L3 Data Centre

LEAF and SPINE

991451344

Professor: Felix Carapaica

Timothy Pang

Objectives:

- Build the following network diagram implementing DC EVPN VXLAN.
- There are three VLAN-VXLAN 10, 20, and 30.
- On the tenant side, the three Cisco routers have three VRF matching each VLAN.
- The Cisco routers must reach all the tenant side networks.

1. **R3# show ip route vrf 10**

```
R3#sh ip route vrf 10

Routing Table: 10
Codes: L - local, C - connected, S - static, R - RIP, M - mobile, B - BGP
       D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area
       N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
       E1 - OSPF external type 1, E2 - OSPF external type 2
       i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
       ia - IS-IS inter area, * - candidate default, U - per-user static route
       o - ODR, P - periodic downloaded static route, H - NHRP, l - LISP
       + - replicated route, % - next hop override

Gateway of last resort is not set

      10.0.0.0/8 is variably subnetted, 5 subnets, 2 masks
C        10.0.10.0/24 is directly connected, FastEthernet1/0.10
L        10.0.10.3/32 is directly connected, FastEthernet1/0.10
C        10.10.3.3/32 is directly connected, Loopback10
O        10.10.4.4/32 [110/2] via 10.0.10.4, 00:05:12, FastEthernet1/0.10
O        10.10.5.5/32 [110/2] via 10.0.10.5, 00:05:17, FastEthernet1/0.10
      192.168.3.0/24 is variably subnetted, 2 subnets, 2 masks
C        192.168.3.0/26 is directly connected, FastEthernet0/0.10
L        192.168.3.1/32 is directly connected, FastEthernet0/0.10
      192.168.4.0/26 is subnetted, 1 subnets
O        192.168.4.0 [110/2] via 10.0.10.4, 00:05:12, FastEthernet1/0.10
      192.168.5.0/26 is subnetted, 1 subnets
O        192.168.5.0 [110/2] via 10.0.10.5, 00:05:17, FastEthernet1/0.10
R3#
```

Figure 1.

2. **R3# show ip route vrf 20**

```
R3#sh ip route vrf 20

Routing Table: 20
Codes: L - local, C - connected, S - static, R - RIP, M - mobile, B - BGP
       D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area
       N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
       E1 - OSPF external type 1, E2 - OSPF external type 2
       i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
       ia - IS-IS inter area, * - candidate default, U - per-user static route
       o - ODR, P - periodic downloaded static route, H - NHRP, l - LISP
       + - replicated route, % - next hop override

Gateway of last resort is not set

      10.0.0.0/8 is variably subnetted, 5 subnets, 2 masks
C        10.0.20.0/24 is directly connected, FastEthernet1/0.20
L        10.0.20.3/32 is directly connected, FastEthernet1/0.20
C        10.20.3.3/32 is directly connected, Loopback20
O        10.20.4.4/32 [110/2] via 10.0.20.4, 00:05:34, FastEthernet1/0.20
O        10.20.5.5/32 [110/2] via 10.0.20.5, 00:05:39, FastEthernet1/0.20
      192.168.3.0/24 is variably subnetted, 2 subnets, 2 masks
C        192.168.3.64/26 is directly connected, FastEthernet0/0.20
L        192.168.3.65/32 is directly connected, FastEthernet0/0.20
      192.168.4.0/26 is subnetted, 1 subnets
O        192.168.4.64 [110/2] via 10.0.20.4, 00:05:34, FastEthernet1/0.20
      192.168.5.0/26 is subnetted, 1 subnets
O        192.168.5.64 [110/2] via 10.0.20.5, 00:05:39, FastEthernet1/0.20
R3#
```

Figure 2.

3. R3# show ip route vrf 30

```
R3#sh ip route vrf 30

Routing Table: 30
Codes: L - local, C - connected, S - static, R - RIP, M - mobile, B - BGP
       D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area
       N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
       E1 - OSPF external type 1, E2 - OSPF external type 2
       i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
       ia - IS-IS inter area, * - candidate default, U - per-user static route
       o - ODR, P - periodic downloaded static route, H - NHRP, l - LISP
       + - replicated route, % - next hop override

Gateway of last resort is not set

      10.0.0.0/8 is variably subnetted, 5 subnets, 2 masks
C        10.0.30.0/24 is directly connected, FastEthernet1/0.30
L        10.0.30.3/32 is directly connected, FastEthernet1/0.30
C        10.30.3.3/32 is directly connected, Loopback30
O        10.30.4.4/32 [110/2] via 10.0.30.4, 00:05:47, FastEthernet1/0.30
O        10.30.5.5/32 [110/2] via 10.0.30.5, 00:06:02, FastEthernet1/0.30
      192.168.3.0/24 is variably subnetted, 2 subnets, 2 masks
C        192.168.3.128/26 is directly connected, FastEthernet0/0.30
L        192.168.3.129/32 is directly connected, FastEthernet0/0.30
      192.168.4.0/26 is subnetted, 1 subnets
O        192.168.4.128 [110/2] via 10.0.30.4, 00:05:47, FastEthernet1/0.30
      192.168.5.0/26 is subnetted, 1 subnets
O        192.168.5.128 [110/2] via 10.0.30.5, 00:06:02, FastEthernet1/0.30
R3#
```

Figure 3.

4. R3# show ip ospf database

```
R3#sh ip ospf data

            OSPF Router with ID (10.30.3.3) (Process ID 30)

                Router Link States (Area 0)

Link ID         ADV Router      Age         Seq#       Checksum Link count
10.30.3.3       10.30.3.3       385         0x80000002 0x004836 3
10.30.4.4       10.30.4.4       376         0x80000002 0x007402 3
10.30.5.5       10.30.5.5       386         0x80000002 0x00A0CD 3

                Net Link States (Area 0)

Link ID         ADV Router      Age         Seq#       Checksum
10.0.30.5       10.30.5.5       376         0x80000002 0x00DD67

            OSPF Router with ID (10.20.3.3) (Process ID 20)

                Router Link States (Area 0)

Link ID         ADV Router      Age         Seq#       Checksum Link count
10.20.3.3       10.20.3.3       384         0x80000002 0x006A86 3
10.20.4.4       10.20.4.4       377         0x80000002 0x009652 3
10.20.5.5       10.20.5.5       384         0x80000002 0x00C21E 3

                Net Link States (Area 0)

Link ID         ADV Router      Age         Seq#       Checksum
10.0.20.5       10.20.5.5       377         0x80000002 0x003344

            OSPF Router with ID (10.10.3.3) (Process ID 10)

                Router Link States (Area 0)

Link ID         ADV Router      Age         Seq#       Checksum Link count
10.10.3.3       10.10.3.3       383         0x80000002 0x008CD6 3
10.10.4.4       10.10.4.4       384         0x80000002 0x00B8A2 3
10.10.5.5       10.10.5.5       384         0x80000002 0x00E46E 3

                Net Link States (Area 0)

Link ID         ADV Router      Age         Seq#       Checksum
10.0.10.5       10.10.5.5       384         0x80000001 0x008A20
R3#
```

Figure 4.

5. CL1# net show bgp route

```
cumulus@cumulus:mgmt:~$ net show bgp route
ERROR: Command not found.

    net show bgp route
                ^ Invalid value here.

The following commands contain keyword(s) 'route', 'show', 'bgp'

    net show bgp [l2vpn] evpn route [detail] [json]
    net show bgp [l2vpn] evpn route [detail] type (ead|1|macip|2|multicast|3|es|4|prefix|5) [json]
    net show bgp [l2vpn] evpn route rd <rd> [json]
    net show bgp [l2vpn] evpn route rd <rd> mac <mac> [json]
    net show bgp [l2vpn] evpn route rd <rd> mac <mac> ip <ip> [json]
    net show bgp [l2vpn] evpn route rd <rd> prefix (<ipv4/prefixlen>|<ipv6/prefixlen>) [json]
    net show bgp [l2vpn] evpn route rd <rd> type (ead|1|macip|2|multicast|3|es|4|prefix|5) [json]
    net show bgp [l2vpn] evpn route vni <1-16777215> [json]
    net show bgp [l2vpn] evpn route vni <1-16777215> mac <mac> [json]
    net show bgp [l2vpn] evpn route vni <1-16777215> mac <mac> ip <ip> [json]
    net show bgp [l2vpn] evpn route vni <1-16777215> multicast <ipv4> [json]
    net show bgp [l2vpn] evpn route vni <1-16777215> type (ead|1|macip|2|multicast|3) [json]
    net show bgp [l2vpn] evpn route vni <1-16777215> vtep <ipv4> [json]
    net show bgp [l2vpn] evpn route vni all [detail] [json]
    net show bgp [l2vpn] evpn route vni all [detail] vtep <ipv4> [json]
    net show bgp ipv4 unicast route-leak [json]
    net show bgp ipv6 unicast route-leak [json]
    net show bgp vrf <text> ipv4 unicast route-leak [json]
    net show bgp vrf <text> ipv6 unicast route-leak [json]
    net show route (bgp|connected|kernel|ospf|ospf6|rip|static|supernets-only|table) [json]
    net show route vrf <text> (bgp|connected|kernel|ospf|ospf6|rip|static|supernets-only|table) [json]

cumulus@cumulus:mgmt:~$
```

Figure 5.

Alternate: CL1# net show bgp l2vpn evpn route

CL1 | CL3

```
cumulus@cumulus:mgmt:~$ net show bgp l2vpn evpn route
BGP table version is 2, local router ID is 10.0.0.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete
EVPN type-1 prefix: [1]:[ESI]:[EthTag]:[IPlen]:[VTEP-IP]:[Frag-id]
EVPN type-2 prefix: [2]:[EthTag]:[MAClen]:[MAC]:[IPlen]:[IP]
EVPN type-3 prefix: [3]:[EthTag]:[IPlen]:[OrigIP]
EVPN type-4 prefix: [4]:[ESI]:[IPlen]:[OrigIP]
EVPN type-5 prefix: [5]:[EthTag]:[IPlen]:[IP]

   Network          Next Hop            Metric LocPrf Weight Path
                    Extended Community
Route Distinguisher: 65003:10
*> [2]:[0]:[48]:[ca:01:30:ac:00:1c]
                    10.0.0.3                               0 65003 i
                    RT:65003:10 ET:8
*> [3]:[0]:[32]:[10.0.0.3]
                    10.0.0.3                               0 65003 i
                    RT:65003:10 ET:8
Route Distinguisher: 65003:20
*> [2]:[0]:[48]:[ca:01:30:ac:00:1c]
                    10.0.0.3                               0 65003 i
                    RT:65003:20 ET:8
*> [3]:[0]:[32]:[10.0.0.3]
                    10.0.0.3                               0 65003 i
                    RT:65003:20 ET:8
Route Distinguisher: 65003:30
*> [2]:[0]:[48]:[ca:01:30:ac:00:1c]
                    10.0.0.3                               0 65003 i
                    RT:65003:30 ET:8
*> [3]:[0]:[32]:[10.0.0.3]
                    10.0.0.3                               0 65003 i
                    RT:65003:30 ET:8
Route Distinguisher: 65004:10
*> [2]:[0]:[48]:[ca:02:07:60:00:1c]
                    10.0.0.4                               0 65004 i
                    RT:65004:10 ET:8
*> [3]:[0]:[32]:[10.0.0.4]
                    10.0.0.4                               0 65004 i
                    RT:65004:10 ET:8
Route Distinguisher: 65004:20
*> [2]:[0]:[48]:[ca:02:07:60:00:1c]
                    10.0.0.4                               0 65004 i
                    RT:65004:20 ET:8
*> [3]:[0]:[32]:[10.0.0.4]
                    10.0.0.4                               0 65004 i
                    RT:65004:20 ET:8
Route Distinguisher: 65004:30
*> [2]:[0]:[48]:[ca:02:07:60:00:1c]
                    10.0.0.4                               0 65004 i
                    RT:65004:30 ET:8
*> [3]:[0]:[32]:[10.0.0.4]
                    10.0.0.4                               0 65004 i
                    RT:65004:30 ET:8
Route Distinguisher: 65005:10
*> [2]:[0]:[48]:[ca:03:2f:28:00:1c]
                    10.0.0.5                               0 65005 i
                    RT:65005:10 ET:8
*> [3]:[0]:[32]:[10.0.0.5]
                    10.0.0.5                               0 65005 i
                    RT:65005:10 ET:8
```

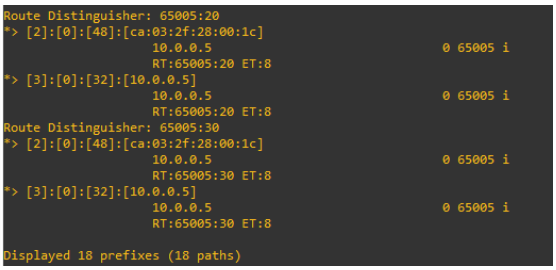


Figure 6.

Alternatively: CL1# net show bgp

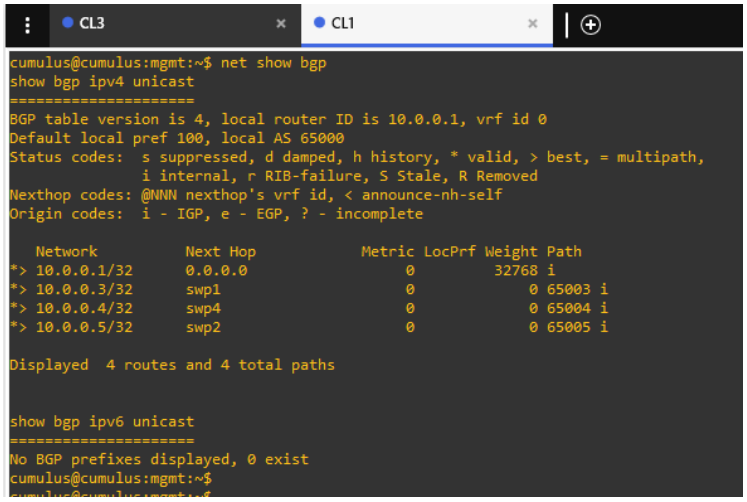


Figure 7.

6. CL1# net show bgp l2vpn evpn route vni 10

Figure 8.

CL3# net show bgp l2vpn evpn route vni 10

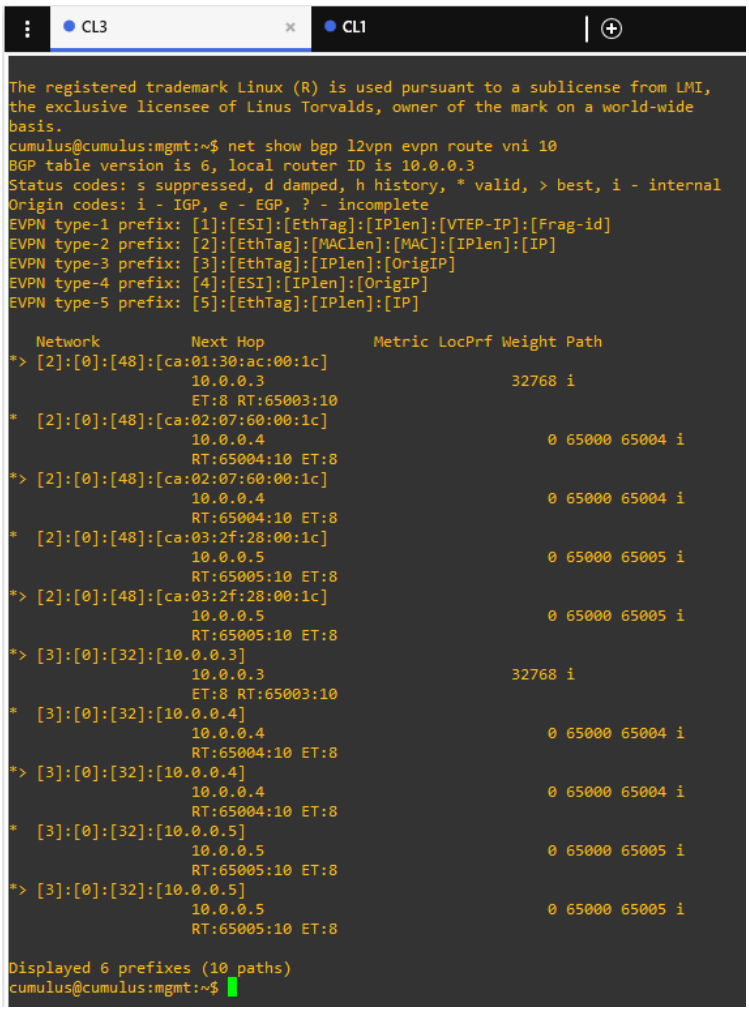


Figure 9.

7. CL2# net show bgp l2vpn evpn route vni 20

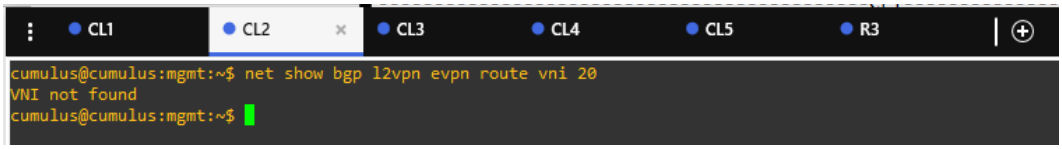


Figure 10.

CL3# net show bgp l2vpn evpn route vni 20

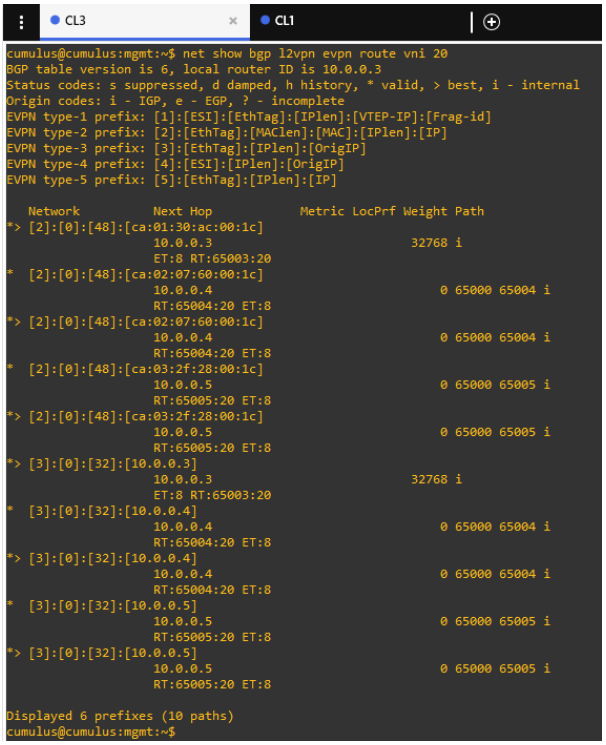


Figure 11.

8. CL3# net show bgp l2vpn evpn route vni 30

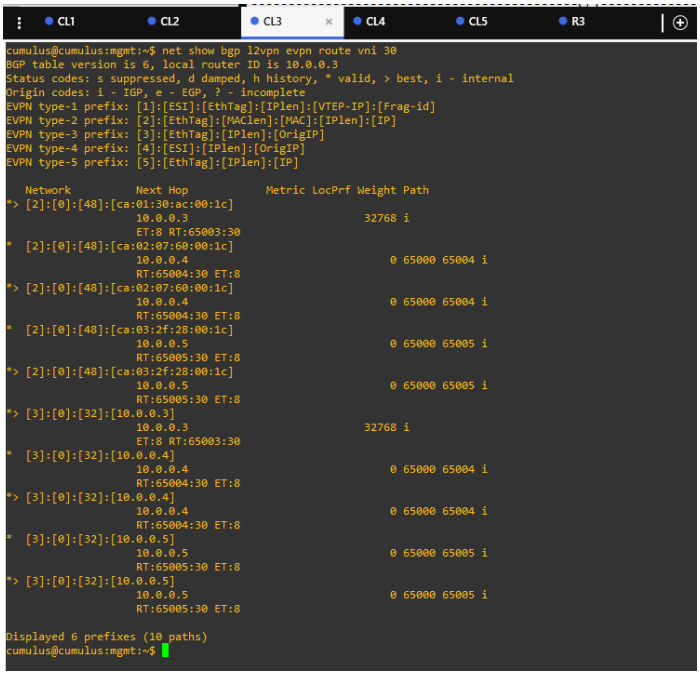


Figure 12.

The analysis of the BGP conversation between CL3 and CL1 from a sniff-capture. The conversation must include BGP OPEN and BGP UPDATE messages only.

The first is the general bgp OPEN and UPDATE messages:

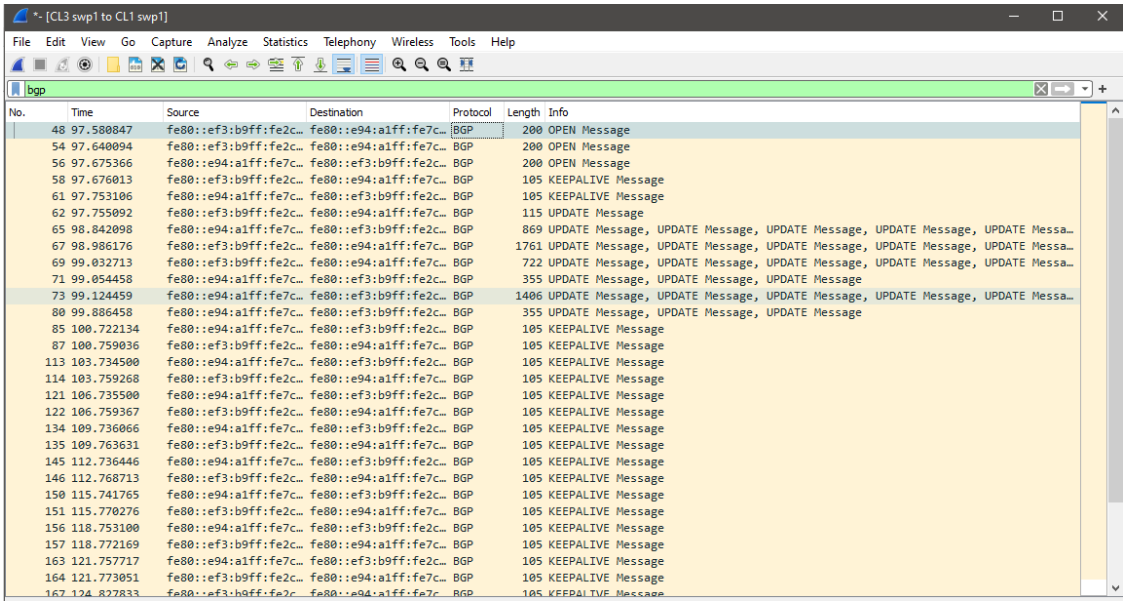


Figure 13.

When we go into the first OPEN message, we get the protocols and in the bgp OPEN message header we get the AS number and the bgp id number. The source and destination address is IPv6 because that is what is configured for the SPINES and the LEAFS communicate with each other.

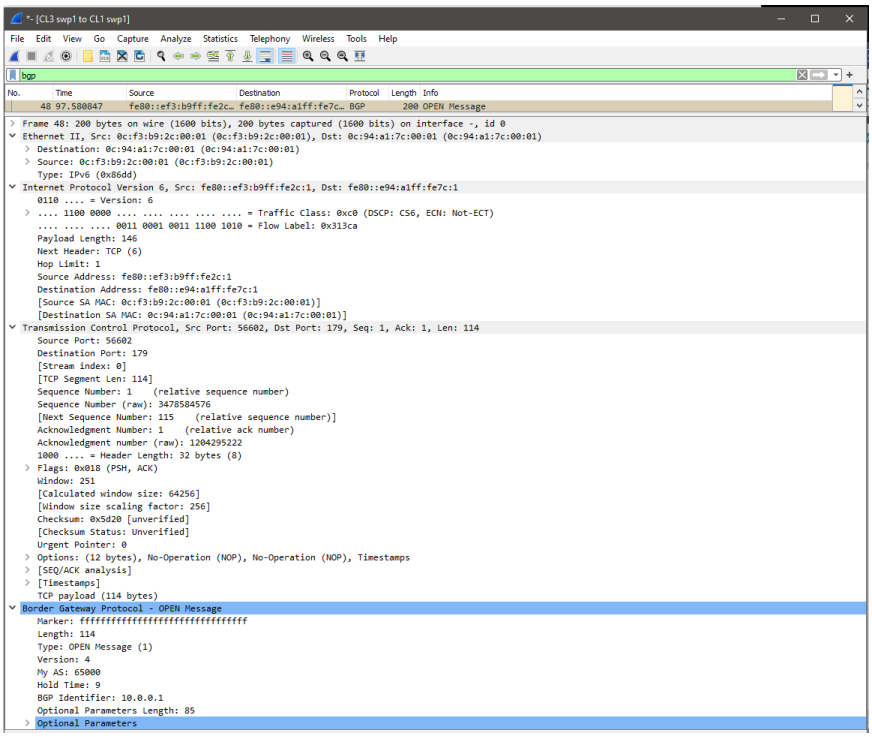


Figure 14.

When we open the UPDATE message, we see the MAC address IPv6 instead of IPv4 because the peer-group LEAF and SPINE are both running IPv6 to communicate with one another.

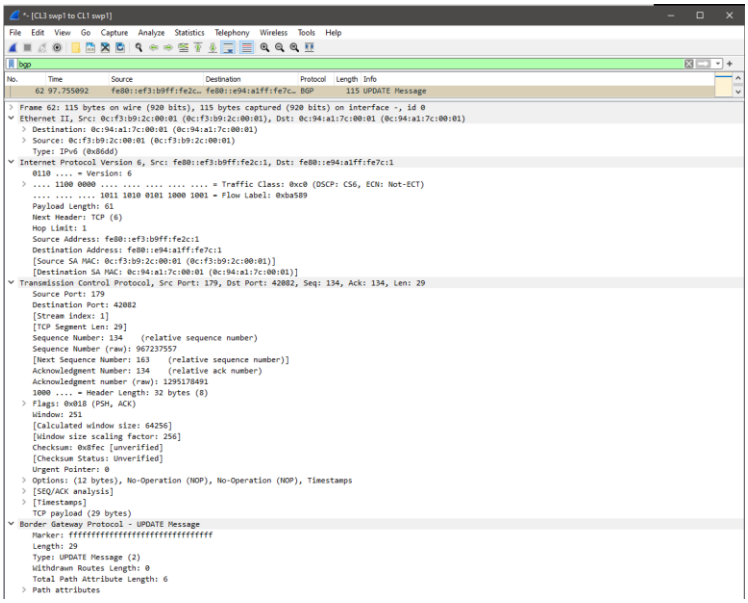


Figure 12.

When we open the next UPDATE message packet, we see multiple BGP update messages within the first UPDATE packet.

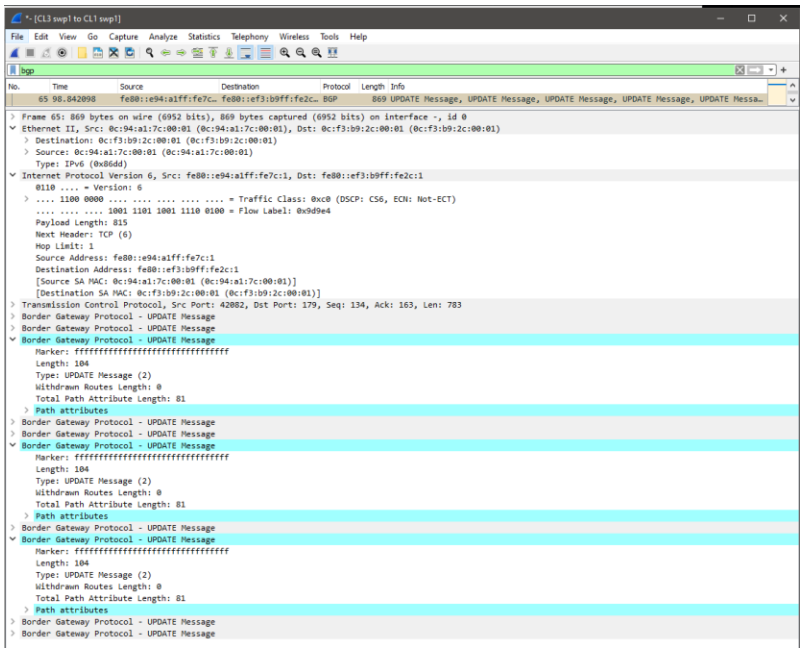


Figure 15.

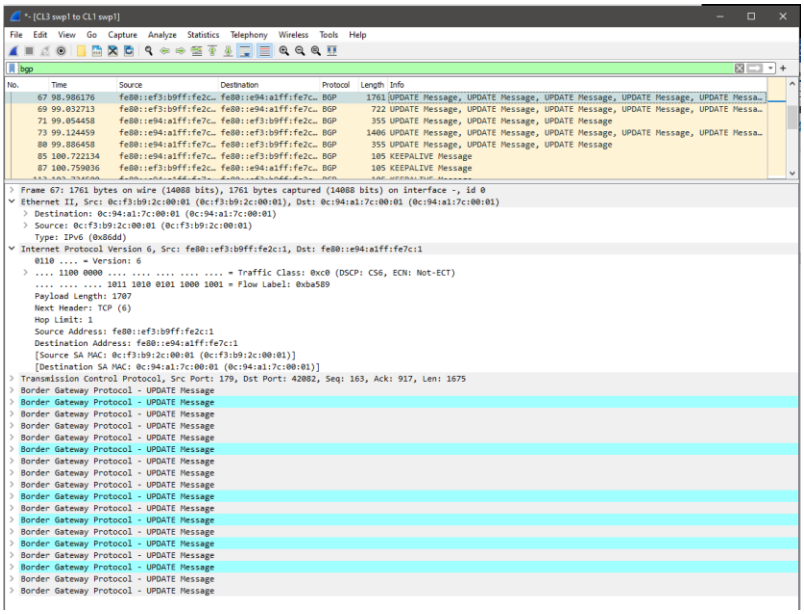


Figure 16.

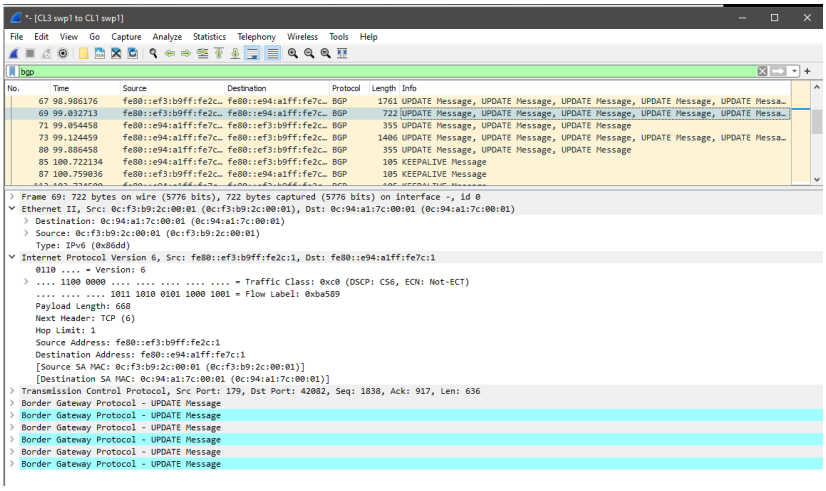


Figure 17.

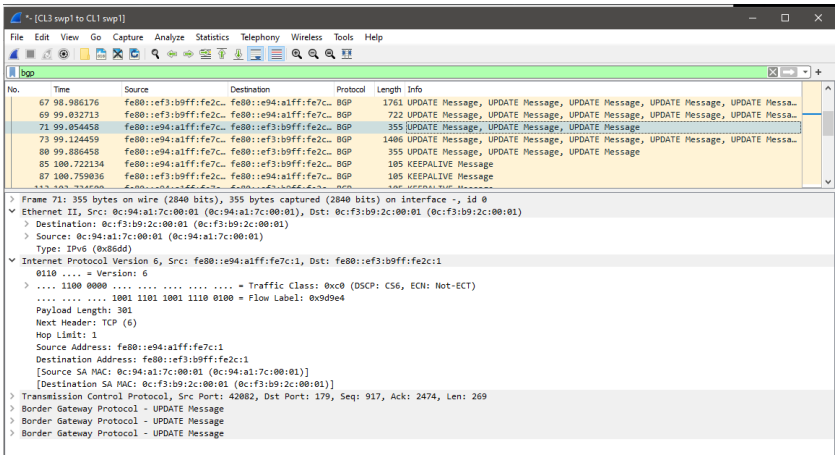


Figure 18.

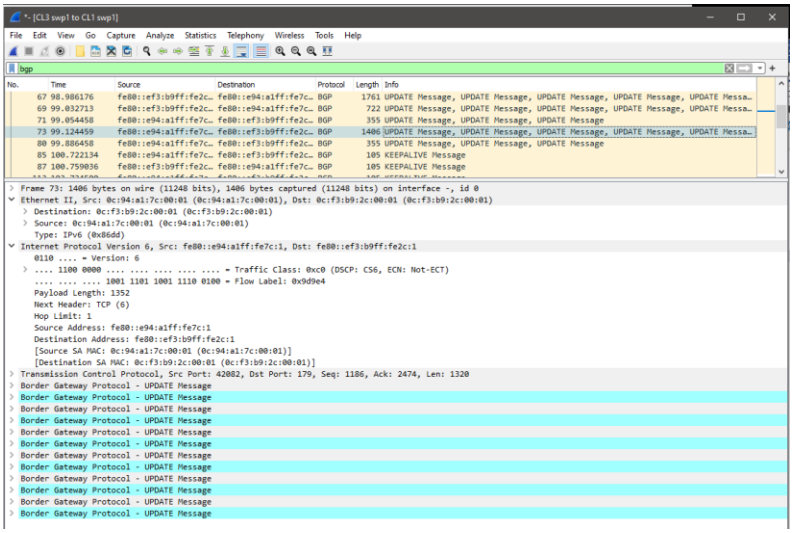


Figure 19.

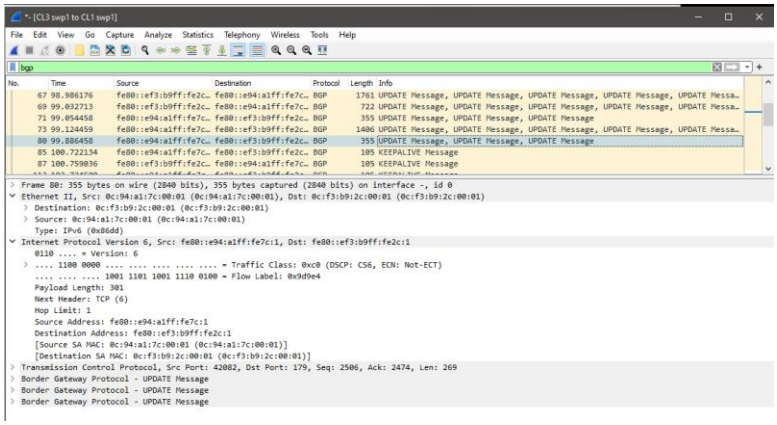


Figure 20.

The detailed explanation of the encapsulation of a ping from an VLAN-VRF in R3 to a destination in R5. The sniff is to be done on the interfaces out of CL3.

Out of CL3 to CL2:

28 10.006465	10.0.20.3	192.168.5.65	ICMP	164 Echo (ping) request	id=0x0002, seq=0/0, ttl=255 (reply in 29)
29 10.017739	192.168.5.65	10.0.20.3	ICMP	164 Echo (ping) reply	id=0x0002, seq=0/0, ttl=255 (request in 28)
30 10.028648	10.0.20.3	192.168.5.65	ICMP	164 Echo (ping) request	id=0x0002, seq=1/256, ttl=255 (reply in 31)
31 10.038719	192.168.5.65	10.0.20.3	ICMP	164 Echo (ping) reply	id=0x0002, seq=1/256, ttl=255 (request in 30)
32 10.049256	10.0.20.3	192.168.5.65	ICMP	164 Echo (ping) request	id=0x0002, seq=2/512, ttl=255 (reply in 33)
33 10.060362	192.168.5.65	10.0.20.3	ICMP	164 Echo (ping) reply	id=0x0002, seq=2/512, ttl=255 (request in 32)
34 10.070762	10.0.20.3	192.168.5.65	ICMP	164 Echo (ping) request	id=0x0002, seq=3/768, ttl=255 (reply in 35)
35 10.081644	192.168.5.65	10.0.20.3	ICMP	164 Echo (ping) reply	id=0x0002, seq=3/768, ttl=255 (request in 34)
36 10.092251	10.0.20.3	192.168.5.65	ICMP	164 Echo (ping) request	id=0x0002, seq=4/1024, ttl=255 (reply in 37)
37 10.103879	192.168.5.65	10.0.20.3	ICMP	164 Echo (ping) reply	id=0x0002, seq=4/1024, ttl=255 (request in 36)

Figure 21.

From CL3 to CL2 looking at the general ping, the icmp packets are sent multiple times back and forth, both request and reply from both parties.

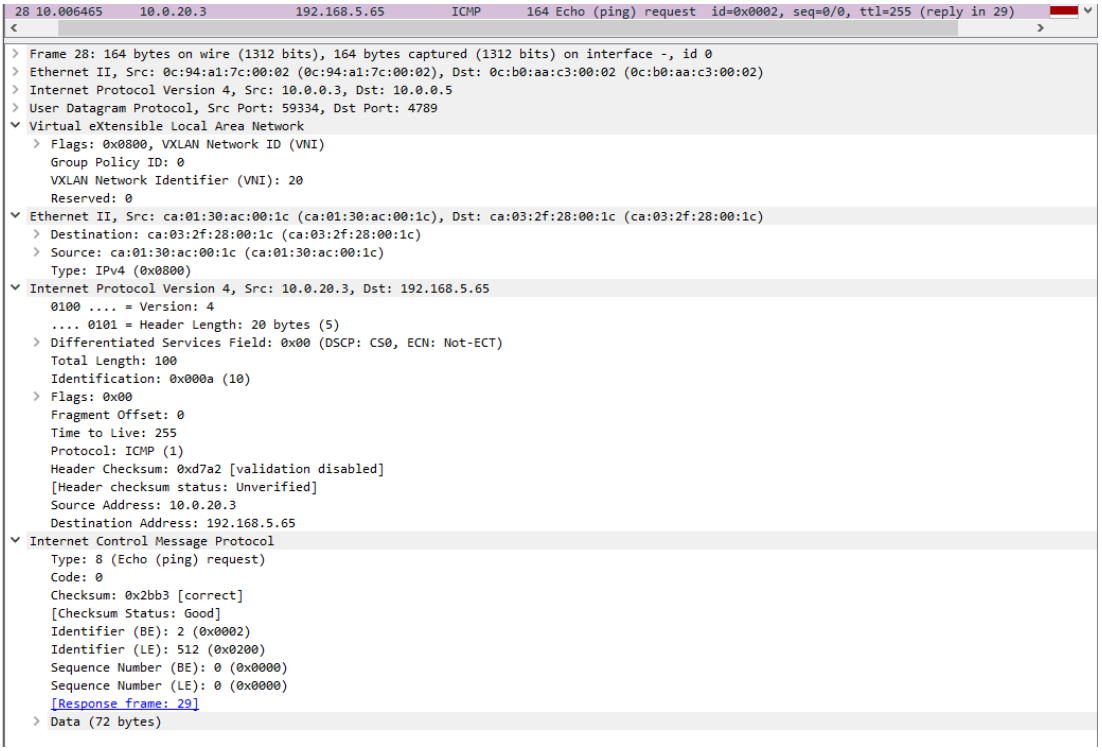


Figure 22.

In the first packet, which is the request packet, the layers are Ethernet, IPv4, UDP, VXLAN, Ethernet, IPv4, and ICMP. It contains the source and destination IP as well as the VXLAN id that was configured in CL3. The VXLAN is added below UDP and above Ethernet II.

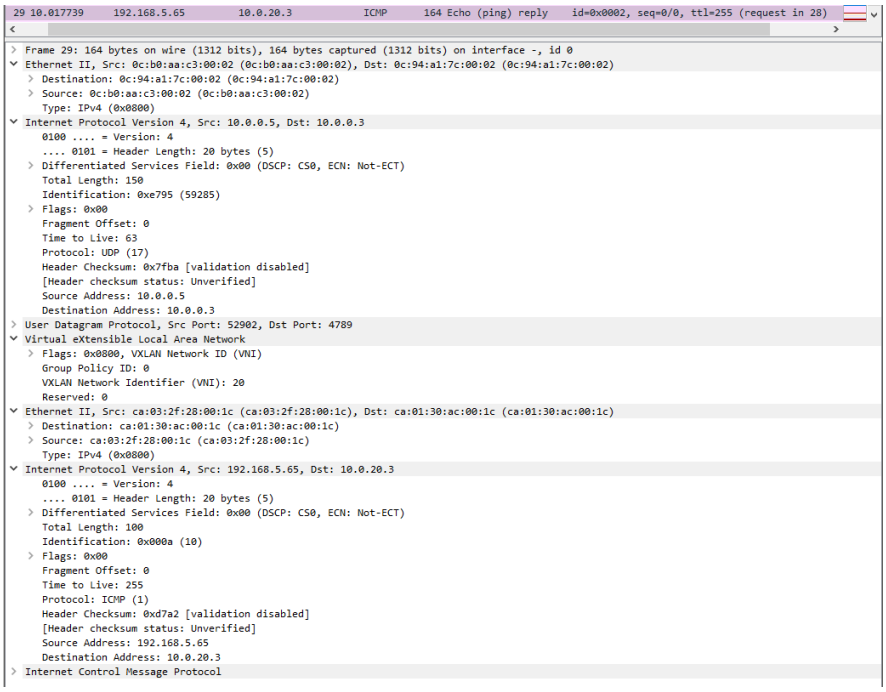


Figure 23.

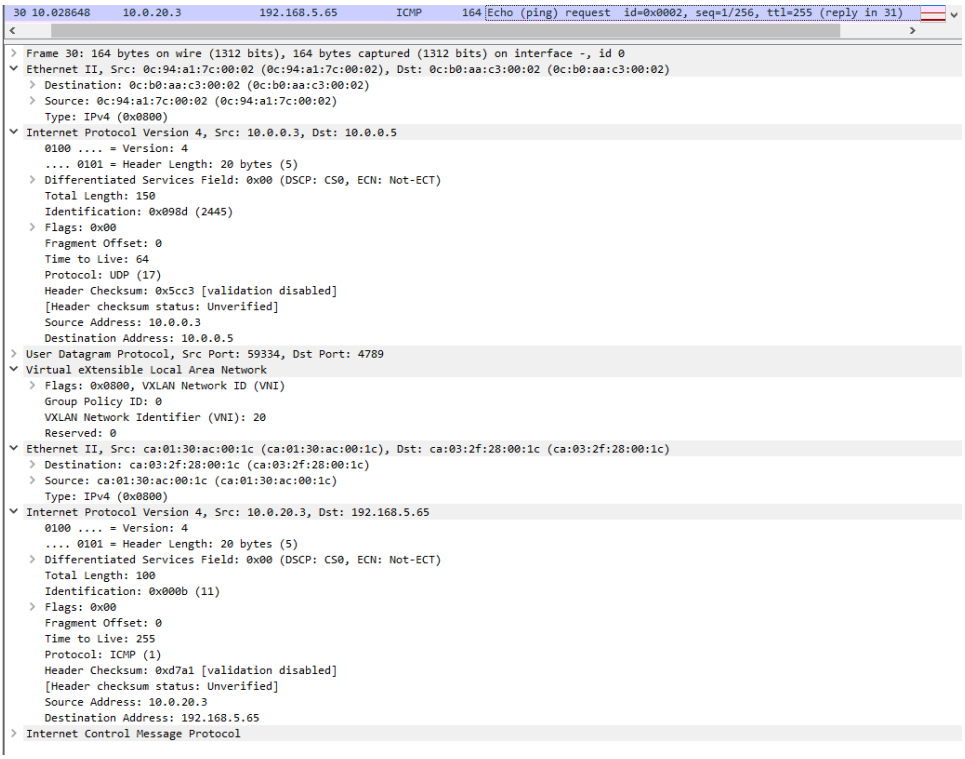


Figure 24.

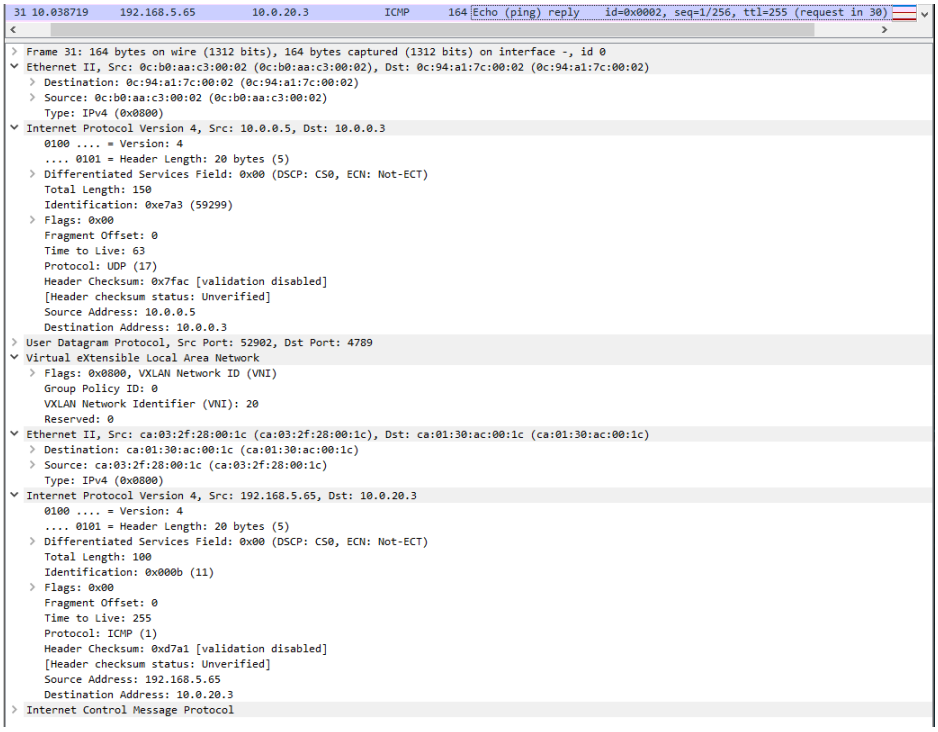


Figure 25.

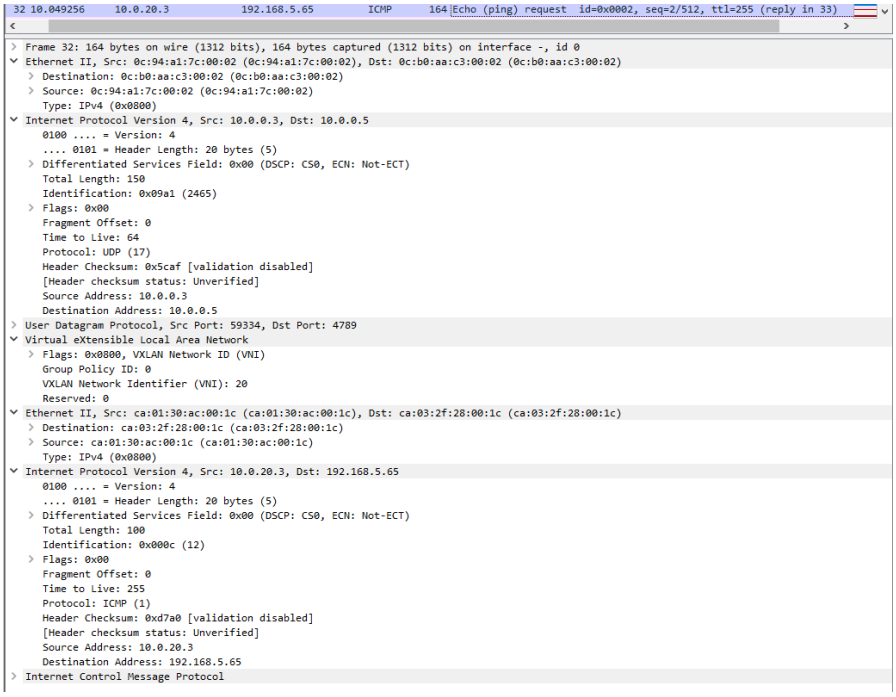


Figure 26.

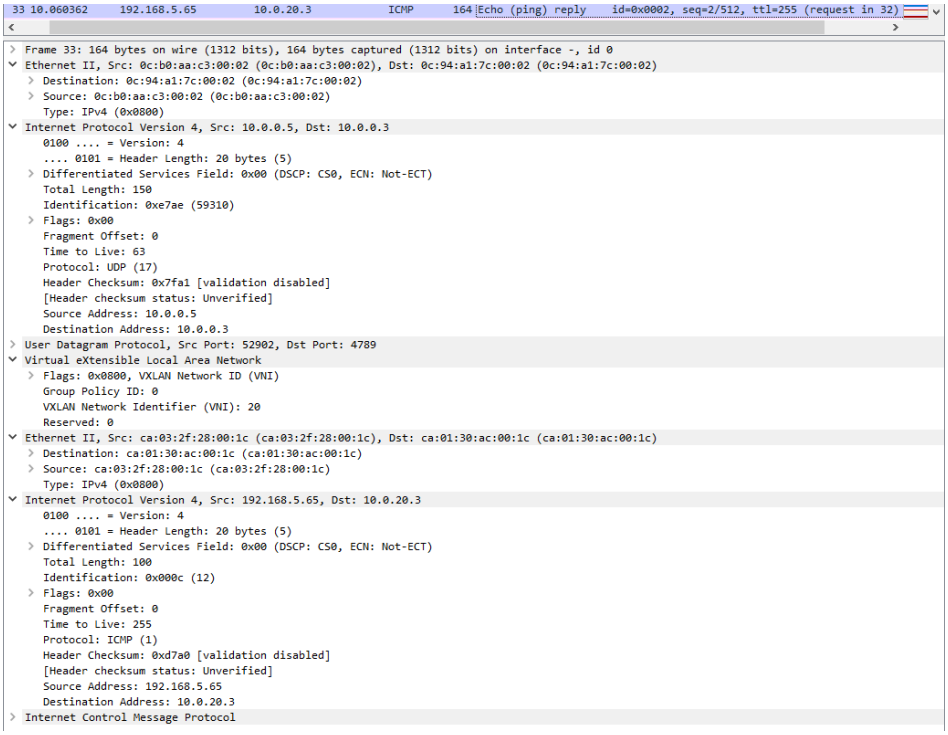


Figure 27.

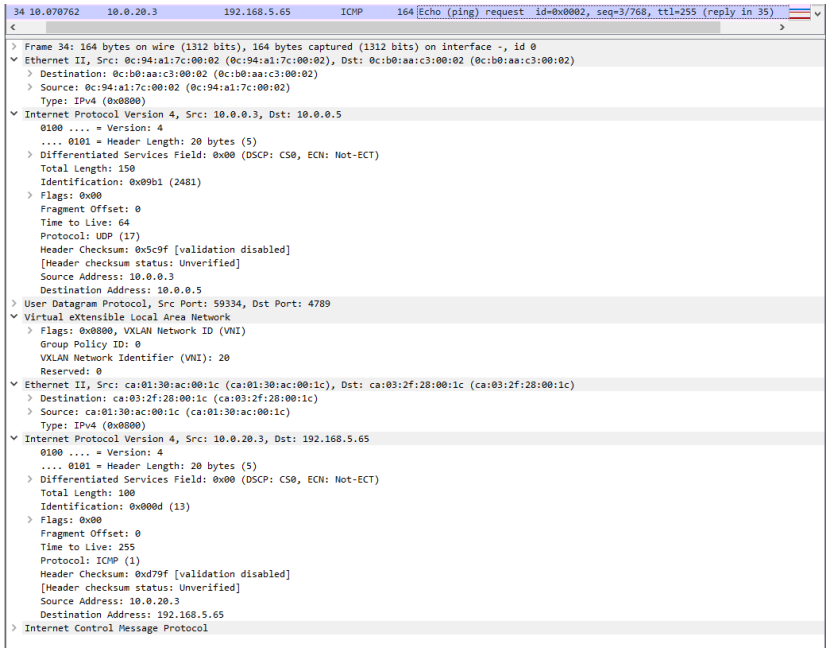


Figure 28.

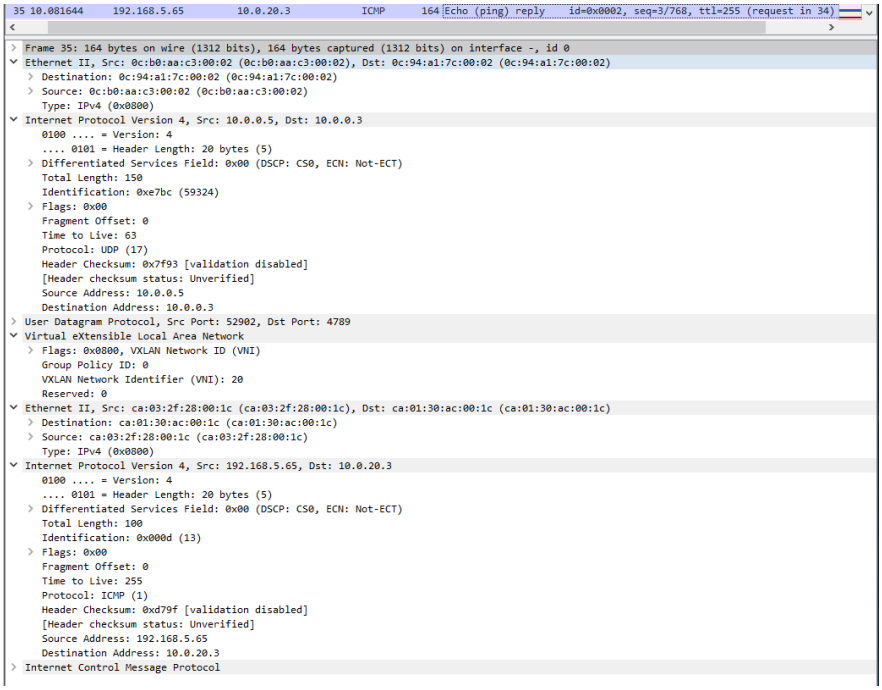


Figure 29.

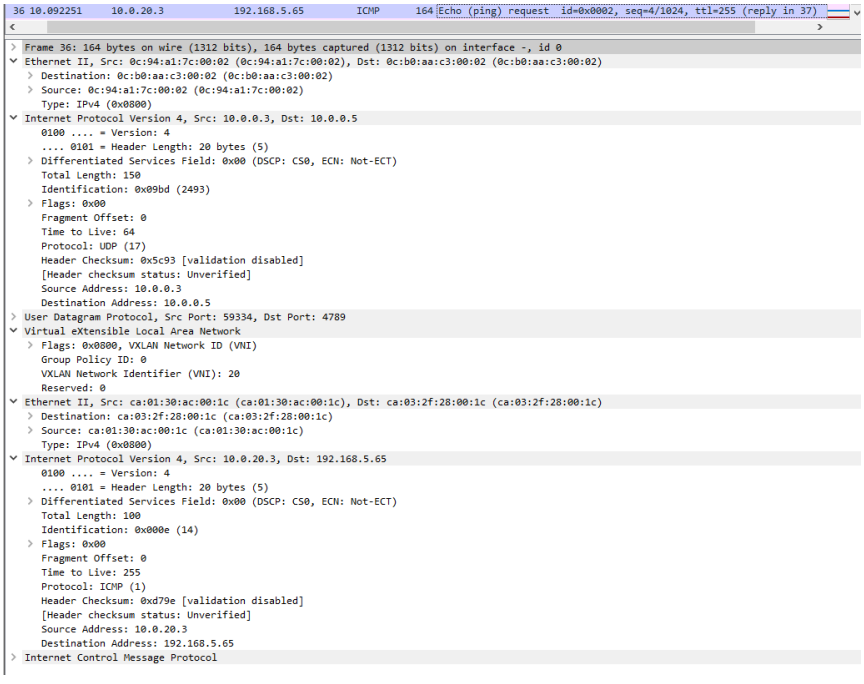


Figure 30.

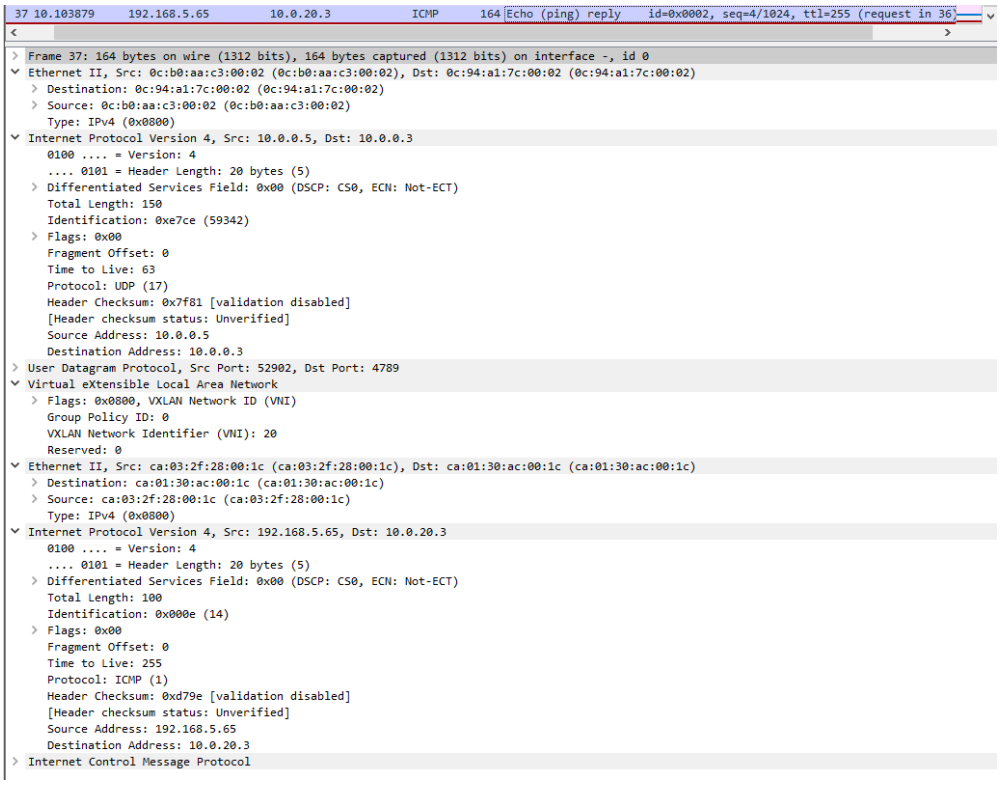


Figure 31.

Out of CL3 to CL1:

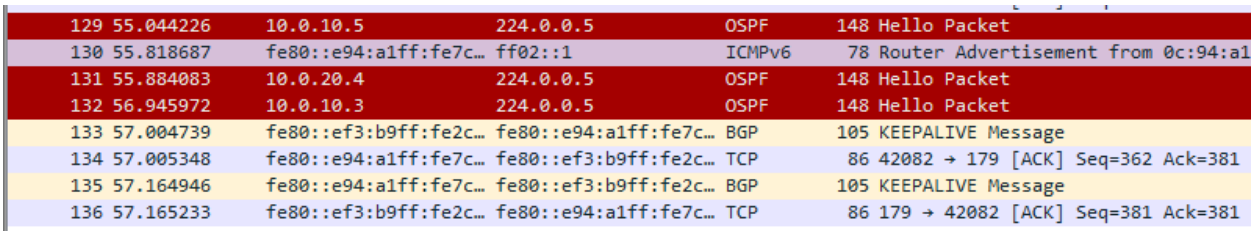


Figure 32.

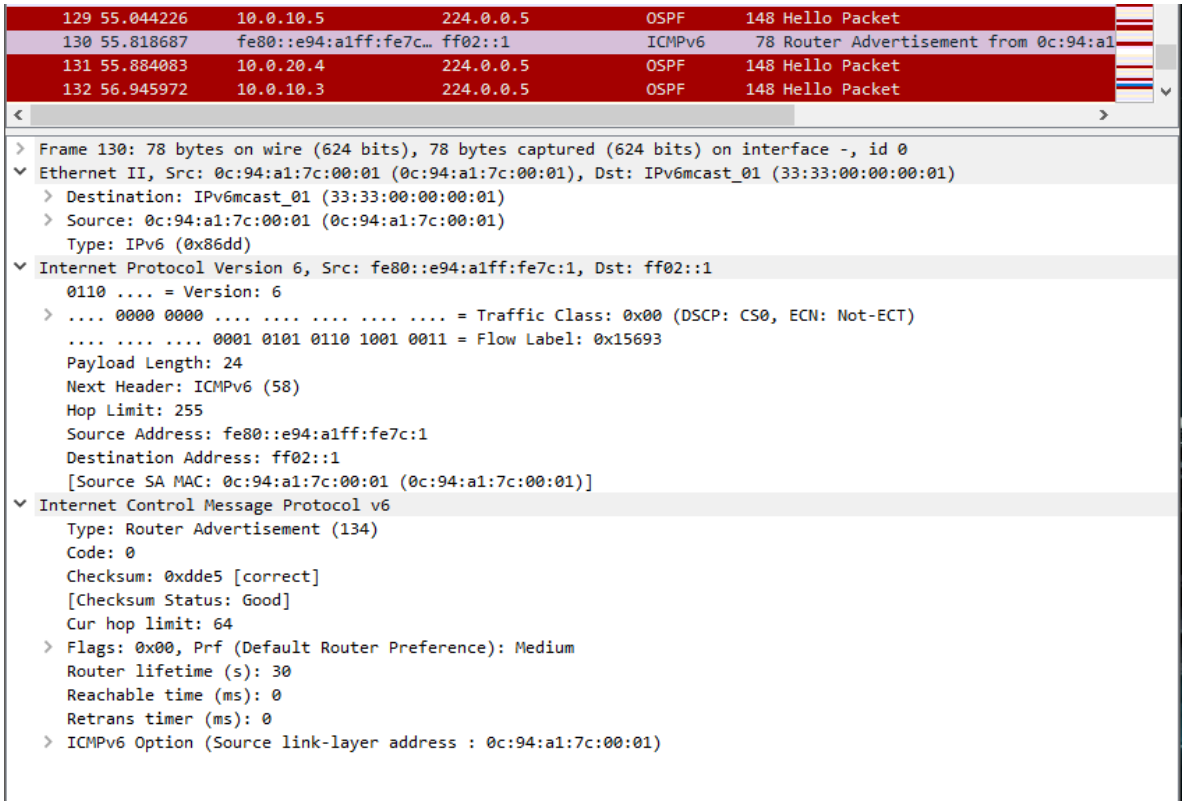
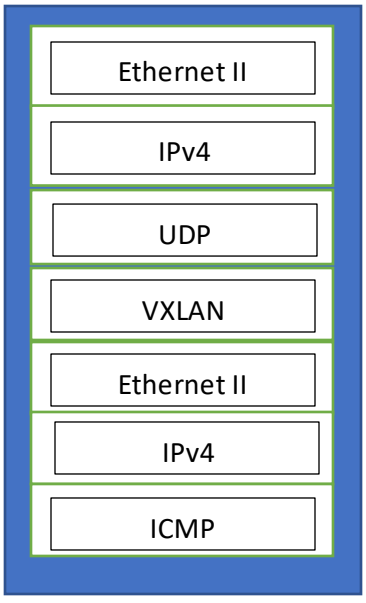


Figure 33.

Looking at the Wireshark capture at the time that the ping from R3 to one of the vrf's in R5 was delivered, in the output of CL3 to CL2, we can see that there is where the request and reply packets are initiated, whereas looking at the captures for CL3 to CL1 the only packet seen during that time is an ICMPv6 packet, as well as some hello packets. Within each of the request and reply packets going between CL3 and CL2, the structure inside is all the same even with multiples of the same packet.



The packet is structured as followed above for each of the request and reply packets. Most of the ping work goes through CL3 and CL2 rather than CL3 to CL1. The SPINES collects all the routing information from the LEAF routers and is then redistributed to each of the LEAF routers.