Troubleshooting & Debugging Openstack Neutron
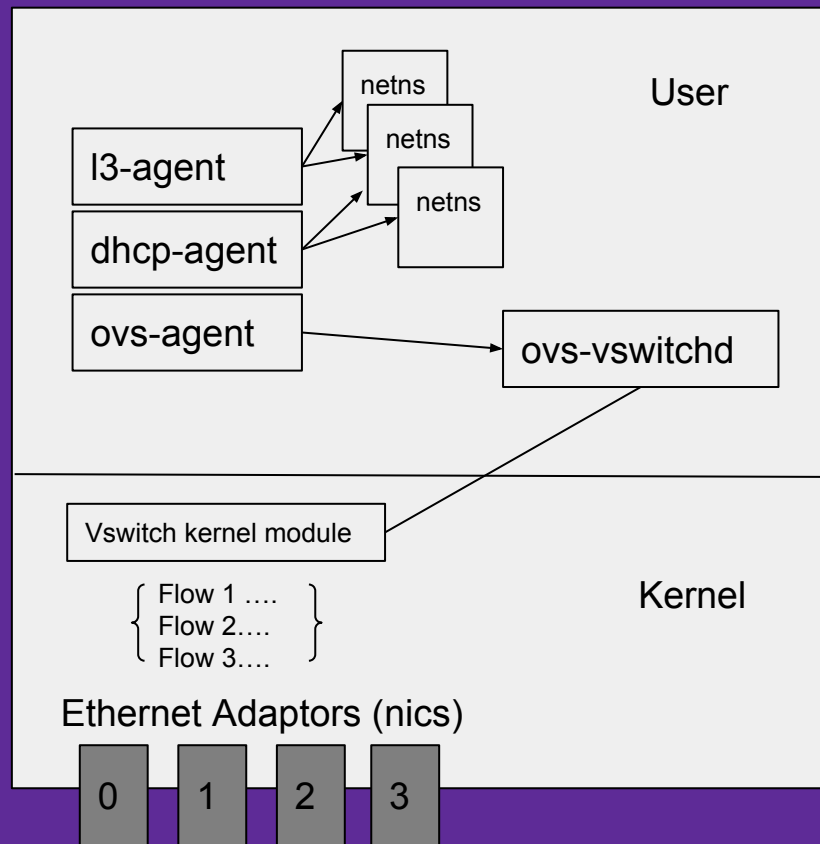
v.1

Check openstack pre-requisites first

- selinux/apparmor
- Ntp (Network Time Protocol)
- Firewall/networkmanager
- Kernel modules (loaded or not?)
- Sysctl configuration check
- Check if ovs is installed & is running - ovs-vsctl show
- MTU (Max. Transmission Unit)
- Server (physical or virtual?)
- System too slow? Check top/htop, ps aux

Neutron components

Ovs-agent
Dhcp-agent
L3-agent

Types of Networks supported (Type Drivers)

Flat networks

VLAN

VXLAN tunneling

GRE tunneling

GENEVE tunneling…

…

Types of L3 architectures

Non DVR (Distributed Virtual Routing)

DVR (Distributed Virtual Routing)

What to check first… IP connectivity

ip route show

```
[arun@arun-centos71-dev ~]$ ip route show
default via 10.4.0.1 dev ens33  proto static   metric 100
10.1.10.3 via 10.4.0.1 dev ens33  proto dhcp   metric 100
10.4.0.0/16 dev ens33  proto kernel  scope link  src 10.4.8.203  metric 100
172.17.0.0/16 dev docker0  proto kernel  scope link  src 172.17.0.1
192.168.122.0/24 dev virbr0  proto kernel  scope link  src 192.168.122.1
```

Ping the default gw ip (Eg: 10.4.0.1)

Ping google DNS (8.8.8.8)

Ping google using hostname (www.google.com)

Ping the other host ips. (Both mgmt traffic & data traffic interfaces/ips)

Check ovs switch connections.

Arun Sriraman

OVS Switches will be connected with each other through patch cables. Check if these cables exist.

```
root@ubuntu01:~# ovs-vsctl show
e8cbdc8e-95f3-46e5-a836-31af58d208b9
    Bridge br-ex
        Port "qg-2dd860c5-f0"
            Interface "qg-2dd860c5-f0"
                type: internal
        Port br-ex
            Interface br-ex
                type: internal
        Port phy-br-ex
            Interface phy-br-ex
                type: patch
                options: {peer=int-br-ex}
        Port "eth2"
            Interface "eth2"
    Bridge br-vlan
        Port phy-br-vlan
            Interface phy-br-vlan
                type: patch
                options: {peer=int-br-vlan}
```

```
        Port "eth1"
            Interface "eth1"
        Port br-vlan
            Interface br-vlan
                type: internal
    Bridge br-int
        fail_mode: secure
        Port "tap2b68f018-74"
            tag: 1
            Interface "tap2b68f018-74"
                type: internal
        Port "qvo69b95a39-b2"
            tag: 1
            Interface "qvo69b95a39-b2"
        Port int-br-ex
            Interface int-br-ex
                type: patch
                options: {peer=phy-br-ex}
        Port int-br-vlan
            Interface int-br-vlan
                type: patch
                options: {peer=phy-br-vlan}
```

```
        Port br-int
            Interface br-int
                type: internal
        Port patch-tun
            Interface patch-tun
                type: patch
                options: {peer=patch-int}
    Bridge br-tun
        fail_mode: secure
        Port patch-int
            Interface patch-int
                type: patch
                options: {peer=patch-tun}
        Port "vxlan-0a0465fd"
            Interface "vxlan-0a0465fd"
                type: vxlan
                options: {df_default="true", in_key=flow,
local_ip="10.4.213.120", out_key=flow,
remote_ip="10.4.101.253"}
        Port br-tun
            Interface br-tun
                type: internal
    ovs_version: "2.5.0"
```

Check if all of the phy ifaces are UP

```
root@ubuntu01:~# ip a
1: lo: <LOOPBACK,UP,LOWER_UP> mtu 65536 qdisc noqueue state UNKNOWN group default
    link/loopback 00:00:00:00:00:00 brd 00:00:00:00:00:00
    inet 127.0.0.1/8 scope host lo
        valid_lft forever preferred_lft forever
    inet6 ::1/128 scope host
        valid_lft forever preferred_lft forever
2: eth0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc pfifo_fast state UP group default qlen
1000
    link/ether fa:16:3e:ae:c7:b4 brd ff:ff:ff:ff:ff:ff
    inet 10.4.213.120/16 brd 10.4.255.255 scope global eth0
        valid_lft forever preferred_lft forever
    inet6 fe80::f816:3eff:feae:c7b4/64 scope link
        valid_lft forever preferred_lft forever
3: eth1: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc pfifo_fast master ovs-system state UP
group default qlen 1000
    link/ether 00:50:56:8b:9c:3d brd ff:ff:ff:ff:ff:ff
    inet6 fe80::250:56ff:fe8b:9c3d/64 scope link
        valid_lft forever preferred_lft forever
4: eth2: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc pfifo_fast master ovs-system state UP
group default qlen 1000
    link/ether 00:50:56:8b:af:86 brd ff:ff:ff:ff:ff:ff
    inet6 fe80::250:56ff:fe8b:af86/64 scope link
        valid_lft forever preferred_lft forever
```

What are the symptoms.....?

1. Instance creation fails
   a. AMQP Error in nova compute logs.. Oops this could be anything :(
   b. Port creation fails (neutron logs), check ovs-agent logs. (VIF binding failed? Possibly ovs-agent down.. Nova timeout after 5 mins)
   c. Retried exception?
2. VM cannot ping (Very helpful information right…?)
   a. Check if VM has an IP. (Check if VIF binding failed? Probably a DHCP issue? - Try dhclient manually)
   b. VM has an IP and can't ping another VM on same network? (Try same host)
   c. VM on same host can't ping each other? (Oops, possibly a security group thing)
   d. VM on same L2 but different hosts can't ping? (Well well well, is it the underlying network? Try tcpdump - lets take this up later, in the mean time try pinging DHCP server)
   e. VM on different L2 network (Is router connected, can I ping the router gw interface?)
   f. VM cannot ping Google. (Can I ping router interface? Can I ping google from router namespace? - lets take a look at this later as well)
   g. I have a FIP and everything seems fine, just can't get into the VM? (Check security grp)
   h. I have a FIP and all hell broke loose (VM is kind of dead on the network?) - Advanced debugging…
   i. ……………
3. ……………

Deep dive into network namespaces

Listing namespaces on a linux box:

```
root@ubuntu01:~# ip netns
snat-203dc92c-18c1-4799-8091-0bcd1f81f3f4
qrouter-203dc92c-18c1-4799-8091-0bcd1f81f3f4
Qdhcp-d188b0f8-9f03-4d3a-bf74-9e39de33168e
```

Run commands within a network namespace: `ip netns exec <namespace> <command>`
- NS has its own net stack (ip route, interfaces, iptables …)

SNAT namespace:
- 2+x ifaces: 1 lo, 1 public network (qg-xx), x tenant n/w Ifaces (sg-xx)
- qg-xx interface is part of external bridge and sg-xx is part of integration bridge in the root namespace.
- Ping VM from snat namespace if tenant n/w is connected to the router & ping to 8.8.8.8 does not work.
- Ping External gateway from within this namespace (check external connectivity)

For floating IPs:
- Fip namespace for DVR. For Non-DVR qrouter-xxx namespace.

Continuing namespaces….

The qdhcp namespace:
- Responsible for per. tenant DHCP.
- Dnsmasq process serves a particular tenant subnet with ip & dns services.
- Dhcp server only responds to IP-mac pair already present in the host file. Does not give out IPs to any dhcp requests.

```
nobody    20506    1  0 Jun06 ?         00:00:00 dnsmasq --no-hosts --no-resolv --strict-order
--except-interface=lo --pid-file=/var/opt/pf9/neutron/dhcp/d188b0f8-9f03-4d3a-bf74-9e39de33168e/pid
--dhcp-hostsfile=/var/opt/pf9/neutron/dhcp/d188b0f8-9f03-4d3a-bf74-9e39de33168e/host
--addn-hosts=/var/opt/pf9/neutron/dhcp/d188b0f8-9f03-4d3a-bf74-9e39de33168e/addn_hosts
--dhcp-optsfile=/var/opt/pf9/neutron/dhcp/d188b0f8-9f03-4d3a-bf74-9e39de33168e/opts
--dhcp-leasefile=/var/opt/pf9/neutron/dhcp/d188b0f8-9f03-4d3a-bf74-9e39de33168e/leases
--dhcp-match=set:ipxe,175 --bind-interfaces --interface=tap2b68f018-74
--dhcp-range=set:tag0,192.168.2.0,static,86400s --dhcp-lease-max=256 --conf-file= --server=8.8.8.8
--domain=openstacklocal
```

All dhcp/dns information you need can be found in the host file. To see if neutron pushed port information (new VMs IP & mac) to the dhcp server, check the hosts file. DHCP uuid is same the network UUID.

```
Eg:    root@ubuntu01:~# cat /var/opt/pf9/neutron/dhcp/d188b0f8-9f03-4d3a-bf74-9e39de33168e/host
       fa:16:3e:3f:ef:6d,host-192-168-2-3.pf9-demo.com.,192.168.2.3
       fa:16:3e:52:10:38,host-192-168-2-17.pf9-demo.com.,192.168.2.17
       fa:16:3e:f8:b7:f4,vm03.pf9-demo.com.,192.168.2.21
       fa:16:3e:77:24:ab,vm01.pf9-demo.com.,192.168.2.23
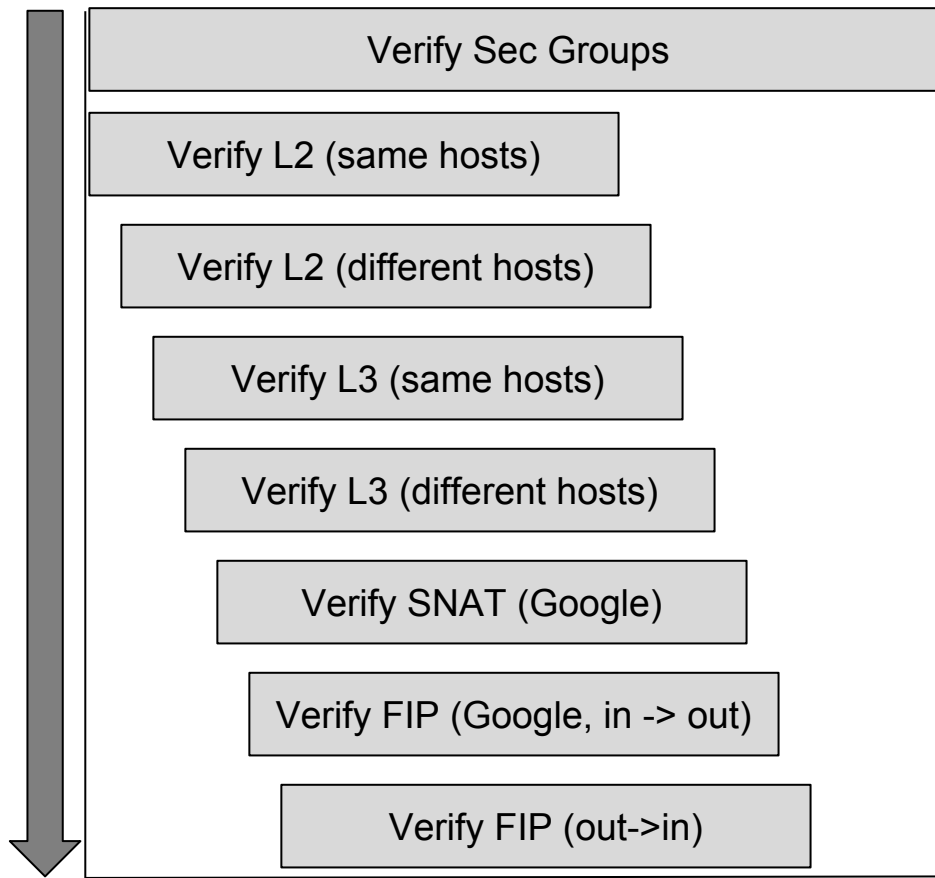```

# DHCP does not work….. :( what next?

- Try dhclient manually, does it get an IP? (Hope it does..)
- Keep dhclient running. Check if the data interface (phy) can see your bootp request packets
  - `tcpdump -i eth0 -n -e -vv | grep boot`
- Check if the bootp request packet reaches the host running dnsmasq. (tcpdump on phy of host2)
- Check if the packet is seen with

- Repeat: For bootp reply packet

## *Find that "blackhole"!*

# Can't figure out the issue? Start from the top!

Verify Sec Groups

Verify L2 (same hosts)

Verify L2 (different hosts)

Verify L3 (same hosts)

Verify L3 (different hosts)

Verify SNAT (Google)

Verify FIP (Google, in -> out)

Verify FIP (out->in)

Great Blog talking about Neutron troubleshooting.

**http://abregman.com/2016/01/06/openstack-neutron-troubleshooting-and-solving-common-problems/**

Avoiding content duplication here :-)

# Retrospect - Most common root causes

- Neutron physical prerequisites setup issues. (Bridge not connected properly. Interface not up)
- Core-network infra configuration (VLANs, upstream bonds, phy interface down)
- MTU problems (hostagent communication, tunneling needs > MTU)
- Arp-cache problem with upstream router disguised as heat stack issue.
- SELinux & firewall problems
- Kernel parameter configuration
- Loop problems in the way of configuring neutron routers. (Education)
- Configuration issues (Need for different ways of networking phy servers)

# Questions