Red Hat
Enterprise Linux

Measuring Geneve Tunnel Throughput for Hardware Accelerated DataPath with OVN!!

Pradipta Sahoo
Senior Software Engineer

Haresh Khandelwal
Principal Software Engineer

Red Hat

Agenda

- To provide an operator's outlook on Geneve hw-offload overlay network performance with ml2/OVN core service provided by OpenStack.
- Explore Trex software traffic-generator to simulate encapsulated traffic with Geneve protocol.
- To identify the challenges hardware offload network while using Geneve hw offload test with OVN.
- Measure the line rate throughput numbers in Perf scenario
- Share the observation / findings

Possible Test Methodology

1. **Virtual Trex Traffic generator on Tenant Layer:**
   - Configured trafficgen emulator in VM running in the same tenant network.
   - During the test, we observed virtual trafficgen provide upto 6-7 Mpps due lack of hardware acceleration.

2. **Bare Metal Trex node with Intermediate Gateway host:**
   - Gateway do the encap/dcap packet between Trafficgen and Tenant VM and expected to take some cycle to process the packets.
   - No changes required in existing trafficgen code.
   - Since the OVN gateway is software, it can not sustain line rate from traffic generator.

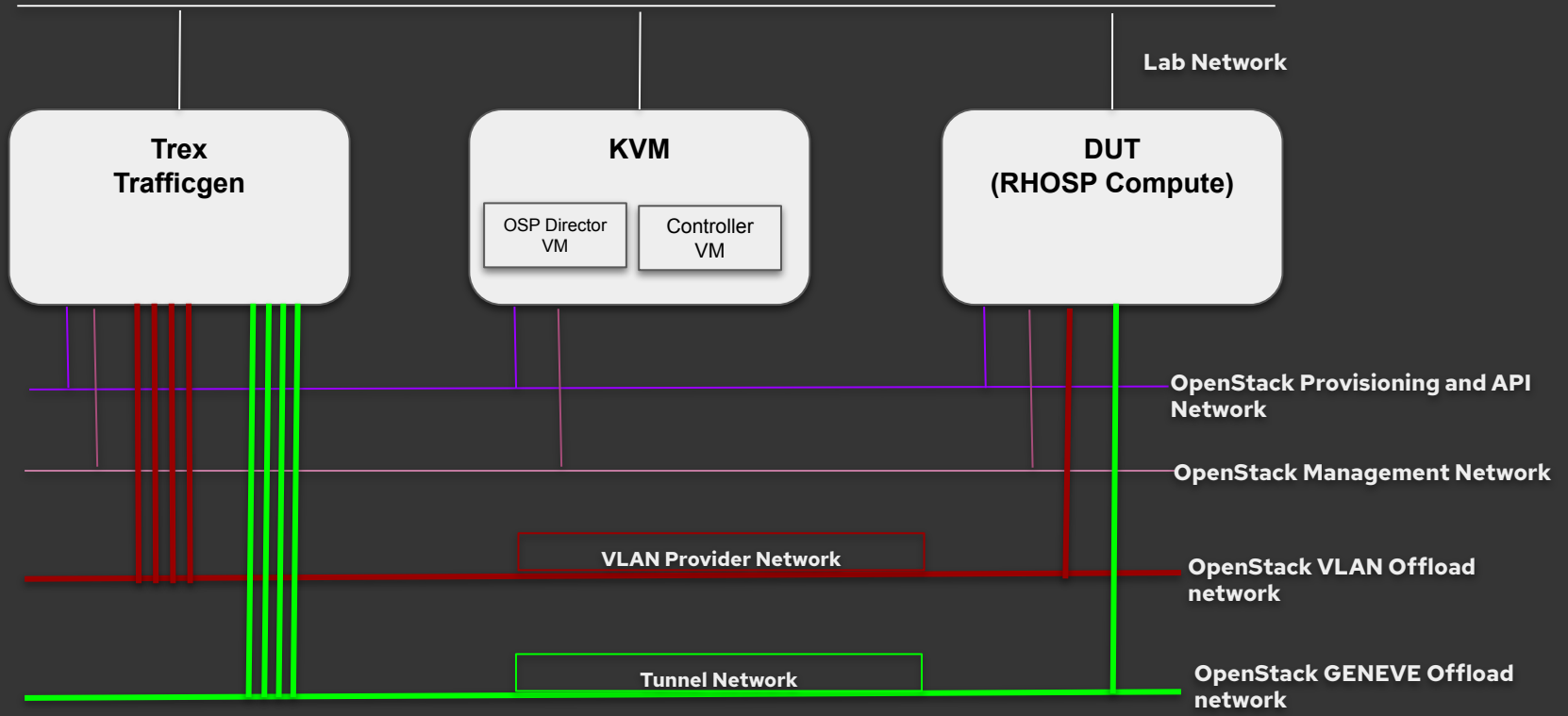
3. **Bare Metal Trex host without Intermediate Gateway Host:**
   - Using existing Trex's library to compose encapsulated packet with GENEVE header.
   - Transmit the encapsulated packets with linerate without additional processing through TEP gateway host.
   - Able to perform the L2-test.
   - Traffic generator established on outside of OpenStack cluster.
   - Ability to to compose packets for multiple vms.

Geneve Protocol Requirements

- Geneve RFC Requirements: https://datatracker.ietf.org/doc/html/rfc8926
  - Outer MAC Header (Source, Destination)
  - Outer VLAN Header (Dot1Q)
  - Outer IP Header (Source, Destination)
  - UDP Protocol (Source, Destination)
  - GENEVE Header (option_length, protocol, VNI, options_class, options_type, options_len, options_data)
  - Internal MAC Header
  - Internal IP Header
  - Internal Protocol
  - Payload

```
Ether(src='40:a6:b7:0b:e9:b1',dst='fa:16:3e:40:4d:7d')/
Dot1Q(vlan=177,type='IPv4')/
IP(proto=17,src='172.17.2.163',dst='172.17.2.60')/
UDP(sport=57025,dport=6081)/
GENEVE(optionlen=2,proto=0x6558,vni=6,options_class=0x0102,options_type=0x80,options_len=1,options_data=b"\x00\x03\x00\x02")/
Ether(src='fa:16:3e:7f:a1:49',dst='fa:16:3e:a0:1f:28',type=0x0800)/
IP(src='192.168.2.1',dst='192.168.2.181')/
UDP(dport=1000,sport=2000)/pad
```

OpenStack Network Topology



Red Hat

Software Specifications

1. **Red Hat OpenStack Platform release 16.2.0 Beta (Train)**
   - OS: Red Hat Enterprise Linux release 8.4 (4.18.0-305.el8.x86_64)
   - OpenvSwitch:
     - rhosp-network-scripts-openvswitch-2.15-4.el8ost.1.noarch
     - openvswitch2.15-2.15.0-15.el8fdp.x86_64
     - rhosp-openvswitch-2.15-4.el8ost.1.noarch
     - openvswitch-selinux-extra-policy-1.0-28.el8fdp.noarch
     - Network-scripts-openvswitch2.15-2.15.0-15.el8fdp.x86_64
   - OVN:
     - rhosp-ovn-2021-4.el8ost.1.noarch
     - ovn-2021-host-21.03.0-21.el8fdp.x86_64
     - ovn-2021-21.03.0-21.el8fdp.x86_64
     - Rhosp-ovn-host-2021-4.el8ost.1.noarch
2. **Trex**
   - v2.89 : https://trex-tgn.cisco.com/trex/release/v2.89.tar.gz
   - OS: Red Hat Enterprise Linux release 8.4 (4.18.0-305.el8.x86_64)

Geneve Scapy library status in Trex

- Challenges on existing Geneve scapy library (v2.4.3)

  (external_libs/scapy-2.4.3/scapy/contrib/geneve.py)

  - Geneve options fields not work for OVN class with metadata fields
- Enhanced the Geneve headers as per OVN metadata requirement.
- Raised the PR in upstream community: https://github.com/secdev/scapy/pull/3329
- Build the Traffic STL profile GENEVE scapy library and build Scapy packet with Geneve headers.

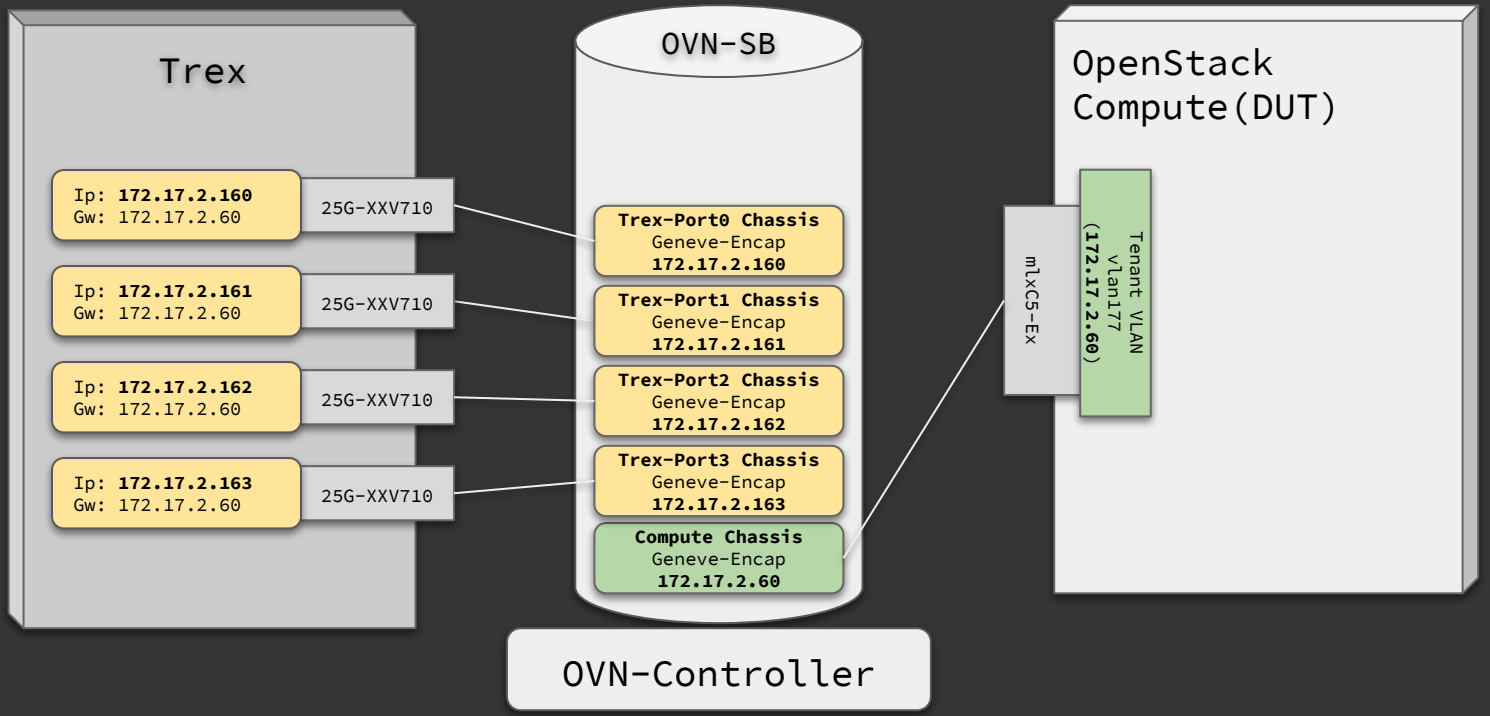
```
>>> packet.show2()
###[ GENEVE ]###
  version   = 0
  optionlen = 2
  oam       = 0
  critical  = 0
  reserved  = 0
  proto     = TEB
  vni       = 0x4
  reserved2 = 0x0
  \options   \
   |###[ Geneve Options ]###
   |  classid   = Open Virtual Networking (OVN)
   |  type      = 0x80
   |  reserved  = 0
   |  length    = 1
   |  data      = '\x00\x04\x00\x05'
```

Challenges and Solution

- Initial Challenges:
  - OVN controller discarded packets with any of below reason:
    - Incorrect outer headers with VLAN.
    - Incorrect Protocol
    - Incorrect Geneve metadata information.
    - Incorrect inner packet headers

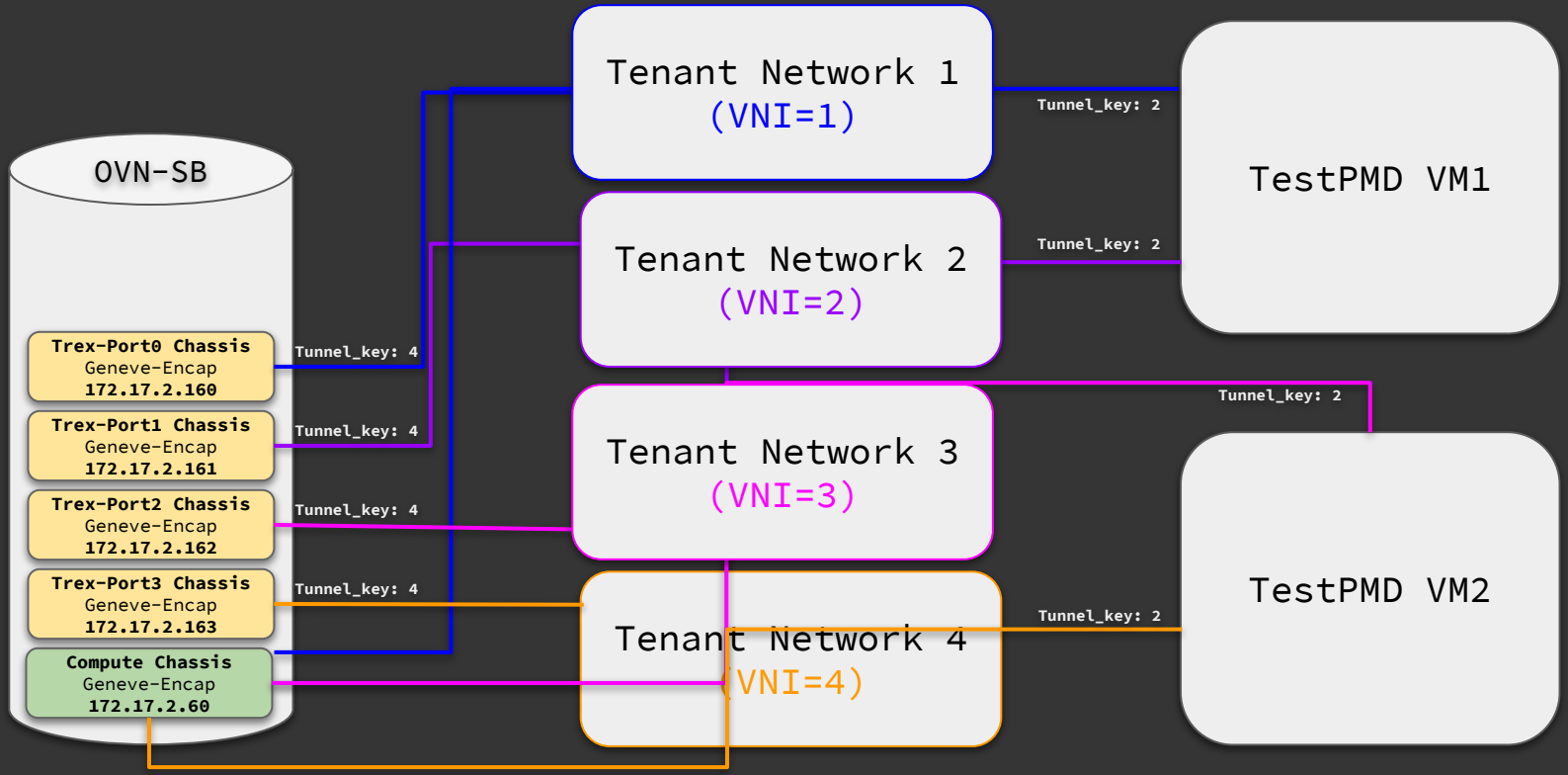
- Solution
  - Trex Server:
    - Outer header Src MAC and IPs can be same subnet of OpenStack Tenant network.
    - Outer header Dst MAC and IPs belongs OpenStack Compute (DUT) tenant network interface with same Tenant VLAN id.
    - Traffic profile update with correct metadata fields.
  - OVN Controller:
    - To recognize outside Geneve packets respective interface of Trex host need to register as a OVN SouthBound Chassis.
    - Map logical port of each tenant network to respective chassis and collect the tunel_key for source and destination traffic.
  - No changes need in OpenStack Compute (DUT) and OpenStack services.
  - TestPMD VM can be run as default IO mode.

Configuration topology



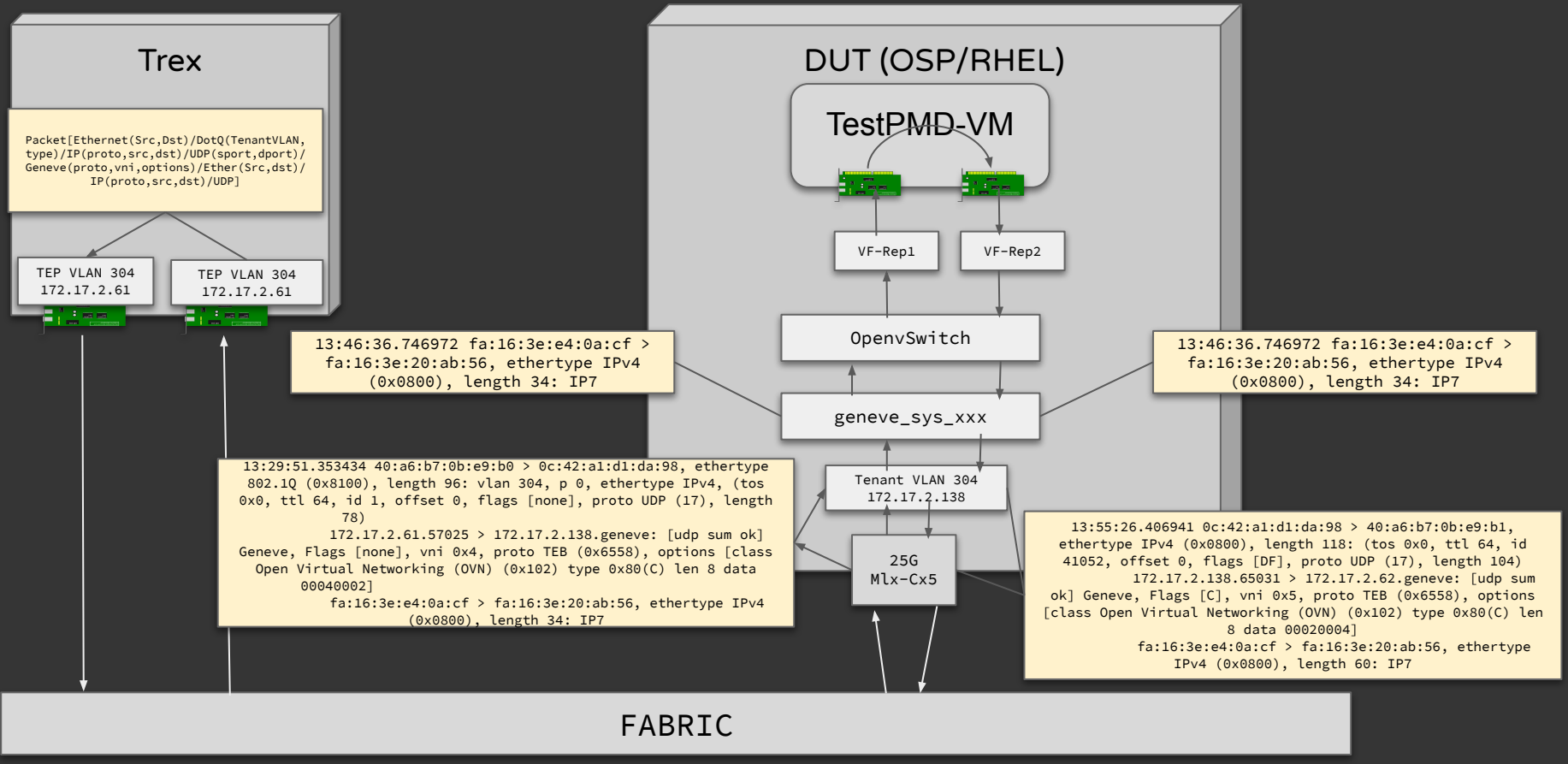
Red Hat

Logical View of Tenant Network Mapping



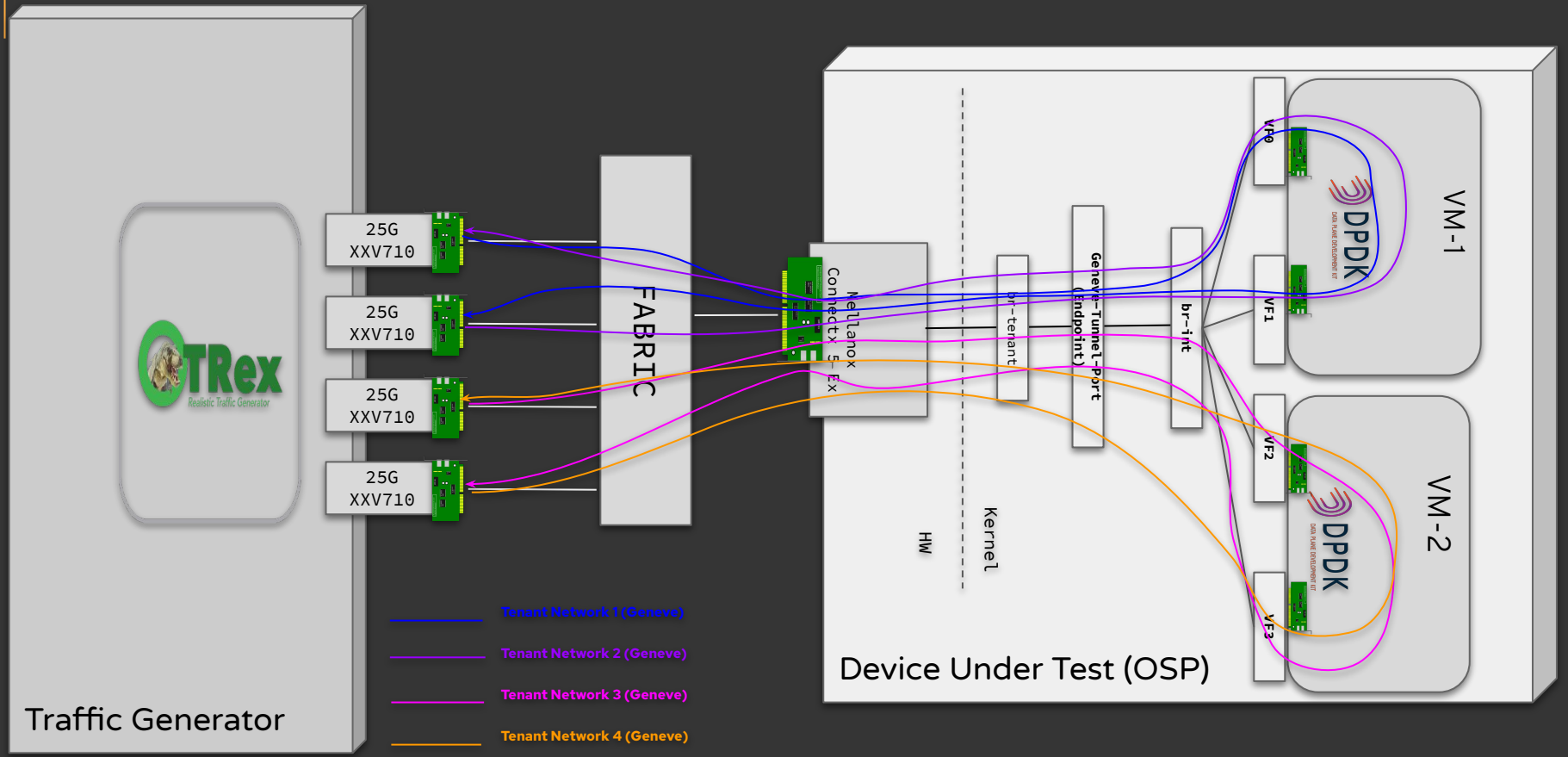
Red Hat

Geneve End-to-End Packet Journey using Trex



Red Hat

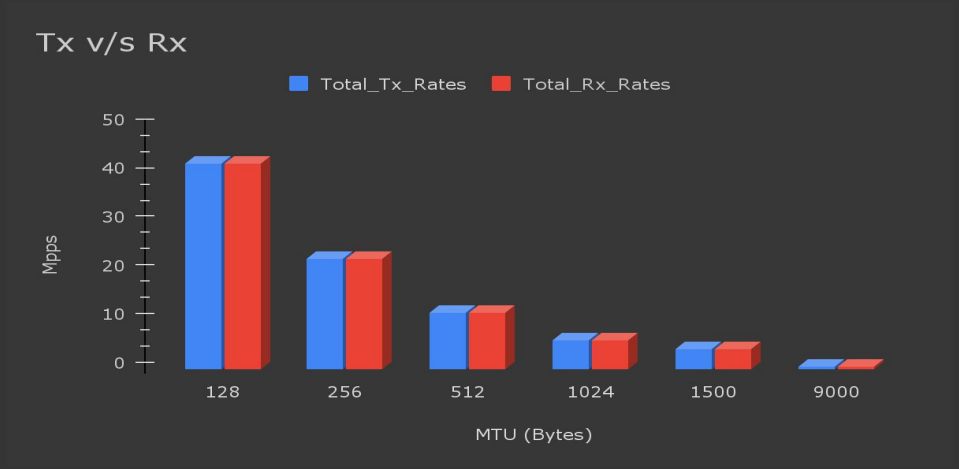
Geneve Bi-Directional Traffic Flow using Tenant Network



Red Hat

Performance Test results

Frame Size (Bytes)	Duration (min)	Traffic Mode	Traffic	Total_Tx_L1 (Gbps)	Total_Rx (Gbps)	Total_Tx_Rates (Mpps)	Total_Rx_Rates (Mpps)	Drop (PPS)	Cpu_Util (%)
128	10	PVP	Bi-Direction	50.02	41.91	42.25	42.25	0	8.12
256	10	PVP	Bi-Direction	50.05	50.05	22.67	22.67	0	4.37
512	10	PVP	Bi-Direction	50.11	47.85	11.77	11.77	0	2.24
1024	10	PVP	Bi-Direction	50	48.85	5.99	5.99	0	1.02
1500	10	PVP	Bi-Direction	50.06	49.3	4.12	4.12	0	4.67
9000	10	PVP	Bi-Direction	50.01	49.88	0.69307	0.69307	0	0.17



Red Hat

Conclusion

- Achieved linerate utilisation without intermediate host (ovn gw).
- Suffice the throughput requirement with traffic profiles.
- No changes required in OpenStack Compute and TestPMD. Can be used with others Infrastructure i.e Openshift
- Very minimal operator level changes in ovn-controller suffice the tunnel requirement

- Next Goal:
  - Integration with Trafficgen to run with binary-search support

Thank you

Red Hat is the world's leading provider of enterprise open source software solutions. Award-winning support, training, and consulting services make
Red Hat a trusted adviser to the Fortune 500.


linkedin.com/company/red-hat


youtube.com/user/RedHatVideos


facebook.com/redhatinc


twitter.com/RedHat

Red Hat