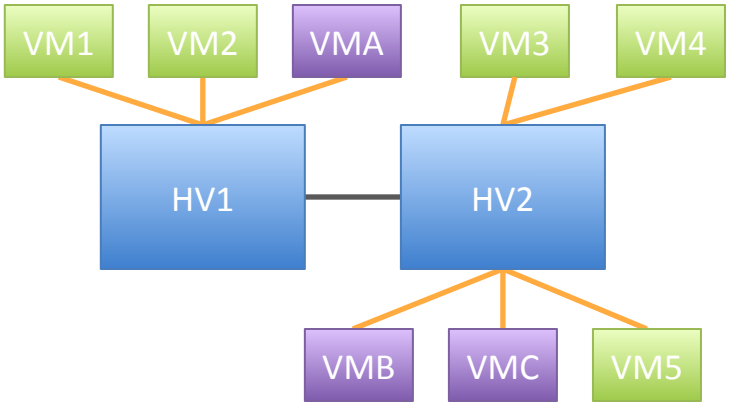
OVN L3 Gateways

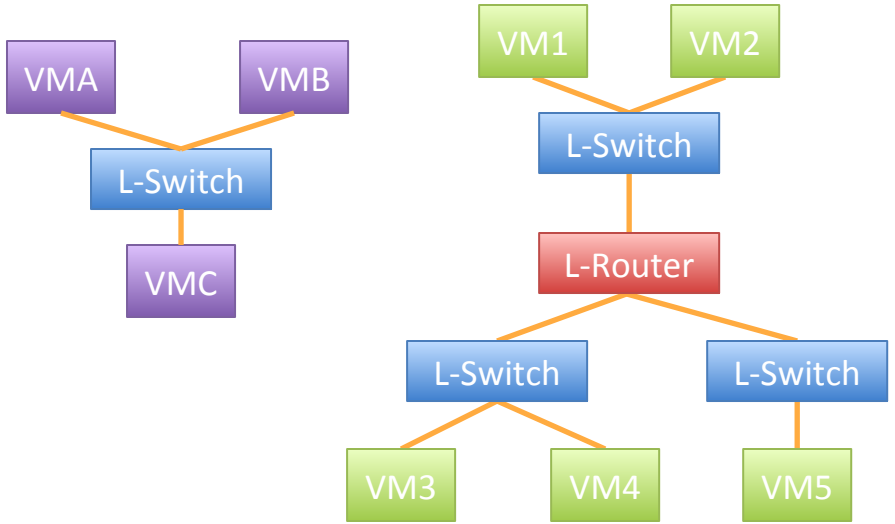
Russell Bryant (@russellbryant)

Virtual Networking Overview

Provides a logical network abstraction on top of a physical network

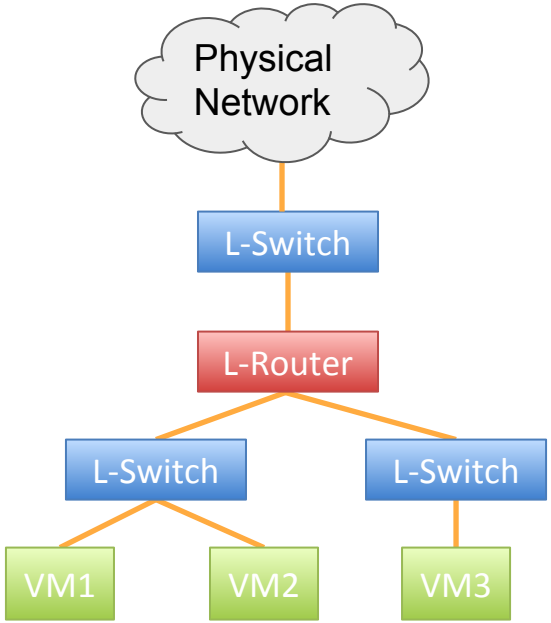


Physical

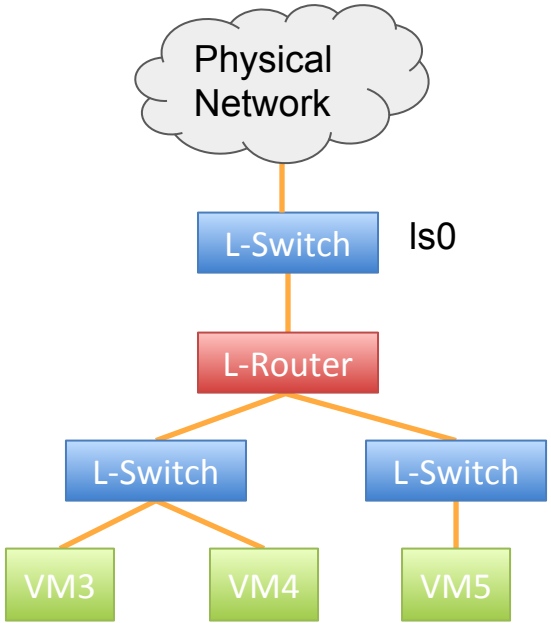


Logical

Modeling Gateways



Specify Logical Switch “localnet” connectivity



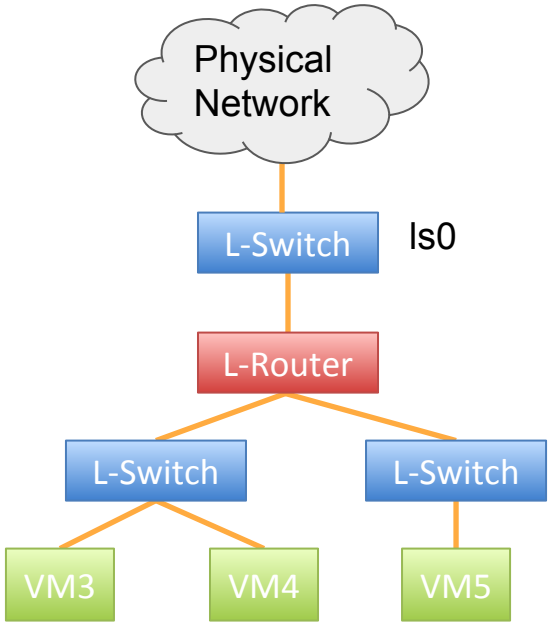

```
$ ovn-nbctl lsp-add ls0 ln_port

$ ovn-nbctl lsp-set-addresses ln_port unknown

$ ovn-nbctl lsp-set-type ln_port localnet

$ ovn-nbctl lsp-set-options network_name=physnet1
```

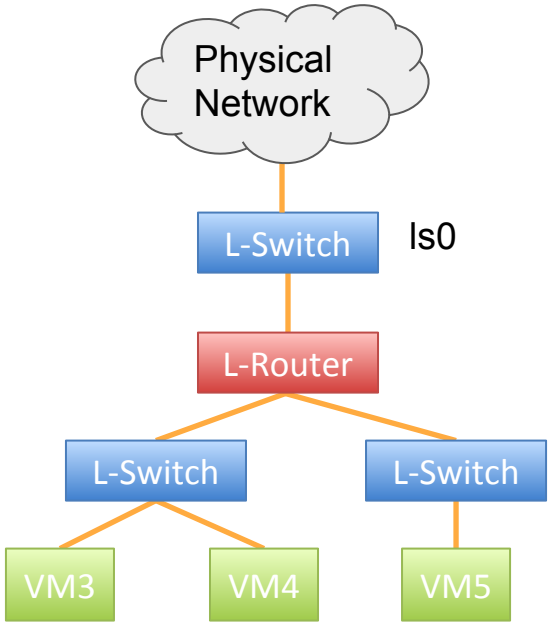
Define Bridge Mappings



(On every node that may host L3 gateways to a physical network)

```
$ ovs-vsctl set open . \
> external-ids:ovn-bridge-mappings=physnet1:br-ex
```

Set up SNAT

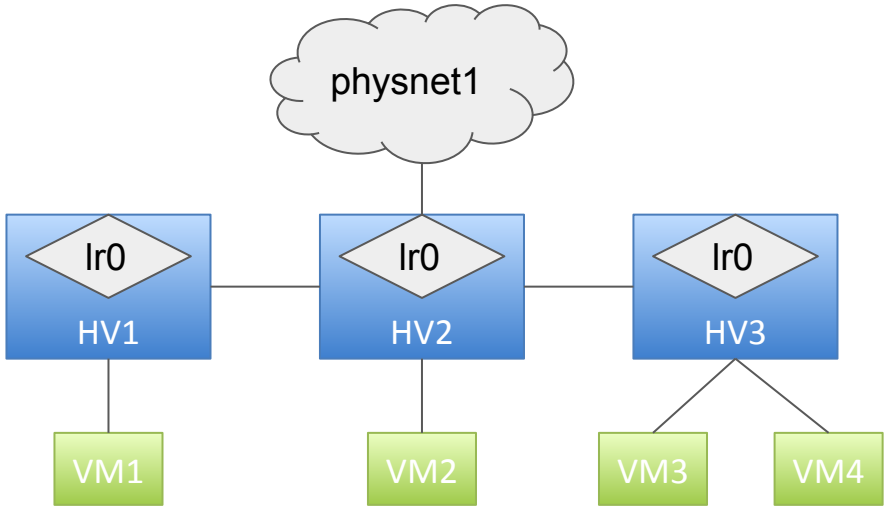



```
$ ovn-nbctl set Logical_Router_Port lrp0 \
> options:redirect-chassis=hv2

$ ovn-nbctl lr-nat-add lr0 \
> snat 192.168.0.1 10.0.0.0/24
```

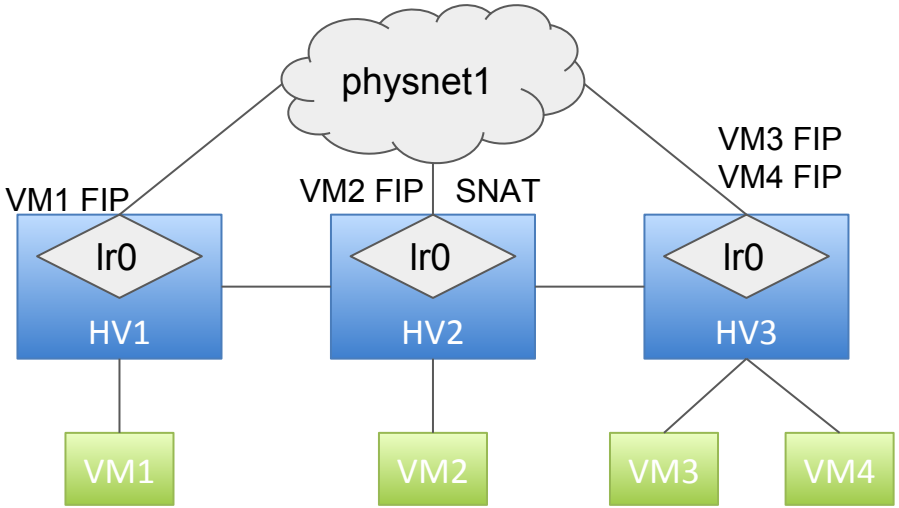
L3 Routing and Gateway

East-West L3 Routing is Distributed. SNAT traffic goes through HV2.



L3 Routing and Gateway

- Add physnet1 bridge mappings to all hypervisors.
- Floating IP SNAT/DNAT on each hypervisor.
- SNAT (without a floating IP) still centralized.

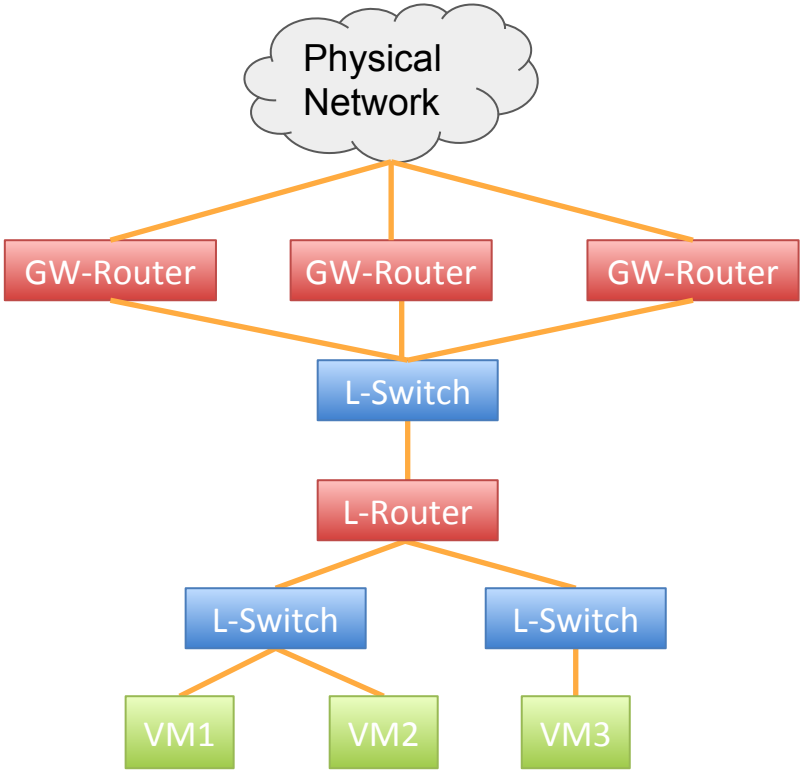



```
$ ovn-nbctl lr-nat-add lr0 snat \ >
192.168.0.1 10.0.0.0/24

$ ovn-nbctl lr-nat-add lr0 \
> dnat_and_snat 192.168.0.3 \
> 10.0.0.3 VM1 f0:00:00:00:00:03
```

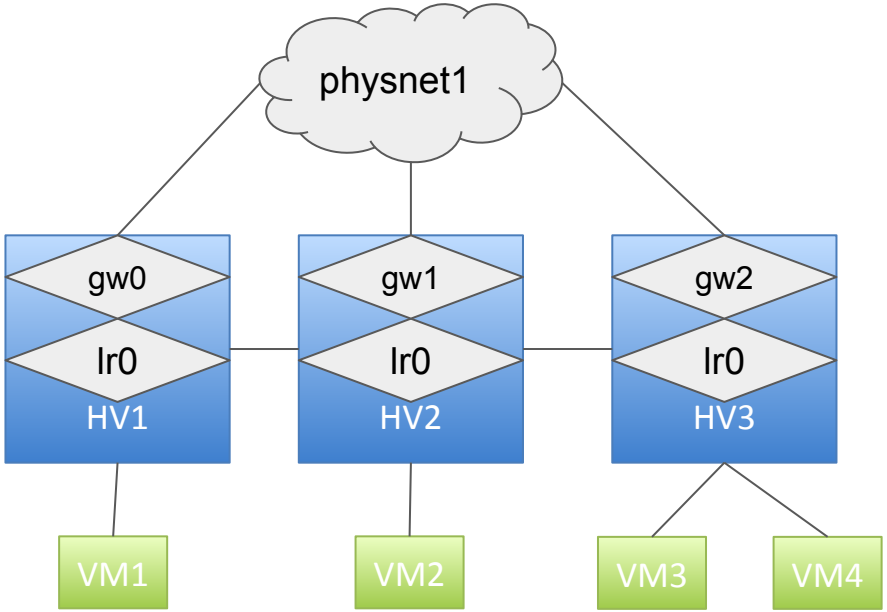
Multiple SNAT Gateways

- Separate distributed East-West from per-node gateways
- Use source based routing on lr0 to choose gateway



Multiple SNAT Gateways

- Separate distributed East-West from per-node gateways
- Use source based routing on lr0 to choose gateway




```
$ ovn-nbctl --policy="src-ip" \
> lr-route-add lr0 <VM1-IP> <gw0-IP>

$ ovn-nbctl --policy="src-ip" \
> lr-route-add lr0 <VM2-IP> <gw1-IP>

$ ovn-nbctl --policy="src-ip" \
> lr-route-add lr0 <VM3-IP> <gw2-IP>

$ ovn-nbctl --policy="src-ip" \
> lr-route-add lr0 <VM4-IP> <gw2-IP>
```

L3 Gateways and OpenStack

- Only centralized SNAT is supported so far
- Gateway scheduling done by networking-ovn
  - Determines all hosts with a bridge mapping to physical network
  - Chooses host with least number of gateways already scheduled there
- Still to integrate in networking-ovn
  - Integration of hypervisor-local floating IPs
  - Support for multiple SNAT gateways on a network

Thank you!

Russell Bryant (@russellbryant)

redhat.

IPv6 Support in OVN

Present status

Numan Siddique (@numansiddique)

redhat.

What does IPv6 support mean

- Support in underlay IPv6 network
- Router Advertisements ?
- Various addressing modes ? - slaac, DHCPv6
- E-W tenant network routing support ?
- N-S tenant network routing support
  - Prefix delegation, GUA routing support

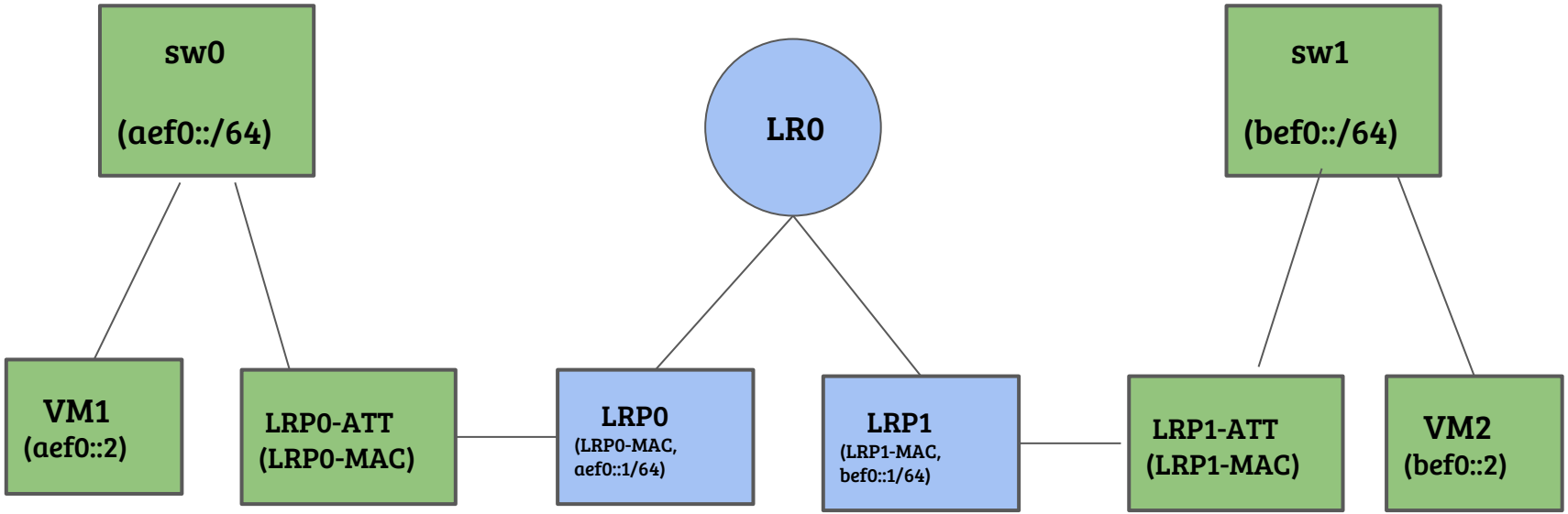
Does OVN services support underlay IPv6 network ?

- Yes
- ovsdb-server can be configured to listen on IPv6 addresses
- ovn-northd and ovn-controller can connect to the DB servers using IPv6 addresses.
- Geneve tunnels work.

```
$ sudo ovs-vsctl show
e62ca127-e486-428c-a201-50d8e3e015c8
    Bridge br-int
        Port "ovn-fd902d-0"
            Interface "ovn-fd902d-0"
                type: geneve
                options: {csum="true", key=flow, remote_ip="aef0::5054:ff:fe48:469e"}
```

L3 (EW) IPv6 routing in OVN

- Similar to IPv4 routing in OVN



Router Advertisements

- Presently OVN doesn't
  - send periodic RA packets
  - Doesn't respond to the Router Solicitation requests

- But, there are patches up for review which responds to the Router solicitation requests*.

- Periodic RA packet support needs to be supported.

- Hopeful that we will have full RA support in the next release.

* Thanks to Zong Kai Lai for the patches

How RA would be supported

Logical flow:

table=5 (lr_in_rs_rsp), priority=50 , match=(inport == "lrp0_ip6" && ip6.dst == ff02::2 && nd_rs),
action=(nd_ra{put_nd_ra_addr_mode("slaac");
put_nd_opt_sll(00:00:00:00:af:01);
put_nd_opt_prefix(aef0::/64);
eth.src = 00:00:00:00:af:01; ip6.src = fe80::200:ff:fe00:af01;
outport = inport; flags.loopback = 1; output;};)

OF Flow:

table=21,
priority=50,icmp6,reg14=0x2,metadata=0x3,ipv6_dst=ff02::2,nw_ttl=255,icmp_type=133,icmp_code=0
actions=controller(userdata=00.00.00.07.00.00.00.00.00.24.00.08.00.01.00.00.09.00.06.00.00.00.00.af.02.00.0b.00.11.40.be.f0.00.00.00.00.00.00.00.00.00.00.00.00.00.00.00.19.00.10.80.00.08.06.00.00.00.00.af.02.00.00.00.19.00.18.80.00.34.10.fe.80.00.00.00.00.00.00.02.00.00.ff.fe.00.af.02.ff.ff.00.18.00.00.23.20.00.06.00.20.00.00.00.00.00.01.1c.04.00.01.1e.04.ff.ff.00.18.00.00.23.20.00.07.00.00.00.01.14.04.00.00.00.00.00.00.00.01.ff.ff.00.10.00.00.23.20.00.0e.ff.f8.20.00.00.00)

Comparing neutron l3 agent with OVN

L3 agent

- Radvd daemon is run in the router namespace
- It sends periodic RA packets
- It responds to the RS packets

OVN

- RS packet is sent to ovn-controller
- ovn-controller sends the RA packet in response
- ovn-controller would send the periodic RA packets*

*No patches yet to support this

Native DHCPv6 support

- OVN supports native DHCPv6
- CMS configures the DHCPv6 options for each logical port
- ovn-northd adds logical flows to support DHCPv6
- ovn-controller responds to the DHCPv6 request packets

Neighbor solicitation support

- OVN has NS responder support similar to ARP responder support.
- ovn-northd adds logical flows in the logical switch ingress pipeline - ARP_ND_RESP
- Local ovn-controller responds to the NS request with the NA packet

Example.

table=10(ls_in_arp_rsp), priority=50 , match=(nd_ns && ip6.dst == {aef0::5264:ff:fe00:2, ff02::1:ff00:2} && nd.target == aef0::5264:ff:fe00:2), action=(**nd_na** { eth.src = 50:64:00:00:00:02; ip6.src = aef0::5264:ff:fe00:2; nd.target = aef0::5264:ff:fe00:2; nd.tll = 50:64:00:00:00:02; outport = inport; flags.loopback = 1; output; };)

N-S routing support

- No SNAT and DNAT support for IPv6
- When the IPv6 packet goes out of the logical gateway port, eth.dst is set to 00:00:00:00:00:00
- OVN should send a Neighbor solicitation for the next hop and this is missing.

Prefix delegation and GUA routing

- No native support for prefix delegation in OVN
- Neutron OVN plugin doesn't support this either
- Can be supported in the Neutron OVN plugin
- GUA routing will not work unless the issues mentioned earlier are not fixed

Summary

- No RA support yet
- Need to send NS request for next hop to have fully functional Gateway support

Thank you!

Numan Siddique (@numansiddique)