name: eip-dnat
```

```
spec:  
  ovnEip: eip-dnat  
  protocol: tcp  
  internalPort: "22  
  externalPort: "22  
  vpc: vpc1  
  v4Ip: 192.168.0.3
```

externalSubnet extra

OvnDnatRule IptablesDnatRule

```
# kubectl get oeip eip-dnat
NAME      V4IP          V6IP      MAC                TYPE    READY
eip-dnat  10.5.49.4    00:00:00:4D:CE:49  dnat    true

# kubectl get odnat
NAME      EIP          PROTOCOL  V4EIP        V4IP        INTERNALPORT  EXTERNALPORT  IPNAME
eip-dnat eip-dnat     tcp       10.5.49.4  192.168.0.3  22           22           vpc-1-busybox01.vpc1
                                                 READY
                                                 true
```

4.2 ovn-dnat vip dnat

```
kind: OvnDnatRule
apiVersion: kubeovn.io/v1
metadata:
  name: eip-dnat
spec:
  ipType: vip #      dnat      pod ip
  ovnIp: eip-dnat
  ipName: test-dnat-vip
  protocol: tcp
  internalPort: "22"
  externalPort: "22"

---
#          VPC      ip

kind: OvnDnatRule
apiVersion: kubeovn.io/v1
metadata:
  name: eip-dnat
spec:
  ipType: vip #      dnat      pod ip
  ovnIp: eip-dnat
  ipName: test-dnat-vip
  protocol: tcp
  internalPort: "22"
  externalPort: "22"
  vpc: vpc1
  v4Ip: 192.168.0.4
```

OvnDnatRule IptablesDnatRule

```
# kubectl get vip test-dnat-vip
NAME          V4IP           PV4IP   MAC                   PMAC    V6IP   PV6IP   SUBNET      READY
test-dnat-vip 192.168.0.4     00:00:00:D0:C0:B5   vpc1-subnet1 true

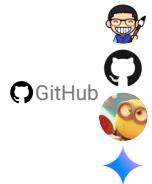
# kubectl get oeip eip-dnat
NAME          V4IP       V6IP   MAC                   TYPE    READY
eip-dnat     10.5.49.4  00:00:00:4D:CE:49  dnat    true

# kubectl get ednat eip-dnat
NAME          EIP        PROTOCOL  V4EIP      V4IP           INTERNALPORT  EXTERNALPORT  IPNAME      READY
eip-dnat     eip-dnat   tcp       10.5.49.4  192.168.0.4  22            22           test-dnat-vip true
```



⌚2025 9 30

⌚2023 3 3



5.7.6

5.8 OVN SNAT ECMP BFD L3 HA

VPC OVN SNAT ECMP Gateway Node ovnnext0

- bfd
- hash

```
graph LR
pod --> vpc-subnet --> vpc --> snat --> ecmp --> external-subnet --> gw-node1-ovnnext0 --> node1-external-switch
external-subnet --> gw-node2-ovnnext0 --> node2-external-switch
external-subnet --> gw-node3-ovnnext0 --> node3-external-switch
```

[ovn-eip-fip-snat.md](#) install.sh provider-network vlan subnet

lsp ovn-eip vpc enable_bfd bfd ecmp

5.8.1 1.

1.1 underlay

1.2 vpc eip_snat

1.3 VPC eip snat fip

[ovn-eip-fip-snat.md](#) VPC ecmp bfd

VPC 2 ovn-eip

```
# cat gw-node-eip.yaml
---
kind: OvnEip
apiVersion: kubeovn.io/v1
metadata:
  name: pc-node-1
spec:
  externalSubnet: external204
  type: lsp

---
kind: OvnEip
apiVersion: kubeovn.io/v1
metadata:
  name: pc-node-2
spec:
  externalSubnet: external204
  type: lsp

---
kind: OvnEip
apiVersion: kubeovn.io/v1
metadata:
  name: pc-node-3
spec:
  externalSubnet: external204
  type: lsp
```

vpc ecmp vpc bfd enable_bfd lrp ovn_eip bfd

5.8.2 2. vpc ecmp bfd L3 HA

```
# cat 01-vpc-ecmp-enable-external-bfd.yml
kind: Vpc
apiVersion: kubeovn.io/v1
metadata:
  name: vpc1
spec:
  namespaces:
  - vpc1
  enableExternal: true
  enableBfd: true # bfd
  #enableBfd: false
```

```
# cat 02-subnet.yml
apiVersion: kubeovn.io/v1
kind: Subnet
metadata:
  name: vpc1-subnet1
spec:
  cidrBlock: 192.168.0.0/24
  default: false
  disableGatewayCheck: false
  disableInterConnection: true
  enableEcmp: true # ecmp
  gatewayNode: ""
  gatewayType: distributed
  #gatewayType: centralized
  natOutgoing: false
  private: false
  protocol: IPv4
  provider: ovn
  vpc: vpc1
  namespaces:
    - vpc1
```

:

1. vpc ecmp ecmp bfd vpc enableBfd subnet enableEcmp ecmp bfd
- 2.
3. VPC VPC VPC snat
4. vpc subnet enableEcmp gatewayType
5. EnableExternal vpc
6. EnableExternal EnableBfd

```
#       ovn
#   vpc
# k get vpc
NAME      ENABLEEXTERNAL  ENABLEBFD  STANDBY  SUBNETS          NAMESPACES
ovn-cluster  true        true      true     ["external204","join","ovn-default"]
vpc1       true        true      true     ["vpc1-subnet1"]      ["vpc1"]

#   vpc  ENABLEBFD
#   vpc

# 1.   bfd
# k ko nbctl list bfd
_uuid           : be7df545-2c4c-4751-878f-b3507987f050
detect_mult    : 3
dst_ip         : "10.5.204.121"
external_ids   : {}
logical_port   : vpc1-external204
min_rx         : 100
min_tx         : 100
options         : {}
status          : up

_uuid           : 684c4489-5b59-4693-8d8c-3beab93f8093
detect_mult    : 3
dst_ip         : "10.5.204.109"
external_ids   : {}
logical_port   : vpc1-external204
min_rx         : 100
min_tx         : 100
options         : {}
status          : up

_uuid           : f0f62077-2ae9-4e79-b4f8-a446ec6e784c
detect_mult    : 3
dst_ip         : "10.5.204.108"
external_ids   : {}
logical_port   : vpc1-external204
min_rx         : 100
min_tx         : 100
options         : {}
status          : up

###   status      up

# 2.   bfd
# k ko nbctl lr-route-list vpc1
IPv4 Routes
Route Table <main>:
  192.168.0.0/24      10.5.204.108 src-ip ecmp ecmp-symmetric-reply bfd
  192.168.0.0/24      10.5.204.109 src-ip ecmp ecmp-symmetric-reply bfd
  192.168.0.0/24      10.5.204.121 src-ip ecmp ecmp-symmetric-reply bfd
```

```

# 3.

# kubectl find Logical_Router_Static_Route policy=src-ip options=ecmp_symmetric_reply="true"
_uuid          : 3aacb384-d5ee-4b14-aebf-59e8c11717ba
bfd            : 684c4489-5b59-4693-8d8c-3beab93f8093
external_ids   : {}
ip_prefix      : "192.168.0.0/24"
nexthop        : "10.5.204.109"
options        : {ecmp_symmetric_reply="true"}
output_port    : []
policy         : src-ip
route_table    : ""

_uuid          : 18bcc585-bc05-430b-925b-ef673c8e1aef
bfd            : f0f62077-2ae9-4e79-b4f8-a446ec6e784c
external_ids   : {}
ip_prefix      : "192.168.0.0/24"
nexthop        : "10.5.204.108"
options        : {ecmp_symmetric_reply="true"}
output_port    : []
policy         : src-ip
route_table    : ""

_uuid          : 7d0a4e6b-cde0-4110-8176-fbaf19738498
bfd            : be7df545-2c4c-4751-878f-b3507987f050
external_ids   : {}
ip_prefix      : "192.168.0.0/24"
nexthop        : "10.5.204.121"
options        : {ecmp_symmetric_reply="true"}
output_port    : []
policy         : src-ip
route_table    : ""

# [root@pc-node-1 ~]# ip netns exec ovnnext bash ip a
/usr/sbin/ip: /usr/sbin/ip: cannot execute binary file
[root@pc-node-1 ~]#
[root@pc-node-1 ~]# ip netns exec ovnnext 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
1541: ovnnext0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1400 qdisc noqueue state UNKNOWN group default qlen 1000
    link/ether 00:00:00:ab:bd:87 brd ff:ff:ff:ff:ff:ff
    inet 10.5.204.108/24 brd 10.5.204.255 scope global ovnnext0
        valid_lft forever preferred_lft forever
    inet6 fe80::200:ff:feab:bd87/64 scope link
        valid_lft forever preferred_lft forever
[root@pc-node-1 ~]#
[root@pc-node-1 ~]# ip netns exec ovnnext route -n
Kernel IP routing table
Destination     Gateway         Genmask        Flags Metric Ref  Use Iface
0.0.0.0         10.5.204.254   0.0.0.0       UG    0      0      0 ovnnext0
10.5.204.0      0.0.0.0        255.255.255.0  U     0      0      0 ovnnext0

##           internal port unerlay     pod      ns

[root@pc-node-1 ~]# ip netns exec ovnnext bfdd-control status
There are 1 sessions:
Session 1
  id=1 local=10.5.204.108 (p) remote=10.5.204.122 state=Up

##     lrp bfd           lrp ecmp

[root@pc-node-1 ~]# ip netns exec ovnnext ping -c1 223.5.5.5
PING 223.5.5.5 (223.5.5.5) 56(84) bytes of data.
64 bytes from 223.5.5.5: icmp_seq=1 ttl=115 time=21.6 ms

#

```

ovnnext ns

```

# tcpdump -i ovnnext0 host 223.5.5.5 -netvv
dropped privs to tcpdump
tcpdump: listening on ovnnext0, link-type EN10MB (Ethernet), capture size 262144 bytes
^C
0 packets captured
0 packets received by filter
0 packets dropped by kernel
[root@pc-node-1 ~]# exit
[root@pc-node-1 ~]# ssh pc-node-2
Last login: Thu Feb 23 09:21:08 2023 from 10.5.32.51
[root@pc-node-2 ~]# ip netns exec ovnnext bash
[root@pc-node-2 ~]# tcpdump -i ovnnext0 host 223.5.5.5 -netvv
dropped privs to tcpdump
tcpdump: listening on ovnnext0, link-type EN10MB (Ethernet), capture size 262144 bytes
^C
0 packets captured
0 packets received by filter

```

```

0 packets dropped by kernel
[root@pc-node-2 ~]# exit
[root@pc-node-2 ~]# logout
Connection to pc-node-2 closed.
[root@pc-node-1 ~]# ssh pc-node-3
Last login: Thu Feb 23 08:32:41 2023 from 10.5.32.51
[root@pc-node-3 ~]# ip netns exec ovnnext bash
[root@pc-node-3 ~]# tcpdump -i ovnnext0 host 223.5.5.5 -netvv
dropped privs to tcpdump
tcpdump: listening on ovnnext0, link-type EN10MB (Ethernet), capture size 262144 bytes
00:00:00:2d:f8:ce > 00:00:00:fd:b2:a4, ethertype IPv4 (0x0800), length 98: (tos 0x0, ttl 63, id 57978, offset 0, flags [DF], proto ICMP (1), length 84)
    10.5.204.102 > 223.5.5.5: ICMP echo request, id 22, seq 71, length 64
00:00:00:fd:b2:a4 > dc:ef:80:5a:44:1a, ethertype IPv4 (0x0800), length 98: (tos 0x0, ttl 62, id 57978, offset 0, flags [DF], proto ICMP (1), length 84)
    10.5.204.102 > 223.5.5.5: ICMP echo request, id 22, seq 71, length 64
^C
2 packets captured
2 packets received by filter
0 packets dropped by kernel
[root@pc-node-3 ~]#
#       down      pod
#       3

```

5.8.3 3. bfd

vpc enable_eip_snat

```

# cat 01-vpc-ecmp-enable-external-bfd.yml
kind: Vpc
apiVersion: kubeovn.io/v1
metadata:
  name: vpc2
spec:
  namespaces:
  - vpc2
  enableExternal: true
  #enableBfd: true
  enableBfd: false

##   bfd

# k ko nbctl lr-route-list vpc2
IPv4 Routes
Route Table <main>:
          0.0.0.0/0           10.5.204.254 dst-ip

#
#   nbctl list bfd      lrp      bfd
#   ovnnext ns      bfd
#       vpc subnet     ping      ( )
#       ( )

```



⌚2025 9 10

⌚2023 3 3



5.8.4

5.9 VPC

VPC VPC VPC NAT

5.9.1

1. VPC
2. VPC CIDR
3. VPC VPC

5.9.2

VPC VPC Subnet Subnet CIDR

```
kind: Vpc
apiVersion: kubeovn.io/v1
metadata:
  name: vpc-1
spec: {}
---
kind: Subnet
apiVersion: kubeovn.io/v1
metadata:
  name: net1
spec:
  vpc: vpc-1
  cidrBlock: 10.0.0.0/16
---
kind: Vpc
apiVersion: kubeovn.io/v1
metadata:
  name: vpc-2
spec: {}
---
kind: Subnet
apiVersion: kubeovn.io/v1
metadata:
  name: net2
spec:
  vpc: vpc-2
  cidrBlock: 172.31.0.0/16
```

VPC vpcPeerings

```
kind: Vpc
apiVersion: kubeovn.io/v1
metadata:
  name: vpc-1
spec:
  vpcPeerings:
    - remoteVpc: vpc-2
      localConnectIP: 169.254.0.1/30
  staticRoutes:
    - cidr: 172.31.0.0/16
      nextHopIP: 169.254.0.2
      policy: policyDst
---
kind: Vpc
apiVersion: kubeovn.io/v1
metadata:
  name: vpc-2
spec:
  vpcPeerings:
    - remoteVpc: vpc-1
      localConnectIP: 169.254.0.2/30
  staticRoutes:
    - cidr: 10.0.0.0/16
      nextHopIP: 169.254.0.1
      policy: policyDst
```

- remoteVpc : VPC
- localConnectIP: IP CIDR IP CIDR
- cidr Subnet CIDR
- nextHopIP VPC localConnectIP

Subnet Pod

```
apiVersion: v1
kind: Pod
metadata:
  annotations:
    ovn.kubernetes.io/logical_switch: net1
  name: vpc-1-pod
spec:
  containers:
    - name: vpc-1-pod
      image: docker.io/library/nginx:alpine
---
apiVersion: v1
kind: Pod
metadata:
  annotations:
    ovn.kubernetes.io/logical_switch: net2
  name: vpc-2-pod
spec:
  containers:
    - name: vpc-2-pod
      image: docker.io/library/nginx:alpine
```

```
# kubectl exec -it vpc-1-pod -- ping $(kubectl get pod vpc-2-pod -o jsonpath='{.status.podIP}')
PING 172.31.0.2 (172.31.0.2): 56 data bytes
64 bytes from 172.31.0.2: seq=0 ttl=62 time=0.655 ms
64 bytes from 172.31.0.2: seq=1 ttl=62 time=0.886 ms
64 bytes from 172.31.0.2: seq=2 ttl=62 time=0.098 ms
^C
--- 172.31.0.2 ping statistics ---
3 packets transmitted, 3 packets received, 0% packet loss
round-trip min/avg/max = 0.086/0.279/0.655 ms
# kubectl exec -it vpc-2-pod -- ping $(kubectl get pod vpc-1-pod -o jsonpath='{.status.podIP}')
PING 10.0.0.2 (10.0.0.2): 56 data bytes
64 bytes from 10.0.0.2: seq=0 ttl=62 time=0.594 ms
64 bytes from 10.0.0.2: seq=1 ttl=62 time=0.093 ms
64 bytes from 10.0.0.2: seq=2 ttl=62 time=0.088 ms
^C
--- 10.0.0.2 ping statistics ---
3 packets transmitted, 3 packets received, 0% packet loss
round-trip min/avg/max = 0.088/0.258/0.594 ms
```

[PDF](#)

[Slack](#)

[Support](#)

⌚2025 7 2

⌚2022 5 24



5.9.3

6.

6.1 kubectl

Kube-OVN kubectl OVN OVN OVS tcpdump

6.1.1

Kube-OVN kubectl

kubectl-ko

```
wget https://raw.githubusercontent.com/kubeovn/kube-ovn/release-1.14/dist/images/kubectl-ko
```

\$PATH

```
mv kubectl-ko /usr/local/bin/kubectl-ko
```

```
chmod +x /usr/local/bin/kubectl-ko
```

```
# kubectl plugin list
The following compatible plugins are available:
/usr/local/bin/kubectl-ko
```

6.1.2

kubectl ko

```
# kubectl ko
kubectl ko {subcommand} [option...]
Available Subcommands:
[nb|sb] [status|kick|backup|dbstatus|restore]      ovn-db operations show cluster status, kick stale server, backup database, get db consistency status or
restore ovn nb db when met 'inconsistent data' error
nbctl [ovn-nbctl options ...]   invoke ovn-nbctl
sbctl [ovn-sbctl options ...]   invoke ovn-sbctl
vsctl {node_name} [ovs-vsctl options ...]   invoke ovs-vsctl on the specified node
ofctl {node_name} [ovs-ofctl options ...]   invoke ovs-ofctl on the specified node
dpctl {node_name} [ovs-dpctl options ...]   invoke ovs-dpctl on the specified node
appctl {node_name} [ovs-appctl options ...]   invoke ovs-appctl on the specified node
tcpdump {namespace/podname} [tcpdump options ...]   capture pod traffic
{trace|ovn-trace} ...   trace ovn microflow of specific packet
  {trace|ovn-trace} {namespace/podname} {target ip address} [target mac address] {icmp|tcp|udp} [target tcp/udp port]   trace ICMP/TCP/UDP
  {trace|ovn-trace} {namespace/podname} {target ip address} [target mac address] arp {request|reply}   trace ARP request/reply
  {trace|ovn-trace} {node//node_name} {target ip address} [target mac address] {icmp|tcp|udp} [target tcp/udp port]   trace ICMP/TCP/UDP
  {trace|ovn-trace} {node//node_name} {target ip address} [target mac address] arp {request|reply}   trace ARP request/reply
diagnose {all|node|subnet|IPPorts} [{node_name|subnet_name|proto1}-{IP1}-{Port1},{proto2}-{IP2}-{Port2}]   diagnose connectivity of all nodes or a specific
node or specify subnet's ds pods or IPPorts like 'tcp-172.18.0.2-53,udp-172.18.0.3-53'
env-check   check the environment configuration
reload   restart all kube-ovn components
log {kube-ovn|ovn|ovs|linux|all}   save log to ./kubectl-ko-log/
perf [image] performance test default image is docker.io/kubeovn/test:v1.13.0
icnbctl [ovn-nbctl options ...]   invoke ovn-ic-nbctl
icsbctl [ovn-sbctl options ...]   invoke ovn-ic-sbctl
```

[nb | sb] [status | kick | backup | dbstatus | restore]

OVN

OVN leader ovs-appctl cluster/status :

```
# kubectl ko nb status
306b
Name: OVN_Northbound
Cluster ID: 9a87 (9a872522-3e7d-47ca-83a3-d74333e1a7ca)
Server ID: 306b (306b256b-b5e1-4eb0-be91-4ca96adf6bad)
Address: tcp:[172.18.0.2]:6643
Status: cluster member
Role: leader
Term: 1
Leader: self
Vote: self

Last Election started 280309 ms ago, reason: timeout
Last Election won: 280309 ms ago
Election timer: 5000
Log: [139, 139]
Entries not yet committed: 0
Entries not yet applied: 0
Connections: <-8723 ->8723 <-85d6 ->85d6
Disconnections: 0
Servers:
  85d6 (85d6 at tcp:[172.18.0.4]:6643) next_index=139 match_index=138 last msg 763 ms ago
  8723 (8723 at tcp:[172.18.0.3]:6643) next_index=139 match_index=138 last msg 763 ms ago
  306b (306b at tcp:[172.18.0.2]:6643) (self) next_index=2 match_index=138
status: ok
```

Server match_index last msg Server

OVN 172.18.0.3 :

```
# kubectl ko nb kick 8723
started removal
```

```
# kubectl ko nb status
306b
Name: OVN_Northbound
Cluster ID: 9a87 (9a872522-3e7d-47ca-83a3-d74333e1a7ca)
Server ID: 306b (306b256b-b5e1-4eb0-be91-4ca96adf6bad)
Address: tcp:[172.18.0.2]:6643
Status: cluster member
Role: leader
Term: 1
Leader: self
Vote: self

Last Election started 324356 ms ago, reason: timeout
Last Election won: 324356 ms ago
Election timer: 5000
Log: [140, 140]
Entries not yet committed: 0
Entries not yet applied: 0
Connections: <-85d6 ->85d6
Disconnections: 2
Servers:
  85d6 (85d6 at tcp:[172.18.0.4]:6643) next_index=140 match_index=139 last msg 848 ms ago
  306b (306b at tcp:[172.18.0.2]:6643) (self) next_index=2 match_index=139
status: ok
```

OVN

```
# kubectl ko nb backup
tar: Removing leading '/' from member names
backup ovn-nb db to /root/ovnnb_db.060223191654183154.backup
```

```
# kubectl ko nb dbstatus
status: ok
```

inconsistent data

inconsistent data

```
# kubectl ko nb restore
deployment.apps/ovn-central scaled
ovn-central original replicas is 3
first nodeIP is 172.18.0.5
ovs-ovn pod on node 172.18.0.5 is ovs-ovn-8jxv9
ovs-ovn pod on node 172.18.0.3 is ovs-ovn-sjzb6
ovs-ovn pod on node 172.18.0.4 is ovs-ovn-t87zk
backup nb db file
restore nb db file, operate in pod ovs-ovn-8jxv9
deployment.apps/ovn-central scaled
finish restore nb db file and ovn-central replicas
recreate ovs-ovn pods
pod "ovs-ovn-8jxv9" deleted
pod "ovs-ovn-sjzb6" deleted
pod "ovs-ovn-t87zk" deleted
```

[nbctl | sbctl] [options ...]

OVN

leader

ovn-nbctl ovn-sbctl

OVN

ovn-nbctl(8) ovn-sbctl(8)

```
# kubectl ko nbctl show
switch c7cd17e8-ceee-4a91-9bb3-e5a313fe1ece (snat)
  port snat-ovn-cluster
    type: router
    router-port: ovn-cluster-snat
switch 20e0c6d0-023a-4756-aec5-200e0c60f95d (join)
  port node-liumengxin-ovn3-192.168.137.178
    addresses: ["00:00:00:64:FF:A8 100.64.0.4"]
  port node-liumengxin-ovn1-192.168.137.176
    addresses: ["00:00:00:AF:98:62 100.64.0.2"]
  port node-liumengxin-ovn2-192.168.137.177
    addresses: ["00:00:00:D9:58:B8 100.64.0.3"]
  port join-ovn-cluster
    type: router
    router-port: ovn-cluster-join
switch 0191705c-f827-427b-9de3-3c3b7d971ba5 (central)
  port central-ovn-cluster
    type: router
    router-port: ovn-cluster-central
switch 2a45ff05-388d-4f85-9daf-e6fccd5833dc (ovn-default)
  port alertmanager-main-0.monitoring
    addresses: ["00:00:00:6C:DF:A3 10.16.0.19"]
  port kube-state-metrics-5d6885d89-4nf8h.monitoring
    addresses: ["00:00:00:6F:02:1C 10.16.0.15"]
  port fake-kubelet-67c55dfdf89-pv86k.kube-system
    addresses: ["00:00:00:5C:12:E8 10.16.19.177"]
  port ovn-default-ovn-cluster
    type: router
    router-port: ovn-cluster-ovn-default
router 212f73dd-d63d-4d72-864b-a537e9afbee1 (ovn-cluster)
  port ovn-cluster-snat
    mac: "00:00:00:7A:82:8F"
    networks: ["172.22.0.1/16"]
  port ovn-cluster-join
    mac: "00:00:00:F8:18:5A"
    networks: ["100.64.0.1/16"]
  port ovn-cluster-central
    mac: "00:00:00:4D:8C:F5"
    networks: ["192.168.0.1/16"]
  port ovn-cluster-ovn-default
    mac: "00:00:00:A3:F8:18"
    networks: ["10.16.0.1/16"]
```

vsctl {nodeName} [options ...]

nodeName ovs-ovn

ovs-vsctl

vswitchd

OVS

ovs-vsctl(8)

```
# kubectl ko vsctl kube-ovn-01 show
0d4c4675-c9cc-440a-8c1a-878e17f81b88
Bridge br-int
  fail_mode: secure
  datapath_type: system
  Port a2c1a8a8b83a_h
    Interface a2c1a8a8b83a_h
    Port "4fa5c4cbb1a5_h"
      Interface "4fa5c4cbb1a5_h"
  Port ovn-eef07d-0
    Interface ovn-eef07d-0
      type: stt
      options: {csum="true", key=flow, remote_ip="192.168.137.178"}
  Port ovn0
```

```

Interface ovn0
    type: internal
Port mirror0
    Interface mirror0
        type: internal
Port ovn-efa253-0
    Interface ovn-efa253-0
        type: stt
        options: {csum="true", key=flow, remote_ip="192.168.137.177"}
Port br-int
    Interface br-int
        type: internal
ovs_version: "2.17.2"

```

ofctl {nodeName} [options ...]

nodeName	ovs-ovn	ovs-ofctl	OpenFlow	OVS	ovs-ofctl(8)
----------	---------	-----------	----------	-----	--------------

```

# kubectl ko ofctl kube-ovn-01 dump-flows br-int
NXST_FLOW reply (xid=0x4): flags=[more]
cookie=0xcf3429e6, duration=671791.432s, table=0, n_packets=0, n_bytes=0, idle_age=65534, hard_age=65534, priority=100, in_port=2 actions=load:0x4->NXM_NX_REG13[],load:0x9->NXM_NX_REG11[],load:0xb->NXM_NX_REG12[],load:0x4->0XM_OF_METADATA[],load:0x1->NXM_NX_REG14[],resubmit(.8)
cookie=0xc91413c6, duration=671791.431s, table=0, n_packets=9978275, n_bytes=9978275, idle_age=0, hard_age=65534, priority=100, in_port=7 actions=load:0x1->NXM_NX_REG13[],load:0x9->NXM_NX_REG11[],load:0xb->NXM_NX_REG12[],load:0x4->0XM_OF_METADATA[],load:0x4->NXM_NX_REG14[],resubmit(.8)
cookie=0xf180459, duration=671791.431s, table=0, n_packets=17348582, n_bytes=2667811214, idle_age=0, hard_age=65534, priority=100, in_port=6317 actions=load:0xa->NXM_NX_REG13[],load:0x9->NXM_NX_REG11[],load:0xb->NXM_NX_REG12[],load:0x4->0XM_OF_METADATA[],load:0x9->NXM_NX_REG14[],resubmit(.8)
cookie=0x7806dd90, duration=671791.431s, table=0, n_packets=3235428, n_bytes=833821312, idle_age=0, hard_age=65534, priority=100, in_port=1 actions=load:0xd->NXM_NX_REG13[],load:0x9->NXM_NX_REG11[],load:0xb->NXM_NX_REG12[],load:0x4->0XM_OF_METADATA[],load:0x3->NXM_NX_REG14[],resubmit(.8)
...

```

dpctl {nodeName} [options ...]

nodeName	ovs-ovn	ovs-dpctl	OVS datapath	OVS	ovs-dpctl(8)
----------	---------	-----------	--------------	-----	--------------

```

# kubectl ko dpctl kube-ovn-01 show
system@ovs-system:
lookups: hit:35080505 missed:21983648 lost:7:
flows: 105
masks: hit:1970748791 total:22 hit/pkt:5.29
port 0: ovs-system (internal)
port 1: ovn0 (internal)
port 2: mirror0 (internal)
port 3: br-int (internal)
port 4: stt_sys_7471 (stt: packet_type=ptap)
port 5: eeb4d9e51b5d_h
port 6: a2c1a8a8b83a_h
port 7: 4fa5c4cbb1a5_h

```

appctl {nodeName} [options ...]

nodeName	ovs-ovn	ovs-appctl	daemon	OVS	ovs-appctl(8)
----------	---------	------------	--------	-----	---------------

```

# kubectl ko appctl kube-ovn-01 vlog/list
      console   syslog   file
      -----  -----
backtrace     OFF      ERR      INFO
bfd          OFF      ERR      INFO
bond         OFF      ERR      INFO
bridge        OFF      ERR      INFO
bundle        OFF      ERR      INFO
bundles       OFF      ERR      INFO
...

```

tcpdump {namespace/podname} [tcpdump options ...]

namespace/podname	kube-ovn-cni	tcpdump	veth
-------------------	--------------	---------	------

```

# kubectl ko tcpdump default/ds1-l6n7p icmp
+ kubectl exec -it kube-ovn-cni-wlg4s -n kube-ovn -- tcpdump -nn -i d7176fe7b4e0_h icmp
tcpdump: verbose output suppressed, use -v or -vv for full protocol decode
listening on d7176fe7b4e0_h, link-type EN10MB (Ethernet), capture size 262144 bytes
06:52:36.619688 IP 10.64.0.3 > 10.16.0.4: ICMP echo request, id 2, seq 1, length 64
06:52:36.619746 IP 10.16.0.4 > 100.64.0.3: ICMP echo reply, id 2, seq 1, length 64
06:52:37.619588 IP 100.64.0.3 > 10.16.0.4: ICMP echo request, id 2, seq 2, length 64
06:52:37.619630 IP 10.16.0.4 > 100.64.0.3: ICMP echo reply, id 2, seq 2, length 64
06:52:38.619933 IP 100.64.0.3 > 10.16.0.4: ICMP echo request, id 2, seq 3, length 64
06:52:38.619973 IP 10.16.0.4 > 100.64.0.3: ICMP echo reply, id 2, seq 3, length 64

```

trace [arguments ...]

Pod	OVN	Openflow
-----	-----	----------

```
kubectl ko trace {namespace/podname} {target ip address} [target mac address] {icmp|tcp|udp} [target tcp/udp port]
kubectl ko trace {namespace/podname} {target ip address} [target mac address] arp {request|reply}
kubectl ko trace {node/nodename} {target ip address} [target mac address] {icmp|tcp|udp} [target tcp/udp port]
kubectl ko trace {node/nodename} {target ip address} [target mac address] arp {request|reply}
```

```
# kubectl ko trace default/ds1-16n7p 8.8.8.8 icmp
+ kubectl exec ovn-central-5bc494cb5-n kube-ovn -- ovn-trace --ct=new ovn-default 'inport == "ds1-16n7p.default" && ip.ttl == 64 && icmp && eth.src == 0a:00:00:10:00:05 && ip4.src == 10.16.0.4 && eth.dst == 00:00:00:B8:CA:43 && ip4.dst == 8.8.8.8'
# icmp,reg14=0xf,vlan_tci=0x0000,d1_src=0a:00:00:10:00:05,d1_dst=00:00:00:B8:CA:43,nw_src=10.16.0.4,nw_dst=8.8.8.8,nw_tos=0,nw_ecn=0,nw_ttl=64,icmp_type=0,icmp_code=0

ingress(dp="ovn-default", inport="ds1-16n7p.default")
-----
0. ls_in_port_sec_l2 (ovn-northd.c:4143): inport == "ds1-16n7p.default" && eth.src == {0a:00:00:10:00:05}, priority 50, uuid 39453393
    next;
1. ls_in_port_sec_ip (ovn-northd.c:2898): inport == "ds1-16n7p.default" && eth.src == 0a:00:00:10:00:05 && ip4.src == {10.16.0.4}, priority 90, uuid 81bcd485
    next;
3. ls_in_pre_acl (ovn-northd.c:3269): ip, priority 100, uuid 7b4f4971
    reg0[0] = 1;
    next;
5. ls_in_pre_stateful (ovn-northd.c:3396): reg0[0] == 1, priority 100, uuid 36cd577
    ct_next;

ct_next(ct_state=new|trk)
-----
6. ls_in_acl (ovn-northd.c:3759): ip && (!ct.est || (ct.est && ct_label.blocked == 1)), priority 1, uuid 7608af5b
    reg0[1] = 1;
    next;
10. ls_in_stateful (ovn-northd.c:3995): reg0[1] == 1, priority 100, uuid 2aba1b90
    ct_commit(ct_label=0x01);
    next;
16. ls_in_l2_1kup (ovn-northd.c:4470): eth.dst == 00:00:00:B8:CA:43, priority 50, uuid 5c9c3c9f
    output = "ovn-default-ovn-cluster";
    output;
...
...
```

trace	Underlay	Mac
-------	----------	-----

```
kubectl ko trace default/virt-handler-7lvm1 8.8.8.8 82:7c:9f:83:8c:01 icmp
```

diagnose {all|node|subnet|IPPorts} [nodename|subnetName|{proto1}-{IP1}-{Port1},{proto2}-{IP2}-{Port2}]

kube-ovn-pinger

```
# kubectl ko diagnose all
switch c7cd17e8-ceee-4a91-9bb3-e5a313fe1ece (snat)
  port snat-ovn-cluster
    type: router
    router-port: ovn-cluster-snat
switch 20e0c6d0-023a-4756-aec5-200e0c60f95d (join)
  port node-liumengxin-ovn3-192.168.137.178
    addresses: ["00:00:00:64:FF:A8 100.64.0.4"]
  port node-liumengxin-ovn1-192.168.137.176
    addresses: ["00:00:00:AF:98:62 100.64.0.2"]
  port join-ovn-cluster
    type: router
    router-port: ovn-cluster-join
switch 0191705c-f827-427b-9de3-3c3b7d971ba5 (central)
  port central-ovn-cluster
    type: router
    router-port: ovn-cluster-central
switch 2a45ff05-388d-4f85-9daf-e6fccd5833dc (ovn-default)
  port ovn-default-ovn-cluster
    type: router
    router-port: ovn-cluster-ovn-default
  port prometheus-k8s-1.monitoring
    addresses: ["00:00:00:AA:37:DF 10.16.0.23"]
router 212f73dd-d63d-4d72-864b-a537e9afbee1 (ovn-cluster)
  port ovn-cluster-snats
    mac: "00:00:00:7A:82:8F"
    networks: ["172.22.0.1/16"]
  port ovn-cluster-join
    mac: "00:00:00:F8:18:5A"
    networks: ["100.64.0.1/16"]
  port ovn-cluster-central
```

```

mac: "00:00:00:4D:8C:F5"
networks: ["192.101.0.1/16"]
port ovn-cluster-ovn-default
  mac: "00:00:00:A3:F8:18"
  networks: ["10.16.0.1/16"]
Routing Policies
  31000 ip4.dst == 10.16.0.0/16 allow
  31000 ip4.dst == 100.64.0.0/16 allow
  30000 ip4.dst == 192.168.137.177 reroute 100.64.0.3
  30000 ip4.dst == 192.168.137.178 reroute 100.64.0.4
  29000 ip4.src == $ovn.default.fake.6_ip4 reroute 100.64.0.22
  29000 ip4.src == $ovn.default.fake.7_ip4 reroute 100.64.0.21
  29000 ip4.src == $ovn.default.fake.8_ip4 reroute 100.64.0.23
  29000 ip4.src == $ovn.default.liumengxin.ovn3.192.168.137.178_ip4 reroute 100.64.0.4
  20000 ip4.src == $ovn.default.liumengxin.ovn1.192.168.137.176_ip4 & ip4.dst != $ovn.cluster.overlay.subnets.IPv4 reroute 100.
64.0.2
  20000 ip4.src == $ovn.default.liumengxin.ovn2.192.168.137.177_ip4 & ip4.dst != $ovn.cluster.overlay.subnets.IPv4 reroute 100.
64.0.3
  20000 ip4.src == $ovn.default.liumengxin.ovn3.192.168.137.178_ip4 & ip4.dst != $ovn.cluster.overlay.subnets.IPv4 reroute 100.
64.0.4
IPv4 Routes
Route Table <main>:
  0.0.0.0/0      100.64.0.1 dst-ip
UUID          LB           PROTO      VIP          IPs
e9bcfd9d-793e-4431-9073-6dec96b75d71  cluster-tcp-load  tcp        10.100.209.132:10660  192.168.137.176:10660
                                         tcp        10.101.239.192:6641   192.168.137.177:6641
                                         tcp        10.101.240.101:3000   10.16.0.7:3000
                                         tcp        10.103.184.186:6642   192.168.137.177:6642
35d2b7a5-e3a7-485a-a4b7-b4970eb0e63b  cluster-tcp-sess  tcp        10.100.158.128:8080  10.16.0.10:8080,10.16.0.5:8080,10.16.63.30:8080
                                         tcp        10.107.26.215:8080  10.16.0.19:8080,10.16.0.20:8080,10.16.0.21:8080
                                         tcp        10.107.26.215:9093  10.16.0.19:9093,10.16.0.20:9093,10.16.0.21:9093
                                         tcp        10.98.187.99:8080  10.16.0.22:8080,10.16.0.23:8080
                                         tcp        10.98.187.99:9090  10.16.0.22:9090,10.16.0.23:9090
f43303e4-89aa-4d3e-a3dc-278a552fe27b  cluster-udp-load  udp        10.96.0.10:53    10.16.0.4:53,10.16.0.9:53
_uuid          : 06776304-5a96-43ed-90c4-c4854c251699
addresses       : []
external_ids     : {vendor=kube-ovn}
name            : node_liumengxin_ovn2_192.168.137.177_underlay_v6

_uuid          : 62690625-87d5-491c-8675-9fd83b1f433c
addresses       : []
external_ids     : {vendor=kube-ovn}
name            : node_liumengxin_ovn1_192.168.137.176_underlay_v6

_uuid          : b03a9bae-94d5-4562-b34c-b5f6198e180b
addresses       : ["10.16.0.0/16", "100.64.0.0/16", "172.22.0.0/16", "192.101.0.0/16"]
external_ids     : {vendor=kube-ovn}
name            : ovn.cluster.overlay.subnets.IPv4

_uuid          : e1056f3a-24cc-4666-8a91-75ee6c3c2426
addresses       : []
external_ids     : {vendor=kube-ovn}
name            : ovn.cluster.overlay.subnets.IPv6

_uuid          : 3e5d5fff-e670-47b2-a2f5-a39f4698a8c5
addresses       : []
external_ids     : {vendor=kube-ovn}
name            : node_liumengxin_ovn3_192.168.137.178_underlay_v6
_uuid          : 2d85dbdc-d0db-4abe-b19e-cc886d32b492
action          : drop
direction        : from-lport
external_ids     : {}
label            : 0
log              : false
match            : "inport==@ovn.sg.kubeovn_deny_all && ip"
meter            : []
name             : []
options          : {}
priority         : 2003
severity         : []

_uuid          : de790cc8-f155-405f-bb32-5a51f30c545f
action          : drop
direction        : to-lport
external_ids     : {}
label            : 0
log              : false
match            : "outport==@ovn.sg.kubeovn_deny_all && ip"
meter            : []
name             : []
options          : {}
priority         : 2003
severity         : []

Chassis "e15ed4d4-1780-4d50-b09e-ea8372ed48b8"
  hostname: liumengxin-ovn1-192.168.137.176
  Encap stt
    ip: "192.168.137.176"
    options: {csum="true"}
  Port_Binding node-liumengxin-ovn1-192.168.137.176
  Port_Binding perf-6vxkn.default
  Port_Binding kube-state-metrics-5d6885d89-4nf8h.monitoring
  Port_Binding alertmanager-main-0.monitoring
  Port_Binding kube-ovn-pinger-6ftdf.kube-system
  Port_Binding fake-kubelet-67c55dfd89-pv86k.kube-system

```

```

Port_Binding prometheus-k8s-0.monitoring
Chassis "eef07da1-f8ad-4775-b14d-bd6a3b4eb0d5"
  hostname: liumengxin-ovn3-192.168.137.178
  Encap stt
    ip: "192.168.137.178"
    options: {csum="true"}
  Port_Binding kube-ovn-pinger-7twb4.kube-system
  Port_Binding prometheus-adapter-86df476d87-r188g.monitoring
  Port_Binding prometheus-k8s-1.monitoring
  Port_Binding node-liumengxin-ovn3-192.168.137.178
  Port_Binding perf-ff475.default
  Port_Binding alertmanager-main-1.monitoring
  Port_Binding blackbox-exporter-676d976865-tvsjd.monitoring
Chassis "efa253c9-494d-4719-83ae-b48ab0f11c03"
  hostname: liumengxin-ovn2-192.168.137.177
  Encap stt
    ip: "192.168.137.177"
    options: {csum="true"}
  Port_Binding grafana-6c4c6b8fb7-pzd2c.monitoring
  Port_Binding node-liumengxin-ovn2-192.168.137.177
  Port_Binding alertmanager-main-2.monitoring
  Port_Binding coredns-6789c94dd8-9jqsz.kube-system
  Port_Binding coredns-6789c94dd8-25d4r.kube-system
  Port_Binding prometheus-operator-7bbc99fc8b-wgjm4.monitoring
  Port_Binding prometheus-adapter-86df476d87-gdxmc.monitoring
  Port_Binding perf-fjnws.default
  Port_Binding kube-ovn-pinger-vh2xg.kube-system
ds kube-proxy ready
kube-proxy ready
deployment ovn-central ready
deployment kube-ovn-controller ready
ds kube-ovn-cni ready
ds ovs-ovn ready
deployment coredns ready
ovn-nb leader check ok
ovn-sb leader check ok
ovn-northd leader check ok
### kube-ovn-controller recent log

### start to diagnose node liumengxin-ovn1-192.168.137.176
##### ovn-controller log:
2022-06-03T00:56:44.897Z|16722|inc_proc_eng|INFO|User triggered force recompute.
2022-06-03T01:06:44.912Z|16723|inc_proc_eng|INFO|User triggered force recompute.
2022-06-03T01:16:44.925Z|16724|inc_proc_eng|INFO|User triggered force recompute.
2022-06-03T01:26:44.936Z|16725|inc_proc_eng|INFO|User triggered force recompute.
2022-06-03T01:36:44.959Z|16726|inc_proc_eng|INFO|User triggered force recompute.
2022-06-03T01:46:44.974Z|16727|inc_proc_eng|INFO|User triggered force recompute.
2022-06-03T01:56:44.988Z|16728|inc_proc_eng|INFO|User triggered force recompute.
2022-06-03T02:06:45.001Z|16729|inc_proc_eng|INFO|User triggered force recompute.
2022-06-03T02:16:45.025Z|16730|inc_proc_eng|INFO|User triggered force recompute.
2022-06-03T02:26:45.040Z|16731|inc_proc_eng|INFO|User triggered force recompute.

##### ovs-vsctl log:
2022-06-02T23:03:00.137Z|00079|dpif(handler)|WARN|system@ovs-system: execute ct(commit,zone=14,label=0/0x1,nat(src)),8 failed (Invalid argument) on packet
  icmp,vlan_tci=0x0000,dl_src=00:00:00:f8:07:c8,dl_dst=00:00:00:fa:1e:50,nw_src=10.16.0.5,nw_dst=10.16.0.10,nw_tos=0,nw_ecn=0,nw_ttl=64,icmp_type=8,icmp_code=0
  icmp_csum=f0d1
  with metadata skb_priority(0),tunnel(tun_id=0x160017000004,src=192.168.137.177,dst=192.168.137.176,ttl=64,tp_src=38881,tp_dst=7471,flags(csum|key)),skb_mark(0),ct_state(0x21),ct_zone(0xe),ct_tuple4(src=10.16.0.5,dst=10.16.0.10,proto=1,tp_src=8,tp_dst=0),in_port(4) mtu 0
  2022-06-02T23:23:31.840Z|00080|dpif(handler)|WARN|system@ovs-system: execute ct(commit,zone=14,label=0/0x1,nat(src)),8 failed (Invalid argument) on packet
  icmp,vlan_tci=0x0000,dl_src=00:00:00:f8:07:c8,dl_dst=00:00:00:fa:1e:50,nw_src=10.16.0.5,nw_dst=10.16.0.10,nw_tos=0,nw_ecn=0,nw_ttl=64,icmp_type=8,icmp_code=0
  icmp_csum=15b2
  with metadata skb_priority(0),tunnel(tun_id=0x160017000004,src=192.168.137.177,dst=192.168.137.176,ttl=64,tp_src=38881,tp_dst=7471,flags(csum|key)),skb_mark(0),ct_state(0x21),ct_zone(0xe),ct_tuple4(src=10.16.0.5,dst=10.16.0.10,proto=1,tp_src=8,tp_dst=0),in_port(4) mtu 0
  2022-06-03T00:09:15.659Z|00081|dpif(handler)|WARN|system@ovs-system: execute ct(commit,zone=14,label=0/0x1,nat(src)),8 failed (Invalid argument) on packet
  icmp,vlan_tci=0x0000,dl_src=00:00:00:dc:a3:63,dl_dst=00:00:00:fa:1e:50,nw_src=10.16.63.30,nw_dst=10.16.0.10,nw_tos=0,nw_ecn=0,nw_ttl=64,icmp_type=8,icmp_code=0
  16.0.10,nw_tos=0,nw_ecn=0,nw_ttl=64,icmp_type=8,icmp_code=0 icmp_csum=e5a5
  with metadata skb_priority(0),tunnel(tun_id=0x150017000004,src=192.168.137.178,dst=192.168.137.176,ttl=64,tp_src=9239,tp_dst=7471,flags(csum|key)),skb_mark(0),ct_state(0x21),ct_zone(0xe),ct_tuple4(src=10.16.0.5,dst=10.16.0.10,proto=1,tp_src=8,tp_dst=0),in_port(4) mtu 0
  2022-06-03T00:30:13.409Z|00064|dpif(handler2)|WARN|system@ovs-system: execute ct(commit,zone=14,label=0/0x1,nat(src)),8 failed (Invalid argument) on packet
  icmp,vlan_tci=0x0000,dl_src=00:00:00:f8:07:c8,dl_dst=00:00:00:fa:1e:50,nw_src=10.16.0.5,nw_dst=10.16.0.10,nw_tos=0,nw_ecn=0,nw_ttl=64,icmp_type=8,icmp_code=0
  icmp_csum=6b4a
  with metadata skb_priority(0),tunnel(tun_id=0x160017000004,src=192.168.137.177,dst=192.168.137.176,ttl=64,tp_src=38881,tp_dst=7471,flags(csum|key)),skb_mark(0),ct_state(0x21),ct_zone(0xe),ct_tuple4(src=10.16.0.5,dst=10.16.0.10,proto=1,tp_src=8,tp_dst=0),in_port(4) mtu 0
  2022-06-03T02:03:23.832Z|00082|dpif(handler)|WARN|system@ovs-system: execute ct(commit,zone=14,label=0/0x1,nat(src)),8 failed (Invalid argument) on packet
  icmp,vlan_tci=0x0000,dl_src=00:00:00:f8:07:c8,dl_dst=00:00:00:fa:1e:50,nw_src=10.16.0.5,nw_dst=10.16.0.10,nw_tos=0,nw_ecn=0,nw_ttl=64,icmp_type=8,icmp_code=0
  icmp_csum=a819
  with metadata skb_priority(0),tunnel(tun_id=0x160017000004,src=192.168.137.177,dst=192.168.137.176,ttl=64,tp_src=38881,tp_dst=7471,flags(csum|key)),skb_mark(0),ct_state(0x21),ct_zone(0xe),ct_tuple4(src=10.16.0.5,dst=10.16.0.10,proto=1,tp_src=8,tp_dst=0),in_port(4) mtu 0

##### ovs-vsctl show results:
0d4c4675-c9cc-448a-8c1a-878e17f81b88
Bridge br-int
  fail_mode: secure
  datapath_type: system
  Port a2c1a8a8b83a_h
    Interface a2c1a8a8b83a_h
      "4fa5c4ccb1a5_h"
    Interface "4fa5c4ccb1a5_h"
  Port ovn-eef07d-0
    Interface ovn-eef07d-0
      type: stt
      options: {csum="true", key=flow, remote_ip="192.168.137.178"}
  Port ovn0
    Interface ovn0

```

```

        type: internal
Port "04d03360e9a0_h"
    Interface "04d03360e9a0_h"
Port eeb4d9e51b5d_h
    Interface eeb4d9e51b5d_h
Port mirror0
    Interface mirror0
        type: internal
Port "8e5d887ccdb0_h"
    Interface "8e5d887ccdb0_h"
Port ovn-efa253-0
    Interface ovn-efa253-0
        type: stt
        options: {csum="true", key=flow, remote_ip="192.168.137.177"}
Port "17512d5be1f1_h"
    Interface "17512d5be1f1_h"
Port br-int
    Interface br-int
        type: internal
ovs_version: "2.17.2"

#### pinger diagnose results:
I0603 10:35:04.349404 17619 pinger.go:19]

Kube-OVN:
Version: v1.14.23
Build: 2022-04-24_08:02:50
Commit: git-73f9d15
Go Version: go1.17.8
Arch: amd64

I0603 10:35:04.376797 17619 config.go:166] pinger config is &{KubeConfigFile: KubeClient:0xc000493380 Port:8080 DaemonSetNameSpace:kube-system DaemonSetName:kube-ovn-pinger Interval:5 Mode:job ExitCode:0 InternalDNS:kubernetes.default ExternalDNS: NodeName:liumengxin-ovn1-192.168.137.176 HostIP:192.168.137.176 PodName:kube-ovn-pinger-6ftdf PodIP:10.16.0.10 PodProtocols:[IPv4] ExternalAddress: NetworkMode:kube-ovn PollTimeout:2 PollInterval:15 SystemRunDir:/var/run/openvswitch DatabaseVswitchName:Open_vSwitch DatabaseVswitchSocketRemote:unix:/var/run/openvswitch/db.sock DatabaseVswitchFilePath:/etc/openvswitch/conf.db DatabaseVswitchFileLogPath:/var/log/openvswitch/ovsdb-server.log DatabaseVswitchFilePidPath:/var/run/openvswitch/ovsdb-server.pid DatabaseVswitchFileSystemIDPath:/etc/openvswitch/system-id.conf ServiceVswitchchFileLogPath:/var/log/openvswitch/ovs-vswitchd.log ServiceVswitchchFilePidPath:/var/run/openvswitch/ovs-vswitchd.pid ServiceOvnControllerFileLogPath:/var/log/ovn/ovn-controller.log ServiceOvnControllerFilePidPath:/var/run/ovn-controller.pid}
I0603 10:35:04.449166 17619 exporter.go:75] liumengxin-ovn1-192.168.137.176: exporter connect successfully
I0603 10:35:04.554011 17619 ovn.go:21] ovs-vswitchd and ovsvdb are up
I0603 10:35:04.651293 17619 ovn.go:33] ovn_controller is up
I0603 10:35:04.651342 17619 ovn.go:39] start to check port binding
I0603 10:35:04.749613 17619 ovn.go:135] chassis id is 1d7f3d6c-eec5-4b3c-adca-2969d9cdfd80
I0603 10:35:04.763487 17619 ovn.go:49] port in sb is [node-liumengxin-ovn1-192.168.137.176 perf-6vxkn.default kube-state-metrics-5d6805d89-4nf8h.monitoring alertmanager-main-0.monitoring kube-ovn-pinger-6ftdf.kube-system fake-kubelet-67c55dfd89-pv86k.kube-system prometheus-k8s-0.monitoring]
I0603 10:35:04.763583 17619 ovn.go:61] ovs and ovn-sb binding check passed
I0603 10:35:05.049309 17619 ping.go:259] start to check apiserver connectivity
I0603 10:35:05.053666 17619 ping.go:268] connect to apiserver success in 4.27ms
I0603 10:35:05.053786 17619 ping.go:129] start to check pod connectivity
I0603 10:35:05.249590 17619 ping.go:159] ping pod: kube-ovn-pinger-6ftdf 10.16.0.10, count: 3, loss count 0, average rtt 16.30ms
I0603 10:35:05.354135 17619 ping.go:159] ping pod: kube-ovn-pinger-7wb4 10.16.63.30, count: 3, loss count 0, average rtt 1.81ms
I0603 10:35:05.458460 17619 ping.go:159] ping pod: kube-ovn-pinger-vh2xg 10.16.0.5, count: 3, loss count 0, average rtt 1.92ms
I0603 10:35:05.458523 17619 ping.go:83] start to check node connectivity

```

diagnose	subnet	subnet	daemonset	kube-ovn-pinger	daemonset	pod	daemonset
diagnose	IPPorts			kube-ovn-pinger		IP Port	

reload

Kube-OVN

```

# kubectl ko reload
pod "ovn-central-8684dd94bd-vzgr" deleted
Waiting for deployment "ovn-central" rollout to finish: 0 of 1 updated replicas are available...
deployment "ovn-central" successfully rolled out
pod "ovs-ovn-bsnvz" deleted
pod "ovs-ovn-m9b98" deleted
pod "kube-ovn-controller-8459db5ff4-64c62" deleted
Waiting for deployment "kube-ovn-controller" rollout to finish: 0 of 1 updated replicas are available...
deployment "kube-ovn-controller" successfully rolled out
pod "kube-ovn-cni-2klhn" deleted
pod "kube-ovn-cni-t2jz4" deleted
Waiting for daemon set "kube-ovn-cni" rollout to finish: 0 of 2 updated pods are available...
Waiting for daemon set "kube-ovn-cni" rollout to finish: 1 of 2 updated pods are available...
daemon set "kube-ovn-cni" successfully rolled out
pod "kube-ovn-pinger-ln72z" deleted
pod "kube-ovn-pinger-w8lrk" deleted
Waiting for daemon set "kube-ovn-pinger" rollout to finish: 0 of 2 updated pods are available...
Waiting for daemon set "kube-ovn-pinger" rollout to finish: 1 of 2 updated pods are available...
daemon set "kube-ovn-pinger" successfully rolled out
pod "kube-ovn-monitor-7fb67d5488-7q6zb" deleted
Waiting for deployment "kube-ovn-monitor" rollout to finish: 0 of 1 updated replicas are available...
deployment "kube-ovn-monitor" successfully rolled out

```

log

kube-ovn Kube-OVN OVN Open vSwitch log linux debug

```
# kubectl ko log all
Collecting kube-ovn logging files
Collecting ovn logging files
Collecting openvswitch logging files
Collecting linux dmesg files
Collecting linux iptables-legacy files
Collecting linux iptables-nft files
Collecting linux route files
Collecting linux link files
Collecting linux neigh files
Collecting linux memory files
Collecting linux top files
Collecting linux sysctl files
Collecting linux netstat files
Collecting linux addr files
Collecting linux ipset files
Collecting linux tcp files
Collected files have been saved in the directory /root/kubectl-ko-log
```

```
# tree kubectl-ko-log/
kubectl-ko-log/
|-- kube-ovn-control-plane
|   |-- kube-ovn
|   |   |-- kube-ovn-cni.log
|   |   |-- kube-ovn-monitor.log
|   |   '-- kube-ovn-pinger.log
|   '-- linux
|       |-- addr.log
|       |-- dmesg.log
|       |-- ipset.log
|       |-- iptables-legacy.log
|       |-- iptables-nft.log
|       |-- link.log
|       |-- memory.log
|       |-- neigh.log
|       |-- netstat.log
|       |-- route.log
|       |-- sysctl.log
|       |-- tcp.log
|       '-- top.log
|   '-- openvswitch
|       |-- ovs-vswitchd.log
|       '-- ovsdb-server.log
|   '-- ovn
|       |-- ovn-controller.log
|       |-- ovn-northd.log
|       '-- ovsdb-server-nb.log
|           '-- ovsdb-server-sb.log
```

perf [image]

Kube-OVN

- 1.
2. Hostnetwork
- 3.
4. OVN-NB, OVN-SB, OVN-Northd leader

image Pod kubeovn/test:v1.12.0

```
# kubectl ko perf
===== Preparing Performance Test Resources =====
pod/test-client created
pod/test-host-client created
pod/test-server created
pod/test-host-server created
service/test-server created
pod/test-client condition met
pod/test-host-client condition met
pod/test-host-server condition met
pod/test-server condition met
=====
===== Start Pod Network Unicast Performance Test =====
Size      TCP Latency    TCP Bandwidth   UDP Latency     UDP Lost Rate   UDP Bandwidth
64        82.8 us        97.7 Mbits/sec  67.6 us        (0%)          8.42 Mbits/sec
128       85.4 us        167 Mbits/sec   67.2 us        (0%)          17.2 Mbits/sec
```

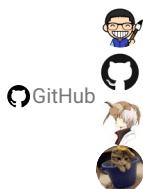
```

512      85.8 us    440 Mbits/sec  68.7 us      (0%)      68.4 Mbits/sec
1k       85.1 us    567 Mbits/sec  68.7 us      (0%)      134 Mbits/sec
4k       138 us     826 Mbits/sec  78.1 us     (1.4%)     503 Mbits/sec
=====
===== Start Host Network Performance Test =====
Size      TCP Latency   TCP Bandwidth   UDP Latency   UDP Lost Rate   UDP Bandwidth
64        49.7 us      120 Mbits/sec  37.9 us      (0%)      18.6 Mbits/sec
128       49.7 us      200 Mbits/sec  38.1 us      (0%)      35.5 Mbits/sec
512       51.9 us      588 Mbits/sec  38.9 us      (0%)      142 Mbits/sec
1k        51.7 us      944 Mbits/sec  37.2 us      (0%)      279 Mbits/sec
4k        74.9 us      1.66 Gbits/sec 39.9 us      (0%)      1.20 Gbits/sec
=====
===== Start Service Network Performance Test =====
Size      TCP Latency   TCP Bandwidth   UDP Latency   UDP Lost Rate   UDP Bandwidth
64        111 us       96.3 Mbits/sec  88.4 us      (0%)      7.59 Mbits/sec
128       83.7 us      150 Mbits/sec  69.2 us      (0%)      16.9 Mbits/sec
512       87.4 us      374 Mbits/sec  75.8 us      (0%)      60.9 Mbits/sec
1k        88.2 us      521 Mbits/sec  73.1 us      (0%)      123 Mbits/sec
4k        148 us       813 Mbits/sec  77.6 us     (0.0044%)    451 Mbits/sec
=====
===== Start Pod Multicast Network Performance Test =====
Size      UDP Latency   UDP Lost Rate   UDP Bandwidth
64        0.014 ms     (0.17%)      5.80 Mbits/sec
128      0.012 ms     (0%)         11.4 Mbits/sec
512      0.016 ms     (0%)         46.1 Mbits/sec
1k       0.023 ms     (0.073%)    89.8 Mbits/sec
4k       0.035 ms     (1.3%)       126 Mbits/sec
=====
===== Start Host Multicast Network Performance =====
Size      UDP Latency   UDP Lost Rate   UDP Bandwidth
64        0.007 ms     (0%)         9.95 Mbits/sec
128      0.005 ms     (0%)         21.8 Mbits/sec
512      0.008 ms     (0%)         86.8 Mbits/sec
1k       0.013 ms     (0.045%)    168 Mbits/sec
4k       0.010 ms     (0.31%)      242 Mbits/sec
=====
===== Start Leader Recover Time Test =====
Delete ovn central nb pod
pod "ovn-central-5cb9c67d75-tlz9w" deleted
Waiting for ovn central nb pod running
===== OVN nb Recovery takes 3.305236803 s =====
Delete ovn central sb pod
pod "ovn-central-5cb9c67d75-szx4c" deleted
Waiting for ovn central sb pod running
===== OVN sb Recovery takes 3.462698535 s =====
Delete ovn central northd pod
pod "ovn-central-5cb9c67d75-zqmqv" deleted
Waiting for ovn central northd pod running
===== OVN northd Recovery takes 2.691291403 s =====
=====
===== Remove Performance Test Resource =====
rm -f unicast-test-client.log
rm -f unicast-test-host-client.log
rm -f unicast-test-client.log
kubectl ko nbctl lb-del test-server
rm -f multicast-test-server.log
kubectl exec ovs-ovn-gxdrf -n kube-system -- ip maddr del 01:00:5e:00:00:64 dev eth0
kubectl exec ovs-ovn-h57bf -n kube-system -- ip maddr del 01:00:5e:00:00:64 dev eth0
rm -f multicast-test-host-server.log
pod "test-client" deleted
pod "test-host-client" deleted
pod "test-host-server" deleted
pod "test-server" deleted
service "test-server" deleted
=====
```

[PDF](#)[Slack](#)[Support](#)

⌚2025 9 10

⌚2022 5 24



6.1.3

6.2



6.2.1

```

# kubectl drain kube-ovn-worker --ignore-daemonsets --force
node/kube-ovn-worker cordoned
WARNING: ignoring DaemonSet-managed Pods: kube-system/kube-ovn-cni-zt74b, kube-system/kube-ovn-pinger-5rxf, kube-system/kube-proxy-jpmnm, kube-system/ovs-ovn-v2k11
evicting pod kube-system/coredns-64897985d-qsgpt
evicting pod local-path-storage/local-path-provisioner-5ddd94ff66-llss6
evicting pod kube-system/kube-ovn-controller-8459db5ff4-941xb
pod/kube-ovn-controller-8459db5ff4-941xb evicted
pod/coredns-64897985d-qsgpt evicted
pod/local-path-provisioner-5ddd94ff66-llss6 evicted
node/kube-ovn-worker drained
  
```

6.2.2 kubelet docker

```

ovs-ovn          ovn-central

systemctl stop kubelet
systemctl stop docker

CRI containerd      ovs-ovn

crictl rm -f $(crictl ps | grep openvswitch | awk '{print $1}')
  
```

6.2.3 Node

```

rm -rf /var/run/openvswitch
rm -rf /var/run/ovn
rm -rf /etc/origin/openvswitch/
rm -rf /etc/origin/ovn/
rm -rf /etc/cni/net.d/00-kube-ovn.conflist
rm -rf /etc/cni/net.d/01-kube-ovn.conflist
rm -rf /var/log/openvswitch
rm -rf /var/log/ovn
  
```

6.2.4 kubectl

```

kubectl delete no kube-ovn-01
  
```

6.2.5 ovn-sb

```

kube-ovn-worker

# kubectl ko sbctl show
Chassis "b0564934-5a0d-4804-a4c0-476c93596a17"
  hostname: kube-ovn-worker
  Encap geneve
    ip: "172.18.0.2"
    options: {csum="true"}
  Port_Binding kube-ovn-pinger-5rxf.kube-system
Chassis "6a29de7e-d731-4eaf-bacd-2f239ee52b28"
  hostname: kube-ovn-control-plane
  Encap geneve
    ip: "172.18.0.3"
    options: {csum="true"}
  Port_Binding coredns-64897985d-nbfln.kube-system
  Port_Binding node-kube-ovn-control-plane
  Port_Binding local-path-provisioner-5ddd94ff66-h4tn9.local-path-storage
  Port_Binding kube-ovn-pinger-hf2p6.kube-system
  Port_Binding coredns-64897985d-fhwlw.kube-system
  
```

6.2.6 chassis

uuid Chassis id

```
# kubectl ko sbctl chassis-del b0564934-5a0d-4804-a4c0-476c93596a17
# kubectl ko sbctl show
Chassis "6a29de7e-d731-4eaf-bacd-2f239ee52b28"
  hostname: kube-ovn-control-plane
  Encap geneve
    ip: "172.18.0.3"
    options: {csum="true"}
  Port_Binding coredns-64897985d-nbf1n.kube-system
  Port_Binding node-kube-ovn-control-plane
  Port_Binding local-path-provisioner-5dd94ff66-h4tn9.local-path-storage
  Port_Binding kube-ovn-pinger-hf2p6.kube-system
  Port_Binding coredns-64897985d-fhw1w.kube-system
```

[PDF](#)[Slack](#)[Support](#)

⌚2023 4 25

⌚2022 6 3

GitHub 

6.2.7

6.3 ovn-central

ovn-central ovn-nb ovn-sb etcd Raft ovn-central

6.3.1 ovn-central

kube-ovn-control-plane2 ovn-central

```
# kubectl -n kube-system get pod -o wide | grep central
ovn-central-6bf58cbc97-2cdhg      1/1   Running   0          21m   172.18.0.3   kube-ovn-control-plane   <none>       <none>
ovn-central-6bf58cbc97-crmfp      1/1   Running   0          21m   172.18.0.5   kube-ovn-control-plane2  <none>       <none>
ovn-central-6bf58cbc97-lxmpl      1/1   Running   0          21m   172.18.0.4   kube-ovn-control-plane3  <none>       <none>
```

ovn-nb

ID

```
# kubectl ko nb status
1b9a
Name: OVN_Northbound
Cluster ID: 32ca (32ca07fb-739b-4257-b510-12fa18e7cce8)
Server ID: 1b9a (1b9a5d76-e69b-410c-8085-39943d0cd38c)
Address: tcp:[172.18.0.3]:6643
Status: cluster member
Role: leader
Term: 1
Leader: self
Vote: self

Last Election started 2135194 ms ago, reason: timeout
Last Election won: 2135188 ms ago
Election timer: 5000
Log: [135, 135]
Entries not yet committed: 0
Entries not yet applied: 0
Connections: <-d64b ->d64b <-4984 ->4984
Disconnections: 0
Servers:
  4984 (4984 at tcp:[172.18.0.4]:6643) next_index=135 match_index=134 last msg 1084 ms ago
  1b9a (1b9a at tcp:[172.18.0.3]:6643) (self) next_index=2 match_index=134
    d64b (d64b at tcp:[172.18.0.5]:6643) next_index=135 match_index=134 last msg 1084 ms ago
status: ok
```

kube-ovn-control-plane2 IP 172.18.0.5 ID d64b ovn-nb

```
# kubectl ko nb kick d64b
started removal
```

```
# kubectl ko nb status
1b9a
Name: OVN_Northbound
Cluster ID: 32ca (32ca07fb-739b-4257-b510-12fa18e7cce8)
Server ID: 1b9a (1b9a5d76-e69b-410c-8085-39943d0cd38c)
Address: tcp:[172.18.0.3]:6643
Status: cluster member
Role: leader
Term: 1
Leader: self
Vote: self

Last Election started 2297649 ms ago, reason: timeout
Last Election won: 2297643 ms ago
Election timer: 5000
Log: [136, 136]
Entries not yet committed: 0
Entries not yet applied: 0
Connections: <-4984 ->4984
Disconnections: 2
Servers:
  4984 (4984 at tcp:[172.18.0.4]:6643) next_index=136 match_index=135 last msg 1270 ms ago
  1b9a (1b9a at tcp:[172.18.0.3]:6643) (self) next_index=2 match_index=135
status: ok
```

ovn-sb

ovn-sb ID

```
kubectl ko sb status
3722
Name: OVN_Southbound
Cluster ID: d4bd (d4bd37a4-0400-499f-b4df-b4fd389780f0)
Server ID: 3722 (3722d5ae-2ced-4820-a6b2-8b744d11fb3e)
Address: tcp:[172.18.0.3]:6644
Status: cluster member
Role: leader
Term: 1
Leader: self
Vote: self

Last Election started 2395317 ms ago, reason: timeout
Last Election won: 2395316 ms ago
Election timer: 5000
Log: [130, 130]
Entries not yet committed: 0
Entries not yet applied: 0
Connections: <-e9f7 ->e9f7 <-6e84 ->6e84
Disconnections: 0
Servers:
  e9f7 (e9f7 at tcp:[172.18.0.5]:6644) next_index=130 match_index=129 last msg 1006 ms ago
  6e84 (6e84 at tcp:[172.18.0.4]:6644) next_index=130 match_index=129 last msg 1004 ms ago
  3722 (3722 at tcp:[172.18.0.3]:6644) (self) next_index=2 match_index=129
status: ok
```

kube-ovn-control-plane2 IP 172.18.0.5 ID e9f7 ovn-sb

```
# kubectl ko sb kick e9f7
started removal
```

```
# kubectl ko sb status
3722
Name: OVN_Southbound
Cluster ID: d4bd (d4bd37a4-0400-499f-b4df-b4fd389780f0)
Server ID: 3722 (3722d5ae-2ced-4820-a6b2-8b744d11fb3e)
Address: tcp:[172.18.0.3]:6644
Status: cluster member
Role: leader
Term: 1
Leader: self
Vote: self

Last Election started 2481636 ms ago, reason: timeout
Last Election won: 2481635 ms ago
Election timer: 5000
Log: [131, 131]
Entries not yet committed: 0
Entries not yet applied: 0
Connections: <-6e84 ->6e84
Disconnections: 2
Servers:
  6e84 (6e84 at tcp:[172.18.0.4]:6644) next_index=131 match_index=130 last msg 642 ms ago
  3722 (3722 at tcp:[172.18.0.3]:6644) (self) next_index=2 match_index=130
status: ok
```

ovn-central

ovn-central NODE_IPS

```
kubectl label node kube-ovn-control-plane2 kube-ovn/role-
kubectl scale deployment -n kube-system ovn-central --replicas=2
kubectl set env deployment/ovn-central -n kube-system NODE_IPS="172.18.0.3,172.18.0.4"
kubectl rollout status deployment/ovn-central -n kube-system
```

ovn-central

ovs-ovn

```
# kubectl set env daemonset/ovs-ovn -n kube-system OVN_DB_IPS="172.18.0.3,172.18.0.4"
daemonset.apps/ovs-ovn env updated
# kubectl delete pod -n kube-system -lapp=ovs
pod "ovs-ovn-4f6jc" deleted
```

```

pod "ovs-ovn-csn2w" deleted
pod "ovs-ovn-mpbmb" deleted

kube-ovn-controller

# kubectl set env deployment/kube-ovn-controller -n kube-system OVN_DB_IPS="172.18.0.3,172.18.0.4"
deployment.apps/kube-ovn-controller env updated

# kubectl rollout status deployment/kube-ovn-controller -n kube-system
Waiting for deployment "kube-ovn-controller" rollout to finish: 1 of 3 updated replicas are available...
Waiting for deployment "kube-ovn-controller" rollout to finish: 2 of 3 updated replicas are available...
deployment "kube-ovn-controller" successfully rolled out

```

kube-ovn-control-plane2

```
rm -rf /etc/origin/ovn
```

Kubernetes

6.3.2 ovn-central

Kubernetes ovn-central

```
/etc/origin/ovn ovnnb_db.db ovnsb_db.db
```

```
rm -rf /etc/origin/ovn
```

ovn-central

ovn-central

```

# kubectl ko nb status
1b9a
Name: OVN_Northbound
Cluster ID: 32ca (32ca07fb-739b-4257-b510-12fa18e7cce8)
Server ID: 1b9a (1b9a5d76-e69b-410c-8085-39943d0cd38c)
Address: tcp:[172.18.0.3]:6643
Status: cluster member
Role: leader
Term: 44
Leader: self
Vote: self

Last Election started 1855739 ms ago, reason: timeout
Last Election won: 1855729 ms ago
Election timer: 5000
Log: [147, 147]
Entries not yet committed: 0
Entries not yet applied: 0
Connections: ->4984 <-4984
Disconnections: 0
Servers:
  4984 (4984 at tcp:[172.18.0.4]:6643) next_index=147 match_index=146 last msg 367 ms ago
  1b9a (1b9a at tcp:[172.18.0.3]:6643) (self) next_index=140 match_index=146
status: ok

# kubectl ko sb status
3722
Name: OVN_Southbound
Cluster ID: d4bd (d4bd37a4-0400-499f-b4df-b4fd389780f0)
Server ID: 3722 (3722d5ae-2ced-4820-a6b2-8b744d11fb3e)
Address: tcp:[172.18.0.3]:6644
Status: cluster member
Role: leader
Term: 33
Leader: self
Vote: self

Last Election started 1868589 ms ago, reason: timeout
Last Election won: 1868579 ms ago
Election timer: 5000
Log: [142, 142]
Entries not yet committed: 0

```

```
Entries not yet applied: 0
Connections: ->6e84 <-6e84
Disconnections: 0
Servers:
  6e84 (6e84 at tcp:[172.18.0.4]:6644) next_index=142 match_index=141 last msg 728 ms ago
  3722 (3722 at tcp:[172.18.0.3]:6644) (self) next_index=134 match_index=141
status: ok
```

ovn-central **NODE_IPS**

```
kubectl label node kube-ovn-control-plane2 kube-ovn/role=master
kubectl scale deployment -n kube-system ovn-central --replicas=3
kubectl set env deployment/ovn-central -n kube-system NODE_IPS="172.18.0.3,172.18.0.4,172.18.0.5"
kubectl rollout status deployment/ovn-central -n kube-system
```

ovn-central**ovs-ovn**

```
# kubectl set env daemonset/ovs-ovn -n kube-system OVN_DB_IPS="172.18.0.3,172.18.0.4,172.18.0.5"
daemonset.apps/ovs-ovn env updated
# kubectl delete pod -n kube-system -lapp=ovs
pod "ovs-ovn-4f6jc" deleted
pod "ovs-ovn-csn2w" deleted
pod "ovs-ovn-mpbmb" deleted
```

kube-ovn-controller

```
# kubectl set env deployment/kube-ovn-controller -n kube-system OVN_DB_IPS="172.18.0.3,172.18.0.4,172.18.0.5"
deployment.apps/kube-ovn-controller env updated

# kubectl rollout status deployment/kube-ovn-controller -n kube-system
Waiting for deployment "kube-ovn-controller" rollout to finish: 1 of 3 updated replicas are available...
Waiting for deployment "kube-ovn-controller" rollout to finish: 2 of 3 updated replicas are available...
deployment "kube-ovn-controller" successfully rolled out
```

[PDF](#)[Slack](#)[Support](#)

⌚2025 9 10

⌚2022 5 24



6.3.3

6.4 OVN

6.4.1

```
kubectl backup
```

```
# kubectl ko nb backup
tar: Removing leading `/' from member names
backup ovn-nb db to /root/ovnnb_db.060223191654183154.backup

# kubectl ko sb backup
tar: Removing leading `/' from member names
backup ovn-sb db to /root/ovnsb_db.060223191654183154.backup
```

6.4.2

```
/var/log/ovn/ovn-northd.log
```

```
* ovn-northd is not running
ovsdb-server: ovsdb error: error reading record 2739 from OVN_Northbound log: record 2739 advances commit index to 6308 but last log index is 6307
* Starting ovsdb-nb
```

OVN_Northbound	OVN_Southbound	OVN_Northbound	ovn-nb
----------------	----------------	----------------	--------

```
# kubectl ko nb status
9182
Name: OVN_Northbound
Cluster ID: e75f (e75fa340-49ed-45ab-990e-26cb865ebc85)
Server ID: 9182 (9182e8dd-b5b0-4dd8-8518-598cc1e374f3)
Address: tcp:[10.0.128.61]:6643
Status: cluster member
Role: leader
Term: 1454
Leader: self
Vote: self

Last Election started 1732603 ms ago, reason: timeout
Last Election won: 1732587 ms ago
Election timer: 1000
Log: [7332, 12512]
Entries not yet committed: 1
Entries not yet applied: 1
Connections: ->f080 <-f080 <-e631 ->e631
Disconnections: 1
Servers:
  f080 (f080 at tcp:[10.0.129.139]:6643) next_index=12512 match_index=12510 last msg 63 ms ago
  9182 (9182 at tcp:[10.0.128.61]:6643) (self) next_index=10394 match_index=12510
  e631 (e631 at tcp:[10.0.131.173]:6643) next_index=12512 match_index=0
```

```
kubectl ko nb kick e631
```

```
mv /etc/origin/ovn/ovnnb_db.db /tmp
```

```
ovn-central Pod
```

```
kubectl delete pod -n kube-system ovn-central-xxxx
```

6.4.3

leader

ovn-central

ovn-central ovn-central

```
kubectl scale deployment -n kube-system ovn-central --replicas=0
```

ovsdb-tool cluster-to-standalone

ovn-central NODE_IPS

/etc/origin/ovn

```
docker run -it -v /etc/origin/ovn:/etc/ovn kubeovn/kube-ovn:v1.14.23 bash
cd /etc/ovn/
ovsdb-tool cluster-to-standalone ovnnb_db_standalone.db ovnnb_db.db
ovsdb-tool cluster-to-standalone ovnsb_db_standalone.db ovnsb_db.db
```

ovn-central

```
mv /etc/origin/ovn/ovnnb_db.db /tmp
mv /etc/origin/ovn/ovnsb_db.db /tmp
```

ovnnb_db.db ovnsb_db.db ovn-central NODE_IPS /etc/origin/ovn/

```
mv /etc/origin/ovn/ovnnb_db_standalone.db /etc/origin/ovn/ovnnb_db.db
mv /etc/origin/ovn/ovnsb_db_standalone.db /etc/origin/ovn/ovnsb_db.db
```

ovn-central

```
kubectl scale deployment -n kube-system ovn-central --replicas=3
kubectl rollout status deployment/ovn-central -n kube-system
```

[PDF](#)

[Slack](#)

[Support](#)

⌚2025 9 10

⌚2022 5 24



6.4.4

6.5 CIDR

CIDR

CIDR Pod

CIDR

Join CIDR

Join CIDR

6.5.1

```
kubectl edit cidrBlock gateway excludeIps
```

```
kubectl edit subnet test-subnet
```

6.5.2 Namespace Pod

test Namespace

```
for pod in $(kubectl get pod --no-headers -n "$ns" --field-selector spec.restartPolicy=Always -o custom-columns=NAME:.metadata.name,HOST:spec.hostNetwork | awk '{if ($2!="true") print $1}'); do
    kubectl delete pod "$pod" -n test --ignore-not-found
done
```

host Pod

```
for ns in $(kubectl get ns --no-headers -o custom-columns=NAME:.metadata.name); do
    for pod in $(kubectl get pod --no-headers -n "$ns" --field-selector spec.restartPolicy=Always -o custom-columns=NAME:.metadata.name,HOST:spec.hostNetwork | awk '{if ($2!="true") print $1}'); do
        kubectl delete pod "$pod" -n "$ns" --ignore-not-found
    done
done
```

6.5.3

CIDR kube-ovn-controller Deployment

```
args:
- --default-cidr=10.17.0.0/16
- --default-gateway=10.17.0.1
- --default-exclude-ipss=10.17.0.1
```

 PDF

 Slack

 Support

 2025 9 10

 2022 5 24



6.5.4

6.6 Join CIDR

Join CIDR

| Join CIDR Pod

6.6.1 Join

```
kubectl patch subnet join --type='json' -p '[{"op": "replace", "path": "/metadata/finalizers", "value": []}]'
kubectl delete subnet join
```

6.6.2

```
kubectl annotate node ovn.kubernetes.io/allocated=false --all --overwrite
```

6.6.3 Join

kube-ovn-controller Join

```
kubectl edit deployment -n kube-system kube-ovn-controller
```

```
args:
- --node-switch-cidr=100.51.0.0/16
```

kube-ovn-controller join

```
kubectl delete pod -n kube-system -lapp=kube-ovn-controller
```

Join

```
# kubectl get subnet
NAME      PROVIDER   VPC      PROTOCOL  CIDR      PRIVATE  NAT    DEFAULT  GATEWAYTYPE  V4USED  V4AVAILABLE  V6USED  V6AVAILABLE
EXCLUDEIPS
join      ovn        ovn-cluster  IPv4     100.51.0.0/16  false    false   false   distributed   2       65531        0       0
["100.51.0.1"]
ovn-default  ovn        ovn-cluster  IPv4     10.17.0.0/16   false    true    true   distributed   5       65528        0       0
["10.17.0.1"]
```

6.6.4 ovn0

ovn0 kube-ovn-cni

```
kubectl delete pod -n kube-system -l app=kube-ovn-cni
```

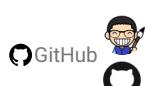
[PDF](#)

[Slack](#)

[Support](#)

⌚2025 9 10

⌚2022 5 24



6.6.5

6.7

kube-ovn.yaml

```
vi kube-ovn.yaml
# ...
- name: kube-ovn-controller
  image: "docker.io/kubeovn/kube-ovn:v1.14.23"
  imagePullPolicy: IfNotPresent
  args:
    - /kube-ovn/start-controller.sh
    - --v=3
# ...
#
```

[PDF](#)

[Slack](#)

[Support](#)

⌚2025 9 10

⌚2023 4 4



⌚ GitHub



6.7.1

6.8

Kube-OVN

6.8.1

1.

- Pod Pod
- Pod
- Pod Service
- Pod
- Pod
- kube-ovn-pinger

2.

- Pod
- kube-ovn-cni CNI
- ovs-ovn OVS
- kube-ovn-controller
- ovn-central OVN
- dmesg
- netstat -s

3. CPU, IO

```
kubectl ko logs
```

6.8.2 Pod IP

```
Pod     Running    kubectl describe Pod      duplicate IPv4 address <ip> found on logical switch port <port>
```

1. Pod IP Pod IP IP IP IP Pod
2. kube-ovn-controller IP
3. kube-ovn-controller IP
4. kubectl ko nbctl show OVN IP
5. OVN Kubernetes IP kubectl ko nbctl del-port <port>

6.8.3 Pod ping gateway failed

```
Pod     Running    kubectl describe Pod      network <ip> with gateway <gw ip> is not ready for interface eth0 after 30 checks
```

1. kubectl ko sbctl show Pod
2. ovn-central ovs-ovn ovn-central ovs-ovn

3. Pod
4. Underlay [Underlay](#)

6.8.4 Pod

VPC Pod

1. kubectl ko trace OVN ACL
2. ACL ACL
3. ACL Subnet stats
4. Subnet Spec
5. kube-ovn-controller

6.8.5 Pod IP CIDR

Pod IP CIDR

1. /etc/cni/net.d/ Kube-OVN CNI
- 2.
3. kubelet Pod

6.8.6 Debug Pod

kubectl debug Pod ContainerCreating Pod Event network not ready no address allocated

debug Pod yaml yaml Annotation

```
ovn.kubernetes.io/ip_address
ovn.kubernetes.io/mac_address
ovn.kubernetes.io/allocated
ovn.kubernetes.io/routed
```

debug Pod yaml debug Pod

6.8.7 ARM

ARM Offload

netstat

```
# netstat -us
IcmpMsg:
  InType0: 22
  InType3: 24
  InType8: 117852
  OutType0: 117852
```

```

OutType3: 29
OutType8: 22
Udp:
  3040636 packets received
  0 packets to unknown port received.
  4 packet receive errors
  602 packets sent
  0 receive buffer errors
  0 send buffer errors
  InCsumErrors: 4
UdpLite:
IpExt:
  InBcastPkts: 10244
  InOctets: 4446320361
  OutOctets: 1496815600
  InBcastOctets: 3095950
  InNoECTPkts: 7683903

```

InCsumErrors

tx offload TCP

ethtool -K eth0 tx off

CentOS 7

4.19.90-25.16.v2101

6.8.8 Pod Service

Pod Service dmesg

```

netlink: Unknown conntrack attr (type=6, max=5)
openvswitch: netlink: Flow actions may not be safe on all matching packets.

```

OVS NAT

1. OVS

2. Overlay kube-ovn-controller --enable-lb=false OVN LB kube-proxy Service

6.8.9 ovn-central

v1.11.x 1 Pod OVN NB SB Kube-OVN ovsdb-server/compact

ovn-central compact

```

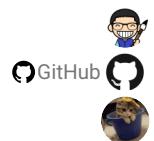
- name: ENABLE_COMPACT
  value: "false"

```

[PDF](#)[Slack](#)[Support](#)

⌚2025 9 10

⌚2022 5 24



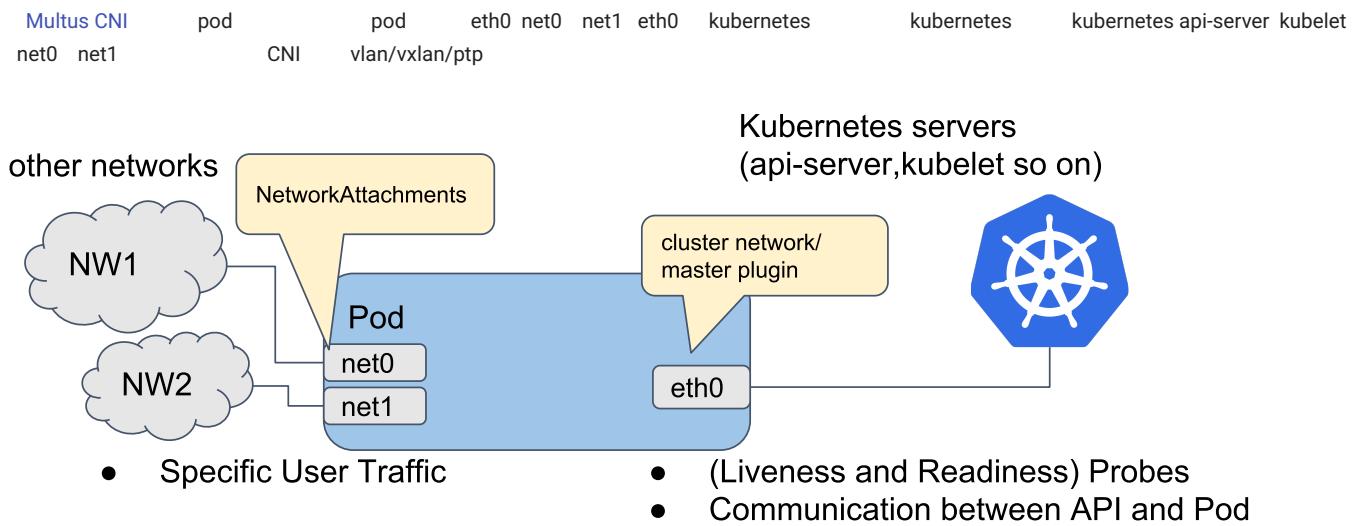
6.8.10

7.

7.1

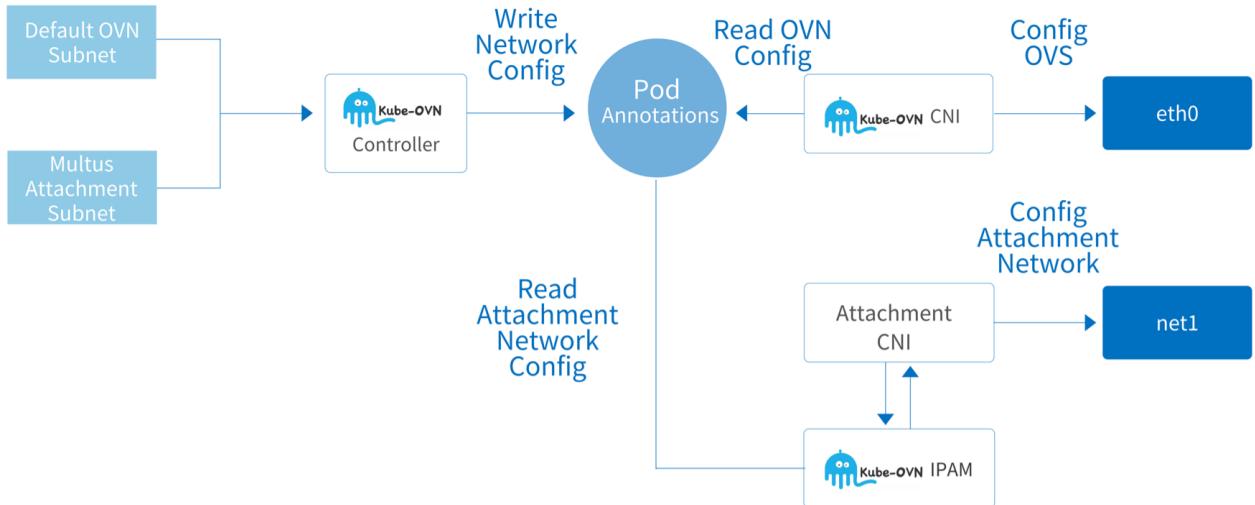
Kube-OVN	CNI	macvlan	vlan	host-device	IPAM	Kube-OVN	IP
Kube-OVN		Kube-OVN					

7.1.1



IPAM

Multus CNI	Pod	IP	Kube-OVN	IPAM	Subnet	IP	CRD	IP	IP



	Kube-OVN	IP	eth0	OVN	net1	CNI	net1	multus-cni	NetworkAttachmentDefinition
Pod	kube-ovn-controller	Pod	Pod	annotation	Subnet	IP	Pod	Pod	Pod annotation
CNI	kube-ovn-cni	ipam	kube-ovn-cni	annotation	Pod annotation	CNI		CNI	CNI

7.1.2

NetworkAttachmentDefinition spec multus defaultConfDir CNI

```
apiVersion: "k8s.cni.cncf.io/v1"
kind: NetworkAttachmentDefinition
metadata:
  name: macvlan-conf-2
```

kube-ovn-controller	NetworkAttachmentDefinition provider	spec	kube-ovn-
controller	provider	spec	IPAM

7.1.3

Kube-OVN Multus

[Kube-OVN](#) [Multus how to use](#) [Kube-OVN Multus-CNI](#)

CNI IPAM

Kube-OVN CNI

NETWORKATTACHMENTDEFINITION

macvlan ipam kube-ovn

```
#   macvlan
sudo modprobe macvlan
```

```

apiVersion: "k8s.cni.cncf.io/v1"
kind: NetworkAttachmentDefinition
metadata:
  name: macvlan
  namespace: default
spec:
  config: '{
    "cniVersion": "0.3.0",
    "type": "macvlan",
    "master": "eth0",
    "mode": "bridge",
    "ipam": {
      "type": "kube-ovn",
      "server_socket": "/run/openvswitch/kube-ovn-daemon.sock",
      "provider": "macvlan.default"
    }
}'

```

- spec.config.ipam.type: kube-ovn kube-ovn
- server_socket: Kube-OVN socket /run/openvswitch/kube-ovn-daemon.sock
- provider: NetworkAttachmentDefinition <name>. <namespace> Kube-OVN Subnet
- master:

Info

provider Underlay ProviderNetwork

KUBE-OVN SUBNET

Kube-OVN Subnet	cidrBlock	exclude_ips	provider	NetworkAttachmentDefinition <name>. <namespace>	macvlan
Subnet					

```

apiVersion: kubeovn.io/v1
kind: Subnet
metadata:
  name: macvlan
spec:
  protocol: IPv4
  provider: macvlan.default
  cidrBlock: 172.17.0.0/16
  gateway: 172.17.0.1
  excludeIps:
  - 172.17.0.0..172.17.0.10

```

gateway	private	nat	provider	ovn	attachment network
---------	---------	-----	----------	-----	--------------------

Pod

Pod	annotation k8s.v1.cni.cncf.io/networks	NetworkAttachmentDefinition <namespace>/<name>
-----	--	--

```

apiVersion: v1
kind: Pod
metadata:
  name: samplepod
  namespace: default
  annotations:
    k8s.v1.cni.cncf.io/networks: default/macvlan
spec:
  containers:
  - name: samplepod
    command: ["/bin/bash", "-c", "trap : TERM INT; sleep infinity & wait"]
    image: docker.io/library/alpine:edge

```

IP Pod

IP Pod <networkAttachmentName>. <networkAttachmentNamespace>. kubernetes.io/ip_address annotation

```

apiVersion: v1
kind: Pod
metadata:
  name: static-ip
  namespace: default
  annotations:
    k8s.v1.cni.cncf.io/networks: default/macvlan
    ovn.kubernetes.io/ip_address: 10.16.0.15
    ovn.kubernetes.io/mac_address: 00:00:00:53:6B:B6

```

```

macvlan.default.kubernetes.io/ip_address: 172.17.0.100
macvlan.default.kubernetes.io/mac_address: 00:00:00:53:6B:BB
spec:
  containers:
  - name: static-ip
    image: docker.io/library/nginx:alpine

```

IP

ippool <networkAttachmentName>.<networkAttachmentNamespace>.kubernetes.io/ip_pool annotation

```

apiVersion: apps/v1
kind: Deployment
metadata:
  namespace: default
  name: static-workload
  labels:
    app: static-workload
spec:
  replicas: 2
  selector:
    matchLabels:
      app: static-workload
  template:
    metadata:
      labels:
        app: static-workload
      annotations:
        k8s.v1.cni.cncf.io/networks: default/macvlan
        ovn.kubernetes.io/ip_pool: 10.16.0.15,10.16.0.16,10.16.0.17
        macvlan.default.kubernetes.io/ip_pool: 172.17.0.200,172.17.0.201,172.17.0.202
  spec:
    containers:
    - name: static-workload
      image: docker.io/library/nginx:alpine

```

macvlan Pod

macvlan	Pod	Pod	annotation	default-route
---------	-----	-----	------------	---------------

```

apiVersion: v1
kind: Pod
metadata:
  name: samplepod-route
  namespace: default
  annotations:
    k8s.v1.cni.cncf.io/networks: '[{
      "name": "macvlan",
      "namespace": "default",
      "default-route": ["172.17.0.1"]
    }]'
spec:
  containers:
  - name: samplepod-route
    command: ["/bin/ash", "-c", "trap : TERM INT; sleep infinity & wait"]
    image: docker.io/library/alpine:edge

```

macvlan Pod

macvlan	Pod	annotation v1.multus-cni.io/default-network	NetworkAttachmentDefinition <namespace>/<name>
---------	-----	---	--

```

apiVersion: v1
kind: Pod
metadata:
  name: samplepod-macvlan
  namespace: default
  annotations:
    v1.multus-cni.io/default-network: default/macvlan
spec:
  containers:
  - name: samplepod-macvlan
    command: ["/bin/ash", "-c", "trap : TERM INT; sleep infinity & wait"]
    image: docker.io/library/alpine:edge

```

KUBE-OVN SUBNET PROVIDER OVN

provider	ovn	subnet	IP	provider	ovn	Kube-OVN Subnet	cidrBlock	exclude_ips	Subnet
----------	-----	--------	----	----------	-----	-----------------	-----------	-------------	--------

```

apiVersion: kubeovn.io/v1
kind: Subnet
metadata:
  name: macvlan
spec:
  protocol: IPv4
  provider: ovn

```

```
cidrBlock: 172.17.0.0/16
gateway: 172.17.0.1
excludeIps:
- 172.17.0.0..172.17.0.10
```

Pod

provider	ovn	subnet	IP	Pod	annotation	k8s.v1.cni.cncf.io/networks <networkAttachmentName>.<networkAttachmentNamespace>.kubernetes.io/logical_switch
-----------------	------------	---------------	-----------	------------	-------------------	---

```
apiVersion: v1
kind: Pod
metadata:
  name: samplepod
  namespace: default
  annotations:
    k8s.v1.cni.cncf.io/networks: default/macvlan
    macvlan.default.kubernetes.io/logical_switch: macvlan
spec:
  containers:
  - name: samplepod
    command: ["/bin/ash", "-c", "trap : TERM INT; sleep infinity & wait"]
    image: docker.io/library/alpine:edge
```

- k8s.v1.cni.cncf.io/networks : NetworkAttachmentDefinition <namespace>/<name>

- macvlan.default.kubernetes.io/logical_switch :

• <networkAttachmentName>.<networkAttachmentNamespace>.kubernetes.io/logical_switch	provider
• ovn	subnet ipam IP Pod IP macvlan Pod
•	macvlan Pod

Kube-OVN**Kube-OVN****NETWORKATTACHMENTDEFINITION**

```
provider ovn

apiVersion: "k8s.cni.cncf.io/v1"
kind: NetworkAttachmentDefinition
metadata:
  name: attachnet
  namespace: default
spec:
  config: '{
    "cniVersion": "0.3.0",
    "type": "kube-ovn",
    "server_socket": "/run/openvswitch/kube-ovn-daemon.sock",
    "provider": "attachnet.default.ovn"
  }'
```

- spec.config.type : kube-ovn CNI Kube-OVN
- server_socket : Kube-OVN socket /run/openvswitch/kube-ovn-daemon.sock
- provider : NetworkAttachmentDefinition <name>.<namespace>.ovn Kube-OVN Subnet ovn

KUBE-OVN SUBNET

Kube-OVN	provider	NetworkAttachmentDefinition	<name>.<namespace>.ovn	ovn	Kube-OVN	Subnet
----------	----------	-----------------------------	------------------------	-----	----------	--------

```
apiVersion: kubeovn.io/v1
kind: Subnet
metadata:
  name: attachnet
spec:
  protocol: IPv4
  provider: attachnet.default.ovn
  cidrBlock: 172.17.0.0/16
  gateway: 172.17.0.1
  excludeIps:
  - 172.17.0.0..172.17.0.10
```

Pod

Pod	annotation k8s.v1.cni.cncf.io/networks	NetworkAttachmentDefinition <namespace>/<name>
-----	--	--

```

apiVersion: v1
kind: Pod
metadata:
  name: samplepod
  namespace: default
  annotations:
    k8s.v1.cni.cncf.io/networks: default/attachnet
spec:
  containers:
  - name: samplepod
    command: ["/bin/ash", "-c", "trap : TERM INT; sleep infinity & wait"]
    image: docker.io/library/alpine:edge

```

Kube-OVN Pod <networkAttachmentName>.<networkAttachmentNamespace>.ovn.kubernetes.io/routes annotation

```

apiVersion: v1
kind: Pod
metadata:
  name: custom-routes-attach
  namespace: default
  annotations:
    k8s.v1.cni.cncf.io/networks: default/attachnet
    attachnet.default.ovn.kubernetes.io/routes: |
      [
        {
          "dst": "192.168.0.101/24",
          "gw": "172.17.0.254"
        },
        {
          "gw": "172.17.0.254"
        }
      ]
spec:
  containers:
  - name: custom-routes-attach
    command: ["/bin/ash", "-c", "trap : TERM INT; sleep infinity & wait"]
    image: docker.io/library/alpine:edge

```

dst

Deployment	DaemonSet	StatefulSet	Annotation	.spec.template.metadata.annotations
------------	-----------	-------------	------------	-------------------------------------

```

apiVersion: apps/v1
kind: Deployment
metadata:
  name: custom-routes-attach
  labels:
    app: nginx
spec:
  replicas: 2
  selector:
    matchLabels:
      app: nginx
  template:
    metadata:
      labels:
        app: nginx
      annotations:
        k8s.v1.cni.cncf.io/networks: default/attachnet
        attachnet.default.ovn.kubernetes.io/routes: |
          [
            {
              "dst": "192.168.0.101/24",
              "gw": "172.17.0.254"
            },
            {
              "gw": "172.17.0.254"
            }
          ]
  spec:
    containers:
    - name: nginx
      image: docker.io/library/nginx:alpine

```

KUBE-OVN SUBNET PROVIDER OVN

provider	ovn	subnet	IP	provider	ovn	Kube-OVN Subnet	cidrBlock	exclude_ips	Subnet
----------	-----	--------	----	----------	-----	-----------------	-----------	-------------	--------

```

apiVersion: kubeovn.io/v1
kind: Subnet
metadata:
  name: attachnet
spec:
  protocol: IPv4
  provider: ovn
  cidrBlock: 172.17.0.0/16

```

```
gateway: 172.17.0.1
excludeIps:
- 172.17.0.0..172.17.0.10
```

Pod

```
provider      ovn    subnet    IP   Pod    annotation k8s.v1.cni.cncf.io/networks
<networkAttachmentName>.<networkAttachmentNamespace>.ovn.kubernetes.io/logical_switch
```

```
apiVersion: v1
kind: Pod
metadata:
  name: samplepod
  namespace: default
  annotations:
    k8s.v1.cni.cncf.io/networks: default/attachnet
    attachnet.default.ovn.kubernetes.io/logical_switch: attachnet
spec:
  containers:
  - name: samplepod
    command: ["/bin/ash", "-c", "trap : TERM INT; sleep infinity & wait"]
    image: docker.io/library/alpine:edge
```

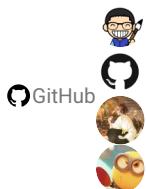
- k8s.v1.cni.cncf.io/networks : NetworkAttachmentDefinition <namespace>/<name>
- attachnet.default.ovn.kubernetes.io/logical_switch :

- <networkAttachmentName>.<networkAttachmentNamespace>.ovn.kubernetes.io/logical_switch provider
- Kube-OVN Pod IP Pod IP macvlan Pod Kube-OVN Pod
- Kube-OVN Pod <networkAttachmentName>.<networkAttachmentNamespace>.ovn.kubernetes.io/routes annotation

[PDF](#)
[Slack](#)
[Support](#)

⌚2025 12 18

⌚2022 5 20



7.1.4

7.2

Kube-OVN

Kube-OVN

7.2.1

1.	CNI	CNI
2.	10ns	20ns
3.	CPU	CPU
4.	Macvlan	SR-IOV

7.2.2

Overlay

- Kubernetes: 1.22.0
- OS: CentOS 7
- Kube-OVN: 1.8.0 *Overlay*
- CPU: Intel(R) Xeon(R) E-2278G
- Network: 2*10Gbps, xmit_hash_policy=layer3+4

```
qperf -t 60 <server ip> -ub -oo msg_size:1 -vu tcp_lat tcp_bw udp_lat udp_bw 1      tcp/udp
```

Type	tcp_lat (us)	udp_lat (us)	tcp_bw (Mb/s)	udp_bw(Mb/s)
Kube-OVN Default	25.7	22.9	27.1	1.59
Kube-OVN Optimized	13.9	12.9	27.6	5.57
HOST Network	13.1	12.4	28.2	6.02

Overlay Underlay

Kube-OVN Overlay Underlay

Environment:

- Kubernetes: 1.22.0
- OS: CentOS 7
- Kube-OVN: 1.8.0
- CPU: AMD EPYC 7402P 24-Core Processor
- Network: Intel Corporation Ethernet Controller XXV710 for 25GbE SFP28

```
qperf -t 60 <server ip> -ub -oo msg_size:1 -vu tcp_lat tcp_bw udp_lat udp_bw
```

Type	tcp_lat (us)	udp_lat (us)	tcp_bw (Mb/s)	udp_bw(Mb/s)
Kube-OVN Overlay	15.2	14.6	23.6	2.65
Kube-OVN Underlay	14.3	13.8	24.2	3.46
HOST Network	16.6	15.4	24.8	2.64

```
qperf -t 60 <server ip> -ub -oo msg_size:1K -vu tcp_lat tcp_bw udp_lat udp_bw
```

Type	tcp_lat (us)	udp_lat (us)	tcp_bw (Gb/s)	udp_bw(Gb/s)
Kube-OVN Overlay	16.5	15.8	10.2	2.77
Kube-OVN Underlay	15.9	14.5	9.6	3.22
HOST Network	18.1	16.6	9.32	2.66

```
qperf -t 60 <server ip> -ub -oo msg_size:4K -vu tcp_lat tcp_bw udp_lat udp_bw
```

Type	tcp_lat (us)	udp_lat (us)	tcp_bw (Gb/s)	udp_bw(Gb/s)
Kube-OVN Overlay	34.7	41.6	16.0	9.23
Kube-OVN Underlay	32.6	44	15.1	6.71
HOST Network	35.9	45.9	14.6	5.59

netfilter kube-proxy netfilter

7.2.3

CPU

CPU

CPU

```
cpupower frequency-set -g performance
```

```
# ethtool -g eno1
Ring parameters for eno1:
Pre-set maximums:
RX:        4096
RX Mini:   0
RX Jumbo:  0
TX:        4096
Current hardware settings:
RX:        255
RX Mini:   0
RX Jumbo:  0
TX:        255
```

```
ethtool -G eno1 rx 4096
ethtool -G eno1 tx 4096
```

tuned

tuned profile

tuned-adm profile network-latency

tuned-adm profile network-throughput

irqbalance CPU CPU

OVN LB

OVN L2 LB	conntrack	recirculate	CPU	20%	CPU	Overlay	kube-proxy	Service
Pod-to-Pod	kube-ovn-controller							

```
command:
- ./kube-ovn/start-controller.sh
args:
...
- --enable-lb=false
...
```

Underlay kube-proxy iptables ipvs LB Service

FastPath

network ns	netfilter	20% CPU	netfilter	FastPath	netfilter CPU
------------	-----------	---------	-----------	----------	---------------

netfilter iptables ipvs nftables

FastPath

insmod kube_ovn_fastpath.ko FastPath dmesg

```
# dmesg
...
[619631.323788] init_module,kube_ovn_fastpath_local_out
[619631.323798] init_module,kube_ovn_fastpath_post_routing
[619631.323800] init_module,kube_ovn_fastpath_pre_routing
[619631.323801] init_module,kube_ovn_fastpath_local_in
...
```

OVS

OVS flow	10% CPU	x86 CPU	popcnt sse4.2	flow CPU
5%				

FastPath

CPU

```
cat /proc/cpuinfo | grep popcnt
cat /proc/cpuinfo | grep sse4_2
```

CENTOS

```
yum install -y gcc kernel-devel-$(uname -r) python3 autoconf automake libtool rpm-build openssl-devel
```

OVS RPM :

```
git clone -b branch-3.5 --depth=1 https://github.com/openvswitch/ovs.git
cd ovs
curl -s https://github.com/kubeovn/ovs/commit/2d2c83c26d4217446918f39d5cd5838e9ac27b32.patch | git apply
./boot.sh
./configure --with-linux=/lib/modules/$(uname -r)/build CFLAGS="-g -O2 -mpopcnt -msse4.2"
make rpm-fedora-kmod
cd rpm/rpmbuild/RPMS/x86_64/
```

RPM

```
rpm -i openvswitch-kmod-3.5.1-1.el7.x86_64.rpm
```

Kube-OVN OVS

UBUNTU

```
apt install -y autoconf automake libtool gcc build-essential libssl-dev
```

OVS

```
apt install -y autoconf automake libtool gcc build-essential libssl-dev

git clone -b branch-3.5 --depth=1 https://github.com/openvswitch/ovs.git
cd ovs
curl -s https://github.com/kubeovn/ovs/commit/2d2c83c26d4217446918f39d5cd5838e9ac27b32.patch | git apply
./boot.sh
./configure --prefix=/usr/ --localstatedir=/var --enable-ssl --with-linux=/lib/modules/$(uname -r)/build
make -j `nproc`
make install
make modules_install

cat > /etc/modprobe.d/openvswitch.conf << EOF
override openvswitch * extra
override vport-* * extra
EOF

depmod -a
cp debian/openvswitch-switch.init /etc/init.d/openvswitch-switch
/etc/init.d/openvswitch-switch force-reload-kmod
```

Kube-OVN OVS

STT



Warning

OpenVswitch 3.6 STT Tunnel

Geneve	Vxlan	UDP	UDP	TCP	offload	TCP
CPU	TCP					
STT	TCP		TCP		TCP	TCP
STT		OVS		OVS		
STT						

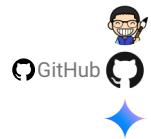
```
kubectl set env daemonset/ovs-ovn -n kube-system TUNNEL_TYPE=stt
```

```
kubectl delete pod -n kube-system -lapp=ovs
```



⌚2025 9 10

⌚2022 5 24



7.2.4

7.3 FastPath

Profile Netfilter

20% CPU FastPath

Netfilter CPU

FastPath

7.3.1

```
git clone --depth=1 https://github.com/kubeovn/kube-ovn.git
```

7.3.2

CentOS

```
yum install -y kernel-devel-$(uname -r) gcc elfutils-libelf-devel
```

7.3.3

3.x

```
cd kube-ovn/fastpath
make all
```

4.x

```
cd kube-ovn/fastpath/4.18
cp ../Makefile .
make all
```

7.3.4

```
kube_ovn_fastpath.ko          /tmp      kube-ovn-cni
```

dmesg

```
# dmesg
[619631.323788] init_module,kube_ovn_fastpath_local_out
[619631.323798] init_module,kube_ovn_fastpath_post_routing
[619631.323800] init_module,kube_ovn_fastpath_pre_routing
[619631.323801] init_module,kube_ovn_fastpath_local_in
```

```
/tmp      kube-ovn-cni
```



[PDF](#)



[Slack](#)



[Support](#)

⌚2025 9 10

⌚2022 5 24



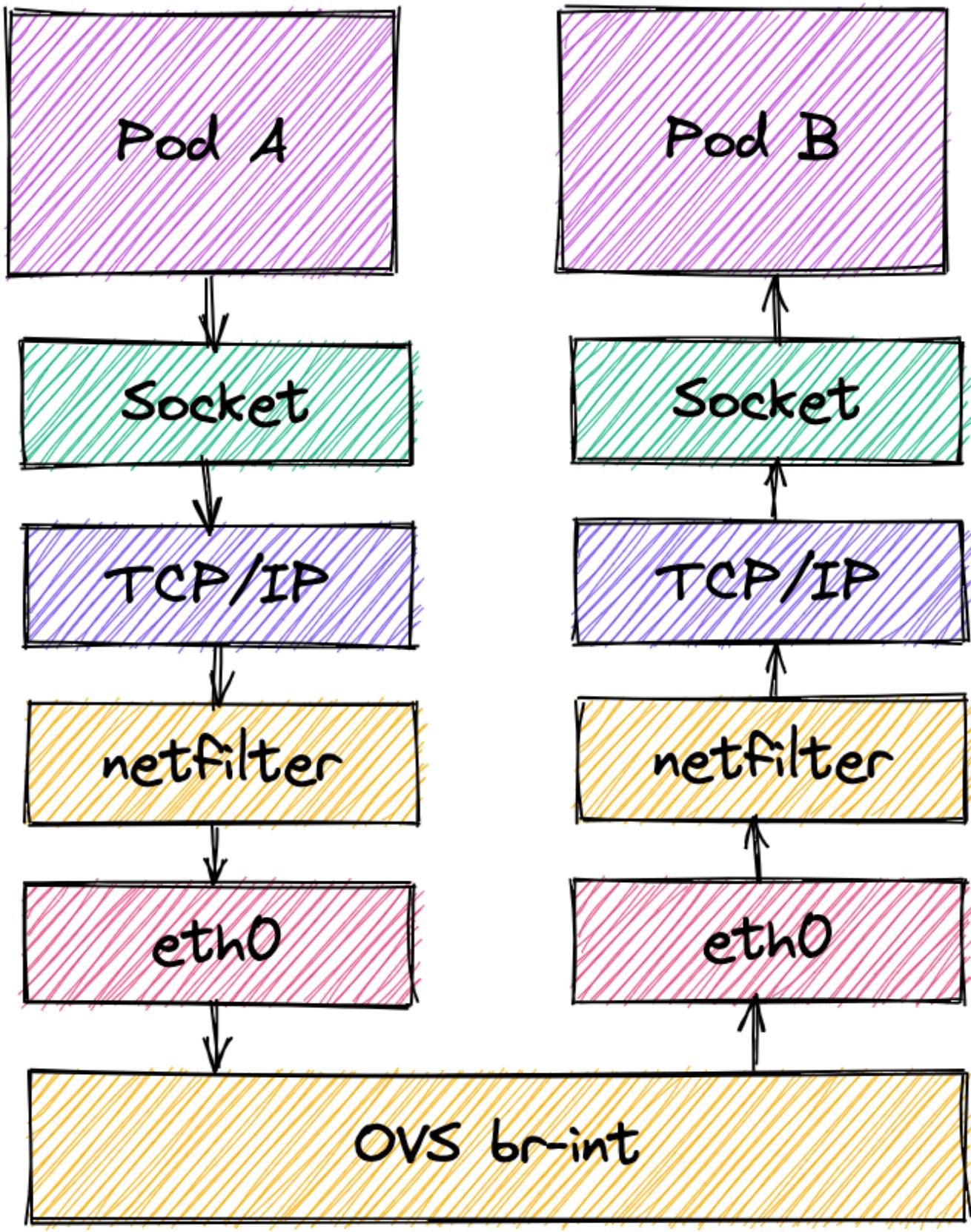
7.3.5

7.4 eBPF TCP

5G Pod TCP Intel **istio-tcpip-bypass** Pod eBPF TCP/IP socket

7.4.1

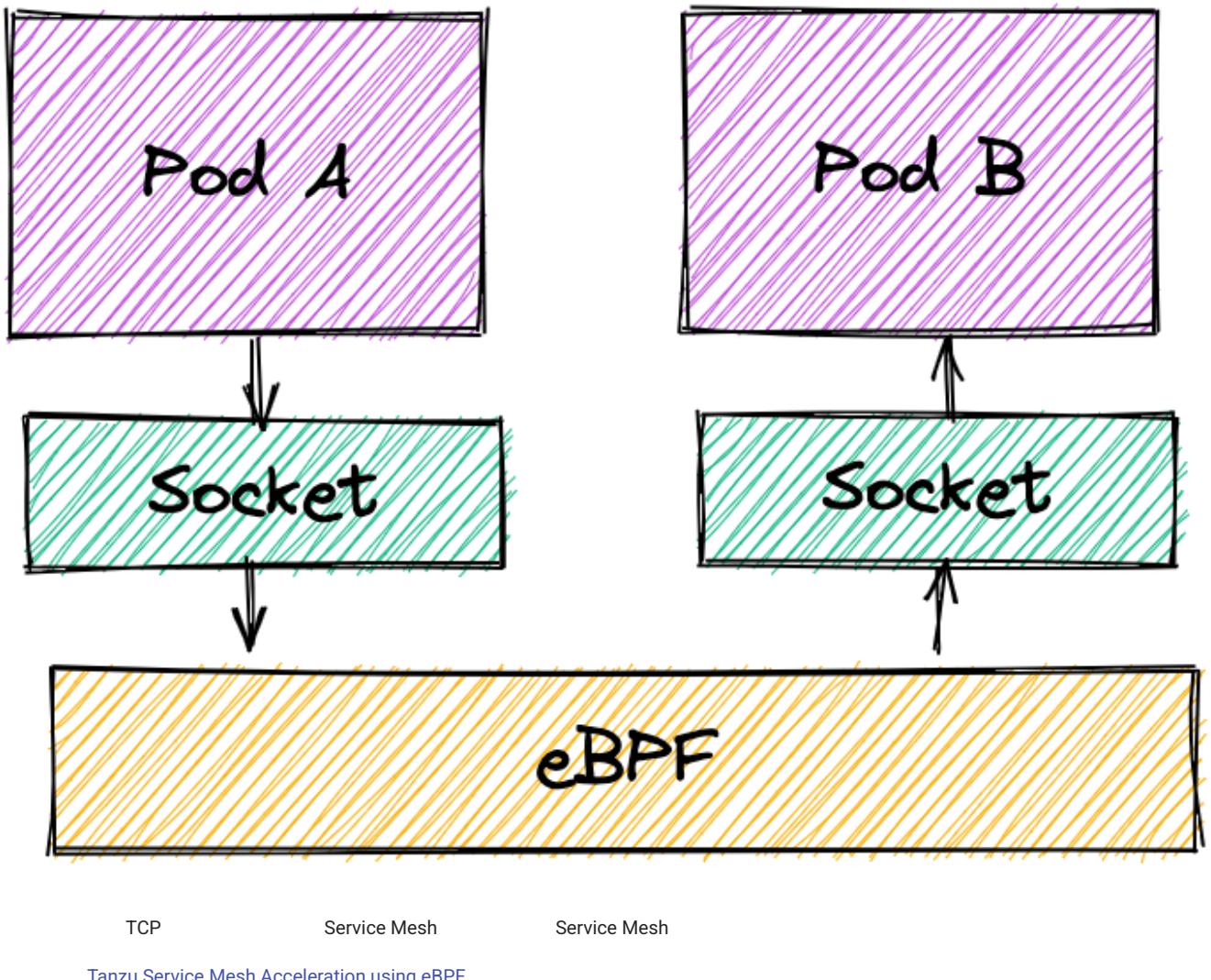
Pod TCP TCP/IP, netfilter, OVS



istio-tcpip-bypass

TCP

socket



7.4.2

eBPF Ubuntu 20.04 Linux 5.4.0-74-generic

7.4.3

Pod nodeSelector

```
# kubectl create deployment perf --image=kubeovn/perf:dev --replicas=2
deployment.apps/perf created
# kubectl get pod -o wide
NAME           READY   STATUS    RESTARTS   AGE     IP          NODE   NOMINATED NODE   READINESS GATES
perf-7697bc6ddf-b2cpv  1/1    Running   0          28s    100.64.0.3  sealos <none>        <none>
perf-7697bc6ddf-p2xpt  1/1    Running   0          28s    100.64.0.2  sealos <none>        <none>
```

Pod qperf server Pod qperf client

```
# kubectl exec -it perf-7697bc6ddf-b2cpv sh
/ # qperf

# kubectl exec -it perf-7697bc6ddf-p2xpt sh
/ # qperf -t 60 100.64.0.3 -ub -oo msg_size:1:16K:*4 -vu tcp_lat tcp_bw
```

istio-tcpip-bypass

```
kubectl apply -f https://raw.githubusercontent.com/intel/istio-tcpip-bypass/main/bypass-tcpip-daemonset.yaml
```

perf client

```
# kubectl exec -it perf-7697bc6ddf-p2xpt sh
/ # qperf -t 60 100.64.0.3 -ub -oo msg_size:1:16K:*4 -vu tcp_lat tcp_bw
```

7.4.4

TCP 40% ~ 60% 1024 40% ~ 80%

Packet Size (byte)	eBPF tcp_lat (us)	Default tcp_lat (us)	eBPF tcp_bw (Mb/s)	Default tcp_bw(Mb/s)
1	20.2	44.5	1.36	4.27
4	20.2	48.7	5.48	16.7
16	19.6	41.6	21.7	63.5
64	18.8	41.3	96.8	201
256	19.2	36	395	539
1024	18.3	42.4	1360	846
4096	16.5	62.6	4460	2430
16384	20.2	58.8	9600	6900

512 eBPF TCP eBPF eBPF TCP

7.4.5

1. [istio-tcpip-bypass](#)
2. [Deep Dive TCP/IP Bypass with eBPF in Service Mesh](#)
3. [Tanzu Service Mesh Acceleration using eBPF](#)

[!\[\]\(a5a1f5de324e965e343fb024aca2e624_img.jpg\) PDF](#)

[!\[\]\(8820d2b7af44a08842bfe1b879d3d4c5_img.jpg\) Slack](#)

[!\[\]\(8ab38bf6d96bb81afe368c33fcb6aa48_img.jpg\) Support](#)

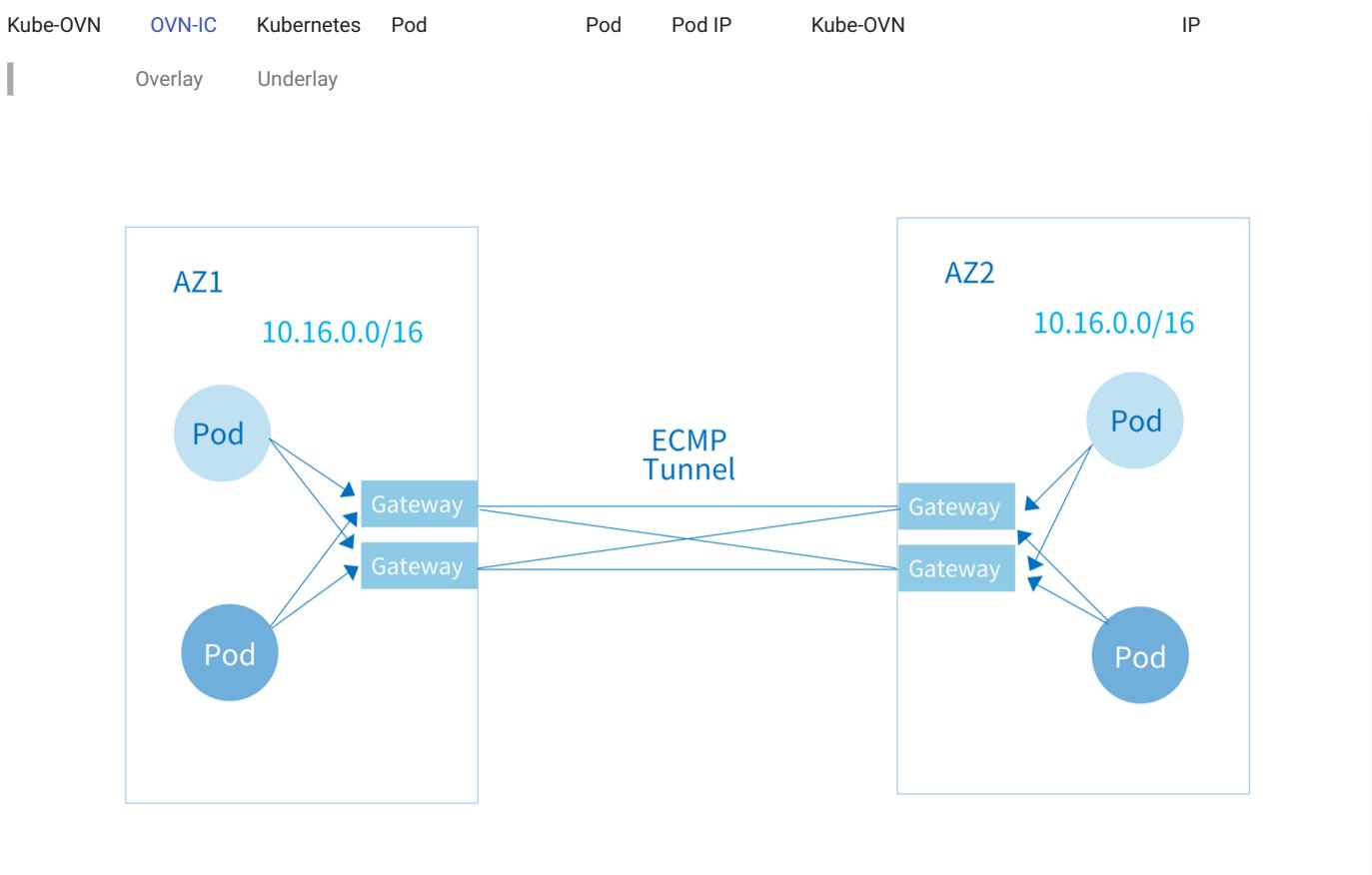
 2025 9 10

 2022 9 27



7.4.6

7.5 OVN-IC



Limitation

OVN-IC Pod IP Service DNS NetworkPolicy Isti

7.5.1

1.1.11.16 install.sh

ENABLE_TC=true

deployment ovn-ic-controller

CIDR

3. kube-ovn-controller IP

4. IP

5. VPC VPC

7.5.2 OVN-IC

1

1 Kube-OVN v1.11.16

" " " " Deployment master master 1 master

```
install-ovn-ic.sh
```

```
wget https://raw.githubusercontent.com/kubeovn/kube-ovn/release-1.14/dist/images/install-ic-server.sh
```

TS_NUM	ECMP Path
--------	-----------

```
sed 's/VERSION=.*/VERSION=v1.14.23/' dist/images/install-ic-server.sh | TS_NUM=3 bash
```

```
deployment.apps/ovn-ic-server created
Waiting for deployment spec update to be observed...
Waiting for deployment "ovn-ic-server" rollout to finish: 0 out of 3 new replicas have been updated...
Waiting for deployment "ovn-ic-server" rollout to finish: 0 of 3 updated replicas are available...
Waiting for deployment "ovn-ic-server" rollout to finish: 1 of 3 updated replicas are available...
Waiting for deployment "ovn-ic-server" rollout to finish: 2 of 3 updated replicas are available...
Waiting for deployment "ovn-ic-server" rollout to finish: 3 of 3 updated replicas are available...
deployment "ovn-ic-server" successfully rolled out
OVN IC Server installed Successfully
```

```
kubectl ko icsbctl show
```

```
kubectl ko icsbctl show
availability-zone az0
gateway 059b5c54-c540-4d77-b009-02d65f181a02
  hostname: kube-ovn-worker
  type: geneve
    ip: 172.18.0.3
  port ts-az0
    transit switch: ts
    address: ["00:00:00:B4:8E:BE 169.254.100.97/24"]
gateway 74ee4b9a-ba48-4a07-861e-1a8e4b9f905f
  hostname: kube-ovn-worker2
  type: geneve
    ip: 172.18.0.2
  port ts1-az0
    transit switch: ts1
    address: ["00:00:00:19:2E:F7 169.254.101.90/24"]
gateway 7e2428b6-344c-4dd5-a0d5-972c1cc581
  hostname: kube-ovn-control-plane
  type: geneve
    ip: 172.18.0.4
  port ts2-az0
    transit switch: ts2
    address: ["00:00:00:EA:32:BA 169.254.102.103/24"]
availability-zone az1
gateway 034da7cb-3826-4318-81ce-6a877a9bf285
  hostname: kube-ovn1-worker
  type: geneve
    ip: 172.18.0.6
  port ts-az1
    transit switch: ts
    address: ["00:00:00:25:3A:B9 169.254.100.51/24"]
gateway 2531a683-283e-4fb8-a619-bdbcb33539b8
  hostname: kube-ovn1-worker2
  type: geneve
    ip: 172.18.0.5
  port ts1-az1
    transit switch: ts1
    address: ["00:00:00:52:87:F4 169.254.101.118/24"]
gateway b0efb0be-e5a7-4323-ad4b-317637a757c4
  hostname: kube-ovn1-control-plane
  type: geneve
    ip: 172.18.0.8
  port ts2-az1
    transit switch: ts2
    address: ["00:00:00:F6:93:1A 169.254.102.17/24"]
```

2

kube-ovn-controller	IP	OVN-IC
---------------------	----	--------

docker		OVN-IC
--------	--	--------

```
docker run --name=ovn-ic-db -d --env "ENABLE_OVN_LEADER_CHECK=false" --network=host --privileged -v /etc/ovn/:/etc/ovn -v /var/run/ovn:/var/run/ovn -v /var/log/ovn:/var/log/ovn kubeovn/kube-ovn:v1.14.23 bash start-ic-db.sh
```

containerd	docker
------------	--------

```
ctr -n k8s.io run -d --env "ENABLE_OVN_LEADER_CHECK=false" --net-host --privileged --mount="type=bind,src=/etc/ovn/,dst=/etc/ovn,options=rbind:rw" --mount="type=bind,src=/var/run/ovn,dst=/var/run/ovn,options=rbind:rw" --mount="type=bind,src=/var/log/ovn,dst=/var/log/ovn,options=rbind:rw" docker.io/kubeovn/kube-ovn:v1.14.23 ovn-ic-db bash start-ic-db.sh
```

7.5.3

VPC	Subnet	CIDR	OVN-IC	Subnet CIDR
-----	--------	------	--------	-------------

kube-system Namespace ovn-ic-config ConfigMap

```
apiVersion: v1
kind: ConfigMap
metadata:
  name: ovn-ic-config
  namespace: kube-system
data:
  enable-ic: "true"
  az-name: "az1"
  ic-db-host: "192.168.65.3"
  ic-nb-port: "6645"
  ic-sb-port: "6646"
  gw-nodes: "az1-gw"
  auto-route: "true"
```

- enable-ic:
- az-name:
- ic-db-host: OVN-IC
- ic-nb-port: OVN-IC 6645
- ic-sb-port: OVN-IC 6646
- gw-nodes:
- auto-route:

ovn-ic-config	ConfigMap	ConfigMap	ConfigMap
---------------	-----------	-----------	-----------

ovn-ic ts

```
# ovn-ic-sbctl show
availability-zone az1
  gateway deee03e0-af16-4f45-91e9-b50c3960f809
    hostname: az1-gw
    type: geneve
      ip: 192.168.42.145
    port ts-az1
      transit switch: ts
      address: ["00:00:00:50:AC:8C 169.254.100.45/24"]
availability-zone az2
  gateway e94cc831-8143-40e3-a478-90352773327b
    hostname: az2-gw
    type: geneve
      ip: 192.168.42.149
    port ts-az2
      transit switch: ts
      address: ["00:00:00:07:4A:59 169.254.100.63/24"]
```

```
# kubectl ko nbctl lr-route-list ovn-cluster
IPv4 Routes
  10.42.1.1      169.254.100.45 dst-ip (learned)
  10.42.1.3      100.64.0.2 dst-ip
  10.16.0.2      100.64.0.2 src-ip
  10.16.0.3      100.64.0.2 src-ip
  10.16.0.4      100.64.0.2 src-ip
  10.16.0.6      100.64.0.2 src-ip
  10.17.0.0/16    169.254.100.45 dst-ip (learned)
  100.65.0.0/16   169.254.100.45 dst-ip (learned)
```

1 Pod ping 2 Pod IP

Subnet disableInterConnection

```
apiVersion: kubeovn.io/v1
kind: Subnet
metadata:
```

```

  name: no-advertise
spec:
  cidrBlock: 10.199.0.0/16
  disableInterConnection: true

```

7.5.4

CIDR

```
kube-system Namespace      ovn-ic-config ConfigMap      auto-route      false
```

```

apiVersion: v1
kind: ConfigMap
metadata:
  name: ovn-ic-config
  namespace: kube-system
data:
  enable-ic: "true"
  az-name: "az1"
  ic-db-host: "192.168.65.3"
  ic-nb-port: "6645"
  ic-sb-port: "6646"
  gw-nodes: "az1-gw"
  auto-route: "false"

```

```

[root@az1 ~]# kubectl ko nbctl show
switch a391d3a1-14a0-4841-9836-4bd930c447fb (ts)
  port ts-az1
    type: router
    router-port: az1-ts
  port ts-az2
    type: remote
    addresses: ["00:00:00:4B:E2:9F 169.254.100.31/24"]

[root@az2 ~]# kubectl ko nbctl show
switch d46138b8-de81-4908-abf9-b2224ec4edf3 (ts)
  port ts-az2
    type: router
    router-port: az2-ts
  port ts-az1
    type: remote
    addresses: ["00:00:00:FB:2A:F7 169.254.100.79/24"]

```

```
az1      az2      169.254.100.31  az2      az1      169.254.100.79
```

```
az1      CIDR  10.16.0.0/24      az2      CIDR  10.17.0.0/24
```

```
az1      az2
```

```
kubectl ko nbctl lr-route-add ovn-cluster 10.17.0.0/24 169.254.100.31
```

```
az2      az1
```

```
kubectl ko nbctl lr-route-add ovn-cluster 10.16.0.0/24 169.254.100.79
```

7.5.5 OVN-IC

1

1 Kube-OVN v1.11.16

1

2

OVN-IC	Raft	3
--------	------	---

OVN-IC	leader
--------	--------

docker

```
docker run --name=ovn-ic-db -d --env "ENABLE_OVN_LEADER_CHECK=false" --network=host --privileged -v /etc/ovn:/etc/ovn -v /var/run/ovn:/var/run/ovn -v /var/log/ovn:/var/log/ovn -e LOCAL_IP="192.168.65.3" -e NODE_IPS="192.168.65.3,192.168.65.2,192.168.65.1" kubeovn/kube-ovn:v1.14.23 bash start-ic-db.sh
```

containerd

```
ctr -n k8s.io run -d --env "ENABLE_OVN_LEADER_CHECK=false" --net-host --privileged --mount="type=bind,src=/etc/ovn/,dst=/etc/ovn,options=rbind:rw" --mount="type=bind,src=/var/run/ovn,dst=/var/run/ovn,options=rbind:rw" --mount="type=bind,src=/var/log/ovn,dst=/var/log/ovn,options=rbind:rw" --env="NODE_IPS=192.168.65.3,192.168.65.2,192.168.65.1" --env="LOCAL_IP=192.168.65.3" docker.io/kubeovn/kube-ovn:v1.14.23 ovn-ic-db bash start-ic-db.sh
```

- LOCAL_IP IP
- NODE_IPS OVN-IC IP
- OVN-IC follower

docker

```
docker run --name=ovn-ic-db -d --network=host --privileged -v /etc/ovn:/etc/ovn -v /var/run/ovn:/var/run/ovn -v /var/log/ovn:/var/log/ovn -e LOCAL_IP="192.168.65.2" -e NODE_IPS="192.168.65.3,192.168.65.2,192.168.65.1" -e LEADER_IP="192.168.65.3" kubeovn/kube-ovn:v1.14.23 bash start-ic-db.sh
```

containerd

```
ctr -n k8s.io run -d --net-host --privileged --mount="type=bind,src=/etc/ovn/,dst=/etc/ovn,options=rbind:rw" --mount="type=bind,src=/var/run/ovn,dst=/var/run/ovn,options=rbind:rw" --mount="type=bind,src=/var/log/ovn,dst=/var/log/ovn,options=rbind:rw" --env="NODE_IPS=192.168.65.3,192.168.65.2,192.168.65.1" --env="LOCAL_IP=192.168.65.2" --env="LEADER_IP=192.168.65.3" docker.io/kubeovn/kube-ovn:v1.14.23 ovn-ic-db bash start-ic-db.sh
```

- LOCAL_IP IP
- NODE_IPS OVN-IC IP
- LEADER_IP: OVN-IC leader IP

ovn-ic-config OVN-IC

```
apiVersion: v1
kind: ConfigMap
metadata:
  name: ovn-ic-config
  namespace: kube-system
data:
  enable-ic: "true"
  az-name: "az1"
  ic-db-host: "192.168.65.3,192.168.65.2,192.168.65.1"
  ic-nb-port: "6645"
  ic-sb-port: "6646"
  gw-nodes: "az1-gw"
  auto-route: "true"
```

7.5.6 ECMP

1

ECMP ECMP path 3 ECMP path

```
kubectl edit deployment ovn-ic-server -n kube-system
```

'TS_NUM' TS_NUM ECMP Path

7.5.7

ovn-ic-config Configmap

```
kubectl -n kube-system delete cm ovn-ic-config
```

ts

```
kubectl ko nbctl ls-del ts
```

7.5.8 az-name

```
kubectl edit ovn-ic-config configmap az-name ovn-cni pod 10
```

```
ovn-appctl -t ovn-controller inc-engine/recompute
```

7.5.9

ovn-ic-config Configmap

```
kubectl -n kube-system delete cm ovn-ic-config
```

ts

```
kubectl ko nbctl ls-del ts
```

OVN-IC

docker

```
docker stop ovn-ic-db
docker rm ovn-ic-db
```

containerd

```
ctr -n k8s.io task kill ovn-ic-db
ctr -n k8s.io containers rm ovn-ic-db
```

deployment ovn-ic-server

```
kubectl delete deployment ovn-ic-server -n kube-system
```

master

DB

```
rm -f /etc/origin/ovn/ovn_ic_nb_db.db
rm -f /etc/origin/ovn/ovn_ic_sb_db.db
```

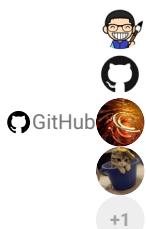
[PDF](#)

[Slack](#)

[Support](#)

⌚2025 6 19

⌚2022 5 24



7.5.10

7.6 Submariner



7.6.1

- Service CIDR CIDR

7.6.2 Submariner

```

subctl

curl -Ls https://get.submariner.io | bash
export PATH=$PATH:~/local/bin
echo export PATH=\$PATH:~/local/bin >> ~/.profile

kubeconfig      submariner-broker

subctl deploy-broker

cluster0      CIDR  10.16.0.0/16 join  CIDR  100.64.0.0/16  cluster1      CIDR  11.16.0.0/16 join  CIDR  100.68.0.0/16

kubeconfig  cluster0  broker  :

subctl join broker-info.subm --clusterid cluster0 --clustercidr 100.64.0.0/16,10.16.0.0/16  --natt=false --cable-driver vxlan --health-check=false
kubectl label nodes cluster0 submariner.io/gateway=true

kubeconfig  cluster1  broker  :

subctl join broker-info.subm --clusterid cluster1 --clustercidr 100.68.0.0/16,11.16.0.0/16  --natt=false --cable-driver vxlan --health-check=false
kubectl label nodes cluster1 submariner.io/gateway=true

join      gateway, routeagent pod  ,  submariner-operator  clusterrole  :

- apiGroups:
  - "apps"
  resources:
  - daemonsets
  verbs:
  - create
  - get
  - list
  - watch
  - update

subnet  ovn-default  centralized  submariner  gateway  subnet

Pod      IP

subctl

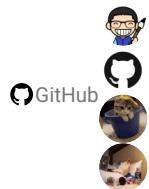
subctl show all
subctl diagnose all

Submariner      Submariner
  
```

[PDF](#) [Slack](#) [Support](#)

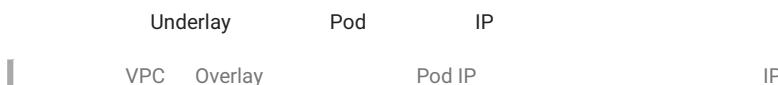
⌚2025 9 10

⌚2022 5 24



7.6.3

7.7 Overlay



7.7.1

- ip_forward
- iptables forward Drop
- ct INVALID

7.7.2

natOutgoing false nat Pod IP

```

apiVersion: kubeovn.io/v1
kind: Subnet
metadata:
  name: routed
spec:
  protocol: IPv4
  cidrBlock: 10.166.0.0/16
  default: false
  excludeIps:
  - 10.166.0.1
  gateway: 10.166.0.1
  gatewayType: distributed
  natOutgoing: false
  
```

Pod

Kubernetes

```
ip route add 10.166.0.0/16 via 192.168.2.10 dev eth0
```

10.166.0.0/16 192.168.2.10 Kubernetes

IP Keepalived VIP VIP

Subnet gatewayType centralized gatewayNode IP

```

apiVersion: kubeovn.io/v1
kind: Subnet
metadata:
  name: routed
spec:
  protocol: IPv4
  cidrBlock: 10.166.0.0/16
  default: false
  excludeIps:
  - 10.166.0.1
  gateway: 10.166.0.1
  gatewayType: centralized
  gatewayNode: "node1"
  natOutgoing: false
  
```

nat VPC NAT

[PDF](#)

[Slack](#)

[Support](#)

⌚2023 12 25

⌚2022 7 6

⌚GitHub 🎉

7.7.3

7.8 BGP



7.8.1 kube-ovn-speaker

kube-ovn-speaker GoBGP

`kube-ovn-speaker`

```
kubectl label nodes speaker-node-1 ovn.kubernetes.io/bgp=true
kubectl label nodes speaker-node-2 ovn.kubernetes.io/bgp=true
```

`kube-ovn-speaker`

ECMP

`yaml:`

```
wget https://raw.githubusercontent.com/kubeovn/kube-ovn/release-1.14/yamls/speaker.yaml
```

`yaml`

```
- --neighbor-address=10.32.32.254
- --neighbor-ipv6-address=2409:AB00:AB00:2000::AFB:8AFE
- --neighbor-as=65030
- --cluster-as=65000
```

```
- --neighbor-address=10.32.32.252,10.32.32.253
- --neighbor-ipv6-address=2409:AB00:AB00:2000::AFB:8AFC,2409:AB00:AB00:2000::AFB:8AFD
- --neighbor-as=65030
- --cluster-as=65000
```

- `neighbor-address`: BGP Peer
- `neighbor-as`: BGP Peer AS
- `cluster-as`: AS

`yaml:`

```
kubectl apply -f speaker.yaml
```

7.8.2 Pod/Subnet

BGP Subnet natOutgoing false Pod IP

`annotation`

```
kubectl annotate pod sample ovn.kubernetes.io/bgp=true
kubectl annotate subnet ovn-default ovn.kubernetes.io/bgp=true
```

`annotation`

```
kubectl annotate pod sample ovn.kubernetes.io/bgp-
kubectl annotate subnet ovn-default ovn.kubernetes.io/bgp-
```

`BGP`

7.8.3 ClusterIP Service

```
Service ClusterIP      kube-ovn-speaker   --announce-cluster-ip   true          BGP
```

annotation

```
kubectl annotate service sample ovn.kubernetes.io/bgp=true
```

annotation

```
kubectl annotate service sample ovn.kubernetes.io/bgp-
```

7.8.4 EIPs

EIPs	VPC NAT Gateway	VpcNatGateway	BGP	BGP Sidecar			
VPC NAT Gateway	BGP	BGP Speaker Sidecar	NetworkAttachmentDefinition	NAD	VPC	Subnet	Sidecar
Kubernetes API	EIPs	VPC	CoreDNS	NAD			

NetworkAttachmentDefinition Subnet provider {nadName}.{nadNamespace}.ovn

```
apiVersion: "k8s.cni.cncf.io/v1"
kind: NetworkAttachmentDefinition
metadata:
  name: api-ovn-nad
  namespace: default
spec:
  config: '{'
    "cniVersion": "0.3.0",
    "type": "kube-ovn",
    "server_socket": "/run/openvswitch/kube-ovn-daemon.sock",
    "provider": "api-ovn-nad.default.ovn"
  '}'
---
apiVersion: kubeovn.io/v1
kind: Subnet
metadata:
  name: vpc-apiserver-subnet
spec:
  protocol: IPv4
  cidrBlock: 100.100.100.0/24
  provider: api-ovn-nad.default.ovn
```

ovn-vpc-nat-config ConfigMap apiNadProvider BGP Speaker :

```
apiVersion: v1
kind: ConfigMap
metadata:
  name: ovn-vpc-nat-config
  namespace: kube-system
data:
  apiNadProvider: api-ovn-nad.default.ovn          # What NetworkAttachmentDefinition provider to use so that the sidecar
                                                    # can access the K8S API, as it can't by default due to VPC segmentation
  bgpSpeakerImage: docker.io/kubeovn/kube-ovn:v1.13.0 # Sets the BGP speaker image used
  image: docker.io/vpc-nat-gateway:v1.13.0
```

ovn-default provider

```
provider: api-ovn-nad.default.ovn
```

VPC NAT Gateway BGP

```
kind: VpcNatGateway
apiVersion: kubeovn.io/v1
metadata:
  name: vpc-natgw
spec:
  vpc: vpc1
  subnet: net1
  lanIp: 10.0.1.10
  bgpSpeaker:
    enabled: true
    asn: 65500
    remoteAsn: 65000
    neighbors:
      - 100.127.4.161
      - fd:01::1
    enableGracefulRestart: true # Optional
```

```

routerId: 1.1.1.1      # Optional
holdTime: 1m           # Optional
password: "password123" # Optional
extraArgs:             # Optional, passed directly to the BGP speaker
  - -v5                # Enables verbose debugging of the BGP speaker sidecar
selector:
  - "kubernetes.io/os: linux"
externalSubnets:
  - ovn-vpc-external-network # Network on which we'll speak BGP and receive/send traffic to the outside world
                                # BGP neighbors need to be on that network

```

BGP EIP

```
kubectl annotate eip sample ovn.kubernetes.io/bgp=true
```

7.8.5

kube-ovn-speaker									
• Cluster:		speaker	Pod IPs/Subnet CIDRs		IP	CIDR	Pod	speaker	Pod
		Pod	Subnet						
• Local:		Pod IPs	Pod	Cluster	Pod				
: Local kube-ovn-speaker Pod speaker									
Cluster Pod/Subnet annotation ovn.kubernetes.io/bgp									
• ovn.kubernetes.io/bgp=cluster		ovn.kubernetes.io/bgp=yes		Cluster					
• ovn.kubernetes.io/bgp=local		Local							
Service	kube-proxy	ClusterIP	Service	Cluster					

7.8.6 BGP

kube-ovn-speaker BGP		
• announce-cluster-ip:	Service	false
• auth-password:	BGP peer	
• holdtime:	BGP	90
• graceful-restart:	BGP Graceful Restart	
• graceful-restart-time:	BGP Graceful restart time	RFC4724 3
• graceful-restart-deferral-time:	BGP Graceful restart deferral time	RFC4724 4.1
• passivemode:	Speaker passive	peer
• ebgp-multipath:	ebgp ttl	1

7.8.7 BGP routes debug

```

# show peer neighbor
gobgp neighbor

# show announced routes to one peer
gobgp neighbor 10.32.32.254 adj-out

```

[PDF](#)
[Slack](#)
[Support](#)

⌚2025 9 10

⌚2022 6 14



7.8.8

7.9 MetalLB Kube-OVN Underlay

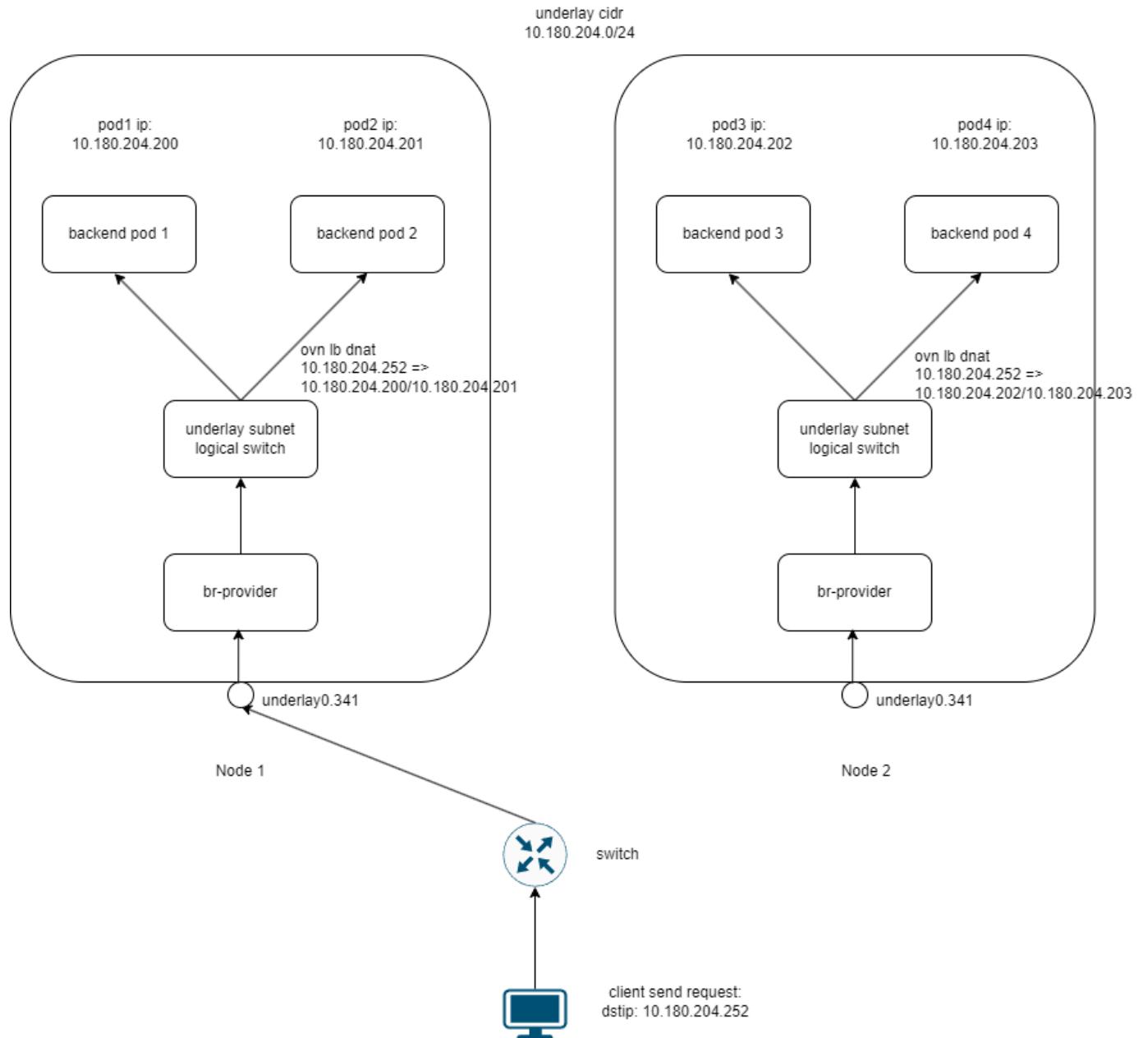


7.9.1

Kube-OVN 1.14.0 MetalLB Underlay

- MetalLB IP
- Pod MetalLB VIP Underlay
- IP SNAT

7.9.2



1 MetalLB VIP Kube-OVN Underlay

MetalLB Kube-OVN Underlay

```

1.          VIP    10.180.204.252   IP  MetalLB    L2           Node1  metallb  VIP
2.          VIP            underlay0.341
3.      br-provider    Underlay
4. br-provider  OpenFlow        OVN
5. underlay subnet      OVN       ovn lb dnat
6. OVN             Pod
10.180.204.0/24  VIP    Pod  IP

```

7.9.3

- Kube-OVN --enable-ovn-lb-prefer-local=true
- Underlay enableExternalLBAddress=true
- Underlay excludeIps MetalLB IP

7.9.4

1. Kube-OVN

```
Kube-OVN  Kube-OVN  --enable-ovn-lb-prefer-local=true  --ls-ct-skip-dst-lport-ips=false
```

```
# kube-ovn-controller Deployment
kubectl edit deployment -n kube-system kube-ovn-controller
```

```
--enable-ovn-lb-prefer-local=true
--ls-ct-skip-dst-lport-ips=false
```

2. Underlay

```
Underlay      LoadBalancer      excludeIps  MetalLB    IP
```

```
apiVersion: kubeovn.io/v1
kind: Subnet
metadata:
  name: underlay-subnet
spec:
  protocol: IPv4
  provider: ovn
  cidrBlock: 10.180.204.0/24 #
  gateway: 10.180.204.1
  excludeIps:
  - 10.180.204.250
  - 10.180.204.251
  - 10.180.204.252 # MetalLB
  natOutgoing: false
  enableExternalLBAddress: true # subnet  cidr  ip      metallb vip
```

3. MetalLB

```
MetalLB      MetalLB
```

```
kubectl apply -f https://raw.githubusercontent.com/metallb/metallb/v0.13.7/config/manifests/metallb-native.yaml
```

```
MetalLB      L2
```

```
apiVersion: metallb.io/v1beta1
kind: IPAddressPool
metadata:
```

```

    name: underlay-pool
    namespace: metallb-system
spec:
  addresses:
  - 10.180.204.250-10.180.204.254 #      VIP 10.180.204.252
---
apiVersion: metallb.io/v1beta1
kind: L2Advertisement
metadata:
  name: l2-advert
  namespace: metallb-system
spec:
  ipAddressPools:
  - underlay-pool

```

4. LoadBalancer Service

LoadBalancer Service Underlay Pod

```

apiVersion: apps/v1
kind: Deployment
metadata:
  labels:
    app: nginx
  name: deploy-16940264
spec:
  replicas: 3
  selector:
    matchLabels:
      app: nginx
  template:
    metadata:
      annotations:
        ovn.kubernetes.io/logical_switch: underlay-subnet
      labels:
        app: nginx
    spec:
      containers:
        - args:
            - netexec
            - --http-port
            - "80"
          image: kubeovn/agnhost:2.47
          imagePullPolicy: IfNotPresent
          name: nginx
---
apiVersion: v1
kind: Service
metadata:
  name: nginx-lb
spec:
  externalTrafficPolicy: Local
  ipFamilies:
  - IPv4
  ipFamilyPolicy: PreferDualStack
  ports:
  - port: 80
    protocol: TCP
    targetPort: 80
  selector:
    app: nginx
  type: LoadBalancer

```

7.9.5

1. Service MetalLB IP

```
kubectl get svc nginx-lb
```

EXTERNAL-IP IP 10.180.204.252

1. Service IP

```
curl http://10.180.204.252
```

1. Pod

Service endpoints Pod

```
#   Service   endpoints
kubectl get endpoints nginx-lb
```

```
# Pod
kubectl get pods -l app=nginx -o wide
```

1. IP

nginx Pod IP IP SNAT IP

```
kubectl exec -it $(kubectl get pods -l app=nginx -o name | head -n1) -- cat /var/log/nginx/access.log
```

7.9.6

IP

MetalLB	Underlay	CIDR	Underlay	excludeIps	IP
---------	----------	------	----------	------------	----

MetalLB	Kube-OVN Underlay	underlay0.341	VLAN	VLAN	ARP	MetalLB VIP
---------	-------------------	---------------	------	------	-----	-------------

- Kube-OVN	--enable-ovn-lb-prefer-local=true	- Service	externalTrafficPolicy: Local
------------	-----------------------------------	-----------	------------------------------

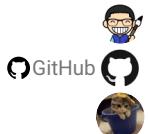
PDF

Slack

Support

2025 9 10

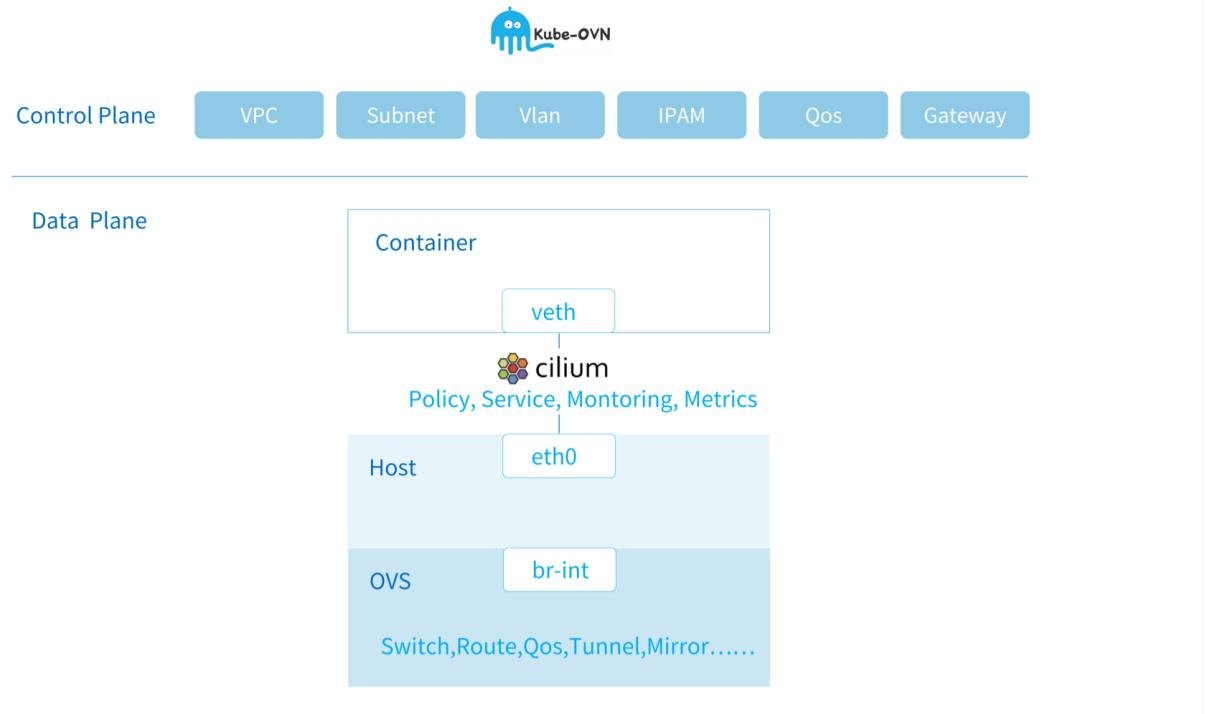
2025 4 2



7.9.7

7.10 Cilium

- Cilium eBPF Kube-OVN CNI Chaining Kube-OVN eBPF
- Cilium Kube-OVN
- Hubble



7.10.1

1. Linux 4.19 eBPF
2. Helm Cilium Helm [Installing Helm](#)

7.10.2 Kube-OVN

Cilium Kube-OVN networkpolicy CNI

```
install.sh
```

```
ENABLE_NP=false
CNI_CONFIG_PRIORITY=10
```

```
kube-ovn-controller networkpolicy
```

```
args:
- --enable-np=false
```

```
kube-ovn-cni CNI
```

```
args:
- --cni-conf-name=10-kube-ovn.conflist
```

Kube-OVN

Cilium

```
mv /etc/cni/net.d/01-kube-ovn.conflist /etc/cni/net.d/10-kube-ovn.conflist
```

7.10.3 Cilium

chaining.yaml Cilium generic-veth

```
apiVersion: v1
kind: ConfigMap
metadata:
  name: cni-configuration
  namespace: kube-system
data:
  cni-config: |-
    {
      "name": "generic-veth",
      "cniVersion": "0.3.1",
      "plugins": [
        {
          "type": "kube-ovn",
          "server_socket": "/run/openvswitch/kube-ovn-daemon.sock",
          "ipam": {
            "type": "kube-ovn",
            "server_socket": "/run/openvswitch/kube-ovn-daemon.sock"
          }
        },
        {
          "type": "portmap",
          "snat": true,
          "capabilities": {"portMappings": true}
        },
        {
          "type": "cilium-cni"
        }
      ]
    }
```

```
kubectl apply -f chaining.yaml
```

Helm Cilium

```
helm repo add cilium https://helm.cilium.io/
helm install cilium cilium/cilium --version 1.11.6 \
  --namespace kube-system \
  --set cni.chainingMode=generic-veth \
  --set cni.customConf=true \
  --set cni.configMap=cni-configuration \
  --set tunnel=disabled \
  --set enableIPv4Masquerade=false \
  --set devices="eth+ ovn0 genev_sys_6081 vxlan_sys_4789" \
  --set enableIdentityMark=false
```

Cilium

```
# cilium status
  /--\
 /--\ /--\   Cilium:      OK
 \--/ \--/   Operator:    OK
 /--\ /--\   Hubble:     disabled
 \--/ \--\   ClusterMesh: disabled
   \--/

DaemonSet      cilium      Desired: 2, Ready: 2/2, Available: 2/2
Deployment     cilium-operator Desired: 2, Ready: 2/2, Available: 2/2
Containers:    cilium      Running: 2
               cilium-operator Running: 2
Cluster Pods: 8/11 managed by Cilium
Image versions cilium      quay.io/cilium/cilium:v1.10.5@sha256:0612218e28288db360c63677c09fafafa2d17edda4f13867bcabf87056046b33bb: 2
               cilium-operator quay.io/cilium/operator-generic:v1.10.5@sha256:2d2f730f219d489ff0702923bf24c0002cd93eb4b47ba344375566202f56d972: 2
```



PDF



Slack



Support

⌚2024 10 8

⌚2022 5 24



7.10.4

7.11 Cilium NetworkPolicy

Kube-OVN	Cilium	Cilium
Cilium	Cilium	Cilium L3 L4

7.11.1

Pod

namespace test yaml test namespace label app=test Pod Pod

```
apiVersion: apps/v1
kind: Deployment
metadata:
  labels:
    app: test
  name: test
  namespace: test
spec:
  replicas: 1
  selector:
    matchLabels:
      app: test
  strategy:
    rollingUpdate:
      maxSurge: 25%
      maxUnavailable: 25%
    type: RollingUpdate
  template:
    metadata:
      labels:
        app: test
    spec:
      containers:
        - image: docker.io/library/nginx:alpine
          imagePullPolicy: IfNotPresent
          name: nginx
```

yaml default namespace label app=dynamic Pod Pod

```
apiVersion: apps/v1
kind: Deployment
metadata:
  labels:
    app: dynamic
  name: dynamic
  namespace: default
spec:
  replicas: 2
  selector:
    matchLabels:
      app: dynamic
  strategy:
    rollingUpdate:
      maxSurge: 25%
      maxUnavailable: 25%
    type: RollingUpdate
  template:
    metadata:
      creationTimestamp: null
    labels:
      app: dynamic
    spec:
      containers:
        - image: docker.io/library/nginx:alpine
          imagePullPolicy: IfNotPresent
          name: nginx
```

Pod Label :

```
# kubectl get pod -o wide --show-labels
NAME           READY   STATUS    RESTARTS   AGE     IP          NODE      NOMINATED NODE   READINESS GATES   LABELS
dynamic-7d8d7874f5-9v5c4   1/1    Running   0          28h    10.16.0.35   kube-ovn-worker   <none>   <none>
template-hash=7d8d7874f5
dynamic-7d8d7874f5-s822n   1/1    Running   0          28h    10.16.0.36   kube-ovn-control-plane   <none>   <none>
template-hash=7d8d7874f5
# kubectl get pod -o wide -n test --show-labels
NAME           READY   STATUS    RESTARTS   AGE     IP          NODE      NOMINATED NODE   READINESS GATES   LABELS
dynamic-7d8d7874f5-6ds96   1/1    Running   0          7h20m  10.16.0.2    kube-ovn-control-plane   <none>   <none>
template-hash=7d8d7874f5
```

```

dynamic-7d8d7874f5-tjgtp  1/1  Running  0   7h46m  10.16.0.42  kube-ovn-worker  <none>    <none>    app=dynamic,pod-
template-hash=7d8d7874f5
label-test1-77b6764857-sq4k  1/1  Running  0   3h43m  10.16.0.12  kube-ovn-worker  <none>    <none>    app=test1,pod-
template-hash=77b6764857

// Pod
test-54c98bc466-mft5s  1/1  Running  0   8h     10.16.0.41  kube-ovn-worker  <none>    <none>    app=test,pod-
template-hash=54c98bc466

```

L3

`yaml CiliumNetworkPolicy :`

```

apiVersion: "cilium.io/v2"
kind: CiliumNetworkPolicy
metadata:
  name: "l3-rule"
  namespace: test
spec:
  endpointSelector:
    matchLabels:
      app: test
  ingress:
  - fromEndpoints:
    - matchLabels:
      app: dynamic

```

default namespace	Pod	Pod	test namespace	Pod
-------------------	-----	-----	----------------	-----

`default namespace :`

```

# kubectl exec -it dynamic-7d8d7874f5-9v5c4 -- bash
bash-5.0# ping -c 3 10.16.0.41
PING 10.16.0.41 (10.16.0.41): 56 data bytes
--- 10.16.0.41 ping statistics ---
3 packets transmitted, 0 packets received, 100% packet loss

```

`test namespace Pod :`

```

# kubectl exec -it -n test dynamic-7d8d7874f5-6dsg6 -- bash
bash-5.0# ping -c 3 10.16.0.41
PING 10.16.0.41 (10.16.0.41): 56 data bytes
64 bytes from 10.16.0.41: seq=0 ttl=64 time=2.558 ms
64 bytes from 10.16.0.41: seq=1 ttl=64 time=0.223 ms
64 bytes from 10.16.0.41: seq=2 ttl=64 time=0.304 ms

--- 10.16.0.41 ping statistics ---
3 packets transmitted, 3 packets received, 0% packet loss
round-trip min/avg/max = 0.223/1.028/2.558 ms

```

Cilium	CiliumNetworkPolicy	Namespace	Cilium
--------	---------------------	-----------	--------

Namespace	Pod	Namespace	Pod
-----------	-----	-----------	-----

Namespace	Namespace
-----------	-----------

`CiliumNetworkPolicy namespace :`

```

ingress:
- fromEndpoints:
  - matchLabels:
    app: dynamic
    k8s.io.kubernetes.pod.namespace: default // Namespace Pod

```

`CiliumNetworkPolicy :`

```

# kubectl get cnp -n test -o yaml l3-rule
apiVersion: cilium.io/v2
kind: CiliumNetworkPolicy
metadata:
  name: l3-rule
  namespace: test
spec:
  endpointSelector:
    matchLabels:
      app: test
  ingress:
  - fromEndpoints:
    - matchLabels:

```

```
app: dynamic
- matchLabels:
  app: dynamic
k8s:io.kubernetes.pod.namespace: default
```

default namespace Pod Pod :

```
# kubectl exec -it dynamic-7d8d7874f5-9v5c4 -n test -- bash
bash-5.0# ping -c 3 10.16.0.41
PING 10.16.0.41 (10.16.0.41): 56 data bytes
64 bytes from 10.16.0.41: seq=0 ttl=64 time=2.383 ms
64 bytes from 10.16.0.41: seq=1 ttl=64 time=0.115 ms
64 bytes from 10.16.0.41: seq=2 ttl=64 time=0.142 ms

--- 10.16.0.41 ping statistics ---
3 packets transmitted, 3 packets received, 0% packet loss
round-trip min/avg/max = 0.115/0.880/2.383 ms
```

Kubernetes	networkpolicy	Cilium	Namespace	Namespace	Pod	Namespace	Namespace	Pod
Kube-OVN	Kube-OVN	k8s	Namespace	Pod	Pod	Namespace	Namespace	Pod
Pod								

L4

yaml L4 :

```
apiVersion: "cilium.io/v2"
kind: CiliumNetworkPolicy
metadata:
  name: "14-rule"
  namespace: test
spec:
  endpointSelector:
    matchLabels:
      app: test
  ingress:
  - fromEndpoints:
    - matchLabels:
      app: dynamic
  toPorts:
  - ports:
    - port: "80"
      protocol: TCP
```

Namespace Pod

```
# kubectl exec -it -n test dynamic-7d8d7874f5-6dsg6 -- bash
bash-5.0# ping -c 3 10.16.0.41
PING 10.16.0.41 (10.16.0.41): 56 data bytes

--- 10.16.0.41 ping statistics ---
3 packets transmitted, 0 packets received, 100% packet loss
bash-5.0#
bash-5.0# curl 10.16.0.41:80
<html>
<head>
  <title>Hello World!</title>
  <link href='//fonts.googleapis.com/css?family=Open+Sans:400,700' rel='stylesheet' type='text/css'>
  <style>
  body {
    background-color: white;
    text-align: center;
    padding: 50px;
    font-family: "Open Sans", "Helvetica Neue", Helvetica, Arial, sans-serif;
  }
  #logo {
    margin-bottom: 40px;
  }
  </style>
</head>
<body>
  <h1>Hello World!</h1>
  <h3>Links found</h3>
  <h3>I am on test-54c98bc466-mft5s</h3>
  <h3>Cookie
    <b>KUBERNETES</b> listening in 443 available at tcp://10.96.0.1:443<br />
    <h3>my name is hanhouchao!</h3>
    <h3> RequestURI='/'</h3>
  </body>
</html>
```

Namespace Pod

```
# kubectl exec -it -n test label-test1-77b6764857-swq4k -- bash
bash-5.0# ping -c 3 10.16.0.41
PING 10.16.0.41 (10.16.0.41): 56 data bytes

--- 10.16.0.41 ping statistics ---
3 packets transmitted, 0 packets received, 100% packet loss
bash-5.0# curl -v 10.16.0.41:80 --connect-timeout 10
* Trying 10.16.0.41:80...
* After 1000ms connect time, move on!
* connect to 10.16.0.41 port 80 failed: Operation timed out
* Connection timeout after 10001 ms
* Closing connection 0
curl: (28) Connection timeout after 10001 ms
```

Namespace L3

L4	ping	TCP
ICMP		L4

L7

chaining	L7	Cilium	Generic Veth Chaining
----------	----	--------	-----------------------

[issue 12454](#)

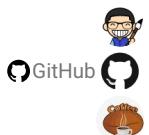
[!\[\]\(b1821d1949cb29f043abc0017902a4f0_img.jpg\) PDF](#)

[!\[\]\(6910059361d054f4732d3bf6459cd67e_img.jpg\) Slack](#)

[!\[\]\(40cad18c4915ebe778453e00aebebcd1_img.jpg\) Support](#)

 2025 9 10

 2022 8 2



7.11.2

7.12 Cilium

```
Kube-OVN      Cilium      Cilium
Cilium          Hubble     Hubble
```

7.12.1 Hubble

```
Cilium      Hubble      Hubble
```

```
helm  Hubble
```

```
helm upgrade cilium cilium/cilium --version 1.11.6 \
--namespace kube-system \
--reuse-values \
--set hubble.relay.enabled=true \
--set hubble.ui.enabled=true
```

```
Hubble      cilium status      Hubble
```

```
# cilium status
  /--\
 /--\_\_/\_--\ Cilium:      OK
 \_\_/\_--\_\_/\ Operator:    OK
 /--\_\_/\_--\ Hubble:      OK
 \_\_/\_--\_\_/\ ClusterMesh: disabled
 \_\_/\

Deployment      hubble-relay      Desired: 1, Ready: 1/1, Available: 1/1
Deployment      cilium-operator   Desired: 2, Ready: 2/2, Available: 2/2
DaemonSet       cilium          Desired: 2, Ready: 2/2, Available: 2/2
Deployment      hubble-ui        Desired: 1, Ready: 1/1, Available: 1/1
Containers:    cilium          Running: 2
                hubble-ui        Running: 1
                hubble-relay     Running: 1
                cilium-operator   Running: 2
Cluster Pods: 16/17 managed by Cilium
Image versions  hubble-relay    quay.io/cilium/hubble-relay:v1.11.6@sha256:fd9034a2d04d5b973f1e8ed44f230ea195b89c37955ff32e34e5aa68f3ed675a: 1
                cilium-operator  quay.io/cilium/operator-generic:v1.11.6@sha256:9f6063c7bcaede801a39315ec7c166389f6a6783e98665f669393cf1701bc17: 2
                cilium          quay.io/cilium/cilium:v1.11.6@sha256:f7f93c26739b6641a3a3d76b1e1605b15989f25d06625260099e01c8243f54c: 2
                hubble-ui        quay.io/cilium/hubble-ui:v0.9.0@sha256:0ef04e9a29212925da6dfdb0a5b581765e41a01fcc30563cef9b30b457fea0: 1
                hubble-ui        quay.io/cilium/hubble-ui-backend:v0.9.0@sha256:000df6b76719f607a9edefb9af94fdf1811a6f1b6a8a9c537cba90bf12df474b: 1
apple@bogon cilium %
```

```
Hubble      Hubble CLI :
```

```
curl -L --fail --remote-name-all https://github.com/cilium/hubble/releases/download/v0.10.0/hubble-linux-amd64.tar.gz
sudo tar xzvfC hubble-linux-amd64.tar.gz /usr/local/bin
```

7.12.2

```
Cilium
```

```
cilium connectivity test Cilium      cilium-test Namespace      cilium-test
```

```
cilium-test namespace
```

```
# kubectl get all -n cilium-test
NAME                  READY  STATUS  RESTARTS  AGE
pod/client-7df6cfbf7b-z5t2j  1/1   Running  0          21s
pod/client2-547996d7d8-nvgxg  1/1   Running  0          21s
pod/echo-other-node-d79544ccf-h14gg  2/2   Running  0          21s
pod/echo-same-node-5d466d5444-m17tc  2/2   Running  0          21s

NAME              TYPE      CLUSTER-IP      EXTERNAL-IP      PORT(S)      AGE
service/echo-other-node  NodePort  10.109.58.126  <none>           8080:32269/TCP  21s
service/echo-same-node  NodePort  10.188.70.32   <none>           8080:32490/TCP  21s

NAME                  READY  UP-TO-DATE  AVAILABLE  AGE
deployment.apps/client  1/1   1          1          21s
deployment.apps/client2  1/1   1          1          21s
deployment.apps/echo-other-node  1/1   1          1          21s
deployment.apps/echo-same-node  1/1   1          1          21s

NAME                  DESIRED  CURRENT  READY  AGE

```

```
replicaset.apps/client-7df6cfbf7b      1     1     1     21s
replicaset.apps/client2-547996d7d8    1     1     1     21s
replicaset.apps/echo-other-node-d79544ccf 1     1     1     21s
replicaset.apps/echo-same-node-5d466d5444 1     1     1     21s
```

7.12.3

Cilium	kube-system namespace	Cilium	pod	hubble observe				
# kubectl get pod -n kube-system -o wide								
NAME	READY	STATUS	RESTARTS	AGE	IP	NODE	NOMINATED NODE	READINESS GATES
cilium-d6h56	1/1	Running	0	2d20h	172.18.0.2	kube-ovn-worker	<none>	<none>
cilium-operator-5887f78bbb-c7sb2	1/1	Running	0	2d20h	172.18.0.2	kube-ovn-worker	<none>	<none>
cilium-operator-5887f78bbb-wj8gt	1/1	Running	0	2d20h	172.18.0.3	kube-ovn-control-plane	<none>	<none>
cilium-tq5xb	1/1	Running	0	2d20h	172.18.0.3	kube-ovn-control-plane	<none>	<none>
kube-ovn-pinger-7lgk8	1/1	Running	0	21h	10.16.0.19	kube-ovn-control-plane	<none>	<none>
kube-ovn-pinger-msvcn	1/1	Running	0	21h	10.16.0.18	kube-ovn-worker	<none>	<none>

```
# kubectl exec -it -n kube-system cilium-d6h56 -- bash
root@kube-ovn-worker:/home/cilium# hubble observe --from-namespace kube-system
Jul 29 03:24:25.551: kube-system/kube-ovn-pinger-msvcn:35576 -> 172.18.0.3:6642 to-stack FORWARDED (TCP Flags: ACK, PSH)
Jul 29 03:24:25.561: kube-system/kube-ovn-pinger-msvcn:35576 -> 172.18.0.3:6642 to-stack FORWARDED (TCP Flags: RST)
Jul 29 03:24:25.561: kube-system/kube-ovn-pinger-msvcn:35576 -> 172.18.0.3:6642 to-stack FORWARDED (TCP Flags: ACK, RST)
Jul 29 03:24:25.572: kube-system/kube-ovn-pinger-msvcn:35578 -> 172.18.0.3:6642 to-stack FORWARDED (TCP Flags: SYN)
Jul 29 03:24:25.572: kube-system/kube-ovn-pinger-msvcn:35578 -> 172.18.0.3:6642 to-stack FORWARDED (TCP Flags: ACK)
Jul 29 03:24:25.651: kube-system/kube-ovn-pinger-msvcn:35578 -> 172.18.0.3:6642 to-stack FORWARDED (TCP Flags: ACK, PSH)
Jul 29 03:24:25.661: kube-system/kube-ovn-pinger-msvcn:35578 -> 172.18.0.3:6642 to-stack FORWARDED (TCP Flags: RST)
Jul 29 03:24:25.661: kube-system/kube-ovn-pinger-msvcn:35578 -> 172.18.0.3:6642 to-stack FORWARDED (TCP Flags: ACK, RST)
Jul 29 03:24:25.761: kube-system/kube-ovn-pinger-msvcn:52004 -> 172.18.0.3:6443 to-stack FORWARDED (TCP Flags: ACK, PSH)
Jul 29 03:24:25.779: kube-system/kube-ovn-pinger-msvcn -> kube-system/kube-ovn-pinger-7lgk8 to-stack FORWARDED (ICMPv4 EchoRequest)
Jul 29 03:24:25.779: kube-system/kube-ovn-pinger-msvcn <- kube-system/kube-ovn-pinger-7lgk8 to-endpoint FORWARDED (ICMPv4 EchoReply)
Jul 29 03:24:25.866: kube-system/hubble-ui-7596f7ffff-7j6f2:55836 <- kube-system/hubble-relay-959988db5-zc5vv:4245 to-stack FORWARDED (TCP Flags: ACK)
Jul 29 03:24:25.866: kube-system/hubble-ui-7596f7ffff-7j6f2:55836 <- kube-system/hubble-relay-959988db5-zc5vv:80 to-endpoint FORWARDED (TCP Flags: ACK)
Jul 29 03:24:25.866: kube-system/hubble-ui-7596f7ffff-7j6f2:55836 -> kube-system/hubble-relay-959988db5-zc5vv:4245 to-stack FORWARDED (TCP Flags: ACK)
Jul 29 03:24:25.866: kube-system/hubble-ui-7596f7ffff-7j6f2:55836 -> kube-system/hubble-relay-959988db5-zc5vv:4245 to-endpoint FORWARDED (TCP Flags: ACK)
Jul 29 03:24:25.975: kube-system/kube-ovn-pinger-7lgk8 -> kube-system/kube-ovn-pinger-msvcn to-endpoint FORWARDED (ICMPv4 EchoRequest)
Jul 29 03:24:25.975: kube-system/kube-ovn-pinger-7lgk8 <- kube-system/kube-ovn-pinger-msvcn to-stack FORWARDED (ICMPv4 EchoReply)
Jul 29 03:24:25.979: kube-system/kube-ovn-pinger-msvcn -> 172.18.0.3 to-stack FORWARDED (ICMPv4 EchoRequest)
Jul 29 03:24:26.037: kube-system/coredns-6d4b75cb6d-lbgjg:36430 -> 172.18.0.3:6443 to-stack FORWARDED (TCP Flags: ACK)
Jul 29 03:24:26.282: kube-system/kube-ovn-pinger-msvcn -> 172.18.0.2 to-stack FORWARDED (ICMPv4 EchoRequest)
```

Hubble Relay Hubble

```
Hubble API          Hubble Service      kubectl port-forward deployment/hubble-relay -n kube-system 4245:4245
```

```
kubectl port-forward
```

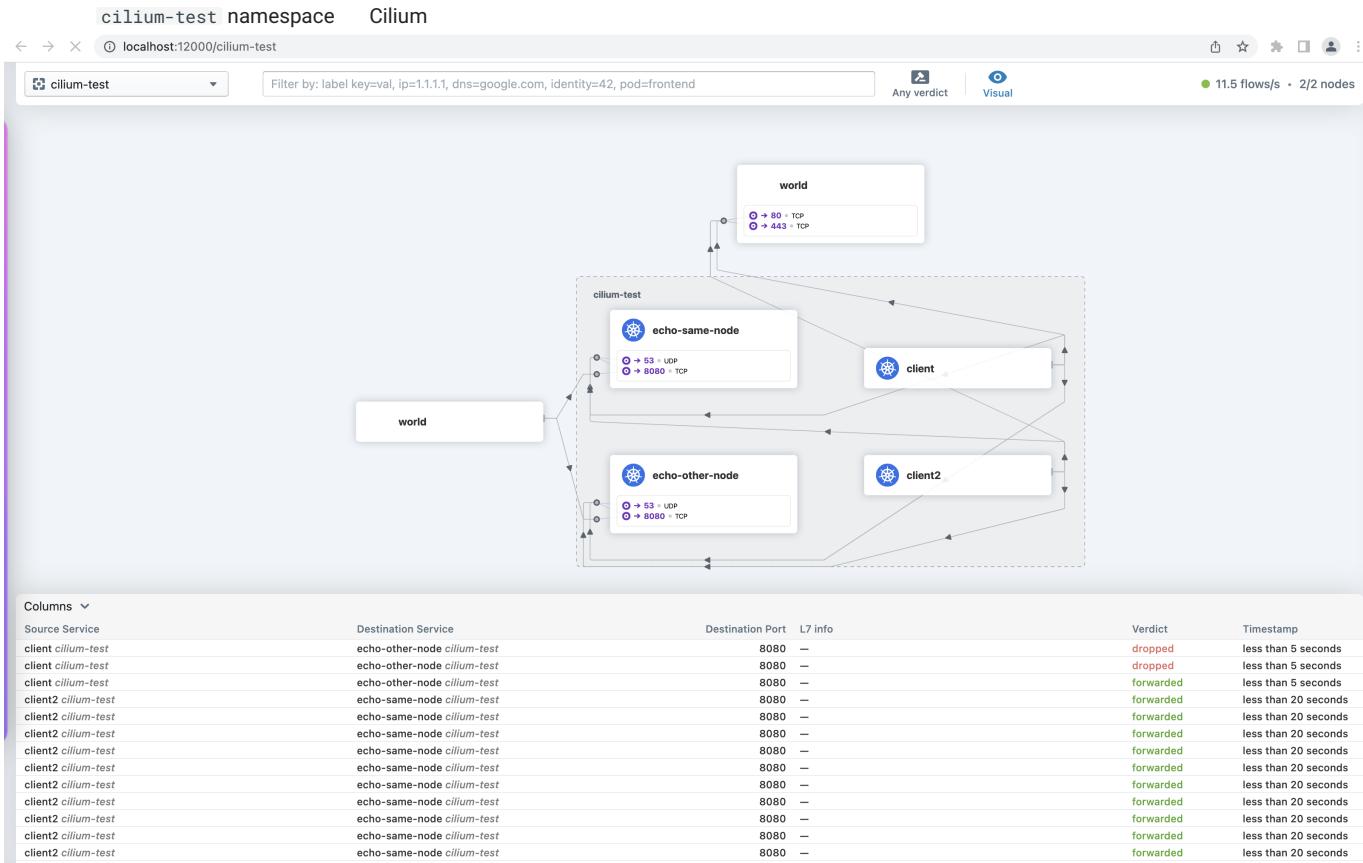
```
hubble status
```

```
# hubble status
Healthcheck (via localhost:4245): Ok
Current/Max Flows: 8,190/8,190 (100.00%)
Flows/s: 22.86
Connected Nodes: 2/2
```

hubble observe **hubble help observe** **Hubble CLI**

7.12.4 UI

```
cilium status      Hubble UI          Hubble           UI  
cilium hubble ui  hubble-ui service  Hubble UI  
http://localhost:12000  UI
```



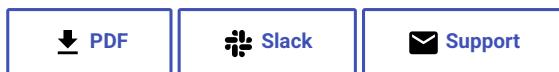
7.12.5 Hubble

Hubble Pod Hubble

```
hubble.metrics.enabled      :  
  
helm upgrade cilium cilium/cilium --version 1.11.6 \  
  --namespace kube-system \  
  --reuse-values \  
  --set hubble.relay.enabled=true \  
  --set hubble.ui.enabled=true \  
  --set hubble.metrics.enabled='{dns.drop.tcp.flow.icmp.http}'
```

kube-system namespace **hubble-metrics** **Endpoints** **Hubble**

```
# curl 172.18.0.2:9091/metrics
# HELP hubble_drop_total Number of drops
# TYPE hubble_drop_total counter
hubble_drop_total{protocol="ICMPv6",reason="Unsupported L3 protocol"} 2
# HELP hubble_flows_processed_total Total number of flows processed
# TYPE hubble_flows_processed_total counter
hubble_flows_processed_total{protocol="ICMPv4",subtype="to-endpoint",type="Trace",verdict="FORWARDED"} 335
hubble_flows_processed_total{protocol="ICMPv4",subtype="to-stack",type="Trace",verdict="FORWARDED"} 335
hubble_flows_processed_total{protocol="ICMPv6",subtype="",type="Drop",verdict="DROPPED"} 2
hubble_flows_processed_total{protocol="TCP",subtype="to-endpoint",type="Trace",verdict="FORWARDED"} 8282
hubble_flows_processed_total{protocol="TCP",subtype="to-stack",type="Trace",verdict="FORWARDED"} 6767
hubble_flows_processed_total{protocol="UDP",subtype="to-endpoint",type="Trace",verdict="FORWARDED"} 1642
hubble_flows_processed_total{protocol="UDP",subtype="to-stack",type="Trace",verdict="FORWARDED"} 1642
# HELP hubble_icmp_total Number of ICMP messages
# TYPE hubble_icmp_total counter
hubble_icmp_total{family="IPv4",type="EchoReply"} 335
hubble_icmp_total{family="IPv4",type="EchoRequest"} 335
hubble_icmp_total{family="IPv4",type="RouterSolicitation"} 2
# HELP hubble_tcp_flags_total TCP flag occurrences
# TYPE hubble_tcp_flags_total counter
hubble_tcp_flags_total{family="IPv4",flag="FIN"} 2043
hubble_tcp_flags_total{family="IPv4",flag="RST"} 301
hubble_tcp_flags_total{family="IPv4",flag="SYN"} 1169
hubble_tcp_flags_total{family="IPv4",flag="SYN-ACK"} 1169
```



⌚2023 3 26

⌚2022 8 2



7.12.6

7.13

Kube-OVN

7.13.1

```
kind: Subnet
apiVersion: kubeovn.io/v1
metadata:
  name: external
spec:
  cidrBlock: 172.31.0.0/16
  gatewayType: centralized
  natOutgoing: false
  externalEgressGateway: 192.168.0.1
  policyRoutingTableID: 1000
  policyRoutingPriority: 1500
```

- natOutgoing: false
- externalEgressGateway
- policyRoutingTableID TableID
- policyRoutingPriority

[!\[\]\(c35531da5557f0569cfc3acf513bb110_img.jpg\) PDF](#)[!\[\]\(646948c7a426ebc69d443a7cc63865bb_img.jpg\) Slack](#)[!\[\]\(b141259fcb519d07f44999390e13667a_img.jpg\) Support](#)

 2022 6 8

 2022 5 24

 GitHub 

7.13.2

7.14 VIP IP

```

VIP   IP      IP      VIP   kube-ovn IP  POD          IP      IP      VIP      Openstack neutron Allowed-Address-Pairs
AAP      Openstack octavia           POD IP      aliyun terway      neutron   IP      VIP      VIP
POD   IP      VIP   IP          OVN   Switch LB      IP   LB   VIP      VIP   OVN
Switch LB Rule      VIP

```

- Allowed-Address-Pairs VIP
- Switch LB rule VIP
- Pod VIP IP

7.14.1 1. Allowed-Address-Pairs VIP

```

IP      Pod
• Kubernetes  Kubernetes  Kubernetes  Underlay  Subnet
• LB          Subnet   IP      Pod
VIP      Allowed-Address-Pairs  IP      IP
• Keepalived  IP

```

1.1 VIP

IP IP yaml

```

apiVersion: kubeovn.io/v1
kind: Vip
metadata:
  name: vip-dynamic-01
spec:
  subnet: ovn-default
  type: ""

```

- subnet: Subnet IP
- type: ipam ip switch_lb_vip vip switch lb vip ip

VIP

```

# kubectl get vip
NAME      V4IP      PV4IP    MAC      PMAC      V6IP      PV6IP    SUBNET    READY
vip-dynamic-01  10.16.0.12  00:00:00:F0:DB:25

```

VIP 10.16.0.12 IP

1.2 VIP

VIP IP yaml

```

apiVersion: kubeovn.io/v1
kind: Vip
metadata:
  name: static-vip01
spec:
  subnet: ovn-default
  v4ip: "10.16.0.121"

```

- subnet: Subnet IP
- v4ip: IP subnet CIDR

VIP

```
# kubectl get vip
NAME      V4IP      PV4IP      MAC          PMAC      V6IP      PV6IP      SUBNET      READY
static-vip01  10.16.0.121           00:00:00:F0:DB:26      ovn-default  true
```

VIP IP

1.3 Pod VIP AAP

Pod	annotation	VIP	AAP	labels	VIP
Pod annotation		VIP			ovn.kubernetes.io/aaps: vip-aap,vip-aap2,vip-aap3
AAP		Pod	AAP	Pod	VIP subnet Port

1.3.1 VIP AAP

```
apiVersion: kubeovn.io/v1
kind: Vip
metadata:
  name: vip-aap
spec:
  subnet: ovn-default
  namespace: default
  selector:
    - "app: aap1"
```

VIP

- namespace : AAP VIP VIP AAP
- selector : AAP VIP Pod Kubernetes NodeSelector

VIP Port

```
# kubectl ko nbctl show ovn-default
switch e32e1d3b-c539-45f4-ab19-be4e33a061f6 (ovn-default)
  port aap-vip
    type: virtual
```

```
apiVersion: v1
kind: Pod
metadata:
  name: busybox
  annotations:
    ovn.kubernetes.io/aaps: vip-aap
  labels:
    app: aap1
spec:
  containers:
    - name: busybox
      image: busybox
      command: ["sleep", "3600"]
      securityContext:
        capabilities:
          add:
            - NET_ADMIN
```

AAP

```
# kubectl ko nbctl list logical_switch_port aap-vip
_uuid      : cd930750-0533-4f06-a6c0-217ddac73272
addresses  : []
dhcpv4_options : []
dhcpv6_options : []
dynamic_addresses : []
enabled   : []
external_ids : {ls=ovn-default, vendor=kube-ovn}
ha_chassis_group : []
mirror_rules : []
name       : aap-vip
options    : {virtual-ip="10.16.0.100", virtual-parents="busybox.default"}
parent_name : []
port_security : []
tag        : []
tag_request : []
type       : virtual
up         : false
```

virtual-ip VIP IP virtual-parents AAP Pod Port

Pod

```
# kubectl exec -it busybox -- ip addr add 10.16.0.100/16 dev eth0
# kubectl exec -it busybox01 -- ip addr show eth0
35: eth0@if36: <BROADCAST,MULTICAST,UP,LOWER_UP,M-DOWN> mtu 1400 qdisc noqueue
    link/ether 00:00:00:e2:ab:0c brd ff:ff:ff:ff:ff:ff
    inet 10.16.0.7/16 brd 10.16.255.255 scope global eth0
        valid_lft forever preferred_lft forever
    inet 10.16.0.100/16 scope global secondary eth0
        valid_lft forever preferred_lft forever
    inet6 fe80::200:ff:fee2:ab0c/64 scope link
        valid_lft forever preferred_lft forever
```

Pod	IP	VIP	IP	subnet	Pod	IP
-----	----	-----	----	--------	-----	----

7.14.2 2. Switch LB rule vip

```
apiVersion: kubeovn.io/v1
kind: Vip
metadata:
  name: slr-01
spec:
  subnet: ovn-default
  type: switch_lb_vip

  • subnet: Subnet IP
  • type: ipam ip switch_lb_vip vip switch lb vip ip
```

7.14.3 3. Pod VIP IP

v1.12

IP

```
apiVersion: kubeovn.io/v1
kind: Vip
metadata:
  name: pod-use-vip
spec:
  subnet: ovn-default
  type: ""
```

annotation VIP Pod

```
apiVersion: v1
kind: Pod
metadata:
  name: static-ip
  annotations:
    ovn.kubernetes.io/vip: pod-use-vip #   vip
    namespace: default
spec:
  containers:
    - name: static-ip
      image: docker.io/library/nginx:alpine
```

3.1 StatefulSet Kubevirt VM VIP

StatefulSet	VM	Pod	VIP
VM	VIP	kube-ovn-controller	keep-vm-ip true Kubevirt VM

[PDF](#)[Slack](#)[Support](#)

2025 9 17

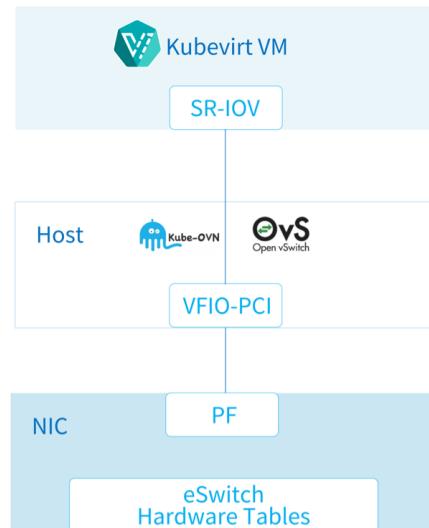
2022 5 24



7.14.4

7.15 Mellanox Offload

Kube-OVN	OVS	CPU	CPU	Mellanox	Accelerated Switching And Packet
Processing (ASAP ²)	OVS	eSwitch	OVS	CPU	



Note

2022

7.15.1

- Mellanox CX5/CX6/CX7/BlueField ASAP²
- CentOS 8 Stream Linux 5.7
- dp_hash hash OVN LB
- bond

7.15.2 SR-IOV Device Plugin

Mellanox	offload	SR-IOV Device Plugin	srivio-network-operator
----------	---------	----------------------	---

SR-IOV Device Plugin

SR-IOV

ID 84:00.0 84.00.1

```
# lspci -nn | grep ConnectX-5
84:00.0 Ethernet controller [0200]: Mellanox Technologies MT27800 Family [ConnectX-5] [15b3:1017]
84:00.1 Ethernet controller [0200]: Mellanox Technologies MT27800 Family [ConnectX-5] [15b3:1017]
```

ID

```
# ls -l /sys/class/net/ | grep 84:00.0
lrwxrwxrwx 1 root root 0 Feb 4 16:16 enp132s0f0np0 -> ../../devices/pci0000:80/0000:80:08.0/0000:84:00.0/net/enp132s0f0np0
# ls -l /sys/class/net/ | grep 84:00.1
lrwxrwxrwx 1 root root 0 Feb 4 16:16 enp132s0f1np1 -> ../../devices/pci0000:80/0000:80:08.0/0000:84:00.1/net/enp132s0f1np1
```

bond

```
enp132s0f0np0  enp132s0f1np1  bond1
```

```
# ip link show enp132s0f0np0 | grep bond
160: enp132s0f0np0: <BROADCAST,MULTICAST,SLAVE,UP,LOWER_UP> mtu 1500 qdisc mq master bond1 state UP mode DEFAULT group default qlen 1000
# ip link show enp132s0f1np1 | grep bond
169: enp132s0f1np1: <BROADCAST,MULTICAST,SLAVE,UP,LOWER_UP> mtu 1500 qdisc mq master bond1 state UP mode DEFAULT group default qlen 1000
```

bond VF

```
ifenslave -d bond1 enp132s0f0np0
ifenslave -d bond1 enp132s0f1np1
echo 0 > /sys/class/net/enp132s0f0np0/device/sriov_numvfs
echo 0 > /sys/class/net/enp132s0f1np1/device/sriov_numvfs
ip link set enp132s0f0np0 down
ip link set enp132s0f1np1 down
```

OVS

SMFS DMFS

- SMFS (software-managed flow steering)
- DMFS (device-managed flow steering)

sysfs devlink API

```
# sysfs
echo <smfs|dmfs> > /sys/class/net/enp132s0f0np0/compat/devlink/steering_mode
echo <smfs|dmfs> > /sys/class/net/enp132s0f1np1/compat/devlink/steering_mode
# devlink
devlink dev param set pci/84:00.0 name flow_steering_mode value smfs cmode runtime
devlink dev param set pci/84:00.1 name flow_steering_mode value smfs cmode runtime
```

VF

```
# cat /sys/class/net/enp132s0f0np0/device/sriov_totalvfs
127
# cat /sys/class/net/enp132s0f1np1/device/sriov_totalvfs
127
```

VF

```
# echo '4' > /sys/class/net/enp132s0f0np0/device/sriov_numvfs
# echo '4' > /sys/class/net/enp132s0f1np1/device/sriov_numvfs
# ip link show enp132s0f0np0
160: enp132s0f0np0: <NO-CARRIER,BROADCAST,MULTICAST,UP> mtu 1500 qdisc mq state DOWN mode DEFAULT group default qlen 1000
    link/ether 00:c0:eb:74:c3:4b brd ff:ff:ff:ff:ff:ff
    vf 0 link/ether 00:00:00:00:00:00 brd ff:ff:ff:ff:ff:ff, spoof checking off, link-state disable, trust off, query_rss off
    vf 1 link/ether 00:00:00:00:00:00 brd ff:ff:ff:ff:ff:ff, spoof checking off, link-state disable, trust off, query_rss off
    vf 2 link/ether 00:00:00:00:00:00 brd ff:ff:ff:ff:ff:ff, spoof checking off, link-state disable, trust off, query_rss off
    vf 3 link/ether 00:00:00:00:00:00 brd ff:ff:ff:ff:ff:ff, spoof checking off, link-state disable, trust off, query_rss off
# ip link show enp132s0f1np1
169: enp132s0f1np1: <NO-CARRIER,BROADCAST,MULTICAST,UP> mtu 1500 qdisc mq state DOWN mode DEFAULT group default qlen 1000
    link/ether 00:c0:eb:74:c3:4b brd ff:ff:ff:ff:ff:ff
    vf 0 link/ether 00:00:00:00:00:00 brd ff:ff:ff:ff:ff:ff, spoof checking off, link-state disable, trust off, query_rss off
    vf 1 link/ether 00:00:00:00:00:00 brd ff:ff:ff:ff:ff:ff, spoof checking off, link-state disable, trust off, query_rss off
    vf 2 link/ether 00:00:00:00:00:00 brd ff:ff:ff:ff:ff:ff, spoof checking off, link-state disable, trust off, query_rss off
    vf 3 link/ether 00:00:00:00:00:00 brd ff:ff:ff:ff:ff:ff, spoof checking off, link-state disable, trust off, query_rss off
# ip link set enp132s0f0np0 up
# ip link set enp132s0f1np1 up
```

VF ID

```
# lspci -nn | grep ConnectX-5 | grep Virtual
84:00.2 Ethernet controller [0200]: Mellanox Technologies MT27800 Family [ConnectX-5 Virtual Function] [15b3:1018]
84:00.3 Ethernet controller [0200]: Mellanox Technologies MT27800 Family [ConnectX-5 Virtual Function] [15b3:1018]
84:00.4 Ethernet controller [0200]: Mellanox Technologies MT27800 Family [ConnectX-5 Virtual Function] [15b3:1018]
```

```
84:00.5 Ethernet controller [0200]: Mellanox Technologies MT27800 Family [ConnectX-5 Virtual Function] [15b3:1018]
84:00.6 Ethernet controller [0200]: Mellanox Technologies MT27800 Family [ConnectX-5 Virtual Function] [15b3:1018]
84:00.7 Ethernet controller [0200]: Mellanox Technologies MT27800 Family [ConnectX-5 Virtual Function] [15b3:1018]
84:01.0 Ethernet controller [0200]: Mellanox Technologies MT27800 Family [ConnectX-5 Virtual Function] [15b3:1018]
84:01.1 Ethernet controller [0200]: Mellanox Technologies MT27800 Family [ConnectX-5 Virtual Function] [15b3:1018]
```

VF

```
echo 0000:84:00.2 > /sys/bus/pci/drivers/mlx5_core/unbind  
echo 0000:84:00.3 > /sys/bus/pci/drivers/mlx5_core/unbind  
echo 0000:84:00.4 > /sys/bus/pci/drivers/mlx5_core/unbind  
echo 0000:84:00.5 > /sys/bus/pci/drivers/mlx5_core/unbind  
echo 0000:84:00.6 > /sys/bus/pci/drivers/mlx5_core/unbind  
echo 0000:84:00.7 > /sys/bus/pci/drivers/mlx5_core/unbind  
echo 0000:84:01.0 > /sys/bus/pci/drivers/mlx5_core/unbind  
echo 0000:84:01.1 > /sys/bus/pci/drivers/mlx5_core/unbind
```

eSwitch

```
devlink dev eswitch set pci@0000:84:00.0 mode switchdev  
devlink dev eswitch set pci@0000:84:00.1 mode switchdev  
ethtool -K enp13s0f0np0 hw-tc-offload on  
ethtool -K enp13s0f1np1 hw-tc-offload on
```

SR-IOV VF

- Active-backup
 - XOR
 - LACP



LACP

```
modprobe bonding mode=802.3ad
ip link set enp132s0f0np0 master bond1
ip link set enp132s0f1np1 master bond1
ip link set enp132s0f0np0 up
ip link set enp132s0f1np1 up
ip link set bond1 up
```



VF

```
echo 0000:84:00.2 > /sys/bus/pci/drivers/mlx5_core/bind  
echo 0000:84:00.3 > /sys/bus/pci/drivers/mlx5_core/bind  
echo 0000:84:00.4 > /sys/bus/pci/drivers/mlx5_core/bind  
echo 0000:84:00.5 > /sys/bus/pci/drivers/mlx5_core/bind  
echo 0000:84:00.6 > /sys/bus/pci/drivers/mlx5_core/bind  
echo 0000:84:00.7 > /sys/bus/pci/drivers/mlx5_core/bind  
echo 0000:84:01.0 > /sys/bus/pci/drivers/mlx5_core/bind  
echo 0000:84:01.1 > /sys/bus/pci/drivers/mlx5_core/bind
```

NetworkManager

NetworkManager

```
systemctl stop NetworkManager  
systemctl disable NetworkManager
```

DEVICE PLUGIN

VF Pod VF SR-IOV Device Plugin

SR-IOV Configmap

```
apiVersion: v1
kind: ConfigMap
metadata:
  name: sriovdp-config
  namespace: kube-system
data:
  config.json: |
```

```
{
  "resourceList": [
    {
      "resourcePrefix": "mellanox.com",
      "resourceName": "cx5_sriov_switchdev",
      "selectors": {
        "vendors": ["15b3"],
        "devices": ["1018"],
        "drivers": ["mlx5_core"]
      }
    }
  ]
}
```

SR-IOV Device Plugin ConfigMap ConfigMap name sriovdp-config

- **selectors:VF**
- **vendors:**
- **devices:**
- **drivers:**

selectors pciAddresses acpiIndexes VF SR-IOV ConfigMap

SR-IOV

```
kubectl apply -f https://raw.githubusercontent.com/k8snetworkplumbingwg/sriov-network-device-plugin/v3.6.2/deployments/sriovdp-daemonset.yaml
```

SR-IOV Kubernetes Node

```
kubectl describe node kube-ovn-01 | grep mellanox

mellanox.com/cx5_sriov_switchdev: 8
mellanox.com/cx5_sriov_switchdev: 8
mellanox.com/cx5_sriov_switchdev 0 0
```

sriov-network-operator SR-IOV Device Plugin

node-feature-discovery

```
kubectl apply -k https://github.com/kubernetes-sigs/node-feature-discovery/deployment/overlays/default?ref=v0.11.3
```

offload annotation:

```
kubectl label nodes [offloadNicNode] feature.node.kubernetes.io/network-sriov.capable=true
```

Operator

```
git clone --depth=1 https://github.com/kubeovn/sriov-network-operator.git
kubectl apply -k sriov-network-operator/deploy
```

Operator

```
# kubectl get -n kube-system all | grep sriov
NAME READY STATUS RESTARTS AGE
pod/sriov-network-config-daemon-bf9nt 1/1 Running 0 8s
pod/sriov-network-operator-54d7545f65-296gb 1/1 Running 0 10s

NAME DESIRED CURRENT READY UP-TO-DATE AVAILABLE NODE
SELECTOR AGE
daemonset.apps/sriov-network-config-daemon 1 1 1 1 1 beta.kubernetes.io/os=linux,feature.node.kubernetes.io/network-sriov.capable=true 8s

NAME READY UP-TO-DATE AVAILABLE AGE
deployment.apps/sriov-network-operator 1/1 1 1 10s

NAME DESIRED CURRENT READY AGE
replicaset.apps/sriov-network-operator-54d7545f65 1 1 1 10s
```

SriovNetworkNodeState node1 Mellanox

```
# kubectl get sriovnetworknodestates.sriovnetwork.openshift.io -n kube-system node1 -o yaml
apiVersion: sriovnetwork.openshift.io/v1
kind: SriovNetworkNodeState
spec: ...
```

```

status:
  interfaces:
    - deviceID: "1017"
      driver: mlx5_core
      mtu: 1500
      pciAddress: "0000:5f:00.0"
      totalvfs: 8
      vendor: "15b3"
      linkSeed: 25000Mb/s
      linkType: ETH
      mac: 08:c0:eb:f4:85:bb
      name: ens41f0np0
    - deviceID: "1017"
      driver: mlx5_core
      mtu: 1500
      pciAddress: "0000:5f:00.1"
      totalvfs: 8
      vendor: "15b3"
      linkSeed: 25000Mb/s
      linkType: ETH
      mac: 08:c0:eb:f4:85:bb
      name: ens41f1np1

```

SriovNetworkNodePolicy nicSelector

```

apiVersion: sriovnetwork.openshift.io/v1
kind: SriovNetworkNodePolicy
metadata:
  name: policy
  namespace: kube-system
spec:
  nodeSelector:
    feature.node.kubernetes.io/network-sriov.capable: "true"
  eSwitchMode: switchdev
  numVfs: 3
  nicSelector:
    pfNames:
      - ens41f0np0
      - ens41f1np1
  resourceName: cx_sriov_switchdev

```

SriovNetworkNodeState status

```

# kubectl get sriovnetworknodestates.sriovnetwork.openshift.io -n kube-system node1 -o yaml

...
spec:
  interfaces:
    - eSwitchMode: switchdev
      name: ens41f0np0
      numVfs: 3
      pciAddress: 0000:5f:00.0
      vfGroups:
        - policyName: policy
          vfRange: 0-2
          resourceName: cx_sriov_switchdev
    - eSwitchMode: switchdev
      name: ens41f1np1
      numVfs: 3
      pciAddress: 0000:5f:00.1
      vfGroups:
        - policyName: policy
          vfRange: 0-2
          resourceName: cx_sriov_switchdev
  status:
    interfaces
      - Vfs:
          - deviceID: 1018
            driver: mlx5_core
            pciAddress: 0000:5f:00.2
            vendor: "15b3"
          - deviceID: 1018
            driver: mlx5_core
            pciAddress: 0000:5f:00.3
            vendor: "15b3"
          - deviceID: 1018
            driver: mlx5_core
            pciAddress: 0000:5f:00.4
            vendor: "15b3"
            deviceID: "1017"
            driver: mlx5_core
            linkSeed: 25000Mb/s
            linkType: ETH
            mac: 08:c0:eb:f4:85:ab
            mtu: 1500
            name: ens41f0np0
            numVfs: 3
            pciAddress: 0000:5f:00.0
            totalvfs: 3
            vendor: "15b3"
      - Vfs:

```

```

- deviceID: 1018
  driver: mlx5_core
  pciAddress: 0000:5f:00.5
  vendor: "15b3"
- deviceID: 1018
  driver: mlx5_core
  pciAddress: 0000:5f:00.6
  vendor: "15b3"
- deviceID: 1018
  driver: mlx5_core
  pciAddress: 0000:5f:00.7
  vendor: "15b3"
deviceID: "1017"
driver: mlx5_core
linkSeed: 25000Mb/s
linkType: ETH
mac: 08:c0:eb:f4:85:bb
mtu: 1500
name: ens41f1np1
numVfs: 3
pciAddress: 0000:5f:00.1
totalvfs: 3
vendor: "15b3"

```

VF

```

# lspci -nn | grep ConnectX
5f:00.0 Ethernet controller [0200]: Mellanox Technologies MT27800 Family [ConnectX-5] [15b3:1017]
5f:00.1 Ethernet controller [0200]: Mellanox Technologies MT27800 Family [ConnectX-5] [15b3:1017]
5f:00.2 Ethernet controller [0200]: Mellanox Technologies MT27800 Family [ConnectX-5 Virtual Function] [15b3:1018]
5f:00.3 Ethernet controller [0200]: Mellanox Technologies MT27800 Family [ConnectX-5 Virtual Function] [15b3:1018]
5f:00.4 Ethernet controller [0200]: Mellanox Technologies MT27800 Family [ConnectX-5 Virtual Function] [15b3:1018]
5f:00.5 Ethernet controller [0200]: Mellanox Technologies MT27800 Family [ConnectX-5 Virtual Function] [15b3:1018]
5f:00.6 Ethernet controller [0200]: Mellanox Technologies MT27800 Family [ConnectX-5 Virtual Function] [15b3:1018]
5f:00.7 Ethernet controller [0200]: Mellanox Technologies MT27800 Family [ConnectX-5 Virtual Function] [15b3:1018]

```

PF

```

# cat /sys/class/net/ens41f0np0/compat/devlink/mode
switchdev

```

7.15.3 Multus-CNI

SR-IOV Device Plugin	ID	Multus-CNI	Kube-OVN	Multus-CNI
----------------------	----	------------	----------	------------

Multus-CNI

```
kubectl apply -f https://raw.githubusercontent.com/k8snetworkplumbingwg/multus-cni/v4.0.2/deployments/multus-daemonset-thick.yml
```

multus	Thin	Thick	SR-IOV	Thick
--------	------	-------	--------	-------

NetworkAttachmentDefinition

```

apiVersion: "k8s.cni.cncf.io/v1"
kind: NetworkAttachmentDefinition
metadata:
  name: sriov
  namespace: default
  annotations:
    k8s.v1.cni.cncf.io/resourceName: mellanox.com/cx5_sriov_switchdev
spec:
  config: '{
    "cniVersion": "0.3.1",
    "name": "kube-ovn",
    "plugins": [
      {
        "type": "kube-ovn",
        "server_socket": "/run/openvswitch/kube-ovn-daemon.sock",
        "provider": "sriov.default.ovn"
      },
      {
        "type": "portmap",
        "capabilities": {
          "portMappings": true
        }
      }
    ]
}'

```

- provider: NetworkAttachmentDefinition {name}.{namespace}.ovn

7.15.4 Overlay

Kube-OVN

```
wget https://raw.githubusercontent.com/kubeovn/kube-ovn/release-1.14/dist/images/install.sh
```

IFACE	IP
-------	----

```
ENABLE_MIRROR=${ENABLE_MIRROR:-false}
HW_OFFLOAD=${HW_OFFLOAD:-true}
ENABLE_LB=${ENABLE_LB:-false}
IFACE="bond1"
#      SR-IoV   Device Plugin          bond    IFACE   bond1   bond    IFACE   enp132s0f0np0   enp132s0f1np1
```

Kube-OVN

```
bash install.sh
```

VF Pod

yaml	VF	Pod:
------	----	------

```
apiVersion: v1
kind: Pod
metadata:
  name: nginx-overlay
  annotations:
    v1.multus-cni.io/default-network: default/sriov
    sriov.default.ovn.kubernetes.io/logical_switch: ovn-default
spec:
  containers:
  - name: nginx-overlay
    image: docker.io/library/nginx:alpine
    resources:
      requests:
        mellanox.com/cx5_sriov_switchdev: '1'
      limits:
        mellanox.com/cx5_sriov_switchdev: '1'
```

- v1.multus-cni.io/default-network: NetworkAttachmentDefinition {namespace}/{name}
- sriov.default.ovn.kubernetes.io/logical_switch: Pod Pod

7.15.5 Underlay

Kube-OVN

```
wget https://raw.githubusercontent.com/kubeovn/kube-ovn/release-1.14/dist/images/install.sh
```

IFACE	IP
-------	----

```
ENABLE_MIRROR=${ENABLE_MIRROR:-false}
HW_OFFLOAD=${HW_OFFLOAD:-true}
ENABLE_LB=${ENABLE_LB:-false}
IFACE=""
#      Underlay   IFACE     PF   IFACE   K8s     PF
```

Kube-OVN

```
bash install.sh
```

VF Pod

yaml	VF	Pod:
------	----	------

```

apiVersion: kubeovn.io/v1
kind: ProviderNetwork
metadata:
  name: underlay-offload
spec:
  defaultInterface: bond1

---
apiVersion: kubeovn.io/v1
kind: Vlan
metadata:
  name: vlan0
spec:
  id: 0
  provider: underlay-offload

---
apiVersion: kubeovn.io/v1
kind: Subnet
metadata:
  name: vlan0
spec:
  protocol: IPv4
  provider: ovn
  cidrBlock: 10.10.204.0/24
  gateway: 10.10.204.254
  vlan: vlan0
  excludeIps:
    - 10.10.204.1..10.10.204.100

---
apiVersion: v1
kind: Pod
metadata:
  name: nginx-underlay
  annotations:
    k8s.v1.cni.cncf.io/networks: '[{
      "name": "sriov",
      "namespace": "default",
      "default-route": ["10.10.204.254"]
    }]'
    ovn.default.ovn.kubernetes.io/logical_switch: vlan0
spec:
  containers:
  - name: nginx-underlay
    image: docker.io/library/nginx:alpine
    resources:
      requests:
        mellanox.com/cx5_sriov_switchdev: '1'
      limits:
        mellanox.com/cx5_sriov_switchdev: '1'

```

• v1.multus-cni.io/default-network: NetworkAttachmentDefinition {namespace}/{name}

	multus	VF	Pod	VF	Pod	VF	Pod	multus
--	--------	----	-----	----	-----	----	-----	--------

yaml VF Pod:

```

apiVersion: v1
kind: Pod
metadata:
  name: nginx-underlay-noVF
  annotations:
    ovn.kubernetes.io/logical_switch: vlan0
spec:
  containers:
  - name: nginx-underlay-noVF
    image: docker.io/library/nginx:alpine

```

	VF	Pod	ovs-kernel	e-switch
--	----	-----	------------	----------

7.15.6

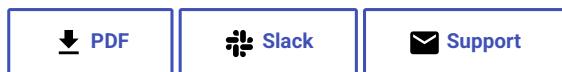
Pod ovs-ovn

```

# ovs-appctl dptctl/dump-flows -m type=offloaded
ufid:91cc45de-e7e9-4935-8f82-1890430b0f66, skb_priority(0/0),skb_mark(0/0),ct_state(0/0x23),ct_zone(0/0),ct_mark(0/0),ct_label(0/0x1),recirc_id(0),dp_hash(0/0),in_port(5b45c61b307e_h),packet_type(ns=0/0,id=0/0),eth(src=00:00:00:c5:6d:4e,dst=00:00:00:e7:16:ce),eth_type(0x0800),ipv4(src=0.0.0.0/0.0.0.0,dst=0.0.0.0/0.0.0.0,proto=0/0,tos=0/0,ttl=0/0,frag=no), packets:941539, bytes:62142230, used:0.260s, offloaded:yes, dp:tc, actions:54235e5753b8_h
ufid:e00768d7-e652-4d79-8182-3291d852b791, skb_priority(0/0),skb_mark(0/0),ct_state(0/0x23),ct_zone(0/0),ct_mark(0/0),ct_label(0/0x1),recirc_id(0),dp_hash(0/0),in_port(54235e5753b8_h),packet_type(ns=0/0,id=0/0),eth(src=00:00:00:e7:16:ce,dst=00:00:00:c5:6d:4e),eth_type(0x0800),ipv4(src=0.0.0.0/0.0.0.0,dst=0.0.0/0.0.0.0,proto=0/0,tos=0/0,ttl=0/0,frag=no), packets:82386659, bytes:115944854173, used:0.260s, offloaded:yes, dp:tc, actions:5b45c61b307e_h

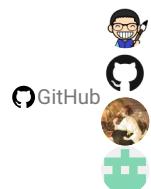
```

offloaded:yes, dp:tc



⌚2025 9 10

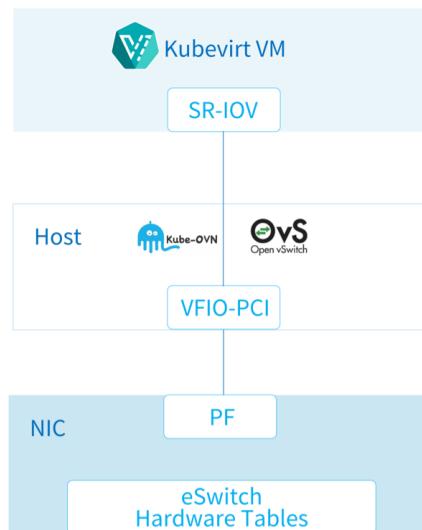
⌚2022 5 24



7.15.7

7.16 Offload

Kube-OVN	OVS	CPU	CPU	Agilio CX	OVS
OVS		CPU			



Note

2022

7.16.1

- Agilio CX
- CentOS 8 Stream Linux 5.7
- dp_hash hash OVN LB

7.16.2 SR-IOV

[Agilio Open vSwitch TC User Guide](#)

```

#!/bin/bash
DEVICE=${1}
DEFAULT_ASSY=scan
ASSY=${2:-$DEFAULT_ASSY}
APP=${3:-flower}

if [ "x${DEVICE}" = "x" -o ! -e /sys/class/net/${DEVICE} ]; then
  echo Syntax: ${0} device [ASSY] [APP]
  echo
  echo This script associates the TC Offload firmware
  echo with a Netronome SmartNIC.
  echo
  echo device: is the network device associated with the SmartNIC

```

```

echo ASSY: defaults to ${DEFAULT_ASSY}
echo APP: defaults to flower. flower-next is supported if updated
echo     firmware has been installed.
exit 1
fi

# It is recommended that the assembly be determined by inspection
# The following code determines the value via the debug interface
if [ "${ASSY}"x = "scanz" ]; then
    ethtool -S ${DEVICE} 0
    DEBUG=$(ethtool -w ${DEVICE} data /dev/stdout | strings)
    SERIAL=$(echo "$DEBUG" | grep '^SN:')
    ASSY=$(echo ${SERIAL} | grep -oE AMDA[0-9]{4})
fi

PCIADDR=$(basename $(readlink -e /sys/class/net/${DEVICE}/device))
FWDIR="/lib/firmware/netronome"

# AMDA0081 and AMDA0097 uses the same firmware
if [ "${ASSY}" = "AMDA0081" ]; then
    if [ ! -e ${FWDIR}/${APP}/nic_AMDA0081.nffw ]; then
        ln -sf nic_AMDA0097.nffw ${FWDIR}/${APP}/nic_AMDA0081.nffw
    fi
fi

FW="${FWDIR}/pci-${PCIADDR}.nffw"
ln -sf "${APP}/nic_${ASSY}.nffw" "${FW}"

# insert distro-specific initramfs section here...

```

```

./agilio-tc-fw-select.sh ens47np0 scan
rmmod nfp
modprobe nfp

```

VF VF

```

# cat /sys/class/net/ens3/device/sriov_totalvfs
65

# echo 4 > /sys/class/net/ens47/device/sriov_numvfs

```

7.16.3 SR-IOV Device Plugin

VF	Pod	VF	SR-IOV Device Plugin
----	-----	----	----------------------

SR-IOV Configmap

```

apiVersion: v1
kind: ConfigMap
metadata:
  name: sriovdp-config
  namespace: kube-system
data:
  config.json: |
    {
      "resourceList": [
        {
          "resourcePrefix": "coragine.com",
          "resourceName": "agilio_sriov",
          "selectors": {
            "vendors": ["19ee"],
            "devices": ["6003"],
            "drivers": ["nfp_netvf"]
          }
        }
      ]
    }

```

SR-IOV :

```
kubectl apply -f https://raw.githubusercontent.com/intel/sriov-network-device-plugin/master/deployments/k8s-v1.16/sriovdp-daemonset.yaml
```

SR-IOV Kubernetes Node

```

kubectl describe no containerserver | grep coragine

coragine.com/agilio_sriov: 4
coragine.com/agilio_sriov: 4
coragine.com/agilio_sriov 0 0

```

7.16.4 Multus-CNI

SR-IOV Device Plugin ID Multus-CNI Kube-OVN Multus-CNI

Multus-CNI

```
kubectl apply -f https://raw.githubusercontent.com/k8snetworkplumbingwg/multus-cni/master/deployments/multus-daemonset.yaml
```

NetworkAttachmentDefinition

```
apiVersion: "k8s.cni.cncf.io/v1"
kind: NetworkAttachmentDefinition
metadata:
  name: default
  namespace: default
  annotations:
    k8s.v1.cni.cncf.io/resourceName: corigine.com/agilio_sriov
spec:
  config: '{
    "cniVersion": "0.3.1",
    "name": "kube-ovn",
    "plugins": [
      {
        "type": "kube-ovn",
        "server_socket": "/run/openvswitch/kube-ovn-daemon.sock",
        "provider": "default.default.ovn"
      },
      {
        "type": "portmap",
        "capabilities": {
          "portMappings": true
        }
      }
    ]
}'

```

- provider: NetworkAttachmentDefinition {name}.{namespace}.ovn

7.16.5 Kube-OVN

```
wget https://raw.githubusercontent.com/kubeovn/kube-ovn/release-1.14/dist/images/install.sh
```

IFACE IP

```
ENABLE_MIRROR=${ENABLE_MIRROR:-false}
HW_OFFLOAD=${HW_OFFLOAD:-true}
ENABLE_LB=${ENABLE_LB:-false}
IFACE="ensp01"
```

Kube-OVN

```
bash install.sh
```

7.16.6 VF Pod

yaml VF Pod:

```
apiVersion: v1
kind: Pod
metadata:
  name: nginx
  namespace: default
  annotations:
    v1.multus-cni.io/default-network: default/default
spec:
  containers:
    - name: nginx
      image: docker.io/library/nginx:alpine
      resources:
        requests:
          corigine.com/agilio_sriov: '1'
```

```

limits:
  coragine.com/agilio_sriov: '1'

• v1.multus-cni.io/default-network:      NetworkAttachmentDefinition {namespace}/{name}

Pod      ovs-ovn

# ovs-appctl dpctl/dump-flows -m type=offloaded
ufid:91cc45de-e7e9-4935-8f82-1890430b0f66, skb_priority(0/0),skb_mark(0/0),ct_state(0/0x23),ct_zone(0/0),ct_mark(0/0),ct_label(0/0x1),recirc_id(0),dp_hash(0/0),in_port(5b45c61b307e_h),packet_type(ns=0/0,id=0/0),eth(src=00:00:00:c5:6d:4e,dst=00:00:00:e7:16:ce),eth_type(0x0800),ipv4(src=0.0.0.0/0.0.0.0,dst=0.0.0/0.0.0.0,proto=0/0,tos=0/0,ttl=0/0,frag=no), packets:941539, bytes:62142230, used:0.260s, offloaded:yes, dp:tc, actions:54235e5753b8_h
ufid:e00768d7-e652-4d79-8182-3291d852b791, skb_priority(0/0),skb_mark(0/0),ct_state(0/0x23),ct_zone(0/0),ct_mark(0/0),ct_label(0/0x1),recirc_id(0),dp_hash(0/0),in_port(54235e5753b8_h),packet_type(ns=0/0,id=0/0),eth(src=00:00:00:e7:16:ce,dst=00:00:00:c5:6d:4e),eth_type(0x0800),ipv4(src=0.0.0.0/0.0.0.0,dst=0.0.0/0.0.0.0,proto=0/0,tos=0/0,ttl=0/0,frag=no), packets:82386659, bytes:115944854173, used:0.260s, offloaded:yes, dp:tc, actions:5b45c61b307e_h

offloaded:yes, dp:tc

```

[PDF](#)[Slack](#)[Support](#)

⌚2025 7 21

⌚2022 5 24

[GitHub](#)

7.16.7

7.17 Offload

Kube-OVN	OVS	CPU	CPU	metaScale	OVS
OVS	CPU				



Note

1. 2024
2. 1.11 Kube-OVN

7.17.1

- metaScale
- MCR
- BIOS SR-IOV VT-d

7.17.2

hw-offload Kube-OVN

1.

```
wget https://github.com/yunsilicon/kube-ovn/blob/release-1.11/dist/images/install.sh
```

1.

```
/opt/ovs-config/ovs-dpdk-config
```

```
# specify log level for ovs dpdk, the value is info or dbg, default is info
VLOG=info
# specify nic offload, the value is true or false, default is true
HW_OFFLOAD=true
# specify cpu mask for ovs dpdk, not specified by default
CPU_MASK=0x02
# specify socket memory, not specified by default
SOCKET_MEM="2048,2048"
# specify encap IP
ENCAP_IP=6.6.208/24
# specify pci device
DPDK_DEV=0000:b3:00.0
# specify mtu, default is 1500
PF_MTU=1500
# specify bond name if bond enabled, not specified by default
BR_PHY_BOND_NAME=bond0
```

1. Kube-OVN

```
bash install.sh
```

SR-IOV

1. metaScale ID b3:00.0:

```
[root@k8s-master ~]# lspci -d 1f67:
b3:00.0 Ethernet controller: Device 1f67:1111 (rev 02)
b3:00.1 Ethernet controller: Device 1f67:1111 (rev 02)
```

1. ID p3p1

```
ls -l /sys/class/net/ | grep b3:00.0
lrwxrwxrwx 1 root root 0 May 7 16:30 p3p1 -> ../../devices/pci0000:b2/0000:b2:00.0/0000:b3:00.0/net/p3p1
```

1. VF

```
cat /sys/class/net/p3p1/device/sriov_totalvfs
512
```

1. VF

```
echo '10' > /sys/class/net/p3p1/device/sriov_numvfs
```

1. VF

```
lspci -d 1f67:
b3:00.0 Ethernet controller: Device 1f67:1111 (rev 02)
b3:00.1 Ethernet controller: Device 1f67:1111 (rev 02)
b3:00.2 Ethernet controller: Device 1f67:1112
b3:00.3 Ethernet controller: Device 1f67:1112
b3:00.4 Ethernet controller: Device 1f67:1112
b3:00.5 Ethernet controller: Device 1f67:1112
b3:00.6 Ethernet controller: Device 1f67:1112
b3:00.7 Ethernet controller: Device 1f67:1112
b3:01.0 Ethernet controller: Device 1f67:1112
b3:01.1 Ethernet controller: Device 1f67:1112
b3:01.2 Ethernet controller: Device 1f67:1112
b3:01.3 Ethernet controller: Device 1f67:1112
```

1. switchdev

```
devlink dev eswitch set pci/0000:b3:00.0 mode switchdev
```

SR-IOV Device Plugin

1. SR-IOV ConfigMap

```
apiVersion: v1
kind: ConfigMap
metadata:
  name: sriovdp-config
  namespace: kube-system
data:
  config.json: |
    {
      "resourceList": [
        {
          "resourceName": "xsc_sriov",
          "resourcePrefix": "yunsilicon.com",
          "selectors": {
            "vendors": ["1f67"],
            "devices": ["1012", "1112"]
          }
        }
      ]
    }
```

1. SR-IOV Device Plugin DevicePlugin

2. SR-IOV

```
# kubectl describe node <node name> | grep yunsilicon.com/xsc_sriov
yunsilicon.com/xsc_sriov: 10
yunsilicon.com/xsc_sriov: 10
yunsilicon.com/xsc_sriov 0 0
```

Multus-CNI

1. Multus-CNI Multus-CNI

```
kubectl apply -f https://raw.githubusercontent.com/k8snetworkplumbingwg/multus-cni/master/deployments/multus-daemonset.yaml
```

1. NetworkAttachmentDefinition

```
apiVersion: "k8s.cni.cncf.io/v1"
kind: NetworkAttachmentDefinition
```

```

metadata:
  name: sriov-net1
  namespace: default
  annotations:
    k8s.v1.cncf.io/resourceName: yunsilicon.com/xsc_sriov
spec:
  config: '{
    "cniVersion": "0.3.1",
    "name": "kube-ovn",
    "plugins": [
      {
        "type": "kube-ovn",
        "server_socket": "/run/openvswitch/kube-ovn-daemon.sock",
        "provider": "sriov-net1.default.ovn"
      },
      {
        "type": "portmap",
        "capabilities": {
          "portMappings": true
        }
      }
    ]
}'

```

SR-IOV Pod

```

apiVersion: v1
kind: Pod
metadata:
  name: nginx
  annotations:
    v1.multus-cni.io/default-network: default/sriov-net1
spec:
  containers:
    - name: nginx
      image: docker.io/library/nginx:alpine
      resources:
        requests:
          yunsilicon.com/xsc_sriov: '1'
        limits:
          yunsilicon.com/xsc_sriov: '1'

```

Offload

Pod ovs-ovn

```

ovs-appctl dptcl/dump-flows type=offloaded
flow-dump from pmd on cpu core: 9
ct_state(-new+est-rel+rpl+trk),ct_mark(0/0x3),recirc_id(0x2d277),in_port(15),packet_type(ns=0,id=0),eth(src=00:00:00:9d:fb:1a,dst=00:00:ce:cf:b9),eth_type(0x0800),ipv4(dst=10.16.0.14,frag=no),packets:6,bytes:588,used:7.276s,actions:ct(zone=4,nat),recirc(0x2d278)
ct_state(-new+est-rel+rpl+trk),ct_mark(0/0x3),recirc_id(0x2d275),in_port(8),packet_type(ns=0,id=0),eth(src=00:00:00:ce:cf:b9,dst=00:00:00:9d:fb:1a),eth_type(0x0800),ipv4(dst=10.16.0.18,frag=no),packets:5,bytes:490,used:7.434s,actions:ct(zone=6,nat),recirc(0x2d276)
ct_state(-new+est-rel-rpl+trk),ct_mark(0/0x1),recirc_id(0x2d276),in_port(8),packet_type(ns=0,id=0),eth(src=00:00:00:ce:cf:b9,dst=00:00:00:9d:fb:1a),eth_type(0x0800),eth_type(0x0800),ipv4(dst=10.16.0.18/255.192.0.0,frag=no),packets:5,bytes:490,used:7.434s,actions:15
recirc_id(0),in_port(15),packet_type(ns=0,id=0),eth(src=00:00:00:9d:fb:1a/01:00:00:00:00:00,dst=00:00:00:ce:cf:b9),eth_type(0x0800),ipv4(dst=10.16.0.14/255.192.0.0,frag=no),packets:6,bytes:588,used:7.277s,actions:ct(zone=6,nat),recirc(0x2d277)
recirc_id(0),in_port(8),packet_type(ns=0,id=0),eth(src=00:00:00:ce:cf:b9/01:00:00:00:00:00,dst=00:00:00:9d:fb:1a),eth_type(0x0800),ipv4(dst=10.16.0.18/255.192.0.0,frag=no),packets:6,bytes:588,used:7.434s,actions:ct(zone=4,nat),recirc(0x2d275)
ct_state(-new+est-rel+rpl+trk),ct_mark(0/0x1),recirc_id(0x2d278),in_port(15),packet_type(ns=0,id=0),eth(dst=00:00:00:ce:cf:b9/01:00:00:00:00:00),eth_type(0x0800),ipv4(frag=no),packets:6,bytes:588,used:7.277s,actions:8

```

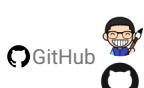
[PDF](#)

[Slack](#)

[Support](#)

⌚2025 7 21

⌚2024 5 29



7.17.3

7.18 Offload



Note

2024

7.18.1

- 2200E
- HADOS
- BIOS SR-IOV

7.18.2

SR-IOV

1. 2200E vendor ID 1f47 ID 00:0a.0 00:0b.0 2200E

```
lspci | grep 1f47
00:0a.0 Ethernet controller: Device 1f47:1001 (rev 10)
00:0b.0 Ethernet controller: Device 1f47:1001 (rev 10)
```

1. VF

```
cat /sys/bus/pci/devices/0000\:00\:0a.0/sriov_totalvfs
256
```

1. VF VF

```
echo 7 > /sys/bus/pci/devices/0000\:00\:0a.0/sriov_numvfs
```

1. VF

```
lspci | grep 1f47
00:0a.0 Ethernet controller: Device 1f47:1001 (rev 10)
00:0a.1 Ethernet controller: Device 1f47:110f (rev 10)
00:0a.2 Ethernet controller: Device 1f47:110f (rev 10)
00:0a.3 Ethernet controller: Device 1f47:110f (rev 10)
00:0a.4 Ethernet controller: Device 1f47:110f (rev 10)
00:0a.5 Ethernet controller: Device 1f47:110f (rev 10)
00:0a.6 Ethernet controller: Device 1f47:110f (rev 10)
00:0a.7 Ethernet controller: Device 1f47:110f (rev 10)
00:0b.0 Ethernet controller: Device 1f47:1001 (rev 10)
```

SR-IOV Device Plugin

1. SR-IOV Configmap SR-IOV Device Plugin VF Pod

```
apiVersion: v1
kind: ConfigMap
metadata:
  name: sriovdp-config
  namespace: kube-system
data:
  config.json: |
    {
      "resourceList": [
        {
          "resourceName": "sriov_dpu",
          "resourcePrefix": "yusur.tech",
          "selectors": {
            "vendors": ["1f47"],
            "devices": ["110f"]
          }
        }
      ]
    }
```

```
    ]  
}
```

1. SR-IOV Device Plugin

```
kubectl apply -f https://raw.githubusercontent.com/k8snetworkplumbingwg/sriov-network-device-plugin/v3.6.2/deployments/sriovdp-daemonset.yaml
```

1. SR-IOV kubernetes Node

```
kubectl describe node node1 | grep yusur  
yusur.tech/sriov_dpu: 7  
yusur.tech/sriov_dpu: 7  
yusur.tech/sriov_dpu 0 0
```

7.18.3 Multus-CNI

Multus-CNI Kube-OCN SRIOV Device ID

```
kubectl apply -f https://raw.githubusercontent.com/k8snetworkplumbingwg/multus-cni/v4.0.2/deployments/multus-daemonset-thick.yaml
```

NetworkAttachmentDefinition

```
apiVersion: "k8s.cni.cncf.io/v1"  
kind: NetworkAttachmentDefinition  
metadata:  
  name: test  
  namespace: kube-system  
  annotations:  
    k8s.v1.cni.cncf.io/resourceName: yusur.tech/sriov_dpu  
spec:  
  config: '{  
    "cniVersion": "0.3.1",  
    "name": "kube-ovn",  
    "plugins": [  
      {  
        "type": "kube-ovn",  
        "server_socket": "/run/openvswitch/kube-ovn-daemon.sock",  
        "provider": "test.kube-system.ovn"  
      },  
      {  
        "type": "portmap",  
        "capabilities": {  
          "portMappings": true  
        }  
      }  
    ]  
  }
```

- provider : NetworkAttachmentDefinition {name}. {namespace}. ovn

7.18.4 Kube-OVN

1.

```
wget https://github.com/kubeovn/kube-ovn/blob/release-1.12/dist/images/install.sh
```

1. IFACE IP

```
ENABLE_MIRROR=${ENABLE_MIRROR:-false}  
HW_OFFLOAD=${HW_OFFLOAD:-true}  
ENABLE_LB=${ENABLE_LB:-false}  
IFACE="p0"
```

1. kube-ovn

```
bash install.sh
```

VF pod

yaml VF Pod

```

apiVersion: v1
kind: Pod
metadata:
  name: nginx
  namespace: default
  annotations:
    v1.multus-cni.io/default-network: kube-system/test
spec:
  containers:
    - name: nginx
      image: docker.io/library/nginx:alpine
      resources:
        requests:
          yusur.tech/sriov_dpu: '1'
        limits:
          yusur.tech/sriov_dpu: '1'

```

- v1.multus-cni.io/default-network : NetworkAttachmentDefinition {namespace}/{name}

Offload

Pod ovs-ovn

```

# ovs-appctl dpctl/dump-flows -m type=offloaded
ufid:67c2e10f-92d4-4574-be70-d072815ff166, skb_priority(0/0),skb_mark(0/0),ct_state(0/0x23),ct_zone(0/0),ct_mark(0/0),ct_label(0/0),recirc_id(0),dp_hash(0/0),in_port(d85b161b6840_h),packet_type(ns=0/0,id=0/0),eth(src=0a:c9:1c:70:01:09,dst=8a:18:a4:22:b7:7d),eth_type(0x0800),ipv4(src=10.0.1.10,dst=10.0.1.6,proto=6,tos=0/0x3,ttl=0/0,frag=no),tcp(src=60774,dst=9001),packets:75021,bytes:109521630,offload_packets:75019,offload_bytes:109521498,used:3.990s,offloaded:yes,dp:tc,actions:set(tunnel(tun_id=0x5,dst=192.168.201.12,ttl=64,tp_dst=6081,geneve({class=0x102,type=0x80,len=4,0xa0006}),flags(csum(key))),genev_sys_6081
ufid:7940666e-a0bd-42a5-8116-1e84e81bb338, skb_priority(0/0),tunnel(tun_id=0x5,src=192.168.201.12,dst=192.168.201.11,ttl=0/0,tp_dst=6081,geneve({class=0x102,type=0x80,len=4,0x6000a}),flags(+key)),skb_mark(0/0),ct_state(0/0),ct_zone(0/0),ct_mark(0/0),ct_label(0/0),recirc_id(0),dp_hash(0/0),in_port(genev_sys_6081),packet_type(ns=0/0,id=0/0),eth(src=8a:18:a4:22:b7:7d,dst=0a:c9:1c:70:01:09),eth_type(0x0800),ipv4(src=10.0.1.6,dst=10.0.1.10,proto=6,tos=0/0,ttl=0/0,frag=no),tcp(src=9001,dst=60774),packets:6946,bytes:459664,offload_packets:6944,offload_bytes:459532,used:4.170s,dp:tc,offloaded:yes,actions:d85b161b6840_h

```



⌚2025 7 21

⌚2024 8 13

GitHub

7.18.5

7.19 DPDK

Kube-OVN OVS-DPDK KubeVirt DPDK

KubeVirt OVS-DPDK patchVhostuser implementation KubeVirt **KVM Device Plugin** OVS-DPDK

7.19.1

- DPDK
 - Hugepages

7.19.2 DPDK

`driverctl` DPDK

```
driverctl set-override 0000:00:0b.0 uio_pci_generic
```

7.19.3

```
kubectl label nodes <node> ovn.kubernetes.io/ovs_dp_type="userspace"
```

ovs-dpdk /opt/ovs-config ovs-dpdk-config

ENCAP_IP=192.168.122.193/24
DPDK_DEV=0000:00:0b.0

- ENCAP_IP :
 - DPDK_DEV : PCI ID

7.19.4 Kube-OVN

```
 wget https://raw.githubusercontent.com/kubeovn/kube-ovn/release-1.14/dist/images/install.sh
```

DPDK

```
bash install.sh --with-hybrid-dpdk
```

7.19.5

vhostuser OVS-DPDK

KVM Device Plugin KVM Device Plugin

```
kubectl apply -f https://raw.githubusercontent.com/kubevirt/kubernetes-device-plugins/master/manifests/kvm-ds.yaml
```

NetworkAttachmentDefinition

```
apiVersion: k8s.cni.cncf.io/v1
kind: NetworkAttachmentDefinition
metadata:
  name: ovn-dpdk
  namespace: default
spec:
  config: >-
    {
      "cniVersion": "0.3.0",
      "tvsne": "kube-ovn"
```

```

        "server_socket": "/run/openvswitch/kube-ovn-daemon.sock",
        "provider": "ovn-dpdk.default.ovn",
        "vhost_user_socket_volume_name": "vhostuser-sockets",
        "vhost_user_socket_name": "sock"
    }
}
```

Dockerfile VM

```

FROM quay.io/kubenvirt/virt-launcher:v0.46.1

# wget http://cloud.centos.org/centos/7/images/CentOS-7-x86_64-GenericCloud.qcow2
COPY CentOS-7-x86_64-GenericCloud.qcow2 /var/lib/libvirt/images/CentOS-7-x86_64-GenericCloud.qcow2

```

```

apiVersion: v1
kind: ConfigMap
metadata:
  name: vm-config
data:
  start.sh: |
    chmod u+w /etc/libvirt/qemu.conf
    echo "hugepages_mount = \"/dev/hugepages\" >> /etc/libvirt/qemu.conf
    virtlogd &
    libvirtd &

    mkdir /var/lock

    sleep 5

    virsh define /root/vm/vm.xml
    virsh start vm

    tail -f /dev/null
  vm.xml: |
    <domain type='kvm'>
      <name>vm</name>
      <uuid>4a9b3f53-fa2a-47f3-a757-dd87720d9d1d</uuid>
      <memory unit='KiB'>2097152</memory>
      <currentMemory unit='KiB'>2097152</currentMemory>
      <memoryBacking>
        <hugepages>
          <page size='2' unit='M' nodeset='0' />
        </hugepages>
      </memoryBacking>
      <vcpu placement='static'>2</vcpu>
      <cpurtune>
        <shares>4096</shares>
        <vcpu pin vcpu='0' cpuset='4' />
        <vcpu pin vcpu='1' cpuset='5' />
        <emulatorpin cpuset='1,3' />
      </cpurtune>
      <os>
        <type arch='x86_64' machine='pc'>hvm</type>
        <boot dev='hd' />
      </os>
      <features>
        <acpi/>
        <apic/>
      </features>
      <cpu mode='host-model'>
        <model fallback='allow' />
        <topology sockets='1' cores='2' threads='1' />
        <numa>
          <cell id='0' cpus='0-1' memory='2097152' unit='KiB' memAccess='shared' />
        </numa>
      </cpu>
      <on_reboot>restart</on_reboot>
      <devices>
        <emulator>/usr/libexec/qemu-kvm</emulator>
        <disk type='file' device='disk'>
          <driver name='qemu' type='qcow2' cache='none' />
          <source file='/var/lib/libvirt/images/CentOS-7-x86_64-GenericCloud.qcow2' />
          <target dev='vda' bus='virtio' />
        </disk>
        <interface type='vhostuser'>
          <mac address='00:00:00:0A:30:89' />
          <source type='unix' path='/var/run/vm.sock' mode='server' />
          <model type='virtio' />
          <driver queues='2'>
            <host mrg_rxbuf='off' />
          </driver>
        </interface>
        <serial type='pty'>
          <target type='isa-serial' port='0'>
            <model name='isa-serial' />
          </target>
        </serial>
        <console type='pty'>
          <target type='serial' port='0' />
        </console>
      </devices>
    </domain>

```

```

</console>
<channel type='unix'>
  <source mode='bind' path='/var/lib/libvirt/qemu/channel/target/domain-1-vm/org.qemu.guest_agent.0' />
  <target type='virtio' name='org.qemu.guest_agent.0' state='connected' />
  <alias name='channel0' />
</channel>

</devices>
</domain>
---
apiVersion: apps/v1
kind: Deployment
metadata:
  name: vm-deployment
  labels:
    app: vm
spec:
  replicas: 1
  selector:
    matchLabels:
      app: vm
  template:
    metadata:
      labels:
        app: vm
      annotations:
        k8s.v1.cni.cncf.io/networks: default/ovn-dpdk
        ovn-dpdk.default.ovn.kubernetes.io/ip_address: 10.16.0.96
        ovn-dpdk.default.ovn.kubernetes.io/mac_address: 00:00:00:0A:30:89
  spec:
    nodeSelector:
      ovn.kubernetes.io/ovs_dp_type: userspace
    securityContext:
      runAsUser: 0
    volumes:
      - name: vhostuser-sockets
        emptyDir: {}
      - name: xml
        configMap:
          name: vm-config
      - name: hugepage
        emptyDir:
          medium: HugePages-2Mi
      - name: libvirt-runtime
        emptyDir: {}
    containers:
      - name: vm
        image: vm-vhostuser:latest
        command: ["bash", "/root/vm/start.sh"]
        securityContext:
          capabilities:
            add:
              - NET_BIND_SERVICE
              - SYS_NICE
              - NET_RAW
              - NET_ADMIN
        privileged: false
        runAsUser: 0
        resources:
          limits:
            cpu: '2'
            devices.kubevirt.io/kvm: '1'
            memory: '8784969729'
            hugepages-2Mi: 2Gi
          requests:
            cpu: 666m
            devices.kubevirt.io/kvm: '1'
            ephemeral-storage: 50M
            memory: '4490002433'
        volumeMounts:
          - name: vhostuser-sockets
            mountPath: /var/run/vm
          - name: xml
            mountPath: /root/vm/
          - mountPath: /dev/hugepages
            name: hugepage
          - name: libvirt-runtime
            mountPath: /var/run/libvirt

```

Pod

```

# virsh set-user-password vm root 12345
Password set successfully for root in vm

# virsh console vm
Connected to domain 'vm'
Escape character is ^] (Ctrl + ])

CentOS Linux 7 (Core)
Kernel 3.10.0-1127.el7.x86_64 on an x86_64

```

```
localhost login: root
Password:
Last login: Fri Feb 25 09:52:54 on ttys0
```

```
ip link set eth0 mtu 1400
ip addr add 10.16.0.96/16 dev eth0
ip ro add default via 10.16.0.1
ping 114.114.114.114
```

[PDF](#)[Slack](#)[Support](#)

⌚2022 7 16

⌚2022 5 24

⌚GitHub 

7.19.6

7.20 OpenStack



7.20.1



1. OpenStack Kubernetes CIDR
- 2.
3. IP
4. Kubernetes OpenStack VPC

OVN-IC

OVN-IC

```
docker run --name=ovn-ic-db -d --network=host -v /etc/ovn/:/etc/ovn -v /var/run/ovn:/var/run/ovn -v /var/log/ovn:/var/log/ovn kubeovn/kube-ovn:v1.14.23 bash start-ic-db.sh
```

Kubernetes

kube-system Namespace ovn-ic-config ConfigMap

```
apiVersion: v1
kind: ConfigMap
metadata:
  name: ovn-ic-config
  namespace: kube-system
data:
  enable-ic: "true"
  az-name: "az1"
  ic-db-host: "192.168.65.3"
  ic-nb-port: "6645"
  ic-sb-port: "6646"
  gw-nodes: "az1-gw"
  auto-route: "true"
```

- enable-ic:
- az-name:
- ic-db-host: OVN-IC
- ic-nb-port: OVN-IC 6645
- ic-sb-port: OVN-IC 6646
- gw-nodes:
- auto-route:

OpenStack

Kubernetes

```
# openstack router create router0
# openstack router list
+-----+-----+-----+-----+
| ID      | Name   | Status | State | Project |
+-----+-----+-----+-----+
| d5b38655-249a-4192-8046-71aa4d2b4af1 | router0 | ACTIVE | UP    | 98a29ab7388347e7b5ff8bdd181ba4f9 |
+-----+-----+-----+-----+
```

OpenStack OVN

```
ovn-nbctl set NB_Global . name=op-az
```

OVN-IC OVN-IC

```
/usr/share/ovn/scripts/ovn-ctl --ovn-ic-nb-db=tcp:192.168.65.3:6645 \
--ovn-ic-sb-db=tcp:192.168.65.3:6646 \
--ovn-northd-nb-db=unix:/run/ovnnb_db.sock \
--ovn-northd-sb-db=unix:/run/ovn/ovnsb_db.sock \
start_ic
```

- ovn-ic-nb-db ovn-ic-sb-db : OVN-IC
- ovn-northd-nb-db ovn-northd-sb-db : OVN

```
ovs-vsctl set open_vswitch . external_ids:ovn-is-interconn=true
```

OpenStack OVN

ts router0

```
ovn-nbctl lrp-add router0 lrp-router0-ts 00:02:ef:11:39:4f 169.254.100.73/24
ovn-nbctl lsp-add ts lsp-ts-router0 -- lsp-set-addresses lsp-ts-router0 router \
-- lsp-set-type lsp-ts-router0 router \
-- lsp-set-options lsp-ts-router0 router-port=lrp-router0-ts
ovn-nbctl lrp-set-gateway-chassis lrp-router0-ts {gateway chassis} 1000
ovn-nbctl set NB_Global . options:ic-route-adv=true options:ic-route-learn=true
```

Kubernetes

```
# ovn-nbctl lr-route-list router0
IPv4 Routes
      10.0.0.22          169.254.100.34 dst-ip (learned)
      10.16.0.0/16        169.254.100.34 dst-ip (learned)
```

router0 Kubernetes Pod

7.20.2 OVN

OpenStack	Kubernetes	OVN	VPC	Subnet			
Kube-OVN	OVN	OpenStack	Neutron	OVN	OpenStack	networking-ovn	Neutron

Neutron

Neutron /etc/neutron/plugins/ml2/ml2_conf.ini

```
[ovn]
...
ovn_nb_connection = tcp:[192.168.137.176]:6641,tcp:[192.168.137.177]:6641,tcp:[192.168.137.178]:6641
ovn_sb_connection = tcp:[192.168.137.176]:6642,tcp:[192.168.137.177]:6642,tcp:[192.168.137.178]:6642
ovn_l3_scheduler = OVN_L3_SCHEDULER
```

- ovn_nb_connection ovn_sb_connection : Kube-OVN ovn-central

OVS

```
ovs-vsctl set open . external_ids:ovn-remote=tcp:[192.168.137.176]:6642,tcp:[192.168.137.177]:6642,tcp:[192.168.137.178]:6642
ovs-vsctl set open . external_ids:ovn-encap-type=geneve
ovs-vsctl set open . external_ids:ovn-encap-ip=192.168.137.200
```

- external-ids:ovn-remote : Kube-OVN ovn-central
- ovn-encap-ip : IP

Kubernetes OpenStack

Kubernetes OpenStack OpenStack Pod



kube-ovn-controller args --enable-external-vpc=true

OpenStack

```
# openstack router list
+-----+-----+-----+
| ID      | Name   | Status | State | Project |
+-----+-----+-----+
| 22040ed5-0598-4f77-bffd-e7fd4db47e93 | router0 | ACTIVE | UP    | 62381a21d569404aa236a5d8712449c |
+-----+-----+-----+
# openstack network list
+-----+-----+
| ID      | Name   | Subnets |
+-----+-----+
| cd59e36a-37db-4c27-b709-d35379a7920f | provider | 01d73d9f-fdaa-426c-9b60-aa34abbfaeae |
+-----+-----+
# openstack subnet list
+-----+-----+-----+
| ID      | Name       | Network          | Subnet |
+-----+-----+-----+
| 01d73d9f-fdaa-426c-9b60-aa34abbfaeae | provider-v4 | cd59e36a-37db-4c27-b709-d35379a7920f | 192.168.1.0/24 |
+-----+-----+-----+
# openstack server list
+-----+-----+-----+-----+-----+
| ID      | Name       | Status | Networks | Image | Flavor |
+-----+-----+-----+-----+-----+
| 8433d622-a8d6-41a7-8b31-49abfd64f639 | provider-instance | ACTIVE | provider=192.168.1.61 | ubuntu | m1 |
+-----+-----+-----+-----+-----+
```

Kubernetes VPC

```
# kubectl get vpc
NAME           STANDBY   SUBNETS
neutron-22040ed5-0598-4f77-bffd-e7fd4db47e93  true      [ "neutron-cd59e36a-37db-4c27-b709-d35379a7920f" ]
ovn-cluster     true      [ "join", "ovn-default" ]
```

neutron-22040ed5-0598-4f77-bffd-e7fd4db47e93 OpenStack VPC

Kube-OVN VPC Subnet Pod

VPC, Subnet Namespace net2 Pod:

```
apiVersion: v1
kind: Namespace
metadata:
  name: net2
---
apiVersion: kubeovn.io/v1
kind: Vpc
metadata:
  creationTimestamp: "2021-06-20T13:34:11Z"
  generation: 2
  labels:
    ovn.kubernetes.io/vpc_external: "true"
  name: neutron-22040ed5-0598-4f77-bffd-e7fd4db47e93
  resourceVersion: "583728"
  uid: 18d4c654-f511-4def-a3a0-a6434d237c1e
spec:
  namespaces:
    - net2
---
kind: Subnet
apiVersion: kubeovn.io/v1
metadata:
  name: net2
spec:
  vpc: neutron-22040ed5-0598-4f77-bffd-e7fd4db47e93
  namespaces:
    - net2
  cidrBlock: 12.0.1.0/24
  natOutgoing: false
---
apiVersion: v1
kind: Pod
metadata:
  name: ubuntu
```

```
namespace: net2
spec:
  containers:
    - image: docker.io/kubeovn/kube-ovn:v1.8.0
      command:
        - "sleep"
        - "604800"
      imagePullPolicy: IfNotPresent
      name: ubuntu
      restartPolicy: Always
```

[!\[\]\(bcaaa3cc7589969bf4de6e41f73e3304_img.jpg\) PDF](#)[!\[\]\(1ba5d29adb699630ac108f62c98be489_img.jpg\) Slack](#)[!\[\]\(807aeb0f6a0e2b74b0719356dbebb138_img.jpg\) Support](#)

 2025 8 4

 2022 5 24



7.20.3

7.21 IPsec

v1.13.0 UDP 500 4500

7.21.1

```
kube-ovn-cni certificatesigningrequest kube-ovn-controller kube-ovn-controller approve kube-ovn-cni ipsec  
ipsec
```

7.21.2 IPsec

```
kube-ovn-controller kube-ovn-cni args --enable-ovn-ipsec=false --enable-ovn-ipsec=true
```

[!\[\]\(8cd29b43af91d7e702ccb3eb7bd5f31a_img.jpg\) PDF](#)[!\[\]\(6dcfcdf465199dda31bcb17c4dcfddf2_img.jpg\) Slack](#)[!\[\]\(ca1f731e7a5d5f6abb630def62f74629_img.jpg\) Support](#)

 2025 9 10

 2023 4 18



 GitHub



7.21.3

7.22 OVN

Pod	GRE/ERSPAN
Kube-OVN	v1.12

7.22.1 Multus-CNI

Multus-CNI Multus

7.22.2

```
apiVersion: "k8s.cni.cncf.io/v1"
kind: NetworkAttachmentDefinition
metadata:
  name: attachnet
  namespace: default
spec:
  config: |
    {
      "cniVersion": "0.3.1",
      "type": "kube-ovn",
      "server_socket": "/run/openvswitch/kube-ovn-daemon.sock",
      "provider": "attachnet.default.ovn"
    }
}
```

provider <NAME>.<NAMESPACE>.ovn

7.22.3 Underlay

MTU LSP/Pod Underlay

Underlay

```
apiVersion: kubeovn.io/v1
kind: ProviderNetwork
metadata:
  name: net1
spec:
  defaultInterface: eth1
---
apiVersion: kubeovn.io/v1
kind: Vlan
metadata:
  name: vlan1
spec:
  id: 0
  provider: net1
---
apiVersion: kubeovn.io/v1
kind: Subnet
metadata:
  name: subnet1
spec:
  protocol: IPv4
  cidrBlock: 172.19.0.0/16
  excludeIps:
  - 172.19.0.2..172.19.0.20
  gateway: 172.19.0.1
  vlan: vlan1
  provider: attachnet.default.ovn
```

provider provider

7.22.4 Pod

Pod

```
apiVersion: v1
kind: Pod
```

```

metadata:
  name: pod1
  annotations:
    k8s.v1.cni.cncf.io/networks: default/attachnet
spec:
  containers:
  - name: bash
    image: docker.io/kubeovn/kube-ovn:v1.14.23
    args:
    - bash
    - -c
    - sleep infinity
    securityContext:
      privileged: true

```

Pod IP

```
$ kubectl get ips | grep pod1
pod1.default          10.16.0.12  00:00:00:FF:34:24  kube-ovn-worker  ovn-default
pod1.default.attachnet.default.ovn  172.19.0.21  00:00:00:A0:30:68  kube-ovn-worker  subnet1
```

IP 172.19.0.21

7.22.5 OVN

OVN

```
kubectl ko nbctl mirror-add mirror1 gre 99 from-lport 172.19.0.21
kubectl ko nbctl lsp-attach-mirror coredns-787d4945fb-gpnkb.kube-system mirror1
```

coredns-787d4945fb-gpnkb.kube-system OVN LSP <POD_NAME>.<POD_NAMESPACE>

OVN

```
ovn-nbctl mirror-add <NAME> <TYPE> <INDEX> <FILTER> <IP>

NAME  - add a mirror with given name
TYPE  - specify TYPE 'gre' or 'erspan'
INDEX - specify the tunnel INDEX value
        (indicates key if GRE, erSPAN_idx if ERSPAN)
FILTER - specify FILTER for mirroring selection
        ('to-lport' / 'from-lport')
IP    - specify Sink / Destination i.e. Remote IP

ovn-nbctl mirror-del [NAME]      remove mirrors
ovn-nbctl mirror-list           print mirrors

ovn-nbctl lsp-attach-mirror PORT MIRROR  attach source PORT to MIRROR
ovn-nbctl lsp-detach-mirror PORT MIRROR  detach source PORT from MIRROR
```

7.22.6 Pod

Pod

```
root@pod1:/kube-ovn# ip link add mirror1 type gretap local 172.19.0.21 key 99 dev net1
root@pod1:/kube-ovn# ip link set mirror1 up
```

Pod

```
root@pod1:/kube-ovn# tcpdump -i mirror1 -nnve
tcpdump: listening on mirror1, link-type EN10MB (Ethernet), snapshot length 262144 bytes
05:13:30.328800 00:00:00:a3:f5:e2 > 00:00:00:97:0f:6e, ethertype ARP (0x0806), length 42: Ethernet (len 6), IPv4 (len 4), Request who-has 10.16.0.7 tell 10.16.0.4, length 28
05:13:30.559167 00:00:00:a3:f5:e2 > 00:00:00:89:d5:cc, ethertype IPv4 (0x0800), length 212: (tos 0x0, ttl 64, id 57364, offset 0, flags [DF], proto UDP (17), length 198)
  10.16.0.4.53 > 10.16.0.6.50472: 34511 NXDomain*- 0/1/1 (170)
05:13:30.560625 00:00:00:a3:f5:e2 > 00:00:00:89:d5:cc, ethertype IPv4 (0x0800), length 212: (tos 0x0, ttl 64, id 57365, offset 0, flags [DF], proto UDP (17), length 198)
  10.16.0.4.53 > 10.16.0.6.45177: 1659 NXDomain*- 0/1/1 (170)
05:13:30.562774 00:00:00:a3:f5:e2 > 00:00:00:89:d5:cc, ethertype IPv4 (0x0800), length 191: (tos 0x0, ttl 64, id 57368, offset 0, flags [DF], proto UDP (17), length 177)
  10.16.0.4.53 > 10.16.0.6.37755: 48737 NXDomain*- 0/1/1 (149)
05:13:30.563523 00:00:00:a3:f5:e2 > 00:00:00:89:d5:cc, ethertype IPv4 (0x0800), length 187: (tos 0x0, ttl 64, id 57369, offset 0, flags [DF], proto UDP (17), length 173)
  10.16.0.4.53 > 10.16.0.6.53887: 45519 NXDomain*- 0/1/1 (145)
05:13:30.564940 00:00:00:a3:f5:e2 > 00:00:00:89:d5:cc, ethertype IPv4 (0x0800), length 201: (tos 0x0, ttl 64, id 57370, offset 0, flags [DF], proto UDP (17),
```

```

length 187)
 10.16.0.4.53 > 10.16.0.6.40846: 25745 NXDomain*- 0/1/1 (159)
05:13:30.565140 00:00:00:a3:f5:e2 > 00:00:00:89:d5:cc, ethertype IPv4 (0x0800), length 201: (tos 0x0, ttl 64, id 57371, offset 0, flags [DF], proto UDP (17),
length 187)
 10.16.0.4.53 > 10.16.0.6.45214: 61875 NXDomain*- 0/1/1 (159)
05:13:30.566023 00:00:00:a3:f5:e2 > 00:00:00:55:e4:4e, ethertype IPv4 (0x0800), length 80: (tos 0x0, ttl 64, id 45937, offset 0, flags [DF], proto UDP (17),
length 66)
 10.16.0.4.44116 > 172.18.0.1.53: 16025+ [1au] AAAA? kube-ovn.io. (38)

```

7.22.7

1.	ERSPAN	OVN	Linux	4.14	ERSPAN	IPv6	Linux	4.16
2.		OVN						

[PDF](#)
[Slack](#)
[Support](#)

⌚2025 7 2

⌚2023 4 20



7.22.8

7.23 DNS Kube-OVN

NodeLocal DNSCache DaemonSet DNS DNS Kube-OVN

7.23.1 DNS

Kubernetes DNS

Kubernetes Nodelocaldnscache

```
#!/bin/bash

localdns=169.254.20.10
domain=cluster.local
kubedns=10.96.0.10

wget https://raw.githubusercontent.com/kubernetes/kubernetes/master/cluster/addons/dns/nodelocaldns/nodelocaldns.yaml
sed -i "s/_PILLAR__LOCAL__DNS__/$localdns/g; s/_PILLAR__DNS__DOMAIN__/$domain/g; s/_PILLAR__DNS__SERVER__//g; s/_PILLAR__CLUSTER__DNS__/$kubedns/g"
nodelocaldns.yaml

kubectl apply -f nodelocaldns.yaml
```

kubelet /var/lib/kubelet/config.yaml clusterDNS DNS IP 169.254.20.10 kubelet

Kube-OVN DNS

Kubernetes Nodelocal DNScache Kube-OVN

UNDERLAY SUBNET U20

Underlay Subnet DNS U20 kubectl edit subnet {your subnet} spec.u2oInterconnection = true , Overlay Subnet

KUBE-OVN-CONTROLLER DNS IP

```
kubectl edit deployment kube-ovn-controller -n kube-system

spec.template.spec.containers.args      --node-local-dns-ip=169.254.20.10
```

POD

Pod /etc/resolv.conf nameserver DNS IP Pod nameserver DNS ClusterIP u2o Pod Pod

7.23.2 DNS

Pod Pod DNS 169.254.20.10

```
# kubectl exec -it pod1 -- nslookup github.com
Server: 169.254.20.10
Address: 169.254.20.10:53
```

```
Name: github.com
Address: 20.205.243.166
```

DNS ovn0 DNS DNS

```
# tcpdump -i any port 53

06:20:00.441889 659246098c56_h P  ifindex 17 00:00:00:73:f1:06 ethertype IPv4 (0x0800), length 75: 10.16.0.2.40230 > 169.254.20.10.53: 1291+ A? baidu.com. (27)
06:20:00.441889 ovn0 In  ifindex 7 00:00:00:50:32:cd ethertype IPv4 (0x0800), length 75: 10.16.0.2.40230 > 169.254.20.10.53: 1291+ A? baidu.com. (27)
06:20:00.441950 659246098c56_h P  ifindex 17 00:00:00:73:f1:06 ethertype IPv4 (0x0800), length 75: 10.16.0.2.40230 > 169.254.20.10.53: 1611+ AAAA?
baidu.com. (27)
06:20:00.441950 ovn0 In  ifindex 7 00:00:00:50:32:cd ethertype IPv4 (0x0800), length 75: 10.16.0.2.40230 > 169.254.20.10.53: 1611+ AAAA? baidu.com. (27)
06:20:00.442203 ovn0 Out ifindex 7 00:00:00:52:99:d8 ethertype IPv4 (0x0800), length 145: 169.254.20.10.53 > 10.16.0.2.40230: 1611* 0/1/0 (97)
```

```

06:20:00.442219 659246098c56_h Out ifindex 17 00:00:00:ea:b3:5e ethertype IPv4 (0x0800), length 145: 169.254.20.10.53 > 10.16.0.2.40230: 1611* 0/1/0 (97)
06:20:00.442273 ovn0 Out ifindex 7 00:00:00:52:99:d8 ethertype IPv4 (0x0800), length 125: 169.254.20.10.53 > 10.16.0.2.40230: 1291* 2/0/0 A 39.156.66.10, A
110.242.68.66 (77)
06:20:00.442278 659246098c56_h Out ifindex 17 00:00:00:ea:b3:5e ethertype IPv4 (0x0800), length 125: 169.254.20.10.53 > 10.16.0.2.40230: 1291* 2/0/0 A 39.
156.66.10, A 110.242.68.66 (77)

```

7.23.3



NetworkPolicy	NetworkPolicy	DNS IP	169.254.20.10	CIDR	NetworkPolicy	DNS	Pod
---------------	---------------	--------	---------------	------	---------------	-----	-----

NetworkPolicy

Pod	DNS	NetworkPolicy
-----	-----	---------------

```

apiVersion: networking.k8s.io/v1
kind: NetworkPolicy
metadata:
  name: allow-local-dns-and-node-cidr
  namespace: default #
spec:
  podSelector: {} # Pod
  policyTypes:
  - Ingress
  - Egress
  egress:
  # DNS
  - to:
    - ipBlock:
      cidr: 169.254.20.10/32
  # CIDR
  - to:
    - ipBlock:
      cidr: 10.0.0.0/8 # CIDR
  ingress:
  # DNS
  - from:
    - ipBlock:
      cidr: 169.254.20.10/32
  # CIDR
  - from:
    - ipBlock:
      cidr: 10.0.0.0/8 # CIDR

```

- 169.254.20.10/32 DNS IP
- 10.0.0.0/8 CIDR

[PDF](#)
[Slack](#)
[Support](#)

⌚2025 7 3

⌚2023 5 5



[GitHub](#)



7.23.4

7.24 VPC NAT

7.24.1

VPC	Overlay	natOutgoing	Subnet	Pod	SNAT	IP	Pod	SNAT
NAT		CIDR	IP		SNAT			

7.24.2

subnet.Spec natOutgoing natOutgoingPolicyRules

```
spec:
  natOutgoing: true
  natOutgoingPolicyRules:
    - action: forward
      match:
        srcIPs: 10.0.11.0/30,10.0.11.254
    - action: nat
      match:
        srcIPs: 10.0.11.128/26
        dstIPs: 114.114.114.114,8.8.8.8
```

NAT

1. IP 10.0.11.0/30 10.0.11.254 SNAT
2. IP 10.0.11.128/26 IP 114.114.114.114 8.8.8.8 SNAT

action	match	action, action	forward	nat	forward	SNAT, nat	SNAT	natOutgoingPolicyRules			
SNAT											
match		srcIPs	dstIPs			IP	IP	match.srcIPs	match.dstIPs	CIDR	IP
match		natOutgoingPolicyRules				action					

[PDF](#)

[Slack](#)

[Support](#)

⌚2023 10 25

⌚2023 6 5



7.24.3

8.

8.1

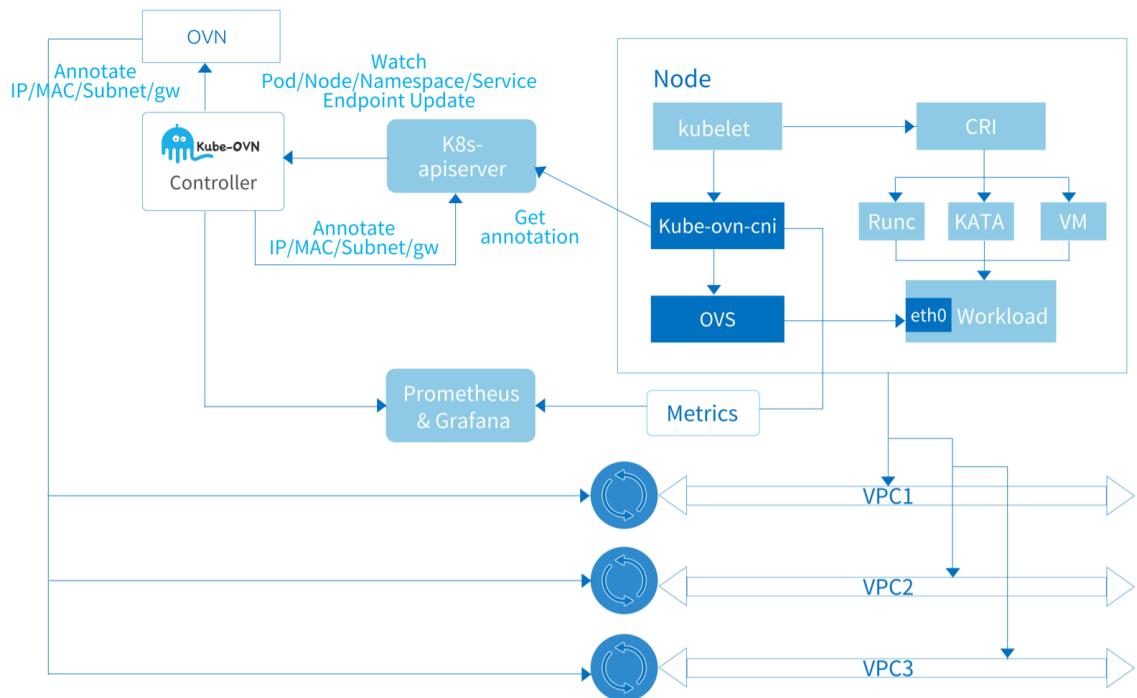
Kube-OVN

Kube-OVN	Kubernetes	OVN	SDN	Kube-OVN	OVN	Kubernetes	CNI Service	Networkpolicy
SDN	VPC	QoS ACL						
Kube-OVN			Cilium Submariner Prometheus KubeVirt					

8.1.1

Kube-OVN

- OVN/OVS
- Agent
-



OVN/OVS

OVN/OVS	Kube-OVN	OVN/OVS	SDN	Kube-OVN	ovn-architecture(7)
OVN	Kube-OVN	OVN	Kubernetes		
OVN/OVS		Kubernetes			

OVN-CENTRAL

```

ovn-central Deployment OVN          ovn-nb , ovn-sb , ovn-northd
  • ovn-nb           API      kube-ovn-controller   ovn-nb
  • ovn-sb           ovn-nb
  • ovn-northd     ovn-nb           ovn-sb
ovn-central       Raft

```

OVS-OVN

```

ovs-ovn  DaemonSet        Pod      openvswitch, ovsdb, ovn-controller    ovn-central  Agent

```

Agent

```
Kube-OVN      OVN  Kubernetes
```

KUBE-OVN-CONTROLLER

Deployment	Kubernetes	OVN	Kube-OVN	kube-ovn-controller	OVN
Pod	Service	Endpoint	Node	NetworkPolicy	VPC
Pod	annotation	IPAM	ovn-central	Subnet	Vlan
CIDR		annotation	kube-ovn-cni		ProviderNetwork

KUBE-OVN-CNI

DaemonSet	CNI	OVS
DaemonSet	kube-ovn	kubelet kube-ovn-cni
		CNI kube-ovn-cni
		/opt/cni/bin

```
kube-ovn-cni
```

1. ovn-controller vswitchd
2. CNI add/del
 - a. veth OVS
 - b. OVS
 - c. iptables/ipset/route
3. QoS.
4. ovn0
5. Vlan/Underlay/EIP
- 6.

Kube-OVN**KUBE-OVN-SPEAKER**

DaemonSet	Pod IP
	BGP

KUBE-OVN-PINGER

DaemonSet	OVS	Kube-OVN
-----------	-----	----------

KUBE-OVN-MONITOR

Deployment	OVN	Kube-OVN
------------	-----	----------

KUBECTL-KO

kubectl

kubectl

 PDF

 Slack

 Support

 2025 9 10

 2022 5 24



 GitHub 



8.1.2

8.2 What's Next

This document lists the features merged into the master branch for the next minor release.

8.2.1 Post-v1.14.0

- Performance: skip conntrack for specific dst CIDRs. [#5821](#)
- NetworkPolicy supports `lax` mode which only deny traffic type of TCP, UDP and SCTP. That means ARP, ICMP and DHCP traffic are always allowed. [#5745](#)
- Remove internal-port type interface code. [#5794](#)
- IPPool
 - Multiple IPPools now can bind to the same Namespace. [#5731](#)
 - Pods in a bound namespace will only get IPs from the bound pool(s), not other ranges in the subnet. [#5731](#)
- AdminNetworkPolicy now supports specify egress peers using FQDNs. [#5703](#)
- Using ARP for IPv4 network ready check: now you don't need ACL allow rules for gateway to make Pod running. [#5716](#)
- Non-primary CNI mode: you can run Kube-OVN as the secondary only network, without annoying unused annotations and logical switch port allocations. [#5618](#)
- VPC NAT Gateway:
 - No default EIP mode: the secondary interface can initialize without a default EIP to avoid the waste. [#5605](#)
 - Custom routes: you can control the route rules within the vpc-nat-gateway Pods to control traffic paths. [#5608](#)
 - Gratuitous ARP: VPC NAT Gateway automatically sends gratuitous ARP packets during initialization to accelerate network convergence. [#5607](#)
- Healthchecks for static endpoints in `SwitchLBRules`: SLR with both selector or endpoints key can support healthchecks. [#5435](#)
- Underlay
 - Node Selectors for `ProviderNetwork`: instead of adding/removing nodes to the `ProviderNetwork` one by one, you can use node selectors to simplify the workflow. [#5518](#)

This document lists the features merged into the master branch for the next minor release.

8.2.2 Post-v1.14.0

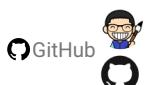
- Subnet with centralized gateway now supports nodeSelectors. [#5956](#)
- Overlay encapsulation NIC selection. [#5946](#)
- Performance: skip conntrack for specific dst CIDRs. [#5821](#)
- NetworkPolicy supports `lax` mode which only deny traffic type of TCP, UDP and SCTP. That means ARP, ICMP and DHCP traffic are always allowed. [#5745](#)
- Remove internal-port type interface code. [#5794](#)
- IPPool
- Multiple IPPools now can bind to the same Namespace. [#5731](#)
- Pods in a bound namespace will only get IPs from the bound pool(s), not other ranges in the subnet. [#5731](#)
- IPPool will create an AddressSet that can be work with VPC Policy Route and ACL. [#5920](#)
- AdminNetworkPolicy now supports specify egress peers using FQDNs. [#5703](#)
- Using ARP for IPv4 network ready check: now you don't need ACL allow rules for gateway to make Pod running. [#5716](#)
- Non-primary CNI mode: you can run Kube-OVN as the secondary only network, without annoying unused annotations and logical switch port allocations. [#5618](#)
- VPC NAT Gateway:

- No default EIP mode: the secondary interface can initialize without a default EIP to avoid the waste. [#5605](#)
- Custom routes: you can control the route rules within the `vpc-nat-gateway` Pods to control traffic paths. [#5608](#)
- Gratuitous ARP: VPC NAT Gateway automatically sends gratuitous ARP packets during initialization to accelerate network convergence. [#5607](#)
- Healthchecks for static endpoints in `SwitchLBRules`: SLR with both selector or endpoints key can support healthchecks. [#5435](#)
- Underlay
- Node Selectors for `ProviderNetwork`: instead of adding/removing nodes to the `ProviderNetwork` one by one, you can use node selectors to simplify the workflow. [#5518](#)
- Different `NetworkProvider`s can now share the same VLAN. [#5471](#)
- Adding `pod_name` and `pod_namespace` labels to interface metrics. [#5463](#)
- IPsec
- Support `cert-manager` to issue certificates. [#5365](#)
- Request new certificate if current certificate is not trusted. [#5710](#)
- `kubectl-ko`
- Collect IPsec and xFRM information. [#5472](#)
- Replace `Endpoint` with `EndpointSlice`. [#5425](#)
- NetworkAttachment caching: reduce APIServer load in large-scale deployments with Multus. [#5386](#)
- Upgrade OVS to 3.5 and OVN to 25.03. [#5537](#)
 - Different `NetworkProvider`s can now share the same VLAN. [#5471](#)
 - Adding `pod_name` and `pod_namespace` labels to interface metrics. [#5463](#)
 - IPsec
 - Support `cert-manager` to issue certificates. [#5365](#)
 - Request new certificate if current certificate is not trusted. [#5710](#)
 - `kubectl-ko`
 - Collect IPsec and xFRM information. [#5472](#)
 - Replace `Endpoint` with `EndpointSlice`. [#5425](#)
 - NetworkAttachment caching: reduce APIServer load in large-scale deployments with Multus. [#5386](#)
 - Upgrade OVS to 3.5 and OVN to 25.03. [#5537](#)

[PDF](#)[Slack](#)[Support](#)

⌚2025 12 1

⌚2025 9 15



8.2.3

8.3

Kube-OVN

8.3.1

Kube-OVN

```
1. default      Pod   IP           CIDR  10.16.0.0/16      10.16.0.1
2. join        Node  Pod          ,     CIDR  100.64.0.0/16    100.64.0.1
```

```
POD_CIDR="10.16.0.0/16"
POD_GATEWAY="10.16.0.1"
JOIN_CIDR="100.64.0.0/16"
EXCLUDE_IPS=""
```

```
EXCLUDE_IP    POD_CIDR          192.168.10.20..192.168.10.30
```

Overlay Service CIDR

[Join](#)

8.3.2 Service

```
kube-proxy  iptables       Kube-OVN      Kube-OVN  Service  CIDR
```

```
SVC_CIDR="10.96.0.0/12"
```

[kube-ovn-controller Deployment](#)

```
args:
- --service-cluster-ip-range=10.96.0.0/12
```

8.3.3 Overlay

Kube-OVN Kubernetes Node IP

```
IFACE=eth1
```

```
ens[a-z0-9]*,eth[a-z0-9]*
```

[kube-ovn-cni DaemonSet](#)

```
args:
- --iface=eth1
```

```
annotation ovn.kubernetes.io/tunnel_interface annotation  iface annotation
```

```
kubectl annotate node no1 ovn.kubernetes.io/tunnel_interface=ethx
```

8.3.4 MTU

```
Overlay          Kube-OVN          MTU          MTU          Overlay    Pod   MTU          MTU - 100 Underlay    Pod
MTU

Overlay      MTU      kube-ovn-cni DaemonSet

args:
- --mtu=1333
```

8.3.5

```
Kube-OVN           mirror0          tcpdump  
=true  
be-ovn-cni DaemonSet      :  
rror=true
```

8.3.6 LB

```
Underlay      kube-proxy          Service           OVN       L2 LB       ClusterIP        Service          LB
              ClusterIP            kube-proxy        kube-proxy

ENABLE_LB=false

kube-ovn-controller Deployment

args:
- --enable-lb=false

LB

Kube-OVN v1.12.0    subnet crd    spec  enableLb  Kube-OVN  LB          LB  kube-ovn-controller
Deployment   enable-lb          load-balancer  enableLb          load-balancer  v1.12.0  enableLb
```

8.3.7 NetworkPolicy

8.3.8 FIP SNAT

EIP SNAT kube-ovn-controller

```
ENABLE_EIP_SNAT=false
```

kube-ovn-controller Deployment

```
args:
- --enable-eip-snat=false
```

EIP SNAT

EIP SNAT

8.3.9 Load Balancer Service

VPC Load Balancer Service

LoadBalancer Service

```
ENABLE_LB_SVC=true
```

kube-ovn-controller Deployment

```
args:
- --enable-lb-svc=true
```

8.3.10 ECMP

ECMP ECMP

kube-ovn-controller Deployment

:

```
args:
- --enable-ecmp=true
```

Kube-OVN v1.12.0	subnet crd	spec enableEcmp	ECMP	ECMP	kube-ovn-controller
Deployment	enable-ecmp	v1.12.0			

8.3.11 Kubevirt VM

Kubevirt VM kube-ovn-controller

StatefulSet Pod

IP

VM

1.10.6

kube-ovn-controller Deployment

```
args:
- --keep-vm-ip=false
```

8.3.12 CNI

Kube-OVN /opt/cni/bin CNI /etc/cni/net.d CNI 01-kube-ovn.conflist CNI

```
CNI_CONF_DIR="/etc/cni/net.d"
CNI_BIN_DIR="/opt/cni/bin"
CNI_CONFIG_PRIORITY="01"
```

kube-ovn-cni DaemonSet Volume

```
volumes:
- name: cni-conf
  hostPath:
    path: "/etc/cni/net.d"
- name: cni-bin
  hostPath:
    path: "/opt/cni/bin"
...
args:
- --cni-conf-name=01-kube-ovn.conflist
```

8.3.13

Kube-OVN Overlay Geneve Vxlan STT

```
TUNNEL_TYPE="vxlan"
```

ovs-ovn DaemonSet

```
env:
- name: TUNNEL_TYPE
  value: "vxlan"
```

STT ovs

8.3.14 SSL

OVN DB API SSL :

```
ENABLE_SSL=true
```

SSL

8.3.15 ip

kube-ovn-controller/kube-ovn-cni/kube-ovn-monitor ip 0.0.0.0 ip

```
ENABLE_BIND_LOCAL_IP=true
```

kube-ovn-monitor pod ip

```
# netstat -tunlp |grep kube-ovn
tcp        0      0 172.18.0.5:10661          0.0.0.0:*              LISTEN      2612/.kube-ovn-mon
```

deployment daemonSet

```
env:
- name: ENABLE_BIND_LOCAL_IP
  value: "false"
```

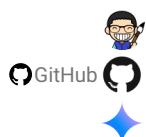
 PDF

 Slack

 Support

⌚2025 12 15

⌚2022 5 24



8.3.16

8.4

Kube-OVN	Minor	Patch	Minor	OVN/OVS	API	Patch	Bug	API
----------	-------	-------	-------	---------	-----	-------	-----	-----

8.4.1

Kube-OVN		master, release-1.12	release-1.11		release-1.12		Bug
backport							
		release-1.11	backport	Bug			

8.4.2

Minor		Patch	Bug	Bug
-------	--	-------	-----	-----

8.4.3 Patch

Patch	hack/release.sh							
1.	Build							
2.	tag	Docker Hub						
3.	tag	Github						
4.								
5.								
6.	Release Note PR							
7.	Release Note ()							
8.	Merge github action	Release Note PR						
9.	Github Release							
10.	Github Release	Release	v1.12.12		Release Note	Release		

8.4.4 Minor

Minor									
1.	Github	release-1.13	()						
2.	VERSION, dist/images/install.sh, charts/kube-ovn/values.yaml	charts/kube-ovn/Chart.yaml		Minor	v1.14.0	()			
3.	tag	Docker Hub ()							
4.	tag	Github ()							
5.	v1.13	mkdocs.yml	version branch ()						
6.	Release Note PR								
7.	Release Note ()								
8.	Merge github action	Release Note PR							
9.	Github Release								
10.	Github Release	Release	v1.13.0		Release Note	Release			
11.	VERSION	Patch	v1.13.1						



PDF



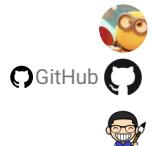
Slack



Support

⌚2024 5 16

⌚2024 5 8

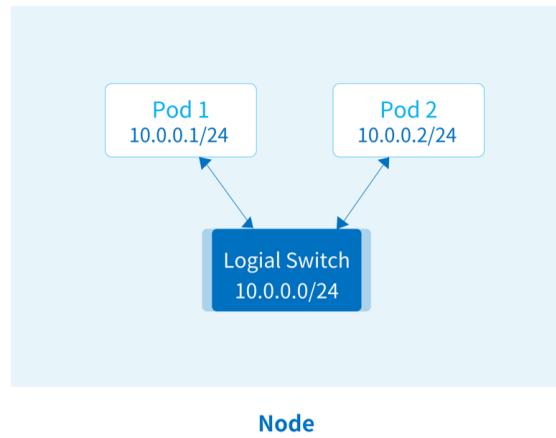


8.4.5

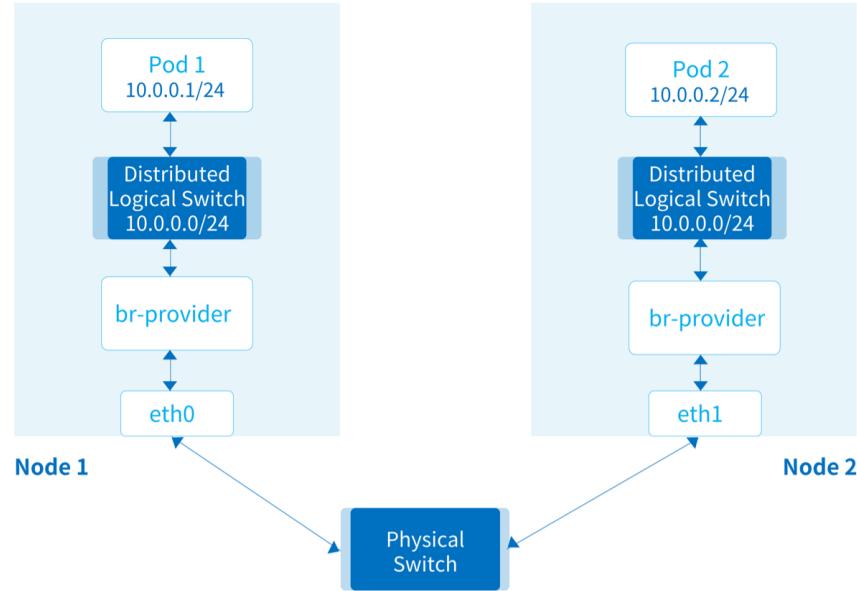
8.5 Underlay

Underlay

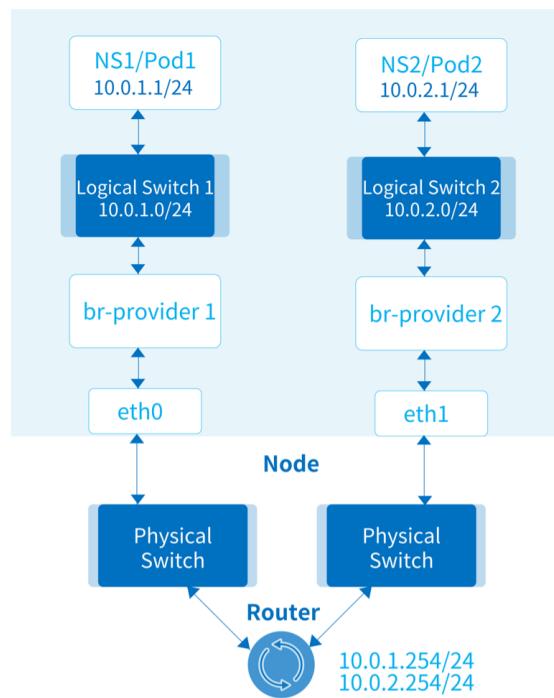
8.5.1



8.5.2



8.5.3

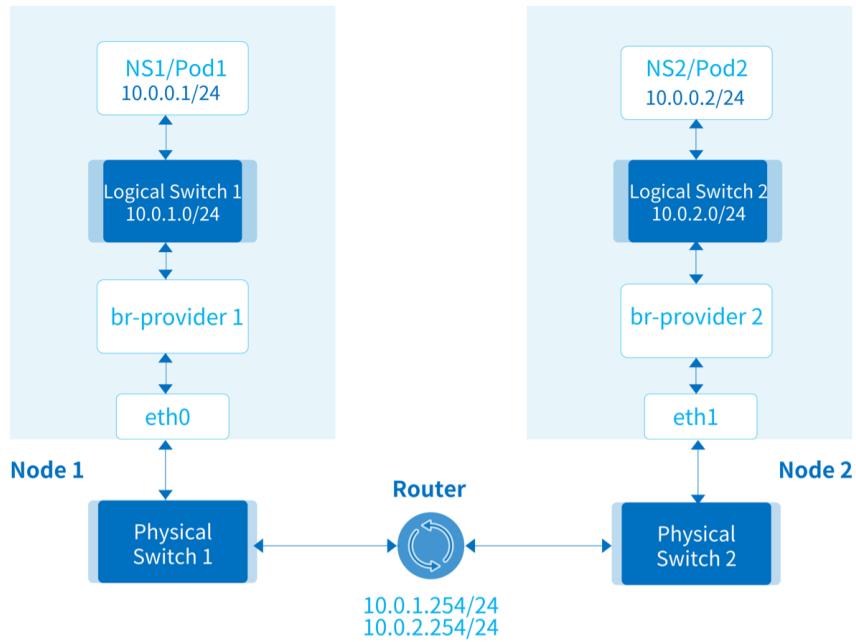


br-provider-1 br-provider-2

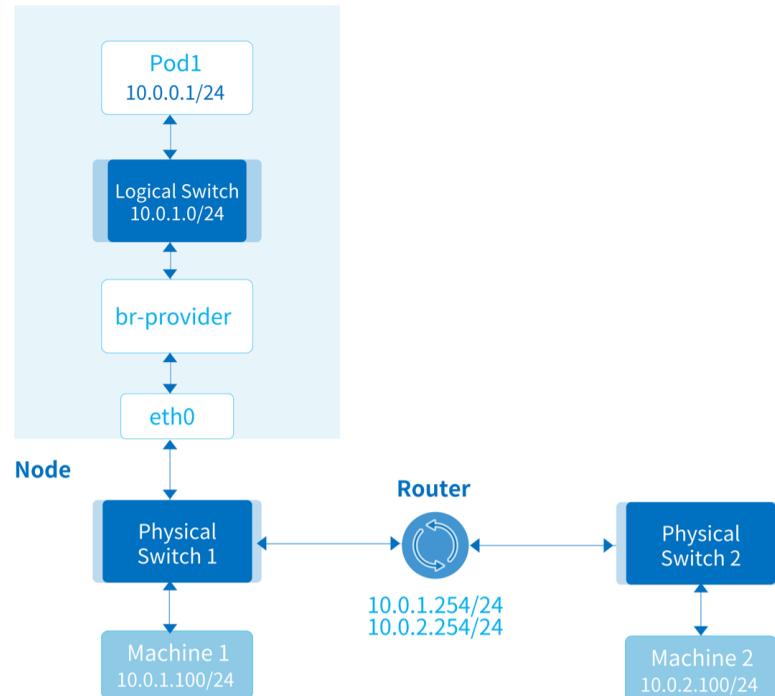
OVS

Provider Network

8.5.4

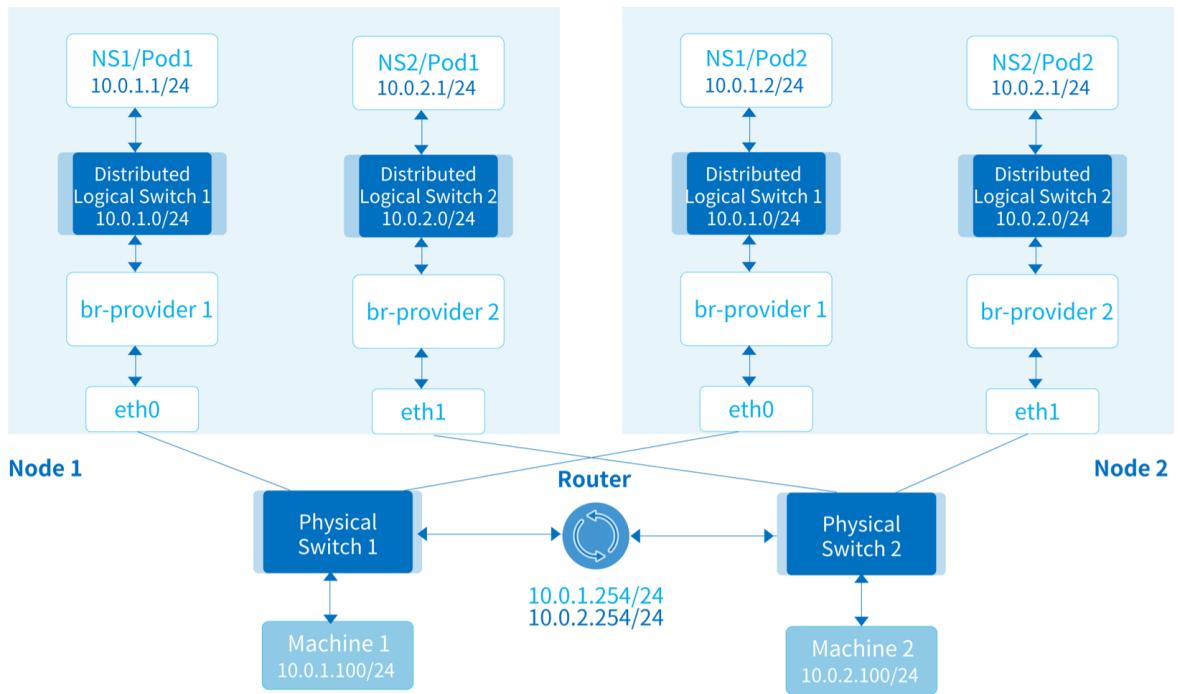


8.5.5

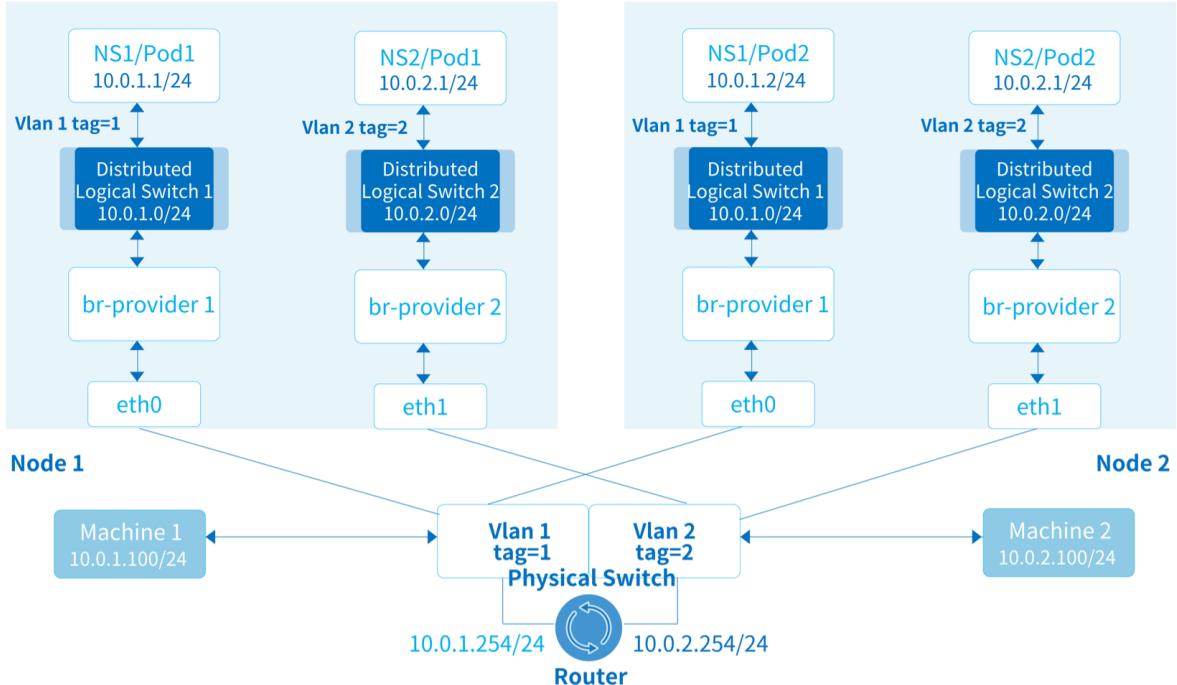


Pod

8.5.6 Vlan Tag

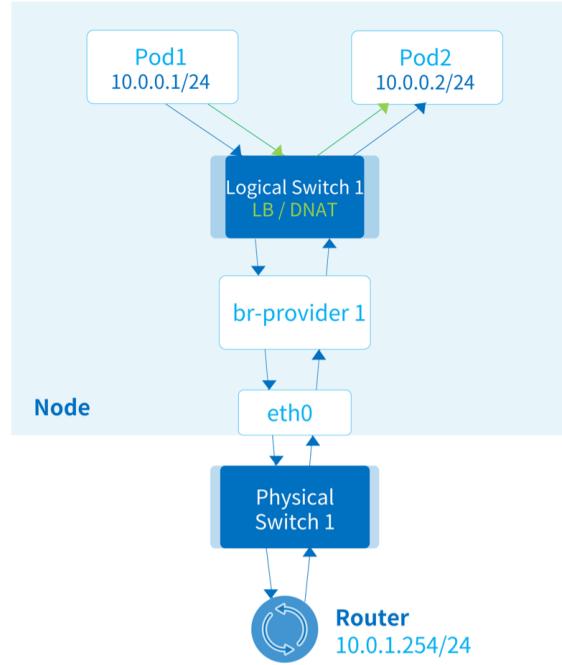
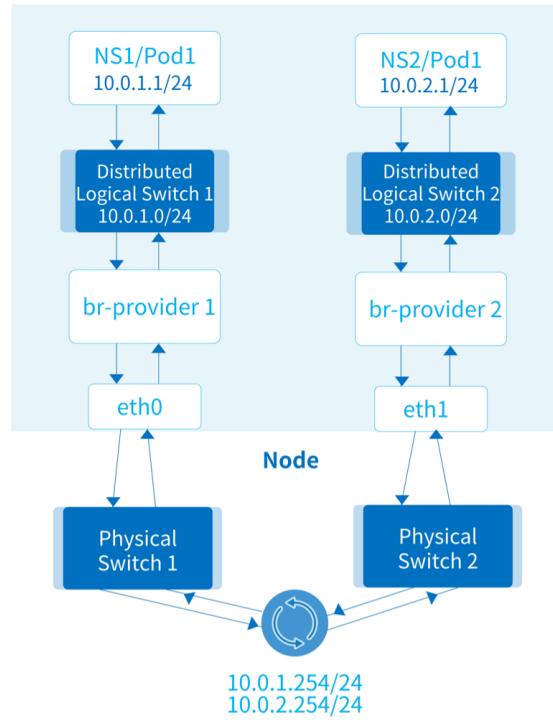


8.5.7 VLAN



8.5.8 Pod Service IP

Kube-OVN	Kubernetes Service	DNAT	IP	Service	Pod Endpoint	Service IP IP	Pod MAC	Service IP MAC	MAC
----------	--------------------	------	----	---------	--------------	---------------	---------	----------------	-----

Service**Pod****Service****Pod**[PDF](#)[Slack](#)[Support](#)

⌚2022 9 23

⌚2022 5 20

⌚GitHub 🎉

8.5.9

8.6 Iptables

Kube-OVN ipset iptables VPC Overlay NAT

ipset

IPv4/IPv6			
ovn40services/ovn60services	hash:net	Service	
ovn40subnets/ovn60subnets	hash:net	Overlay	NodeLocal DNS IP
ovn40subnets-nat/ovn60subnets-nat	hash:net	NatOutgoing	Overlay
ovn40subnets-distributed-gw/ovn60subnets-distributed-gw	hash:net	Overlay	
ovn40other-node/ovn60other-node	hash:net	IP	
ovn40local-pod-ip-nat/ovn60local-pod-ip-nat	hash:ip		
ovn40subnets-nat-policy	hash:net	natOutgoingPolicyRules	
ovn40natpr-418e79269dc5-dst	hash:net	natOutgoingPolicyRules	rule dstIPs
ovn40natpr-418e79269dc5-src	hash:net	natOutgoingPolicyRules	rule srcIPs

iptables IPv4

filter	INPUT	-m set --match-set ovn40services src -j ACCEPT	k8s Service Pod	--	
filter	INPUT	-m set --match-set ovn40services dst -j ACCEPT		--	
filter	INPUT	-m set --match-set ovn40subnets src -j ACCEPT		--	
filter	INPUT	-m set --match-set ovn40subnets dst -j ACCEPT		--	
filter	FORWARD	-m set --match-set ovn40services src -j ACCEPT		--	
filter	FORWARD	-m set --match-set ovn40services dst -j ACCEPT		--	
filter	FORWARD	-m set --match-set ovn40subnets src -j ACCEPT		--	
filter	FORWARD	-m set --match-set ovn40subnets dst -j ACCEPT		--	
filter	FORWARD	-s 10.16.0.0/16 -m comment --comment "ovn-subnet-gateway,ovn-default"	subnet	10.16.0.0/16 subnet cidr comment ovn-subnet-gateway subnet	iptables ovn-default subnet
filter	FORWARD	-d 10.16.0.0/16 -m comment --comment "ovn-subnet-gateway,ovn-default"	subnet		
filter	OUTPUT	-p udp -m udp --dport 6081 -j MARK --set-xmark 0x0	SNAT	UDP: bad checksum on VXLAN interface	
nat	PREROUTING	-m comment --comment "kube-ovn prerouting rules" -j OVN-PREROUTING	OVN-PREROUTING	--	
nat	POSTROUTING	-m comment --comment "kube-ovn postrouting rules" -j OVN-POSTROUTING	OVN-POSTROUTING	--	
nat	OVN-PREROUTING	-i ovn0 -m set --match-set ovn40subnets src -m set --match-set ovn40services dst -j MARK --set-xmark 0x4000/0x4000	Pod Service masquerade		LB
nat	OVN-PREROUTING	-p tcp -m addrtype --dst-type LOCAL -m set --match-		kube-proxy ipvs	

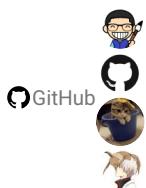
		set KUBE-NODE-PORT- LOCAL-TCP dst -j MARK -- set-xmark 0x80000/0x80000	ExternalTrafficPolicy Local Service TCP		
nat	OVN- PREROUTING	-p udp -m addrtype --dst- type LOCAL -m set --match- set KUBE-NODE-PORT- LOCAL-UDP dst -j MARK -- set-xmark 0x80000/0x80000	ExternalTrafficPolicy Local Service UDP		
nat	OVN- POSTROUTING	-m set --match-set ovn40services src -m set -- match-set ovn40subnets dst -m mark --mark 0x4000/0x4000 -j SNAT -- to-source	Service IP Overlay Pod IP		kube-proxy ipvs
nat	OVN- POSTROUTING	-m mark --mark 0x4000/0x4000 -j MASQUERADE	SNAT	--	
nat	OVN- POSTROUTING	-m set --match-set ovn40subnets src -m set -- match-set ovn40subnets dst -j MASQUERADE	Pod Service SNAT	--	
nat	OVN- POSTROUTING	-m mark --mark 0x80000/0x80000 -m set -- match-set ovn40subnets- distributed-gw dst -j RETURN	ExternalTrafficPolicy Local Service Endpoint SNAT	--	
nat	OVN- POSTROUTING	-m mark --mark 0x80000/0x80000 -j MASQUERADE	ExternalTrafficPolicy Local Service Endpoint SNAT	--	
nat	OVN- POSTROUTING	-p tcp -m tcp --tcp-flags SYN NONE -m conntrack -- ctstate NEW -j RETURN	Pod IP	SNAT	--
nat	OVN- POSTROUTING	-s 10.16.0.0/16 -m set ! -- match-set ovn40subnets dst -j SNAT --to-source 192.168.0.101	Pod NatOutgoing IP SNAT	10.16.0.0/16 IP	192.168.0.101
nat	OVN- POSTROUTING	-m set --match-set ovn40subnets-nat src -m set ! --match-set ovn40subnets dst -j MASQUERADE	Pod NatOutgoing SNAT	--	
nat	OVN- POSTROUTING	-m set --match-set ovn40subnets-nat-policy src -m set ! --match-set ovn40subnets dst -j OVN- NAT-POLICY	Pod natOutgoingPolicyRules SNAT	natOutgoingPolicyRules NAT-POLICY	OVN
nat					

	OVN- POSTROUTING	-m mark --mark 0x90001/0x90001 -j MASQUERADE --random- fully	OVN-NAT-POLICY 0x90001/0x90001	tag SNAT
nat	OVN- POSTROUTING	-m mark --mark 0x90002/0x90002 -j RETURN	OVN-NAT-POLICY 0x90002/0x90002	, tag SNAT
nat	OVN-NAT-POLICY	-s 10.0.11.0/24 -m comment --comment natPolicySubnet-net1 -j OVN-NAT-PSUBNET- aa98851157c5	10.0.11.0/24 net1 CIDR OVN-NAT- PSUBNET-aa98851157c5 natOutgoingPolicyRules	
nat	OVN-NAT- PSUBNET- xxxxxxxxxxxx	-m set --match-set ovn40natpr-418e79269dc5- src src -m set --match-set ovn40natpr-418e79269dc5- dst dst -j MARK --set-xmark 0x90002/0x90002	418e79269dc5 natOutgoingPolicyRules ID status.natOutgoingPolicyRules[index].Rule srcIPs ovn40natpr-418e79269dc5- src dstIPs ovn40natpr-418e79269dc5- dst tag 0x90002	
mangle	OVN-OUTPUT	-d 10.241.39.2/32 -p tcp -m tcp -dport 80 -j MARK --set- xmark 0x90003/0x90003	kubelet tproxy	
mangle	OVN- PREROUTING	-d 10.241.39.2/32 -p tcp -m tcp -dport 80 -j TPROXY -- on-port 8102 --on-ip 172.18.0.3 --tproxy-mark 0x90004/0x90004	kubelet tproxy	

[PDF](#)[Slack](#)[Support](#)

⌚2025 11 20

⌚2022 9 6



8.6.1

8.7

8.7.1

Kube-OVN Github / Github Issue PR Maintainer Review Github Action

8.7.2

Kube-OVN Go Go Modules G0111MODULE="on"

golangci-lint local-installation

Kube-OVN Docker buildx Docker buildx:

```
docker buildx create --use
```

8.7.3

Kube-OVN

```
git clone https://github.com/kubeovn/kube-ovn.git
cd kube-ovn
make release
```

ARM

```
make release-arm
```

8.7.4 base

OVS/OVN base

base Dockerfile dist/images/Dockerfile.base

```
# build x86 base image
make base-amd64

# build arm base image
make base-arm64
```

8.7.5 E2E

Kube-OVN :

- KIND Kubernetes go install sigs.k8s.io/kind@latest
- jinjanator : pip install jinjanator
- Ginkgo go install github.com/onsi/ginkgo/v2/ginkgo; go get github.com/onsi/gomega/...

E2E

```
make kind-init
make kind-install
make e2e
```

Underlay E2E

```
make kind-init
make kind-install-underlay
make e2e-underlay-single-nic
```

ovn vpc nat gw eip, fip, snat, dnat

```
make kind-init
make kind-install
make ovn-vpc-nat-gw-conformance-e2e
```

iptables vpc nat gw eip, fip, snat, dnat

```
make kind-init
make kind-install-vpc-nat-gw
make iptables-vpc-nat-gw-conformance-e2e
```

loadbalancer service

```
make kind-init
make kind-install-lb-svc
make kube-ovn-lb-svc-conformance-e2e
```

```
make kind-clean
```

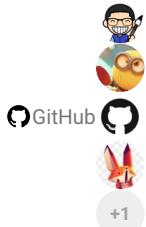
[!\[\]\(b192f47813074671dca41aa92c26b069_img.jpg\) PDF](#)

[!\[\]\(9bd704001d526f9130bb075b2acd795b_img.jpg\) Slack](#)

[!\[\]\(741b706e02930df49ae6f9d28176105c_img.jpg\) Support](#)

 2025 6 29

 2022 5 20



8.7.6

8.8 OVS/OVN

OVN/OVS	SDN	Kubernetes	Kube-OVN	Kube-OVN	OVN/OVS
OVN/OVS	Kube-OVN				

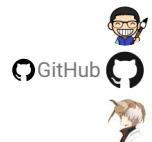
- [4228eab1d7](#) vswitchd ofport_usage
- [54056ea65d](#) timer
- [6b4dccb311f](#) fdb
- [f627b7721e](#) hairpin fdb
- [3f3e3a436f](#) ovsdb-tool join-cluster Server ID
- [a6cb8215a8](#) QoS
- [d4d76ddb2e](#) ovsdb-tool fix-cluster
- [ffd2328d4a](#) netdev CPU
- [d088c5d8c2](#) ovs-router kube-ipvs0
- [1b31f07dc6](#)
- [54b7678229](#) ovs-sandbox docker run
- [9ee66bd91b](#)
- [e889d46924](#) Underlay resubmit
- [f9e97031b5](#) ovn-controller Kube-OVN localnet GARP
- [78cade0187](#) conntrack
- [85aa6263ad](#) northd DNS IP conntrack
- [34dc3e3fcf](#) lflow lport conntrack
- [a297b840c2](#) DNAT lsp
- [03e35ed9c5](#) ovn-controller
- [e7d3ba53cd](#) ACL DNS IP conntrack
- [9286e1fd57](#)
- [e5916eb53a](#) lr-lb DNAT
- [e4e6ea9c5f](#) BFD LRP
- [e76880e792](#) northd nb version_compatibility
- [477695a010](#) northd localnet lrp arp/nd

- [20626ea909](#) LB ACL
- [a2d9ff3cccd](#) Deb



⌚2025 9 10

⌚2022 5 24



8.8.1

8.9

Kube-OVN OVN/OVS Geneve Vxlan STT OVN
[OVN Architecture Design Decision](#)

8.9.1 Geneve

Geneve	Kube-OVN	OVN	Offload	Geneve	24bit
datapath		datapath 32768			
Mellanox	OVS Geneve	5.4	backport		
UDP	TCP over UDP	TCP	CPU		

8.9.2 Vxlan

Vxlan	OVN	Offload	OVN	datapath	4096 datapath
datapath 4096		inport ACL			
Mellanox	OVS Vxlan				
UDP	TCP over UDP	TCP	CPU		

8.9.3 STT

⚠ Warning

OpenVswitch 3.6 STT Tunnel

STT	OVN	TCP	TCP	TCP	OVN	datapath
		OVS				
		OVS				

8.9.4

- [VXLAN vs GENEVE: Understand The Difference](#)
- [OVN FAQ](#)
- [What is Geneve](#)

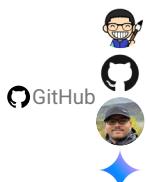
[!\[\]\(b0715a6528e0323b2f3605d0e8bed5e1_img.jpg\) PDF](#)

[!\[\]\(f24436961626030c1441132e4cd6ca3b_img.jpg\) Slack](#)

[!\[\]\(053eb519d8bb52a15049af457fb8ce7c_img.jpg\) Support](#)

 2025 9 10

 2022 6 17



8.9.5

8.10 Kube-OVN

Kube-OVN

8.10.1 ovn-monitor

OVN

Gauge	kube_ovn_ovn_status	OVN	(2) follower	(1) leader,	(0)
Gauge	kube_ovn_failed_req_count	OVN			
Gauge	kube_ovn_log_file_size_bytes	OVN			
Gauge	kube_ovn_db_file_size_bytes	OVN			
Gauge	kube_ovn_chassis_info	OVN chassis	(1)	(0)	
Gauge	kube_ovn_db_status	OVN	, (1)	(0)	
Gauge	kube_ovn_logical_switch_info	OVN logical switch	(1)	logical switch	
Gauge	kube_ovn_logical_switch_external_id	OVN logical switch external_id	(1)	external-id	
Gauge	kube_ovn_logical_switch_port_binding	OVN logical switch logical switch port	(1)		
Gauge	kube_ovn_logical_switch_tunnel_key	OVN logical switch tunnel key			
Gauge	kube_ovn_logical_switch_ports_num	OVN logical switch logical port			
Gauge	kube_ovn_logical_switch_port_info	OVN logical switch port	(1)		
Gauge	kube_ovn_logical_switch_port_tunnel_key	OVN logical switch port tunnel key			
Gauge	kube_ovn_cluster_enabled	(1) OVN	(0) OVN		
Gauge	kube_ovn_cluster_role		(1)		
Gauge	kube_ovn_cluster_status		(1)		
Gauge	kube_ovn_cluster_term	RAFT term			
Gauge	kube_ovn_cluster_leader_self		leader (1)	(0)	
Gauge	kube_ovn_cluster_vote_self		leader (1)	(0)	
Gauge	kube_ovn_cluster_election_timer	election timer			
Gauge	kube_ovn_cluster_log_not_committed	commit RAFT			
Gauge	kube_ovn_cluster_log_not_applied	apply RAFT			
Gauge	kube_ovn_cluster_log_index_start	RAFT			
Gauge	kube_ovn_cluster_log_index_next	RAFT			
Gauge	kube_ovn_cluster_inbound_connections_total				
Gauge	kube_ovn_cluster_outbound_connections_total				
Gauge	kube_ovn_cluster_inbound_connections_error_total				
Gauge	kube_ovn_cluster_outbound_connections_error_total				

8.10.2 ovs-monitor

ovsdb vswitchd

Gauge	ovs_status	OVS	(1)	(0)
Gauge	ovs_info	OVS	(1)	
Gauge	failed_req_count	OVS		
Gauge	log_file_size	OVS		
Gauge	db_file_size	OVS		
Gauge	datapath	Datapath	(1)	
Gauge	dp_total	OVS datapath		
Gauge	dp_if	Datapath	(1)	
Gauge	dp_if_total	datapath port		
Gauge	dp_flows_total	Datapath flow		
Gauge	dp_flows_lookup_hit	Datapath flow		
Gauge	dp_flows_lookup_missed	Datapath flow		
Gauge	dp_flows_lookup_lost	Datapath userspace		
Gauge	dp_masks_hit	Datapath mask		
Gauge	dp_masks_total	Datapath mask		
Gauge	dp_masks_hit_ratio	Datapath mask		
Gauge	interface	OVS (1)		
Gauge	interface_admin_state	(0) down, (1) up, (2)		
Gauge	interface_link_state	(0) down, (1) up, (2)		
Gauge	interface_mac_in_use	OVS Interface MAC		
Gauge	interface_mtu	OVS Interface MTU		
Gauge	interface_of_port	OVS Interface OpenFlow Port ID		
Gauge	interface_if_index	OVS Interface Index		
Gauge	interface_tx_packets	OVS Interface		
Gauge	interface_tx_bytes	OVS Interface		
Gauge	interface_rx_packets	OVS Interface		
Gauge	interface_rx_bytes	OVS Interface		
Gauge	interface_rx_crc_err	OVS Interface		
Gauge	interface_rx_dropped	OVS Interface		
Gauge	interface_rx_errors	OVS Interface		
Gauge	interface_rx_frame_err	OVS Interface		
Gauge	interface_rx_missed_err	OVS Interface miss		
Gauge	interface_rx_over_err	OVS Interface overrun		
Gauge	interface_tx_dropped	OVS Interface		

Gauge	interface_tx_errors	OVS Interface
Gauge	interface_collisions	OVS interface

8.10.3 kube-ovn-pinger

Gauge	pinger_ovs_up	OVS
Gauge	pinger_ovs_down	OVS
Gauge	pinger_ovn_controller_up	ovn-controller
Gauge	pinger_ovn_controller_down	ovn-controller
Gauge	pinger_inconsistent_port_binding	OVN-SB portbinding OVS interface
Gauge	pinger_apiserver_healthy	kube-ovn-pinger apiserver
Gauge	pinger_apiserver_unhealthy	kube-ovn-pinger apiserver
Histogram	pinger_apiserver_latency_ms	kube-ovn-pinger apiserver
Gauge	pinger_internal_dns_healthy	kube-ovn-pinger
Gauge	pinger_internal_dns_unhealthy	kube-ovn-pinger
Histogram	pinger_internal_dns_latency_ms	kube-ovn-pinger
Gauge	pinger_external_dns_health	kube-ovn-pinger
Gauge	pinger_external_dns_unhealthy	kube-ovn-pinger
Histogram	pinger_external_dns_latency_ms	kube-ovn-pinger
Histogram	pinger_pod_ping_latency_ms	kube-ovn-pinger ping Pod
Gauge	pinger_pod_ping_lost_total	kube-ovn-pinger ping Pod
Gauge	pinger_pod_ping_count_total	kube-ovn-pinger ping Pod
Histogram	pinger_node_ping_latency_ms	kube-ovn-pinger ping Node
Gauge	pinger_node_ping_lost_total	kube-ovn-pinger ping Node
Gauge	pinger_node_ping_count_total	kube-ovn-pinger ping Node
Histogram	pinger_external_ping_latency_ms	kube-ovn-pinger ping
Gauge	pinger_external_lost_total	kube-ovn-pinger ping

8.10.4 kube-ovn-controller

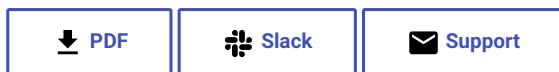
`kube-ovn-controller`

Histogram	rest_client_request_latency_seconds	apiserver
Counter	rest_client_requests_total	apiserver
Counter	lists_total	API list
Summary	list_duration_seconds	API list
Summary	items_per_list	API list
Counter	watches_total	API watch
Counter	short_watches_total	API watch
Summary	watch_duration_seconds	API watch
Summary	items_per_watch	API watch
Gauge	last_resource_version	resource version
Histogram	ovs_client_request_latency_milliseconds	OVN
Gauge	subnet_available_ip_count	IP
Gauge	subnet_used_ip_count	IP

8.10.5 kube-ovn-cni

`kube-ovn-cni`

Histogram	cni_op_latency_seconds	CNI
Counter	cni_wait_address_seconds_total	CNI
Counter	cni_wait_connectivity_seconds_total	CNI
Counter	cni_wait_route_seconds_total	CNI
Histogram	rest_client_request_latency_seconds	apiserver
Counter	rest_client_requests_total	apiserver
Counter	lists_total	API list
Summary	list_duration_seconds	API list
Summary	items_per_list	API list
Counter	watches_total	API watch
Counter	short_watches_total	API watch
Summary	watch_duration_seconds	API watch
Summary	items_per_watch	API watch
Gauge	last_resource_version	resource version
Histogram	ovs_client_request_latency_milliseconds	OVN



⌚2022 10 18

⌚2022 6 21

GitHub 

8.10.6

8.11 Kube-OVN

Kube-OVN Kube-OVN CRD CRD

8.11.1 Condition

type	String			
status	String	True	False	Unknown
reason	String			
message	String			
observedGeneration	Int64			
lastUpdateTime	Time			
lastTransitionTime	Time			

CRD Status Condition

8.11.2

Subnet

SUBNET

apiVersion	String	Kubernetes	kubeovn.io/v1
kind	String	Kubernetes	Subnet
metadata	ObjectMeta	Kubernetes	
spec	SubnetSpec	Subnet	
status	SubnetStatus	Subnet	

SubnetSpec

default	Bool				
vpc	String	VPC	ovn-cluster		
protocol	String	IP	IPv4	IPv6	Dual
namespaces	[]String		namespace		
cidrBlock	String		10.16.0.0/16		
gateway	String		CIDRBlock		
excludelps	[]String				
provider	String	OVN Subnet	NetworkAttachmentDefinition	. Kube-OVN	
gatewayType	String	Overlay	distributed	centralized	
gatewayNode	String		centralized		
natOutgoing	Bool	NAT	externalEgressGateway		
externalEgressGateway	String		natOutgoing		
policyRoutingPriority	Uint32				
policyRoutingTableID	Uint32		TableID		
mtu	Uint32		MTU		
private	Bool				
allowSubnets	[]String				
vlan	String		Vlan		
vips	[]String	virtual	Isp	virtual-ip	
logicalGateway	Bool				
disableGatewayCheck	Bool		Pod		
disableInterConnection	Bool				
enableDHCP	Bool		Isp	dhcp	
dhcpV4Options	String		Isp	dhcpv4_options	DHCP_Options
dhcpV6Options	String		Isp	dhcpv6_options	DHCP_Options
enableIPv6RA	Bool		Irp	ipv6_ra_configs	
ipv6RAConfigs	String		Irp	ipv6_ra_configs	
acls	[]Acl		logical-switch	acls	
allowEWTraffic	Bool				
natOutgoingPolicyRules	[]NatOutgoingPolicyRule	NAT			
u2oInterconnectionIP	String	Underlay/Overlay	IP		
u2oInterconnection	Bool	Overlay/Underlay			
enableLb	*Bool	logical-switch	load-balancer		
enableEcmp	Bool	ECMP			

enableMulticastSnoop	Bool
enableExternalLBAddress	Bool
routeTable	String
namespaceSelectors	[]LabelSelector

Acl

direction	String	Acl	from-lport	to-lport
priority	Int	Acl	0	32767
match	String	Acl		
action	String	Acl	allow-related	allow-stateless
			allow	drop
			reject	

NatOutgoingPolicyRule

match	NatOutGoingPolicyMatch
action	String

NatOutGoingPolicyMatch

srcIPs	String	IP
dstIPs	String	IP

SubnetStatus

conditions	>[]SubnetCondition	Condition	
v4availableIPs	Float64	IPv4 IP	
v4availableIPrange	String	IPv4	
v4usingIPs	Float64	IPv4 IP	
v4usingIPrange	String	IPv4	
v6availableIPs	Float64	IPv6 IP	
v6availableIPrange	String	IPv6	
v6usingIPs	Float64	IPv6 IP	
v6usingIPrange	String	IPv6	
activateGateway	String		
dhcpV4OptionsUUID	String	lsp dhcpv4_options	DHCP_Options
dhcpV6OptionsUUID	String	lsp dhcpv6_options	DHCP_Options
u2oInterconnectionIP	String	Overlay/Underlay	IP
u2oInterconnectionMAC	String	Overlay/Underlay	MAC
u2oInterconnectionVPC	String	Overlay/Underlay	VPC
natOutgoingPolicyRules	[]NatOutgoingPolicyRuleStatus	NAT	
mcastQuerierIP	String	IP	
mcastQuerierMAC	String	MAC	

IP

IP

apiVersion	String	Kubernetes	kubeovn.io/v1
kind	String	Kubernetes	IP
metadata	ObjectMeta	Kubernetes	
spec	IPSpec	IP	

IPSpec

podName	String	Pod	
namespace	String	Pod Namespace	
subnet	String	IP Subnet	
attachSubnets	[]String	IP	
nodeName	String	Pod	
ipAddress	String	IP v4IP v6IP	
v4IpAddress	String	IPv4 IP	
v6IpAddress	String	IPv6 IP	
attachIps	[]String	IP IP	
macAddress	String	Pod MAC	
attachMacs	[]String	IP MAC	
containerID	String	Pod Container ID	
podType	String	Pod StatefulSet VirtualMachine	

Vpc

VPC

apiVersion	String	Kubernetes	kubeovn.io/v1
kind	String	Kubernetes	Vpc
metadata	ObjectMeta	Kubernetes	
spec	VpcSpec	Vpc	
status	VpcStatus	Vpc	

VpcSpec

defaultSubnet	String	
namespaces	[]String	Vpc
staticRoutes	[]StaticRoute	
policyRoutes	[]PolicyRoute	
vpcPeerings	[]VpcPeering	VPC
enableExternal	Bool	
extraExternalSubnets	[]String	
enableBfd	Bool	BFD ()
bfdPort	BFDPort	BFD

StaticRoute

policy	String	
cidr	String	
nextHopIP	String	IP
ecmpMode	String	ECMP
bfdId	String	BFD ID
routeTable	String	

PolicyRoute

priority	Int	
match	String	
action	String	allow drop reroute
nextHopIP	String	IP action reroute

VpcPeering

remoteVpc	String	VPC
localConnectIP	String	IP

BFDPort

enabled	Bool	BFD
ip	String	BFD IP
nodeSelector	LabelSelector	BFD LRP

VpcStatus

conditions	IVpcCondition	Vpc	Condition
standby	Bool	VPC	
default	Bool	VPC	
defaultLogicalSwitch	String		
router	String		
tcpLoadBalancer	String	TCP	
udpLoadBalancer	String	UDP	
sctpLoadBalancer	String	SCTP	
tcpSessionLoadBalancer	String	TCP	
udpSessionLoadBalancer	String	UDP	
sctpSessionLoadBalancer	String	SCTP	
subnets	IString	VPC	
vpcPeerings	IString	VPC	
enableExternal	Bool		
extraExternalSubnets	IString		
enableBfd	Bool	BFD	

8.11.3 Underlay**Vlan**

apiVersion	String	Kubernetes	kubeovn.io/v1
kind	String	Kubernetes	Vlan
metadata	ObjectMeta	Kubernetes	
spec	VlanSpec	Vlan	
status	VlanStatus	Vlan	

VLANSPEC

id	Int	Vlan tag	0~4096
provider	String	Vlan	ProviderNetwork

VLANSTATUS

subnets	>[]String	Vlan	
conflict	Bool		
conditions	>[]VlanCondition	Vlan	Condition

ProviderNetwork

apiVersion	String	Kubernetes	kubeovn.io/v1
kind	String	Kubernetes	ProviderNetwork
metadata	ObjectMeta	Kubernetes	
spec	ProviderNetworkSpec	ProviderNetwork	
status	ProviderNetworkStatus	ProviderNetwork	

PROVIDERNETWORKSPEC

defaultInterface	String		
customInterfaces	>[]CustomInterface		
excludeNodes	>[]String		
exchangeLinkName	Bool	OVS	

CustomInterface

interface	String	Underlay	
nodes	>[]String		

PROVIDERNETWORKSTATUS

ready	Bool		
readyNodes	>[]String		
notReadyNodes	>[]String		
vlans	>[]String	Vlan	
conditions	>[]ProviderNetworkCondition	ProviderNetwork	Condition

8.11.4

SecurityGroup

apiVersion	String	Kubernetes	kubeovn.io/v1
kind	String	Kubernetes	SecurityGroup
metadata	ObjectMeta	Kubernetes	
spec	SecurityGroupSpec	SecurityGroup	
status	SecurityGroupStatus	SecurityGroup	

SECURITYGROUPSPEC

ingressRules	[]SecurityGroupRule
egressRules	[]SecurityGroupRule
allowSameGroupTraffic	Bool

SecurityGroupRule

ipVersion	String	IP	ipv4	ipv6		
protocol	SgProtocol		all	icmp	tcp	udp
priority	Int		1-200			
remoteType	SgRemoteType		address	securityGroup		
remoteAddress	String					
remoteSecurityGroup	String					
portRangeMin	Int		1			
portRangeMax	Int		65535			
policy	SgPolicy		allow	drop		

SECURITYGROUPSTATUS

portGroup	String	
allowSameGroupTraffic	Bool	
ingressMd5	String	MD5
egressMd5	String	MD5
ingressLastSyncSuccess	Bool	
egressLastSyncSuccess	Bool	

8.11.5 IP

Vip

apiVersion	String	Kubernetes	kubeovn.io/v1
kind	String	Kubernetes	Vip
metadata	ObjectMeta	Kubernetes	
spec	VipSpec	Vip	
status	VipStatus	Vip	

VIPSPEC

namespace	String	VIP
subnet	String	VIP
type	String	VIP
v4ip	String	IPv4
v6ip	String	IPv6
macAddress	String	MAC
selector	[]String	
attachSubnets	[]String	

VIPSTATUS

conditions	[]VipCondition	VIP	Condition
type	String	VIP	
v4ip	String	IPv4	
v6ip	String	IPv6	
mac	String	MAC	

SwitchLBRule

apiVersion	String	Kubernetes	kubeovn.io/v1
kind	String	Kubernetes	SwitchLBRule
metadata	ObjectMeta	Kubernetes	
spec	SwitchLBRuleSpec	SwitchLBRule	
status	SwitchLBRuleStatus	SwitchLBRule	

SWITCHLBRULESPEC

vip	String	IP
namespace	String	
selector	[]String	
endpoints	[]String	
sessionAffinity	String	
ports	[]SwitchLBRulePort	

SwitchLBRulePort

name	String	
port	Int32	
targetPort	Int32	
protocol	String	

SWITCHLBRULESTATUS

conditions	[]SwitchLBRuleCondition	SwitchLBRule	Condition
ports	String	SwitchLBRule	
service	String	SwitchLBRule	service

8.11.6 QoS IP**QoS Policy**

apiVersion	String	Kubernetes	kubeovn.io/v1
kind	String	Kubernetes	QoS Policy
metadata	ObjectMeta	Kubernetes	
spec	QoS Policy Spec	QoS Policy	

QOSPOLICYSPEC

bandwidthLimitRules	QoS Policy Bandwidth Limit Rules	
shared	Bool	
bindingType	QoS Policy Binding Type	

IPPool

apiVersion	String	Kubernetes	kubeovn.io/v1
kind	String	Kubernetes	IPPool
metadata	ObjectMeta	Kubernetes	
spec	IPPoolSpec	IPPool	
status	IPPoolStatus	IPPool	

IPPOOLSPEC

subnet	String
namespaces	[]String
ips	[]String
	IP

IPPOOLSTATUS

v4AvailableIPs	BigInt	IPv4	IP
v4AvailableIPRange	String	IPv4	IP
v4UsingIPs	BigInt	IPv4	IP
v4UsingIPRange	String	IPv4	IP
v6AvailableIPs	BigInt	IPv6	IP
v6AvailableIPRange	String	IPv6	IP
v6UsingIPs	BigInt	IPv6	IP
v6UsingIPRange	String	IPv6	IP
conditions	[]IPPoolCondition	IP	Condition

8.11.7 NAT IP

IptablesEIP

apiVersion	String	Kubernetes	kubeovn.io/v1
kind	String	Kubernetes	IptablesEIP
metadata	ObjectMeta	Kubernetes	
spec	IptablesEIPSpec	IptablesEIP	
status	IptablesEIPStatus	IptablesEIP	

IPTABLESEIPSPEC

v4ip	String	IPv4
v6ip	String	IPv6
macAddress	String	MAC
natGwDp	String	NAT
qosPolicy	String	QoS
externalSubnet	String	

IPTABLESEIPSTATUS

ready	Bool	IptablesEIP			
ip	String	IptablesEIP	IP	IPv4	
redo	String	IptablesEIP	CRD		
nat	String	IptablesEIP	fip	snat	dnat
qosPolicy	String	QoS			
conditions	[]IptablesEIPCondition	IptablesEIP		Condition	

OvnEip

apiVersion	String	Kubernetes	kubeovn.io/v1
kind	String	Kubernetes	OvnEip
metadata	ObjectMeta	Kubernetes	
spec	OvnEipSpec	OvnEip	
status	OvnEipStatus	OvnEip	

OVNEIPSPEC

externalSubnet	String	
v4Ip	String	IPv4
v6Ip	String	IPv6
macAddress	String	MAC
type	String	fip lsp nat

IptablesFIPRule

apiVersion	String	Kubernetes	kubeovn.io/v1
kind	String	Kubernetes	IptablesFIPRule
metadata	ObjectMeta	Kubernetes	
spec	IptablesFIPRuleSpec	IptablesFIPRule	

IPTABLESFIPRULESPEC

eip	String	IP
internalIP	String	IP

OvnFip

apiVersion	String	Kubernetes	kubeovn.io/v1
kind	String	Kubernetes	OvnFip
metadata	ObjectMeta	Kubernetes	
spec	OvnFipSpec	OvnFip	
status	OvnFipStatus	OvnFip	

OVNFIPSPEC

ovnEip	String	OVN EIP	
ipType	String	IP	vip ip
ipName	String	IP	
vpc	String	VPC	
v4Ip	String	IPv4	
v6Ip	String	IPv6	
type	String	distributed	centralized

IptablesDnatRule

apiVersion	String	Kubernetes	kubeovn.io/v1
kind	String	Kubernetes	IptablesDnatRule
metadata	ObjectMeta	Kubernetes	
spec	IptablesDnatRuleSpec	IptablesDnatRule	

IPTABLESDNATRULESPEC

eip	String	IP
externalPort	String	
protocol	String	
internalIP	String	IP
internalPort	String	

OvnDnatRule

apiVersion	String	Kubernetes	kubeovn.io/v1
kind	String	Kubernetes	OvnDnatRule
metadata	ObjectMeta	Kubernetes	
spec	OvnDnatRuleSpec	OvnDnatRule	
status	OvnDnatRuleStatus	OvnDnatRule	

OVNDNATRULESPEC

ovnEip	String	OVN EIP
ipType	String	IP vip ip
ipName	String	IP
internalPort	String	
externalPort	String	
protocol	String	
vpc	String	VPC
v4Ip	String	IPv4
v6Ip	String	IPv6

OVNDNATRULESTATUS

vpc	String	VPC	
v4Eip	String	IPv4 EIP	
v6Eip	String	IPv6 EIP	
externalPort	String		
v4Ip	String	IPv4	
v6Ip	String	IPv6	
internalPort	String		
protocol	String		
ipName	String	IP	
ready	Bool	DNAT	
conditions	[]OvnDnatRuleCondition	OVN DNAT	Condition

IptablesSnatRule

apiVersion	String	Kubernetes	kubeovn.io/v1
kind	String	Kubernetes	IptablesSnatRule
metadata	ObjectMeta	Kubernetes	
spec	IptablesSnatRuleSpec	IptablesSnatRule	

IPTABLESSNATRULESPEC

eip	String	IP
internalCIDR	String	CIDR

OvnSnatRule

apiVersion	String	Kubernetes	kubeovn.io/v1
kind	String	Kubernetes	OvnSnatRule
metadata	ObjectMeta	Kubernetes	
spec	OvnSnatRuleSpec	OvnSnatRule	
status	OvnSnatRuleStatus	OvnSnatRule	

OVNSNATRULESPEC

ovnEip	String	OVN EIP
vpcSubnet	String	VPC
ipName	String	IP
vpc	String	VPC
v4IpCidr	String	IPv4 CIDR
v6IpCidr	String	IPv6 CIDR

8.11.8 VPC

VpcNatGateway

apiVersion	String	Kubernetes	kubeovn.io/v1	
kind	String	Kubernetes	VpcNatGateway	
metadata	ObjectMeta	Kubernetes		
spec	VpcNatGatewaySpec	VpcNatGateway		
status	VpcNatGatewayStatus	VpcNatGateway		

VPCNATGATEWAYSPEC

vpc	String	VPC	Pod	VPC
subnet	String	VPC	Pod	
externalSubnets	[]String			
lanIp	String	VPC	Pod	IP
selector	[]String	Kubernetes Selector		
tolerations	[]Toleration	Kubernetes		
affinity	Affinity	Kubernetes		
qosPolicy	String	QoS		
bgpSpeaker	VpcBgpSpeaker	BGP speaker		

VpcBgpSpeaker

enabled	Bool	BGP speaker
asn	Uint32	
remoteAsn	Uint32	
neighbors	[]String	BGP
holdTime	Duration	BGP
routerId	String	BGP ID
password	String	BGP
enableGracefulRestart	Bool	
extraArgs	[]String	

Route

cidr	String	CIDR
nextHopIP	String	IP

VPCNATGATEWAYSTATUS

qosPolicy	String	QoS
externalSubnets	[]String	
selector	[]String	Kubernetes Selector
tolerations	[]Toleration	Kubernetes
affinity	Affinity	Kubernetes

VpcEgressGateway

apiVersion	String	Kubernetes	kubeovn.io/v1
kind	String	Kubernetes	VpcEgressGateway
metadata	ObjectMeta	Kubernetes	
spec	VpcEgressGatewaySpec	VpcEgressGateway	
status	VpcEgressGatewayStatus	VpcEgressGateway	

VPCEGRESSGATEWAYSPEC

vpc	String	VPC
replicas	Int32	
prefix	String	
image	String	
internalSubnet	String	
externalSubnet	String	
internalIPs	[]String	IP
externalIPs	[]String	IP
trafficPolicy	String	

VpcDns

apiVersion	String	Kubernetes	kubeovn.io/v1
kind	String	Kubernetes	VpcDns
metadata	ObjectMeta	Kubernetes	
spec	VpcDNSSpec	VpcDns	
status	VpcDNSStatus	VpcDns	

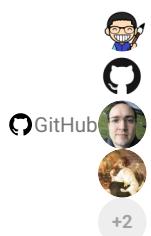
VPCDNSSPEC

replicas	Int32	
vpc	String	VPC
subnet	String	



⌚2025 9 2

⌚2023 2 16



8.11.9

8.12 Annotation

Kube-OVN	Pod	Node Annotation	Annotation
			Annotation

8.12.1 Pod Annotation

Key	Value	Description														
ovn.kubernetes.io/allocated	true or false	Pod														
ovn.kubernetes.io/routed	true or false	Pod	OVN													
ovn.kubernetes.io/routes	String	Pod														
ovn.kubernetes.io/mac_address	String	Pod	Mac	Pod	Annotation	Mac										
ovn.kubernetes.io/ip_address	String	Pod	IP	Pod	Annotation	IP										
ovn.kubernetes.io/cidr	String	Pod	CIDR													
ovn.kubernetes.io/gateway	String	Pod	Gateway													
ovn.kubernetes.io/ip_pool	IP	Pod	Workload		IP											
ovn.kubernetes.io/bgp	true, cluster, local	BGP	Pod													
ovn.kubernetes.io/snat	String	Pod	SNAT													
ovn.kubernetes.io/eip	String	Pod	EIP													
ovn.kubernetes.io/vip	String	Pod	VIP	Annotation	VIP											
ovn.kubernetes.io/aaps	String	Pod	AAPs (Additional Allowed Addresses Pairs)													
ovn.kubernetes.io/virtualmachine	String	Pod	VirtualMachineInstance													
ovn.kubernetes.io/activation_strategy	String	Pod														
ovn.kubernetes.io/logical_router	String	Pod	VPC													
ovn.kubernetes.io/layer2_forward	true or false	Pod	OVN LSP	unknown												
ovn.kubernetes.io/port_security	true or false	Pod	Port Security													
ovn.kubernetes.io/logical_switch	String	Pod														
ovn.kubernetes.io/vlan_id	Int	Pod	Vlan ID													
ovn.kubernetes.io/ingress_rate	Int	Pod	Mbits/s													
ovn.kubernetes.io/egress_rate	Int	Pod	Mbits/s													
ovn.kubernetes.io/security_groups	String	Pod	Security Group													
ovn.kubernetes.io/default_route	true or false															
ovn.kubernetes.io/provider_network	String	Pod	ProviderNetwork													
ovn.kubernetes.io/mirror	true or false	Pod														
ovn.kubernetes.io/north_gateway	String	Pod														
ovn.kubernetes.io/latency	Int	Pod	ms													
ovn.kubernetes.io/limit	Int	Pod	qdisc													
ovn.kubernetes.io/loss	Float	Pod														
ovn.kubernetes.io/jitter	Int	Pod	ms													
ovn.kubernetes.io/generate-hash	true or false	Pod														
ovn.kubernetes.io/attachmentprovider	String	Pod														

8.12.2 Node Annotation

Key	Value	Description		
ovn.kubernetes.io/allocated	true or false	ovn0	join	
ovn.kubernetes.io/mac_address	String	Node ovn0	Mac	
ovn.kubernetes.io/ip_address	String	Node ovn0	IP	
ovn.kubernetes.io/cidr	String	Node ovn0	join	CIDR
ovn.kubernetes.io/gateway	String	Node ovn0	join	Gateway
ovn.kubernetes.io/chassis	String	Node OVN-SouthBoundDB	Chassis ID	
ovn.kubernetes.io/port_name	String	Node ovn0	OVN-NorthboundDB	LSP
ovn.kubernetes.io/logical_switch	String	Node ovn0		
ovn.kubernetes.io/tunnel_interface	String			

8.12.3 Namespace Annotation

Key	Value	Description	
ovn.kubernetes.io/cidr	CIDR	Namespace	CIDR
ovn.kubernetes.io/exclude_ips	excludelPs	Namespace	excludelPs

8.12.4 Subnet Annotation

Key	Value	Description
ovn.kubernetes.io/bgp	true, cluster, local	BGP

8.12.5 Service Annotation

Key	Value	Description		
ovn.kubernetes.io/bgp	true or false	BGP	Service	
ovn.kubernetes.io/switch_lb_vip	String	Service	Kube-OVN	VIP
ovn.kubernetes.io/vpc	String	Service	VPC	
ovn.kubernetes.io/service_external_ip_from_subnet	true or false	Service	IP	
ovn.kubernetes.io/service_health_check	true or false	Service		
ovn.kubernetes.io/lb_svc_img	String			

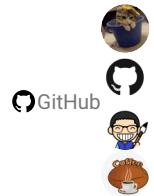
8.12.6 Networkpolicy Annotation

Key	Value	Description
ovn.kubernetes.io/enable_log	true or false	NetworkPolicy
ovn.kubernetes.io/log_acl_actions	"allow,drop,pass"	Action ACL

[PDF](#)[Slack](#)[Support](#)

⌚2025 11 20

⌚2024 6 12



8.12.7

8.13

8.13.1

Bad	Good

Bad	Good
Kube-OVN	1.10 Kube-OVN

Bad	Good
wget 127.0.0.1	wget 127.0.0.1

8.13.2

yaml yaml

Bad	Good
.... apiVersion: kubeovn.io/v1 kind: Subnet metadata: name: attach-subnetyaml apiVersion: kubeovn.io/v1 kind: Subnet metadata: name: attach-subnet

bash

Bad	Good
.... wget 127.0.0.1bash wget 127.0.0.1

#

Bad

```
oilbeater@macdeMac-3 ~ ping 114.114.114.114 -c 3
PING 114.114.114.114 (114.114.114.114): 56 data bytes
64 bytes from 114.114.114.114: icmp_seq=0 ttl=83 time=10.429 ms
64 bytes from 114.114.114.114: icmp_seq=1 ttl=79 time=11.360 ms
64 bytes from 114.114.114.114: icmp_seq=2 ttl=76 time=10.794 ms

--- 114.114.114.114 ping statistics ---
3 packets transmitted, 3 packets received, 0.0% packet loss
round-trip min/avg/max/stddev = 10.429/10.861/11.360/0.383 ms
```

Good

```
# ping 114.114.114.114 -c 3
PING 114.114.114.114 (114.114.114.114): 56 data bytes
64 bytes from 114.114.114.114: icmp_seq=0 ttl=83 time=10.429 ms
64 bytes from 114.114.114.114: icmp_seq=1 ttl=79 time=11.360 ms
64 bytes from 114.114.114.114: icmp_seq=2 ttl=76 time=10.794 ms

--- 114.114.114.114 ping statistics ---
3 packets transmitted, 3 packets received, 0.0% packet loss
round-trip min/avg/max/stddev = 10.429/10.861/11.360/0.383 ms
```

#

Bad

```
# mv /etc/origin/ovn/ovnnb_db.db /tmp
# mv /etc/origin/ovn/ovnsb_db.db /tmp
```

Good

```
mv /etc/origin/ovn/ovnnb_db.db /tmp
mv /etc/origin/ovn/ovnsb_db.db /tmp
```

8.13.3

md

Bad

```
[ ](http://kubeovn.github.io/prepare)
```

Good

```
[ ](../prepare.md)
```

Bad

```
[Kubernetes ](http://kubernetes.io)
```

Good

```
[Kubernetes ](http://kubernetes.io){: target="_blank" }
```

8.13.4

Bad

```
```bash
wget 127.0.0.1
```

```

Good

```
```bash
wget 127.0.0.1
```

```

Bad

```
```bash
wget 127.0.0.1
```

```

Good

```
```bash
wget 127.0.0.1
```

```



PDF



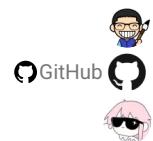
Slack



Support

⌚2025 9 10

⌚2022 7 19



8.13.5

9.



⌚2025 9 10

⌚2022 6 20



9.1
