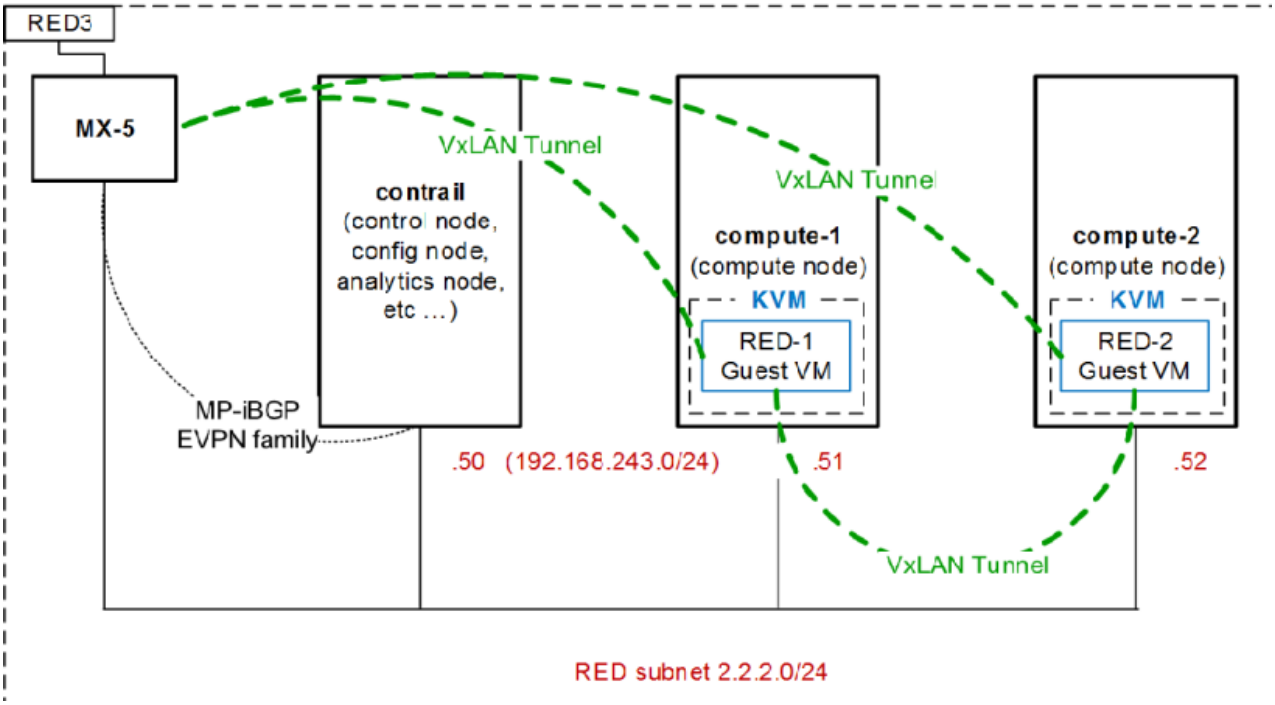
Packet Expert

Kashif Nawaz – JNCIEs (SP & Sec), RHCE and CKA

SDN

Contrail Integration with Bare Metal Devices via EVPN-VxLAN

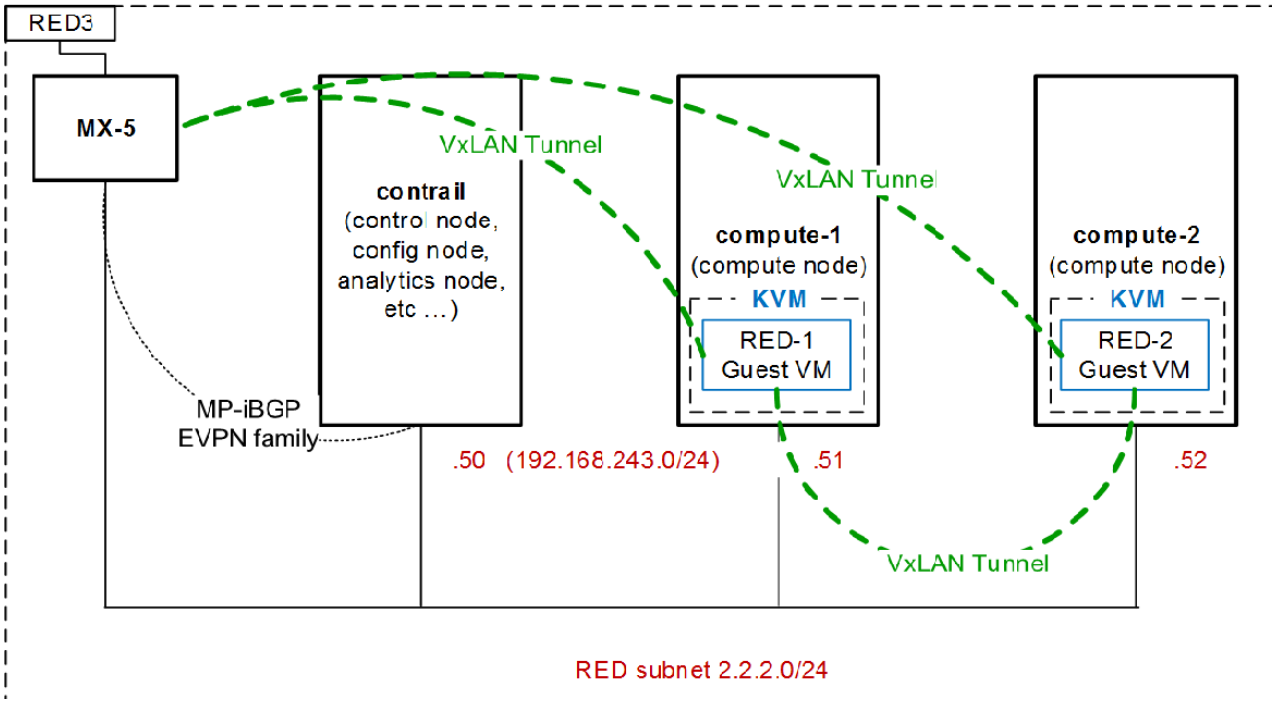


Date: November 24, 2016 Author: packetexpert 5 Comments

In this blog we will discuss how to integrate Bare metal devices with Juniper Contrail (SDN Controller) by using EVPN-VXLAN.

My earlier blogs on Contrail can be viewed on links Blog-1, Blog-2 ,

Reference Topology



Problem statement "Gust VM spawned inside SDN environment needs to communicate with Bare Metal Device (same sub net or different sub net here we will discuss former use case only).

Solution "EVPN based control plane will be established between MX Router and Contrail Controller to exchange ARP entries between them, VxLAN based forwarding plane will be configured for communication between Guest VMs and Bare Metal Devices"

Solution components:-

1. Contrail GUI
   - RED network 2.2.2.0/ is configured and VMs are spawned using open stack "Horizon" Web GUI (not covered in this article)
   - Configure VxLAN as 1st encapsulation method under "Encapsulation Priority Order" go to Configure then Infrastructure then Global Config and click edit button.
   - Select VxLAN Identifier Mode as "user configured"
   - Configure VxLAN ID & Route target community for the desired network
     - Go to Configure then Networking then Networks, select the desired network and click edit
     - Add the VxLAN ID (in this example 2000 is used for RED subnet) under Advanced Option
     - Add the route target community which is 200:1 for the RED network
2. MX Router
   - A Routing Instance "instance type virtual switch" configured "RED-EVPN"
   - Protocol EVPN configured inside routing-instance.
   - Bridge domain 200 and VxLAN VNI ID 2000 under EVPN routing instances
   - Physical port connected with Bare metal devices configured as access port with domain member 200.
   - Route target community on Virtual-Switch routing instance must match with route target community assigned to RED sub net in Contrail GUI
   - Forwarding plane functionality needs valid route pointing to compute nodes in inet.3 routing table of MX gateway. (If we recall VPN stuff in Junos, next-hop

lookup for VRF is always in inet.3.)

- There are two methods to achieve this goal:-
  - Adding route in inet.3 RIB dynamically through GRE tunnel.
  - Adding route in inet.3 RIB statically.
- I choose to go with both methods at same time for following reasons:-
  - Consider that we need to push external routes toward Contrail from MX gateway, in this case MPLS over GRE forwarding plane will be required which is dependent on dynamic GRE tunnel.
  - At same I need EVPN-VXLAN extension from Contrail to MX gateway for integration of bare metal devices connected on MX router, for this use case I wanted to avoid GRE tunnel to be used in forwarding plane, so adding static route in inet.3 RIB helped me.

Configuration

```
admin@ER> show configuration routing-instances RED-EVPN
vtep-source-interface lo0.101;
instance-type virtual-switch;
interface ge-1/1/9.0;                ##interface connected with bare metal device##
route-distinguisher 192.168.240:1;
vrf-target target:200:1;
protocols {
evpn {
encapsulation vxlan;
extended-vni-list 2000;
multicast-mode ingress-replication;
}
}
bridge-domains {
vlan200 {
domain-type bridge;
vlan-id 200;
routing-interface irb.200;    ##Optional
vxlan {
vni 2000;
ingress-node-replication;
}
}
}
```

```
admin@ER> show configuration routing-instances RED

##VRF configuration is optional , this use case is already covered in my earlier bog

instance-type vrf;
interface irb.200;
interface lo0.1;
route-distinguisher 200:1;
vrf-target target:200:1;      ##Route target is matching with virtual switch route target
vrf-table-label;
```

```
admin@ER> show configuration interfaces ge-1/1/9
unit 0 {
family bridge {
interface-mode access;
vlan-id 200;
}
}

admin@ER> show configuration routing-options dynamic-tunnels to-contrail
source-address 101.101.101.101;
gre;
destination-networks {
192.168.243.0/24;
}
admin@ER> show configuration routing-options rib inet.3
static {
route 192.168.243.0/24 receive;
}
admin@ER> show route table inet.3

inet.3: 5 destinations, 6 routes (5 active, 0 holddown, 0 hidden)
+ = Active Route, – = Last Active, * = Both
192.168.243.0/24 *[Static/5] 1d 02:04:03 ##static route added in inet.3 , will be used for
EVPN-VXLAN
Receive
[Tunnel/300] 1d 02:02:07
Tunnel
192.168.243.50/32 *[Tunnel/300] 1d 02:02:07 ##Dynamic routes added in inet.3 , will be
used for MPLS over GRE forwarding plane
> via gr-0/0/0.32770
192.168.243.51/32 *[Tunnel/300] 1d 02:02:07
> via gr-0/0/0.32769
192.168.243.52/32 *[Tunnel/300] 10:58:22
```

```
admin@ER> show bgp summary
Groups: 1 Peers: 1 Down peers: 0
Table          Tot Paths  Act Paths Suppressed    History Damp State    Pending
bgp.rtarget.0
                       2          1          0          0          0          0
inet.0
                       0          0          0          0          0          0
bgp.l3vpn.0
                       2          2          0          0          0          0
bgp.evpn.0
                       5          5          0          0          0          0
Peer                     AS      InPkt     OutPkt    OutQ   Flaps Last Up/Dwn State|#Active/Received/Accepted/Damped...
192.168.243.50          100       3874       4295       0       9  1d 8:08:36 Establ
  bgp.rtarget.0: 1/2/2/0
  bgp.l3vpn.0: 2/2/2/0
  bgp.evpn.0: 5/5/5/0
  RED.inet.0: 2/2/2/0
  RED-EVPn.evpn.0: 5/5/5/0
  __default_evpn__.evpn.0: 0/0/0/0
```

```
admin@ER> show route table RED-EVPn.evpn.0

RED-EVPn.evpn.0: 9 destinations, 9 routes (9 active, 0 holddown, 0 hidden)
+ = Active Route, - = Last Active, * = Both

2:192.168.240.0:1::2000::0c:86:10:1b:2d:c9/304                    ##MAC address of bare metal device"
            *[EVPN/170] 00:00:04
              Indirect
2:192.168.240.0:1::2000::20:4e:71:90:49:f8/304
            *[EVPN/170] 1d 01:54:08
              Indirect
2:192.168.243.51:1::2000::02:f3:7c:3c:85:a2/304
            *[BGP/170] 1d 02:05:45, localpref 100, from 192.168.243.50
              AS path: ?, validation-state: unverified
             > via gr-0/0/0.32769
2:192.168.243.52:1::2000::02:88:03:d5:d4:a2/304
            *[BGP/170] 10:31:27, localpref 100, from 192.168.243.50
              AS path: ?, validation-state: unverified
             > via gr-0/0/0.32771
2:192.168.240.0:1::2000::20:4e:71:90:49:f8::2.2.2.254/304
            *[EVPN/170] 1d 01:54:08
              Indirect
2:192.168.243.51:1::2000::02:f3:7c:3c:85:a2::2.2.2.3/304        ##ARP entry for RED 1 VM coming from compute 1 (192.168.243.51) with VNI
2000
            *[BGP/170] 1d 02:04:35, localpref 200, from 192.168.243.50
              AS path: ?, validation-state: unverified
             > via gr-0/0/0.32769
2:192.168.243.52:1::2000::02:88:03:d5:d4:a2::2.2.2.4/304        ##ARP entry for RED 2 VM coming from compute 2 (192.168.243.51) with VNI
2000
            *[BGP/170] 10:31:10, localpref 200, from 192.168.243.50
              AS path: ?, validation-state: unverified
             > via gr-0/0/0.32771


  admin@ER> show route advertising-protocol bgp 192.168.243.50 table RED-EVPn.evpn.0

  RED-EVPn.evpn.0: 9 destinations, 9 routes (9 active, 0 holddown, 0 hidden)
    Prefix              Nexthop          MED     Lclpref    AS path
    2:192.168.240.0:1::2000::0c:86:10:1b:2d:c9/304                    ##bare metal MAC advertised to contrail as type 2 route
  *                     Self             100     I
    2:192.168.240.0:1::2000::20:4e:71:90:49:f8/304
  *                     Self             100     I
    2:192.168.240.0:1::2000::20:4e:71:90:49:f8::2.2.2.254/304         ##ARP for bridge domain  200 advertised to Contrail
  *                     Self             100     I


  admin@ER> show bridge mac-table

  MAC flags     (S -static MAC, D -dynamic MAC, L -locally learned, C -Control MAC
    O -OVSDB MAC, SE -Statistics enabled, NM -Non configured MAC, R -Remote PE MAC)

  Routing instance : RED-EVPn
   Bridging domain : vlan200, VLAN : 200
     MAC               MAC     Logical            Active
     address           flags   interface          source
     02:88:03:d5:d4:a2  D      vtep.32770         192.168.243.52    ##MAC address for RED-1 learned throguh VTEP
     02:f3:7c:3c:85:a2  D      vtep.32769         192.168.243.51     ##MAC address for RED-2 learned throguh VTEP
     0c:86:10:1b:2d:c9  D      ge-1/1/9.0                            ##MAC address for RED-3 learned throguh physical interface
```

Published by packetexpert

Every new second is coming up with some innovation in the IT industry , the basic and foremost important building block behind all technology innovations and updates is the "PACKET". I always endeavored to understand packet anatomy started from switch access port , securing it and then further traversing through IP/ MPLS network till its destination. During my journey to understand packet anatomy I achieved 2 x JNCIEs (SP and Security) and currently learning Open-stack and SDN besides bit of automation stuff using Python. View all posts by packetexpert

5 thoughts on “Contrail Integration with Bare Metal Devices via EVPN-VxLAN”

Add Comment

1. Pingback: Deep Dive- Contrail Data Center Interconnect – Packet Expert
2. **David Bass says:**
   December 23, 2016 at 7:04 pm
   So this looks like it will work great in data center environments. Have you seen, or played with use cases that involve deployments where you don’t own the default gateway? For instance, a enterprise or provider who wants to try and do bare metal deployments using broadband Internet as the WAN link? Could also be a provider internal L3 MPLS cloud.

   Reply
   1. **packetexpert says:**
      December 23, 2016 at 9:22 pm
      David, I focused on VPC (Virtual private cloud) use case. other use case for Contrail is NFV. http://www.opencontrail.org/opencontrail-architecture-documentation/#section1_1

      Reply
3. **jgrinwis (@jgrinwis) says:**
   March 13, 2017 at 9:39 pm
   is the above setup working?
   We’ve tried to build a similar setup with a vMX. EVPN up, mac’s learnt and gre/vxlan tunnels up:

   netconf@vmx1> show evpn database
   Instance: _contrail_l2_4_vxlan
   VLAN DomainId MAC address Active source Timestamp IP address
   3020 00:50:56:86:d3:10 ge-0/0/1.302 Mar 13 20:30:41
   3020 02:23:7b:d9:12:86 192.168.2.14 Mar 13 20:01:10
   3020 02:4b:9d:0b:8e:12 192.168.2.12 Mar 13 20:01:10
   3020 02:96:90:48:37:49 192.168.2.10 Mar 13 20:01:10

netconf@vmx1> show bridge domain

Routing instance Bridge domain VLAN ID Interfaces
_contrail_l2_4_vxlan bd-3020 304
ge-0/0/1.302
vtep.32768
vtep.32769
vtep.32770

Contrail seeing mac of remote BM host but traffic from bare-metal host never makes it into VM and visa-versa.
Reply

1. **packetexpert says:**
   March 24, 2017 at 3:22 pm
   Hi, you need to change security group setting in open-stack to allow ping

   Reply