About this book

This book is based on my personal blog "The Network Times" (https://nwktimes.blogspot.com/). The first quarter of the book focuses on the Underlay Network solutions used in VXLAN fabric. It start by explaining the operation of OSPF focusing Dijkstra algorithm and Shortest Path Tree calculation process. Then it discusses of the differences between the OSPF and IS-IS routing protocols from the Underlay Network perspective. The first part also introduces three BGP based Underlay Network routing solutions with Single-AS solution, Dual-AS solution and Multi-AS solution. After the Unicast Routing section, this book explains the Multicast Routing solution used for L2VNI specific L2BUM traffic forwarding by introducing the Anycast-RP with PIM and the PIM BiDir solutions. The focus of the second quarter is VXLAN with BGP EVPN Control Plane. This part of the book explains the basic building blocks and configuration needed for both Layer2 and Layer3 services. It also discusses the BGP EVPN Control Plane operation showing how BGP EVPN NLRI information is advertised within a VXLAN fabric and how this information is used in Data Plane. The third quarter discusses of multi-homing solution focusing vPC explaining the following solutions; vPC Multi-homing, vPC and GIR and vPC Peer-Link as a Underlay Network Backup Path. In addition to vPC this section explains how to implement Firewall into VXLAN Fabric. The Last quarter of this book start by explains the standard based EVPN ESI Multi-homing solution. It also discusses the Data Center interconnect (DCI) solution based on the EVPN Multi-Site architecture.

Each chapter includes various configuration and verification examples as well as traffic captures. The only physical device used in labs is my personal computer and the example labs are done by using both Cisco "Virtual Internet Routing Lab" (VIRL) and Emulated Virtual Environment - Next Generation (EVE-NG).

Disclaimers

The content of this book is based on my own experience and testing results. This book is meant to be neither a Design nor Implementation guide but rather it tries to give to readers a basic understanding of VXLAN. After reading this book readers should do their own experiments and testing to verify the technology before using it in production environment.

Chapter 8: BGP EVPN VXLAN Configuration and building blocks.

This chapter introduces the basic building block and their configuration. It start by explaining the Intra-VNI switching and its configuration. Next, it discusses about tenant-specific Inter-VNI routing. Figure 8-1 shows the example topology and addressing scheme used in this chapter.

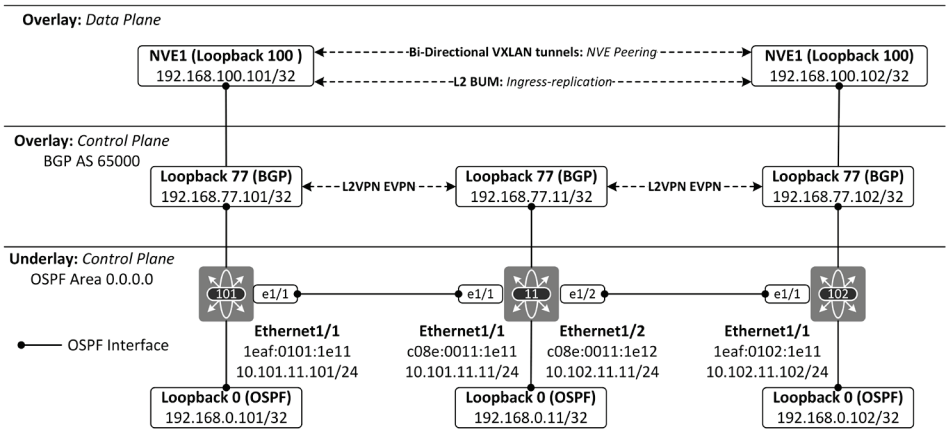

Figure 8-1: *Example topology with Anycast-RP - IP addresses.*

BGP EVPN VXLAN Building Blocks for Intra-VNI switching

This section introduces the basic building blocks of BGP EVPN VXLAN fabric from Intra-VNI switch perspective. The Underlay Network solution used in this section is Unicast-only (no Multicast support) using OSPF for routing exchange. The Underlay Network routing protocol has four main tasks from VXLAN persective:

Offer an IP connectivity between Loopback interfaces used for Overlay Network BGP peering: Virtual Tunnel End Point (VTEP) switches Leaf-101 and Leaf-102 used in this example establish iBGP L2VPN EVPN peering with Spine-11 by using their logical interface Loopback77. Leaf-switches exchanges L2VPN EVPN NLRIs (Network Layer Reachability Information) over these peering. This section introduces the L2VPN EVPN NLRI route-type 3 *"Inclusive Multicast Route"*, which is used by Leaf-switches to introduce their willingness to participate in Ingress-Replication L2 multi-destination tree. This multi-destination tree is used for Intra-VNI L2BUM traffic (mainly ARP and DHCP). In case of a tenant based L3 Multicast routing, the "*Tenant Routed Multicast"* (TRM) solution is required.

Offer an IP connectivity between Loopback interfaces that are used for Overlay Network NVE peering: Network Virtualization Edge (NVE) interface on Leaf-101 and Leaf-102 will establish an NVE peering between each other by using the IP address of interface Loopback100. NVE interfaces are used for VXLAN encapsulation/decapsulation for user data. The first requirement for NVE peering between leaf-switches is an NVE-to-NVE IP reachability. The second requirement that is needed to bringing up the NVE peering is two-way BGP L2VPN EVPN BGP Update messages exchange between leaf-switches. It could be *MAC Advertisement Route,* which carries information about the VM MAC address. In addition, if Ingress-Replication is used for L2BUM traffic with BGP, the trigger for NVE tunnel could be *Inclusive Multicast Route* BGP Update. Both of these route-types includes information about advertising switch and VNI for which the specific Update belongs to and based on that information, switches will build an NVE peering. The IP address of NVE interface is used as a next-hop value in MP-REACH_NLRI PA in BGP update messages. It is also used in the outer IP header in VXLAN encapsulated frames.

Advertise all Inter-Switch links subnets: Inter-switch links IP address has to be reachable for all switches for next-hop resolving.

Advertise the Multicast Rendezvous Point (RP) IP address: In a solution where Multicast is used for L2BUM traffic such as ARP-request forwarding, Multicast routing has to be enabled in an Underlay Network. The RP of given Multicast Group has to reachable for all switches in local VXLAN fabric.

Underlay Network: OSPF configuration

OSPF configuration is straightforward, enable OSPF feature, start the OSPF process and define the OSPF Router-Id. In addition, enable OSPF in Loopback interfaces and in all Inter-Switch link.

```
Leaf-101# sh run ospf
<snipped>
feature ospf

router ospf UNDERLAY-NET
  router-id 192.168.0.101

interface loopback0
  ip router ospf UNDERLAY-NET area 0.0.0.0

interface loopback77
  ip router ospf UNDERLAY-NET area 0.0.0.0

interface loopback100
  ip router ospf UNDERLAY-NET area 0.0.0.0

interface Ethernet1/1
  ip ospf network point-to-point
```

```
  ip router ospf UNDERLAY-NET area 0.0.0.0
```

Example 8-1: *OSPF configuration on Leaf-101.*

Overlay Network: BGP L2VPN EVPN configuration

Example 8-2 illustrates the BGP configuration on Leaf-101. Spine-11 is an iBGP neighbor and Leaf-101 exchange BGP L2VPN EVPN Updates. The "***nv overlay evpn***" –command is required before BGP neighbor can be specified as L2VPN EVPN peer. The configuration is identical in both leaf –switches, excluding the BGP Router-Id value.

```
Leaf-101# sh run bgp
<snipped>
feature bgp
nv overlay evpn
!
router bgp 65000
  router-id 192.168.77.101
  address-family ipv4 unicast
  neighbor 192.168.77.11
    remote-as 65000
    description ** Spine-11 BGP-RR **
    update-source loopback77
    address-family l2vpn evpn
      send-community extended
```

Example 8-2: *BGP configuration on Leaf-101.*

The BGP configuration of Spine-11 follows the same construction with an exception that leaf-switches are defined as route-reflector clients. Output taken from Spine-11 shows that Leaf-101 and Leaf-102 have both established iBGP L2VPN EVPN peering with Spine-11 but neither switch has sent any BGP Updates.

```
Spine-11# sh bgp l2vpn evpn summary
<>
Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down
State/PfxRcd
192.168.77.101  4 65000      21      21        4    0    0 00:15:13 0
192.168.77.102  4 65000      21      21        4    0    0 00:15:15 0
```

Example 8-3: *BGP verification.*

Overlay Network: NVE Peering

Feature ***"nv overlay"*** is required in order to configure NVE-interface. Host reachability information is exchange by using BGP with interface Loopback100 as a source IP address. NVE interface is only needed in Leaf-101 and Leaf-102. The command source-interface hold-down-time can be used for delaying the advertisement of the NVE loopback interface IP address until the Overlay Network Control Plane is coverged. The default value is 180 seconds and the range is 0-1000 secods.

```
feature nv overlay
!
interface nve1
  no shutdown
  host-reachability protocol bgp
  source-interface loopback100
```

Example 8-4: *Interface NVE1 on Leaf-101.*

Example 8-5 show that there Leaf-101 has no NVE peers at the moment is ok.

```
Leaf-101# sh nve peers detail
Leaf-101#
```

Example 8-5: *Leaf-101 NVE peers.*

Example 8-6 verifies that the NVE1 interface is up and it is using IP address 192.168.100.101 as a source address. The packet sends over NVE interface will be encapsulated with a VXLAN header. VXLAN is MAC in IP/UDP encapsulation where the original frames are encapsulated with an outer MAC/IP/UDP/VXLAN header. These headers are explored later when the actual encapsulated frame is sent over NVE interface.

```
Leaf-101# sh nve interface nve1 detail
Interface: nve1, State: Up, encapsulation: VXLAN
 VPC Capability: VPC-VIP-Only [not-notified]
 Local Router MAC: 5000.0001.0007
 Host Learning Mode: Control-Plane
 Source-Interface: loopback100 (primary: 192.168.100.101, secondary:
0.0.0.0)
 Source Interface State: Up
 Virtual RMAC Advertisement: No
 NVE Flags:
 Interface Handle: 0x49000001
 Source Interface hold-down-time: 180
 Source Interface hold-up-time: 30
 Remaining hold-down time: 0 seconds
 Virtual Router MAC: N/A
 Interface state: nve-intf-add-complete
```

Example 8-6: *Interface NVE1 on Lwaf-101.*

Overlay Network: Host Mobility Manager

The VXLAN service in NX-OS has a *Host Mobility Manager (HMM)* component. The HMM keeps track on MAC address moving inside VXLAN fabric switches. The operation of HMM is explained later in the context of the Control Plane learning process. Example 8-7 shows the command that enables the HMM feature.

```
feature fabric forwarding
```

Example 8-7: *Enabling "Host Mobility Manager" function on Leaf-101.*

Overlay Network: Anycast Gateway (AGW)

Virtual Machines (VMs) may move from one leaf switch to another in VXLAN fabric. Each leaf switch operates as an Anycast Gateway for VLAN SVIs (Switched Virtual Interface). In practice, this means that the VLAN specific SVI is configured in every leaf switch where the VLAN exists. In addition, each SVI uses the same Anycast MAC-address regardless of L3 SVI or tenant. Features needed for SVI and Anycast Gateway are shown in Example 8-8

```
feature interface-vlan
!
fabric forwarding anycast-gateway-mac 0001.0001.0001
```

Example 8-8: *Enabling "Anycast Gateway" function on Leaf-101.*

Overlay Network: VLAN based service

The example VXLAN fabric use VLAN based Virtual Network Segments (VN-Segment). The example 8-9 shows the configuration.

```
feature vn-segment-vlan-based
```

Example 8-9: *Enabling "VLAN based VNI-segment" function on Leaf-101.*

Overlay Network: TCAM modification

TCAM (Ternary Content Addressable Memory) is a part of the specialized memory block, which supports fast parallel lookups. The memory slices on TCAM consists of 256 bytes or 512 bytes blocks/banks. There are one ingress TCAM and one egress TCAM whose size varies based on the switching platform. The TCAM modification used in this example (Router-ACL, vPC-Convergence, and arp-ether) are shown in example 8-10.

```
hardware access-list tcam region racl 512
hardware access-list tcam region vpc-convergence 256
hardware access-list tcam region arp-ether 256 double-wide
```

Example 8-10: *Enabling "VLAN based VNI-segment" function on Leaf-101.*

Example 8-11 summarizes the VXLAN configuration required for switches before Layer 2 Intra-VNI or Layer 3 Inter-VNI services can be implemented into VXLAN fabric.

```
nv overlay evpn
feature ospf
```

```
feature bgp
feature fabric forwarding
feature interface-vlan
feature vn-segment-vlan-based
feature nv overlay
!
fabric forwarding anycast-gateway-mac 0001.0001.0001
!
hardware access-list tcam region racl 512
hardware access-list tcam region vpc-convergence 256
hardware access-list tcam region arp-ether 256 double-wide
!
interface nve1
  no shutdown
  host-reachability protocol bgp
  source-interface loopback100
!
interface Ethernet1/1
  no switchport
  mac-address 1eaf.0101.1e11
  medium p2p
  ip address 10.101.11.101/24
  ip ospf network point-to-point
  ip router ospf UNDERLAY-NET area 0.0.0.0
  no shutdown
!
interface loopback0
  description ** RID/Underlay **
  ip address 192.168.0.101/32
  ip router ospf UNDERLAY-NET area 0.0.0.0
!
interface loopback77
  description ** BGP peering **
  ip address 192.168.77.101/32
  ip router ospf UNDERLAY-NET area 0.0.0.0
!
interface loopback100
  description ** VTEP/Overlay **
  ip address 192.168.100.101/32
  ip router ospf UNDERLAY-NET area 0.0.0.0
!
router ospf UNDERLAY-NET
  router-id 192.168.0.101
router bgp 65000
  router-id 192.168.77.101
  address-family ipv4 unicast
  neighbor 192.168.77.11
    remote-as 65000
    description ** Spine-11 BGP-RR **
    update-source loopback77
    address-family l2vpn evpn
      send-community extended
```

Example 8-11: *Complete VXLAN service configuration on Leaf-101.*

Intra-VNI service (L2VNI) in VXLAN Fabric

This section explains the Intra-VNI solutions used in VXLAN Fabric. One way to define the VNI is a broadcast domain (VLAN) stretched over the VXLAN fabric. Intra-VNI implementation starts by creating the L2 VLAN and attaching it to VN-Segment. Example 8-12 show how to create vlan and attached it to VNI 10000.

```
vlan 10
  name VLAN10-mapped-to-VNI10000
  vn-segment 10000
```

Example 8-12: *Adding VLAN and assign it to VNI.*

Next step is to create an EVPN instance for VNI 10000. Route-Distinguisher (RD) is used to differentiating possible overlapping MAC addresses in different VNIs. RD is formed based on BGP RID and base value 32767 plus VLAN Id. Leaf-101 BGP RID is 192.168.77.101 and the VLAN Id used in this example is 10. This gives the RD 192.168.77.101:32777 for all MAC addresses participating in VLAN10/VNI 10000.

Each MAC Advertisement route exported to the BGP process carries VNI specific Route-Target value that is taken from EVPN instance for given VNI. Automatic creation of Route-Target is formed based on AS number: VNI-Id. This gives the VNI value 65000:10000 for EVPN instance VNI 10000. The RT is the same in all leaf-switches where VLAN 10 is used for VNI 10000. Note that VLAN Id does not have to be the same in all VTEP switches. VLAN Id has only local (switch) significance while VNI-Id has Fabric wide significance.

```
evpn
  vni 10000 l2
    rd auto
    route-target import auto
    route-target export auto
```

Example 8-13: *Creating EVPN instance for VNI 10000.*

The last two steps step is to add the VNI10000 to VXLAN process and define the L2BUM traffic replication mode. There are two modes for L2BUM replication, first is the *Multicast Mode* explained in chapters 6 and 7. The second one, *Ingress-Replication* (IR) is explained in this section. In IR mode peers leaf-switches can be defined either manually or by resolving automatically by using BGP. This section introduces the BGP solution. Ingress-Replication differs from Multicast L2BUM replication in two ways. In Multicast Mode, switch sends only one copy of L2BUM traffic to given Multicast group Rendezvous-Point without VXLAN encapsulation. In Ingress-Replication mode, the leaf-switch will send copies of L2BUM traffic to each leaf-switches participating in L2BUM delivery tree over VXLAN tunnel.

```
interface nve1
```

```
  no shutdown
  host-reachability protocol bgp
  source-interface loopback100
  member vni 10000
    ingress-replication protocol bgp
```

Example 8-14: *Adding VNI to VXLAN service and defining the L2BUM handling mode.*

Figure 8-12 illustrates the hardware resources reserved for VLAN10/VNI1000. First, the VLAN 10 is activated and MAC address can be installed into MAC address-table. Next, there is an L2 Routing and Forwarding Instance a.k.a L2RIB where the locally learned MAC address information is installed by L2 Forwarding component (L2FWDR). From the L2RIB the MAC address information is exported into BGP table with VNI specific Path Attributes (PA) values such as a Route-Target (Extended Community PA) and a Route-Distinguisher (MP-REACH-NLRI PA). Because there are no connected hosts in VLAN 10 at this moment, the MAC address-table and L2RIB are empty. Also, there are no route-type 2 routes (MAC Advertisement Route) in BGP table concerning L2VNI 10000 for the same reason. Though, there is a route-type 3 (Inclusive Multicast Route). Leaf-101 use this route-type to inform it willingness to join the L2BUM delivery tree for VNI10000. Switches hosting VNI10000 will import this route int BGP table (BRIB) based on the auto-generated RT value.

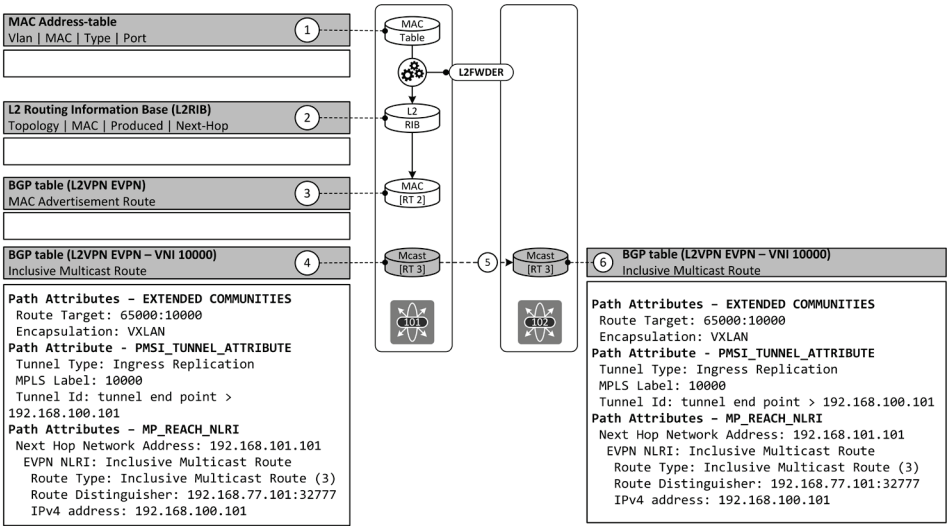

Figure 8-2: *Hardware resources and BGP route-type 3.*

P-Multicast Service Instance (PMSI) Path Attribute shown in figure 8-2 describes the PSMI tunnel end-point for Multi-Destination tree for VNI 10000. The MPLS label describes the Virtual Network Identifier (VNI) for this Multi-destination tree. Capture 8-1 shows the BGP L2VPN EVPN Route-Type 3 BGP Update message send by Leafd-101.

```
Ethernet II, Src: 1e:af:01:01:1e:11 (1e:af:01:01:1e:11), Dst:
c0:8e:00:11:1e:11 (c0:8e:00:11:1e:11)
Internet Protocol Version 4, Src: 192.168.77.101, Dst:
192.168.77.11
Transmission Control Protocol, Src Port: 27001, Dst Port: 179, Seq: 1,
Ack: 1, Len: 100
Border Gateway Protocol - UPDATE Message
    Marker: ffffffffffffffffffffffffffffffff
    Length: 100
    Type: UPDATE Message (2)
    Withdrawn Routes Length: 0
    Total Path Attribute Length: 77
    Path attributes
        Path Attribute - ORIGIN: IGP
        Path Attribute - AS_PATH: empty
        Path Attribute - LOCAL PREF: 100
        Path Attribute - EXTENDED_COMMUNITIES
            Flags: 0xc0, Optional, Transitive, Complete
            Type Code: EXTENDED_COMMUNITIES (16)
            Length: 16
            Carried extended communities: (2 communities)
                Route Target: 65000:10000
                Encapsulation: VXLAN Encapsulation [Transitive Opaque]
        Path Attribute - PMSI_TUNNEL_ATTRIBUTE
            Flags: 0xc0, Optional, Transitive, Complete
            Type Code: PMSI_TUNNEL_ATTRIBUTE (22)
            Length: 9
            Flags: 0
            Tunnel Type: Ingress Replication (6)
            0000 0000 0010 0111 0001 .... = MPLS Label: 625
            Tunnel ID: tunnel end point -> 192.168.100.101
        Path Attribute - MP_REACH_NLRI
            Type Code: MP_REACH_NLRI (14)
            Length: 28
            Address family identifier (AFI): Layer-2 VPN (25)
            Subsequent address family identifier (SAFI): EVPN (70)
            Next hop network address (4 bytes)
            Number of Subnetwork points of attachment (SNPA): 0
            Network layer reachability information (19 bytes)
                EVPN NLRI: Inclusive Multicast Route
                    Route Type: Inclusive Multicast Route (3)
                    Length: 17
                    Route Distinguisher: (192.168.77.101:32777)
                    Ethernet Tag ID: 0
                    IP Address Length: 32
                    IPv4 address: 192.168.100.101
```

Capture 8-1: *Inclusive Multicast Route (route-type 3).*

Figure 8-3 illustrates the VXLAN tunnel that has now been set up. Both Unicast and L2BUM frames originated by the local host will be encapsulated with a tunnel header that includes outer MAC and IP headers, where source MAC address is taken from the egress interface E1/1 and the source IP is taken from the interface NVE1. The destination IP address will be set to an interface NVE1 of destination leaf-switch. After the MAC and IP headers, there is a UDP header where the destination port is always set to 4789 for

VXLAN. The source UDP port is the result of 5-tuple hash. In addition, there is a VXLAN header that describes the VNI where carried frame belongs to. The Control Plane and Data Plane operation are discussed in detail in chapter 9. Note, in the case where the Ingress-Replication mode is static, the NVE peering does not come up until one BGP Update message is sent by both leaf-switches.

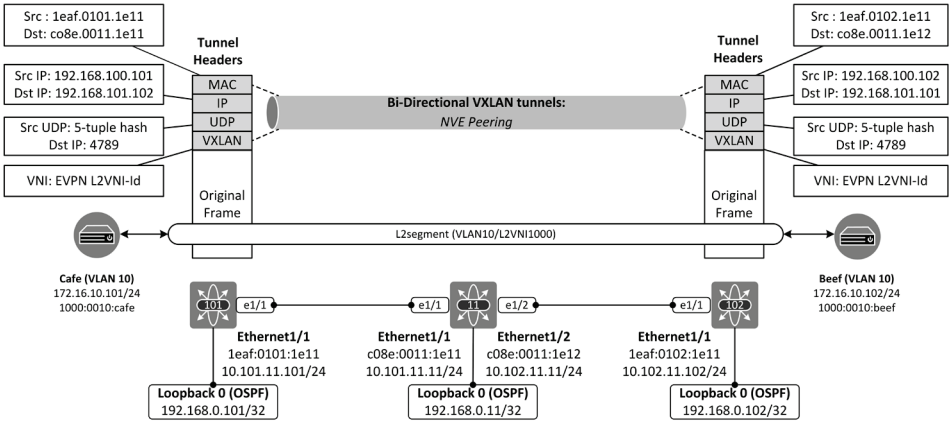

Figure 8-3: *Bi-directional VXLAN tunnel.*

Layer 2 Intra VNI service can be used for example in the case where the VXLAN fabric offers only an L2 switching path between. One example is the solution where the Firewall is used as a gateway between Inter-VNI segments.

Tenant based Inter-VNI Routing (L3VNI) in VXLAN Fabric

The previous section introduces the basic building blocks and configuration for Intra-VNI switching. This section discusses tenant based Inter-VNI routing.

The first step is to set up tenant-based routing domain by configuring VRF context (Virtual Routing and Forwarding instance) for specific customer/application (example 8-15). The VNI used for inter-VLAN routing in this example is 10077. This VNI is used in VXLAN header as tenant identifier. Just like in previous L2VNI examples, auto-generated values for both RD and RT are used. RD is formed based on switch BGP RID:<VRF Id>. Example 8-16 shows that VRF-Id of TENANT77 is three. This gives RD 192.168.77.101:3 for Inter-VNI routes. RT, in turn, is formed based on BGP AS:VNI-Id. This gives RT 192.168.77.101:10077 for Inter-VNI routes.

```
vrf context TENANT77
  vni 10077
  rd auto
  address-family ipv4 unicast
    route-target both auto
    route-target both auto evpn
```

Example 8-15: *configuring VRF context named TEANANT77on Leaf-101.*

```
Leaf-101# sh vrf TENANT77
VRF-Name                           VRF-ID State   Reason
TENANT77                                3 Up      --
```

Example 8-16: *VRF context TEANANT77 VRF-Id on Leaf-101.*

Next step is to define the actual routing interface for Inter-VNI routing. It is done by specifying the routing interface and L2 VLAN for it in every switch that host the VRF. There is no IP address attached to L3 interface because it is used only in Data Plane (it is not advertised in any routing protocol). When a VTEP switch receives an IP packet to another subnet (inside tenant) from its directly connected host, switch routes the packet over the L3 interface. The VNI-Id added to VXLAN header is set to value that is defined under the VRF Context configuration. L2VLAN for routing interface is used for reserving hardware resources. VXLAN is a MAC-in-IP/UDP encapsulation. So there is a need for inner MAC address also for routed Inter-VNI packets. The address of the sending host is naturally out of the question so the inner source MAC address is taken from the L2 VLAN associated to L3 routing interface. MAC address depends on the Underlay Network Inter-switch IP addressing. If Inter-Switch links use Unnumbered interface addressing scheme, the MAC address attaches to L2 VLAN is the switch system MAC, otherwise, the MAC address of the uplink interface is used. The destination MAC address on inner frame is set to destination VTEP switch MAC (in case of Unicast packet). Capture 8-2 shows the ICMP request sent by host Cafe in VLAN 10 (172.16.10.101) connected to Leaf-101 to host Beef in VLAN 20 (172.16.20.102) connected to remote Leaf-102.

```
vlan 77
  name TENANT77
  vn-segment 10077
!
interface Vlan77
  no shutdown
  mtu 9216
  vrf member TENANT77
  ip forward
```

Example 8-17: *L2 VLAN and L3 SVI for Inter-VNI routing inside TEANANT77.*

```
Ethernet II,
 Src: 1e:af:01:01:1e:11, Dst: c0:8e:00:11:1e:11
Internet Protocol Version 4,
 Src: 192.168.100.101, Dst: 192.168.100.102
User Datagram Protocol,
 Src Port: 63500, Dst Port: 4789
Virtual eXtensible Local Area Network
    Flags: 0x0800, VXLAN Network ID (VNI)
    Group Policy ID: 0
    VXLAN Network Identifier (VNI): 10077
    Reserved: 0
Ethernet II,
 Src: 1e:af:01:01:1e:11, Dst: 50:00:00:03:00:07
Internet Protocol Version 4,
 Src: 172.16.10.101, Dst: 172.16.20.102
Internet Control Message Protocol
```

Capture 8-2: *Routed IP packet from VLAN10 to VLAN20 taken from Leaf-101 uplink.*

As a last step, the VNI used for routing is attached into NVE1 interface.

```
interface nve1
  no shutdown
  host-reachability protocol bgp
  source-interface loopback100
  member vni 10000
    ingress-replication protocol bgp
  member vni 10077 associate-vrf
```

Example 8-18: *Associating VNI 10077 into interface NVE1.*

The previous section describes the tenant-based routing. This short section adds an SVI for VLAN 10. Interface VLAN is attached into VRF TENANT77 and ip address 172.16.10.1/24 is attached to it. The example 8-8 defines the tenant wide Anycast Gateway (AGW) MAC address 0001.0001.0001 used in every VLAN. The last command "*fabric forwarding mode anycast-gateway*" enables AGW for VLAN 10.

```
interface Vlan10
  no shutdown
  vrf member TENANT77
  ip address 172.16.10.1/24
  fabric forwarding mode anycast-gateway
```

Example 8-19: *SVI for VLAN 10.*

When host Cafe in VLAN 10 send data packets to host Beef in VLAN 20, the packet is first sent to Leaf-101 that is AGW for network 172.16.10.0/24. Leaf-101 does routing lookup and forwards packet to Leaf-102 by using VNI 10077 in VXLAN header. When Leaf-102 receives the packet it knows that the packet belongs to specific tenant and and does routing lookup of tenant RIB. Then packet is sent to host Beef. Chapter 9 discusses routing model in detail.

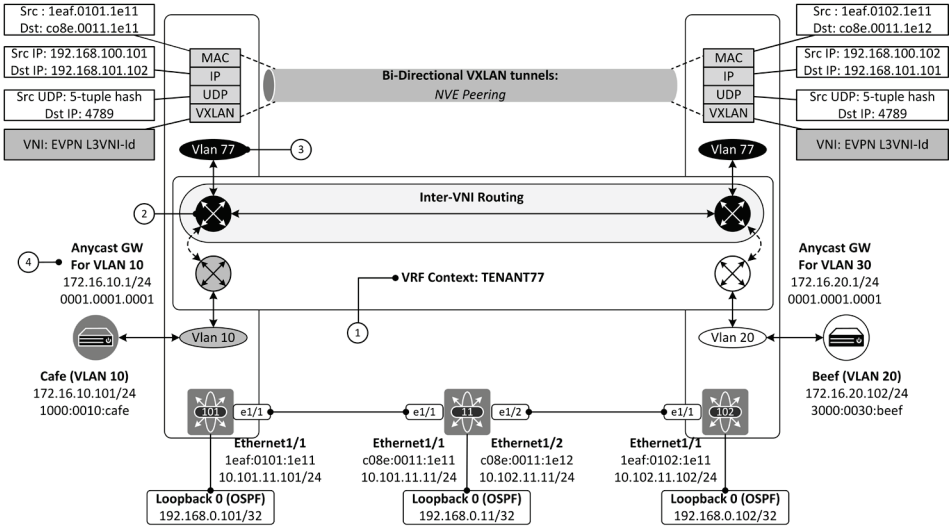

Figure 8-4: *Inter-VNI Routing model.*

This chapter introduces the basic VXLAN building blocks and their configuration. Next chapter introduces the BGP L2VPN EVPN Control Plane operation.

References

Nexus 9000 TCAM Carving: https://www.cisco.com/c/en/us/support/docs/switches/nexus-9000-series-switches/119032-nexus9k-tcam-00.html

Cisco Nexus 9000 Series NX-OS Security Configuration Guide, Release 7.x: https://www.cisco.com/c/en/us/td/docs/switches/datacenter/nexus9000/sw/7-x/security/configuration/guide/b_Cisco_Nexus_9000_Series_NX-OS_Security_Configuration_Guide_7x/b_Cisco_Nexus_9000_Series_NX-OS_Security_Configuration_Guide_7x_chapter_01001.html#concept_846AE66E9B2C4E0EAA3E54FBE51C4A87

Configuring VXLAN BGP EVPN: https://www.cisco.com/c/en/us/td/docs/switches/datacenter/nexus9000/sw/7-x/vxlan/configuration/guide/b_Cisco_Nexus_9000_Series_NX-OS_VXLAN_Configuration_Guide_7x/b_Cisco_Nexus_9000_Series_NX-OS_VXLAN_Configuration_Guide_7x_chapter_0100.pdf

A TCAM-Based Distributed Parallel IP Lookup Scheme and Performance Analysis: https://www.researchgate.net/publication/3335215_A_TCAM-based_distributed_parallel_IP_lookup_scheme_and_performance_analysis