Open vSwitch in Neutron

Performance Challenges and Hardware Offload

Date: Hong Kong, 6th Nov. 2013

Authors: Yongsheng Gong  gongysh@unitedstack.com
         Bo Liang         liang.bo@99cloud.net
197 Deployments

Network Drivers

- Linux Bridge 31%
- Open vSwitch 39%
- Cisco 11%
- Other 19%
- Hyper-V 5%
- Nicira 5%
- Brocade 3%
- big switch 2%
- NEC 2%

Agendas

1. Open vSwitch usages in Neutron
2. Open vSwitch & Openflow
3. Problem Statement
4. Acceleration Solutions
Open vSwitch usages in Neutron
Popular Deployment

**Public Switch 10.10.1.0/24**

- CONTROLLER
  - KEYSONE
  - NOVA-COMPONENTS
  - GLANCE
  - NEUTRON-SERVER
  - CINDER

- NETWORK NODE
  - OVS-AGENT
  - DHCP-AGENT
  - L3-AGENT
  - METADATA-AGENT

- COMPUTE NODE
  - NOVA-COMPUTE
  - OVS-AGENT

**Private Switch 172.16.0.0/24**
VM Traffic Path
VLAN Bridges

Router

br-ex

br-int

br-eth1

br-eth1 (veth)

veth

eth1

10.0.0.1

10.0.0.2

10.0.0.3

10.0.0.4

74.125.31.101

74.125.31.1

74.125.31.254
GRE Bridges

1.1.1.1

br-tun
- gre-2(remote_ip="1.1.1.2")
- patch-int (peer=patch-tun)

br-int
- patch-int (peer=patch-tun)
- qr-c0b1ff-ce(tag=1)
- tap-cfbc01f-a1(tag=1)

br-ex
- qg-9c24ade1c0-c1

eth1

1.1.1.2

br-tun
- gre-2(remote_ip="1.1.1.1")
- patch-int (peer=patch-tun)

br-int
- patch-int (peer=patch-tun)
- tap-99c710cc-ce(tag=1)
- tapa-98bfce2c-11(tag=1)

eth1

Router

10.0.0.1 10.0.0.3
10.0.0.2 10.0.0.4
74.125.31.101
74.125.31.1
74.125.31.101
74.125.31.254
74.125.31.1
Neutron Workflow

1. Start Neutron-Server
2. Start Open vSwitch Agent
3. Start L3-Agent
4. Start DHCP-Agent
5. Start METADATA-Agent
6. Create Networks
7. Create Routers
8. Boot VMs
GRE Tunnel

Host1

Tunnel 1-2

Host2

Tunnel 1-3

Tunnel 2-3

Host3
Open vSwitch & Openflow
Open vSwitch & Openflow

L2 switching: “in_port=*,actions=output:7”
Broadcast: “in_port=1,actions=output:flood”
Flow Tables with L2 Population

Table 10
LEARN_FROM_TUN

Table 2
GRE_TUN_TO_LV

Table 0
in_port=1

dl_dst=00:00:00:00:00:00/01:00:00:00:00:00

Table 1
PATCH_LV_TO_TUN

Table 20
UCAST_TO_TUN

Table 21
FLOOD_TO_TUN

ovs bridge br-tun

1 Patch in port

dl_dst=01:00:00:00:00:00/01:00:00:00:00:00

2 GRE Tunnel Port
Problem Statement
GRE Head

HTTP Server MTU Compatible Problem
MTU Comparison

![MTU Comparison Graph]

Bandwidth | TCP_RR | TCP_CRR | UDP | UDP_RR
---|---|---|---|---
500 | mtu1400 | mtu 1500
1000 | | | |
1500 | | | |
2000 | | | |
2500 | | | |
3000 | | | |
3500 | | | |
4000 | | | |
GRE Tunnel Extensibility

Tunnels: N - 1
Flow Match

- First Packet
- Subsequent

Source → user space → kernel → Destination
Improved Architecture

- Compute Node:
  - OVS/LinuxBridge
  - phynet
  - br-int

Tunnel Between Switches

A Switch Running OVS

- Compute Node:
  - OVS/LinuxBridge
  - phynet
  - br-int
Other Solutions

- Flow Matching
  - DPDK, PF_RING, netmap
- GRE Tunnel
  - VxLAN
OVS Hardware Comparison

Bandwidth | TCP_RR | TCP_CRR | UDP | UDP_RR
---|---|---|---|---
OVS | centec v330

11000 | 8250 | 5500 | 2750 | 0
OVS Debugging Tips

• Test basic connectivity
• Use tcpdump to see if expected packets are on the wire
• Try it without OVS
• Use “ovs-ofctl dump-flows bridge” to see if packets hit the flows
Thanks!
That’s all for today

Date: Hong Kong, 6th Nov. 2013
Authors: Yongsheng Gong, Bo Liang