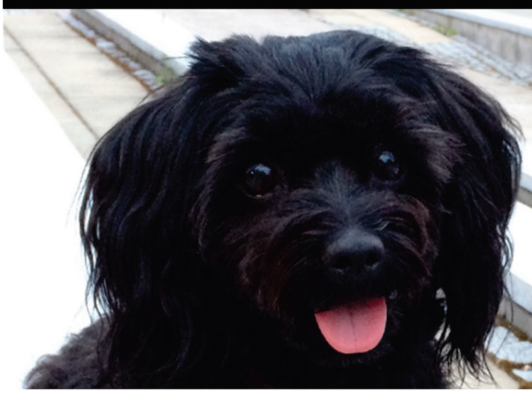
The Network Times
Handbook Series – Part VI

Network Virtualization

LISP, OMP, and BGP EVPN Operation and Interaction



Campus Fabric, SD-WAN, and Datacenter Fabric each have their unique Control-Plane solutions for Network Virtualization. Campus Fabric uses LISP, SD-WAN uses OMP, and Datacenter Fabric uses BGP L2VPN EVPN. This book describes the basics operation of each solution. It also explains how endpoint reachability information is advertised from Campus Fabric via SD-WAN to Datacenter Fabric and another way around.

Toni Pasanen, CCIE#28158

Network Virtualization

LISP, OMP, and BGP EVPN Operation and Interaction

Toni Pasanen, CCIE 28158

Copyright © Toni Pasanen, All rights reserved.

Published - 11 August 2021

About the Editor:

Toni Pasanen. CCIE No. 28158 (RS), Distinguished Engineer at Fujitsu Finland. Toni started his IT carrier in 1998 at Tieto, where he worked as a Service Desk Specialist moving via the LAN team to the Data Center team as a 3rd. Level Network Specialist. Toni joined Teleware (Cisco Learning partner) in 2004, where he spent two years teaching network technologies focusing on routing/switching and MPLS technologies. Toni joined Tieto again in 2006, where he spent the next six years as a Network Architect before joining Fujitsu. In his current role, Toni works closely with customers helping them in selecting the right network solutions not only from the technology perspective but also from the business perspective. He is also the author of books:

Virtual Extensible LAN – VXLAN: The Practical Guide to VXLAN Solution –2019

LISP Control-Plane in Campus Fabric. A Practical Guide to Understand the Operation of Campus Fabric– 2020

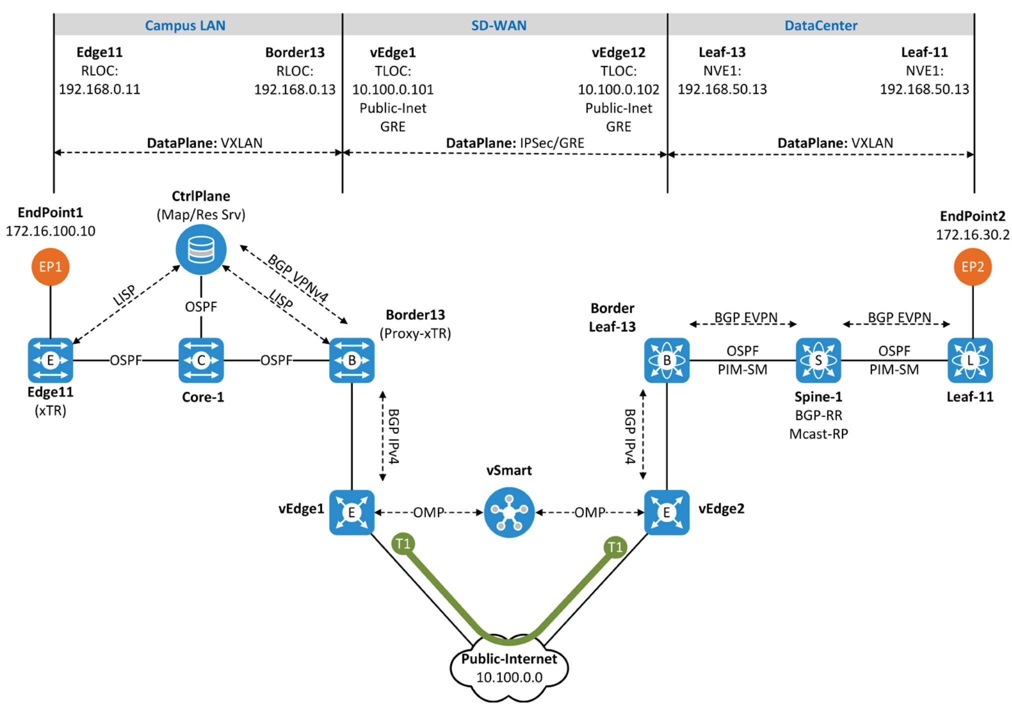
VXLAN Fabric with BGP EVPN Control-Plane. Design Considerations – 2020

Object-Based Approach to Cisco ACI: The Logic Behind the Application Centric Infrastructure - 2020.

Cisco SD-WAN: A Practical Guide to Understand the Basics of Cisco Viptela Based SD-WAN Solution- 2021.

About This Book

I have written a couple of books about Network Virtualization Overlay over Layer 3 (NVO3). My first book was about Datacenter network virtualization based on BGP L2VPN EVPN. After that, I wrote a book about Campus networks based on LISP. In my latest book, I introduced the Cisco SD-WAN solution running OMP in Control-Plane. I wanted to write one more book where I combine these three different NVO3 solutions. I haven't used pictures in the "About This Book" section in my previous books but now I decided to do that because one picture tells more than 1000 words. The figure below combines these three NVO3 solutions and illustrates what is needed to have IP connectivity between EP1 in the LISP domain and EP2 in the BGP EVPN domain. After reading this book you should be able to understand the processes of how IP reachability information about local hosts are advertised from the LISP domain over the SD-WAN to BGP EVPN domain and another way around. I wanted to keep this complex solution as simple as possible. That is why I didn't include any redundancy.



Disclaimers

The content of this book is based only on the author's own experience and testing results. Its content is neither validated nor accepted by Cisco or any other organization or person. This book is meant to be neither design nor an implementation guide. After reading this book, readers should do their technology validation before using it in a production environment.

Table of Contents

Chapter 1: Campus LAN Control-Plane: LISP

EID-to-RLOC Map-Registration Process

Introduction

Figure 1-1 illustrates phases of how Edge-xTR-11 registers the IP/VRF information of its connected host EP1 to the centralized Control Plane device MapSrv-22. First, Edge-xTR-11 verifies that it can reach the MapSrv-22. Then it sends an unreliable Map-Register message concerning EP1 information to MapSrv-22 to which MapSrv-22 responds with Map-Notify message. After that Edge-xTR-11 opens a TCP connection to Map-Srv-22 and sends new, reliable Map-Register message.

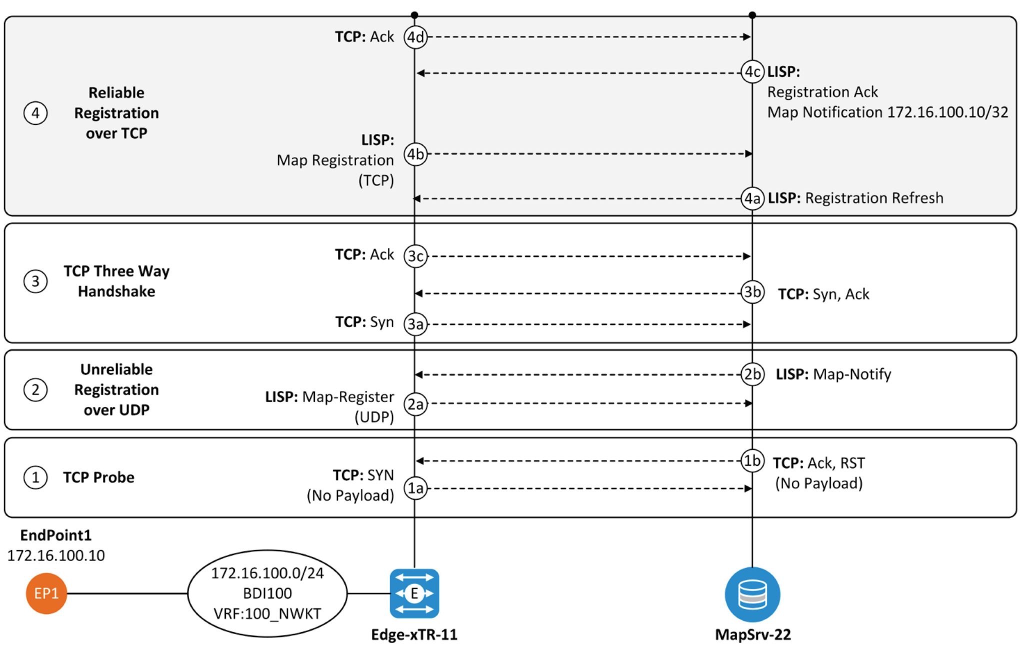


Figure 1-1: Overview of *EID-to-RLOC Map-Register Process.*

Phase-1: Updating Local MAC table and Dynamic EID

Device Edge-xTR-11 in figure 1-2 has two roles, it is an egress Tunnel Router (eTR). This means that it forwards traffic from remote endpoints to its locally connected hosts. In this role, it registers locally connected hosts to Control Plane node MapSrv-22. In its other role, Edge-xTR-11 is an ingress Tunnel Router (iTR). This means that when one of its locally connected endpoints wants to communicate with a remote endpoint, Edge-xTR-11 requests location information from the MapSrv-22 if the information is not yet resolved and stored into local Mapping Cache.

When EP1 in our example boots up, it sends a Gratuitous ARP message to make sure that other devices are not using the IP address assigned to it. When Edge-xTR-11 as a role of iTR receives the massage, it first saves the host MAC/instance (BDI)/ interface information to its MAC address table (example 1-1). Besides, it stores the information into the LISP Dynamic Endpoint Identifier (Dynamic-EID) database (example 1-2). In our example, EP1 belongs to Layer 2 Bridge Domain 100 (subnet 172.16.100.0/24). Subnet 172.16.100.0/24 (VRF 100_NWKT), in turn, is attached to LISP instance 100, which uses RLOC 192.168.0.11 (Loopback 0 IP of Edge-xTR-11).

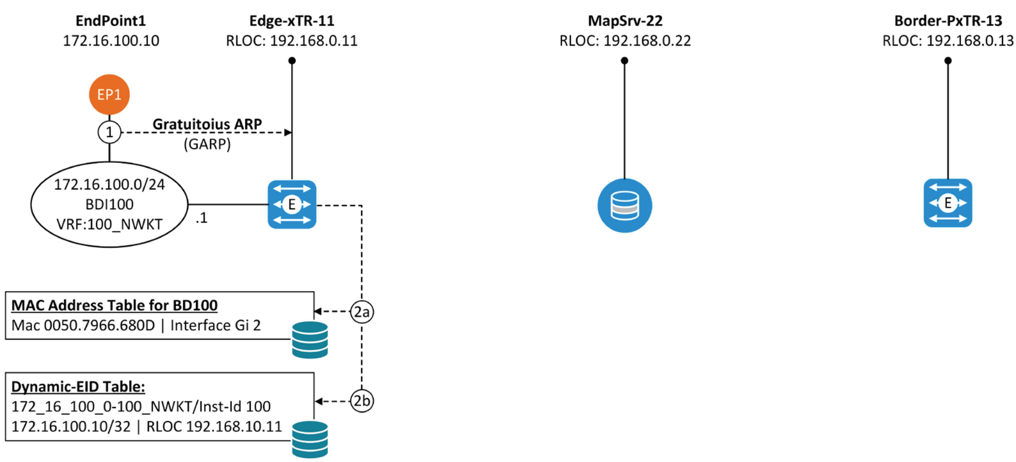


Figure 1-2: *EID-to-RLOC Map-Register – Local Learning.*

```
Edge-xTR-11#show bridge-domain 100
Bridge-domain 100 (2 ports in all)
State: UP                    Mac learning: Enabled
Aging-Timer: 300 second(s)
    BDI100  (up)
    GigabitEthernet2 service instance 100
   AED MAC address    Policy  Tag       Age  Pseudoport
   -   0001.0001.0001 to_bdi  static    0    BDI100
   0   0050.7966.680D forward dynamic   251  GigabitEthernet2.EFP100
```

Example 1-1: *EID-to-RLOC Map Register – Local Learning – Mac Address Table.*

```
Edge-xTR-11#show lisp instance-id 100 ipv4 database 172.16.100.10/32
LISP ETR IPv4 Mapping Database for EID-table vrf 100_NWKT (IID 100), LSBs: 0x1
Entries total 1, no-route 0, inactive 0

172.16.100.10/32, dynamic-eid 172_16_100_0-100_NWKT, inherited from default locator-
set RLOC-SET1
  Locator       Pri/Wgt  Source     State
  192.168.0.11    1/1    cfg-intf   site-self, reachable
```

Example 1-2: *EID-to-RLOC Map-Register – Local Learning– Dynamic EID Database.*

Phase-2: Verifying Mapping Server reachability

After updating its MAC table and Dynamic-EID table, Edge-xTR-11 starts the EID-to-RLOC mapping process (figure 1-2). As a first step, it verifies that MapSrv-22 is reachable. It sends an IP packet without payload and with TCP SYN bit set using destination TCP port 4342 to MapSrv-22 (Capture 1-1). When MapSrv-22 receives an "empty" IP packet with TCP SYN bit set and with the destination TCP port 4342, it knows that this is just a probe. It replies with an empty IP packet with TCP ACK and RST bits set (Capture 1-2).

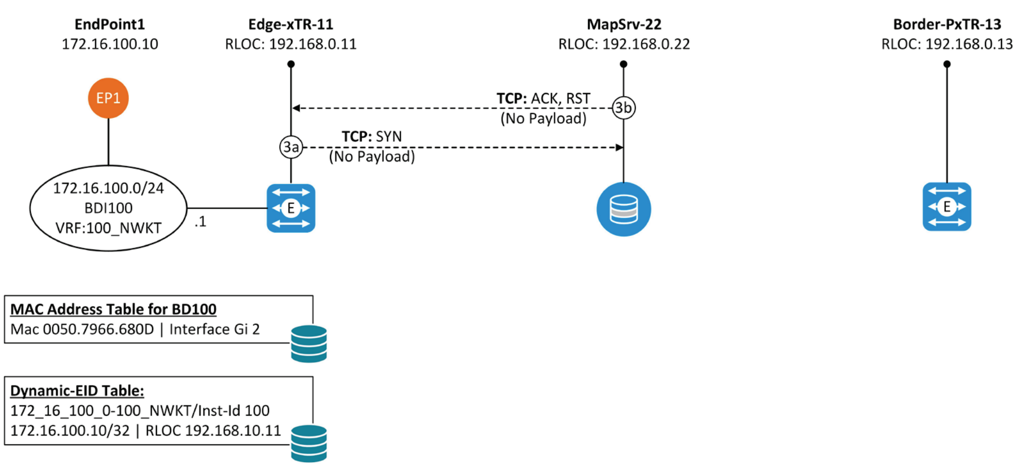


Figure 1-3: *EID-to-RLOC Map-Register – TCP Probes.*

```
Internet Protocol Version 4, Src: 192.168.0.11, Dst: 192.168.0.22
Transmission Control Protocol, Src Port: 42548, Dst Port: 4342, Seq: 0, Len: 0
    Source Port: 42548
    Destination Port: 4342
    [Stream index: 0]
    [TCP Segment Len: 0]
    Sequence number: 0    (relative sequence number)
    Sequence number (raw): 1675383545
    [Next sequence number: 1    (relative sequence number)]
    Acknowledgment number: 0
    Acknowledgment number (raw): 0
    0110 .... = Header Length: 24 bytes (6)
    Flags: 0x002 (SYN)
    <snipped>
```

Capture 1-1: *EID-to-RLOC Map-Register - TCP Probe Sent by Edge-xTR-11.*

```
Internet Protocol Version 4, Src: 192.168.0.22, Dst: 192.168.0.11
Transmission Control Protocol, Src Port: 4342, Dst Port: 42548, Seq: 1, Ack: 1, Len: 0
    Source Port: 4342
    Destination Port: 42548
    [Stream index: 0]
    [TCP Segment Len: 0]
    Sequence number: 1    (relative sequence number)
    Sequence number (raw): 0
    [Next sequence number: 1    (relative sequence number)]
    Acknowledgment number: 1    (relative ack number)
    Acknowledgment number (raw): 1675383546
    0101 .... = Header Length: 20 bytes (5)
    Flags: 0x014 (RST, ACK)
    <snipped>
```

Capture 1-2: *EID-to-RLOC Map-Register – Reply to TCP Probe by MapSrv-22.*

Phase-3: Unreliable EID-to-RLOC registration (UDP)

After verifying the IP connectivity to MapSrv-22, Edge-xTR-11 sends a LISP Map-Register message to MapSrv-22. At this phase, the Map-Register message is sent by using unreliable UDP by using source/destination port 4342 (capture 1-3). Map-Register message describes the EID of EP1. EID is a combination of the LISP Instance Id 100 bind to subnet 172.16.100.0/24 and the host IP address 172.16.100.10/32. The EID itself is bind to location information Routing Locator (RLOC). In other words, Map-Register message tells to MapSrv-22 that EP1 has an IP address 172.16.100.10/32, it belongs to Virtual Network 100, and it is reachable through the Edge-xTR-11 (192.168.0.11).

There are three bits set in the Map-Register message originated by Edge-xTR-11. The *Proxy-Map-Reply* bit, when set, means that Map-Srv-22 can respond to requested endpoint information on behalf of Edge-xTR-11. The *xTR-Id present* bit, when set, describes that this message carries randomly generated xTR-Identifier of sender. The *Want-Map-Notify* bit, when set, in the Map-register message indicates that Edge-xTR-11 wants MapSrv-22 to confirm that it has received the message. This means that even though Edge-xTR-11 uses UDP at this phase, it uses LISP Map-Notify messages like TCP uses Acknowledgement messages. The Authoritative bit, when set, tells that the Edge-xTR-11 is the last-hop router for the EID.

The *Nonce* value carried within the Map-Register message protects against *Map-Register Reply Attacks*. Edge-xTR-11 accepts Map-Notify messages from MapSrv-22 only if the nonce value in the message is the same as what it uses with the Map-Register message. Edge-xTR-11 also increases the nonce value every time it sends a new Map-Register message. MapSrv-22 knows this and it only allows Map-Register messages from Edge-xTR-11 if the nonce value is higher than carried with the previous Map-Register message.

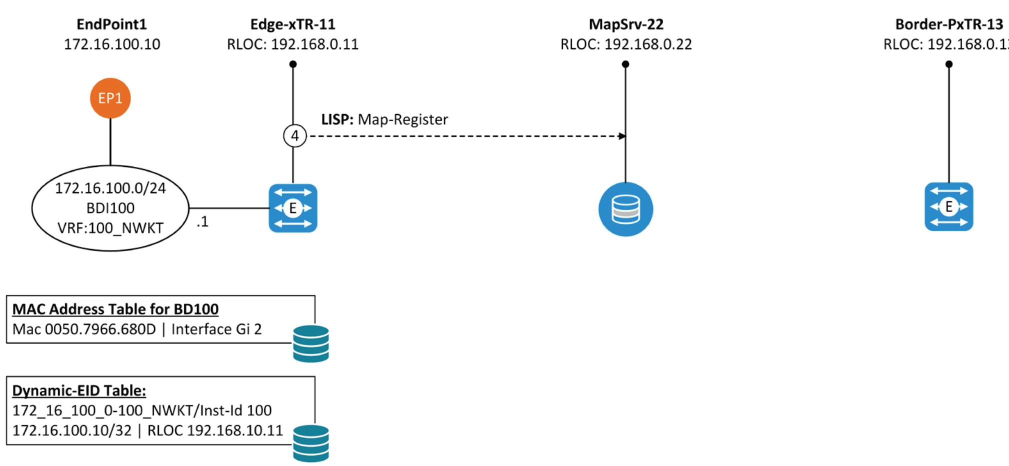


Figure 1-4: *EID-to-RLOC Map-Register - Unreliable Map-Register Message.*

```
Internet Protocol Version 4, Src: 192.168.0.11, Dst: 192.168.0.22
User Datagram Protocol, Src Port: 4342, Dst Port: 4342
Locator/ID Separation Protocol
    0011 .... .... .... .... .... = Type: Map-Register (3)
    .... 1... .... .... .... .... = P bit (Proxy-Map-Reply): Set
    .... .0.. .... .... .... .... = S bit (LISP-SEC capable): Not set
    .... ..1. .... .... .... .... = I bit (xTR-ID present): Set
    .... ...0 .... .... .... .... = R bit (Built for an RTR): Not set
    .... .... 0000 0000 0000 000. = Reserved bits: 0x0000
    .... .... .... .... .... ...1 = M bit (Want-Map-Notify): Set
    Record Count: 1
    Nonce: 0x3e4c8dce8abf56b7
    Key ID: 0x0001
    Authentication Data Length: 20
    Authentication Data: a705477bc22dc0d6842c2462125a35dbfa34b0ee
    Mapping Record 1, EID Prefix: [100] 172.16.100.10/32, TTL: 1440, Action: No-
Action, Authoritative
        Record TTL: 1440
        Locator Count: 1
        EID Mask Length: 32
        000. .... .... .... = Action: No-Action (0)
        ...1 .... .... .... = Authoritative bit: Set
        .... .000 0000 0000 = Reserved: 0x000
        0000 .... .... .... = Reserved: 0x0
        .... 0000 0000 0000 = Mapping Version: 0
        EID Prefix AFI: LISP Canonical Address Format (LCAF) (16387)
        EID Prefix: [100] 172.16.100.10
            LCAF: Instance ID: 100, Address: 172.16.100.10
                LCAF Header: 00000220000a
                Instance ID: 100
                Address AFI: IPv4 (1)
                Address: 172.16.100.10
        Locator Record 1, Local RLOC: 192.168.0.11, Reachable, Priority/Weight: 1/1,
Multicast Priority/Weight: 1/1
    xTR-ID: e5e601fff689f8a9aab82834637ee536
    Site-ID: 0000000000000000
```

Capture 1-3: *EID-to-RLOC Map-Register - Unreliable Map-Register Message.*

Phase-4: Map-Server - EID-RLOC DB Updates

MapSrv-22 has two roles. The first one, the *Mapping Servers* component is responsible for EID-to-RLOC Map-Register message handling. It validates the message by checking the authentication data matches and that the nonce value is higher if this is not the first Map-Register message from Edge-xTR-11. Then, if the EID IP address is within the subnet listed in the site and instance-specific EID-Record, the EOD-to-RLOC is installed into the Mapping Database. The process is shown in example 1-3. Example 1-4 shows the IP address 172.16.100.10/32 in instance 100 (virtual network 100) is reachable through the 192.168.0.11 (Edge-xTR-11). Note that the TTL field in the example is represented as 1 day, while in the original Map-Register message the TTL was represented as 1440 seconds which is 24 hours.

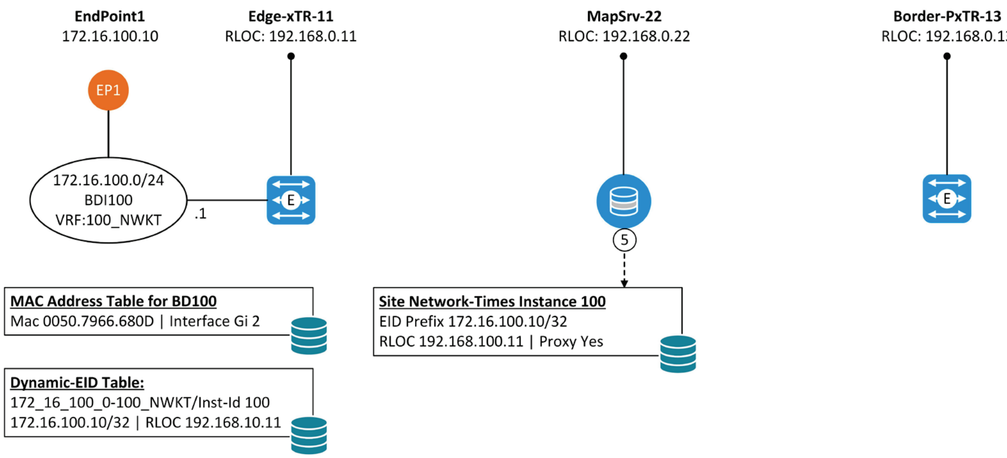


Figure 1-5: *EID-to-RLOC Map-Register - Updating Mapping Database.*

```
00:50:43: LISP: Processing IP Map-Register mapping record for IID 100 172.16.100.10/32
LCAF 2, ttl 1440, action none, authoritative, 1 locator
        192.168.0.11 pri/wei=1/1 LpR

00:50:43: LISP-0: MS Site EID IID 100 prefix 172.16.100.10/32 SVC_IP_IAF_IPv4 site
Network-Times, Created dynamic site EID prefix entry.

00:50:43: LISP-0: MS registration IID 100 prefix 172.16.100.10/32 192.168.0.11
SVC_IP_IAF_IPv4 site Network-Times, Created new registration.
```

Example 1-3: *EID-to-RLOC Map-Register Processing.*

```
MapSrv-22# show lisp site name Network-Times instance-id 100 | sec 172.16.100.10
  EID-prefix: 172.16.100.10/32 instance-id 100
    First registered:     01:47:46
    Last registered:      01:47:42
    Routing table tag:    0
    Origin:               Dynamic, more specific of 172.16.100.0/24
    Merge active:         No
    Proxy reply:          Yes
    TTL:                  1d00h
    State:                complete
    Registration errors:
      Authentication failures:   0
      Allowed locators mismatch: 0
    ETR 192.168.0.11, last registered 01:47:42, proxy-reply, map-notify
                      TTL 1d00h, no merge, hash-function sha1, nonce 0x3E4C8DCE-
0x8ABF56B7
                      state complete, no security-capability
                      xTR-ID 0xE5E601FF-0xF689F8A9-0xAAB82834-0x637EE536
                      site-ID unspecified
                      sourced by reliable transport
      Locator       Local  State      Pri/Wgt  Scope
      192.168.0.11  yes    up            1/1   IPv4 none
```

Example 1-4: *EID-to-RLOC Map-Register - Mapping Database Verification.*

Phase 5: Map Server - LISP Map Notify Message

Because the Wants-Map-Notify bit was set in the Map-Register message, MapSrv-22 confirms that it has received the message by sending a Map-Notify message to Edge-xTR-11. The process is shown in example 1-5. The message contains basically the same information as what we saw in the Map-Register message.

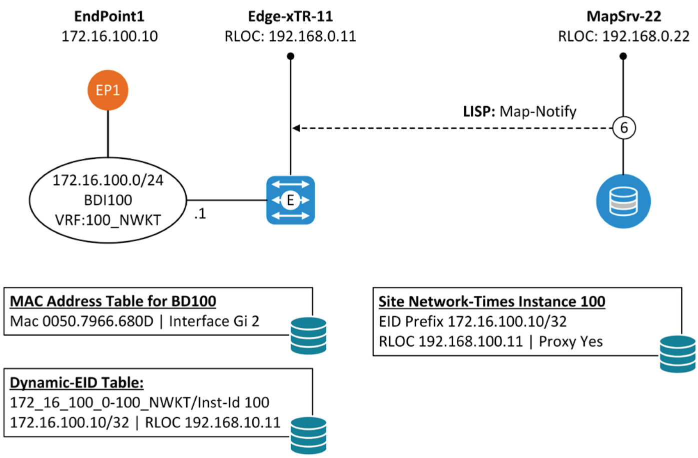


Figure 1-6: *EID-to-RLOC Map-Register - Map-Notify Message Sent By MapSrv-22.*

```
00:50:43: LISP-0: MS EID IID 100 prefix 172.16.100.10/32 SVC_IP_IAF_IPv4 site Network-
Times, Scheduling map notifications for prefix 172.16.100.10/32.

00:50:43: LISP-0: MS EID IID 100 prefix 172.16.100.10/32 SVC_IP_IAF_IPv4 site Network-
Times, Scheduling unreliable map notification for prefix 172.16.100.10/32 to ETR
192.168.0.11 which registered 172.16.100.10/32.

00:50:43: LISP-0: Map-Notify 192.168.0.22:4342->192.168.0.11:4342 xTR-ID 0xB8F66E3B-
0x101CCD7C-0x5B9F516E-0x3526FC6E, sending with 1 prefix, nonce 0xE297E49D-0xEDA98C98
```

Example 1-5: *EID-to-RLOC Map-Register - Map-Notify Message Processing.*

```
Internet Protocol Version 4, Src: 192.168.0.22, Dst: 192.168.0.11
User Datagram Protocol, Src Port: 4342, Dst Port: 4342
Locator/ID Separation Protocol
    0100 .... .... .... .... .... = Type: Map-Notify (4)
    .... 1... .... .... .... .... = I bit (xTR-ID present): Set
    .... .0.. .... .... .... .... = R bit (Built for an RTR): Not set
    .... ..00 0000 0000 0000 0000 = Reserved bits: 0x00000
    Record Count: 1
    Nonce: 0x3e4c8dce8abf56b7
    Key ID: 0x0001
    Authentication Data Length: 20
    Authentication Data: 3fbdd90d102be7e6ce504e295dfb196a52b238f2
    Mapping Record 1, EID Prefix: [100] 172.16.100.10/32, TTL: 1440, Action: No-
Action, Not Authoritative
        <snipped>
        EID Prefix AFI: LISP Canonical Address Format (LCAF) (16387)
        EID Prefix: [100] 172.16.100.10
        <snipped>
        Locator Record 1, RLOC: 192.168.0.11, Reachable, Priority/Weight: 1/1,
Multicast Priority/Weight: 1/1
    xTR-ID: e5e601fff689f8a9aab82834637ee536
    Site-ID: 0000000000000000
```

Capture 1-4: *EID-to-RLOC Map-Register Map-Notify Message Sent by MapSrv-22.*

Phase 6: TCP Connection Between Edge-xTR-11 and MapSrv-22

After receiving Map-Notify message from MapSrv-22, Edge-xTR-11 starts the reliable EID-to-RLOC Mapping process by opening TCP connection with MapSrv-22 (figure 1-7).

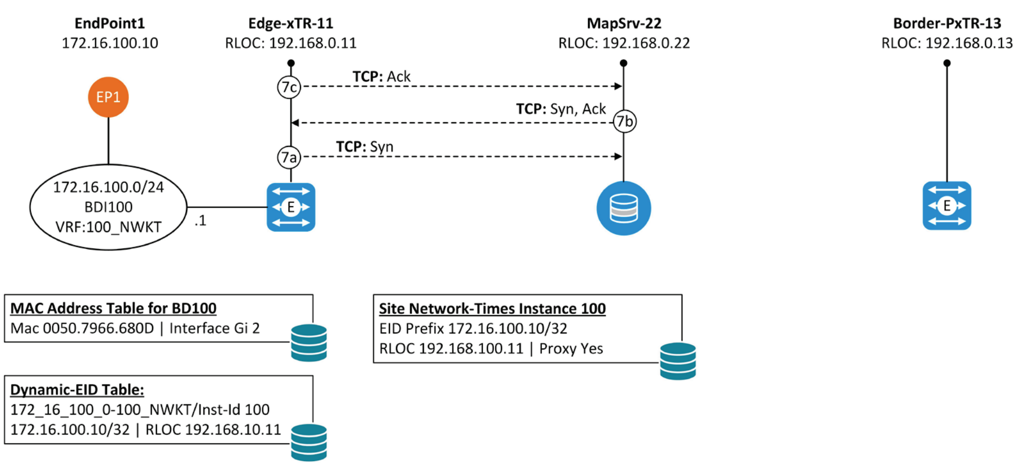


Figure 1-7: *EID-to-RLOC Map-Register - TCP Three Way Handshake.*

Captures 1-5, 1-6, and 1-7 show the TCP three-way handshake process.

```
Internet Protocol Version 4, Src: 192.168.0.11, Dst: 192.168.0.22
Transmission Control Protocol, Src Port: 46666, Dst Port: 4342, Seq: 0, Len: 0
    Source Port: 46666
    Destination Port: 4342
    <snipped>
    Flags: 0x002 (SYN)
<snipped>
```

Capture 1-5: *EID-to-RLOC Map Register - TCP Three Way Handshake TCP SYN.*

```
Internet Protocol Version 4, Src: 192.168.0.22, Dst: 192.168.0.11
Transmission Control Protocol, Src Port: 4342, Dst Port: 46666, Seq: 0, Ack: 1, Len: 0
    Source Port: 4342
    Destination Port: 46666
    <snipped>
    Flags: 0x012 (SYN, ACK)
<snipped>
```

Capture 1-6: *EID-to-RLOC Map Register - TCP Three Way Handshake TCP SYN, ACK.*

```
Internet Protocol Version 4, Src: 192.168.0.11, Dst: 192.168.0.22
Transmission Control Protocol, Src Port: 46666, Dst Port: 4342, Seq: 1, Ack: 1, Len: 0
    Source Port: 46666
    Destination Port: 4342
    Flags: 0x010 (ACK)
<snipped>
```

Capture 1-7: *EID-to-RLOC Map Register - TCP Three Way Handshake TCP ACK.*

Phase 7: LISP Registration Refresh

After successful TCP session establishment, MapSrv-22 asks Edge-xTR-11 to send complete EID-to-RLOC information now by using the reliable method using TCP.

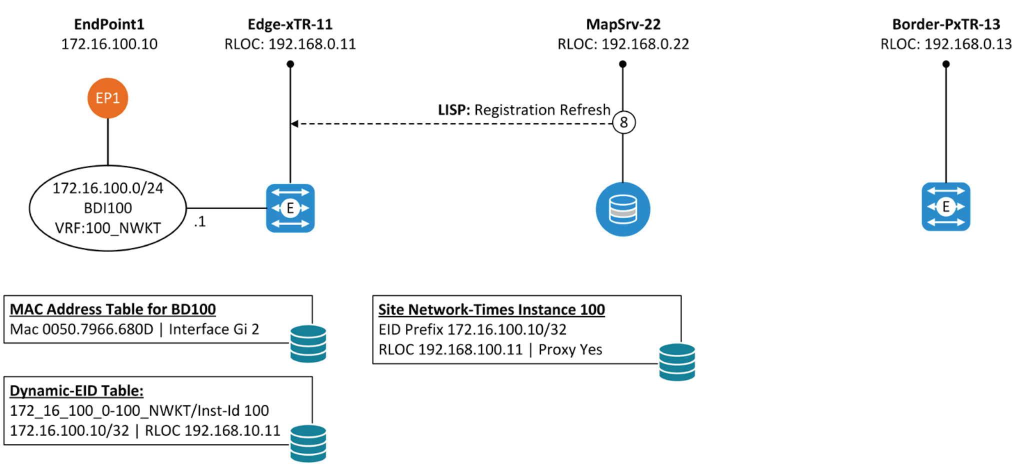


Figure 1-8: *EID-to-RLOC Map Register - LISP Registration Refresh Message.*

Capture 1-8 below shows the LISP Registration Refresh message. The message is sent using TCP.

```
Internet Protocol Version 4, Src: 192.168.0.22, Dst: 192.168.0.11
Transmission Control Protocol, Src Port: 4342, Dst Port: 46666, Seq: 1, Ack:
1, Len: 15
Locator/ID Separation Protocol (Reliable Transport), Msg: 0, Registration
Refresh, Scope: All prefixes under all address families under all EID
instances
    Type: Registration Refresh (20)
    Length: 15
    Message ID: 0
    Registration refresh scope: All prefixes under all address families under
all EID instances (0)
    0... .... .... .... = Rejected only: Not set
    .000 0000 0000 0000 = Reserved: 0x0000
    Message End Marker: 0x9facade9 (correct)
```

Capture 1-8: *EID-to-RLOC Map-Register - LISP Registration Refresh Message*.

Phase 8: Reliable EID-to-RLOC registration (TCP)

As a response to Registration Refresh message, Edge-xTR-11 send the EID-to-RLOC Map-Register message to MapSrv-22 but now using TCP. Because Edge-xTR-11 now uses reliable transport, MapSrv-22 has to confirm that it has received the message by sending the Registration Refresh message among the Map-Notify message (due to Want-Map-Notify). As the last step, Edge-xTR-11 sends an IP packet with the TCP ACK bit set.

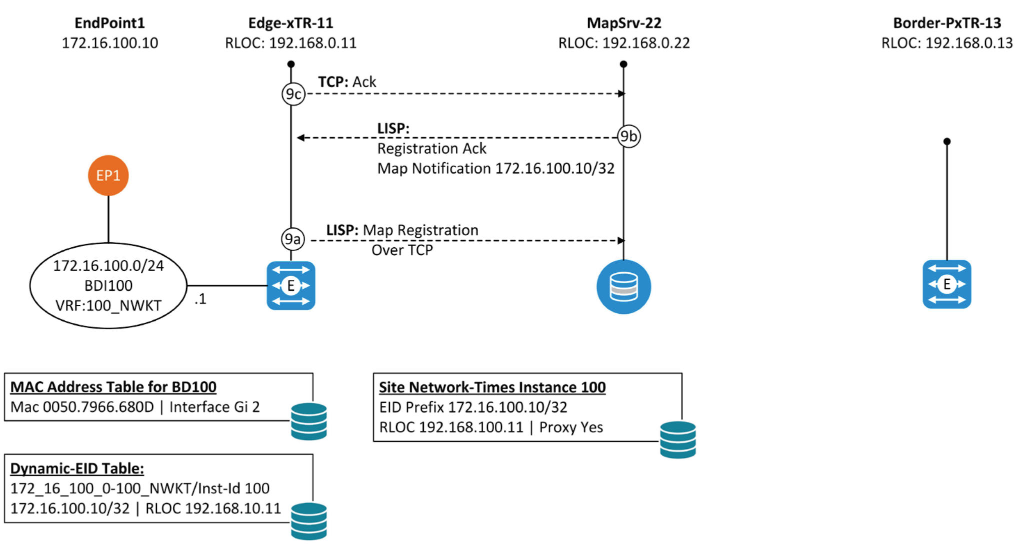


Figure 1-9: *EID-to-RLOC Map-Register - Registration Over TCP.*

Capture 1-9 shows the Map-register message sent by Edge-xTR-11

```
Internet Protocol Version 4, Src: 192.168.0.11, Dst: 192.168.0.22
Transmission Control Protocol, Src Port: 46666, Dst Port: 4342, Seq: 1, Ack: 16, Len:
130
<snipped>
Locator/ID Separation Protocol (Reliable Transport), Msg: 1, Registration for [100]
172.16.100.10/32
    Type: Registration (17)
    Length: 112
    Message ID: 1
    Map-Register
        .... 1... .... .... .... .... = P bit (Proxy-Map-Reply): Set
        .... .0.. .... .... .... .... = S bit (LISP-SEC capable): Not set
        .... ..1. .... .... .... .... = I bit (xTR-ID present): Set
        .... ...0 .... .... .... .... = R bit (Built for an RTR): Not set
        .... .... 0000 0000 0000 000. = Reserved bits: 0x0000
        .... .... .... .... .... ...1 = M bit (Want-Map-Notify): Set
        Record Count: 1
        Nonce: 0x3e4c8dce8abf56b7
        Key ID: 0x0001
        Authentication Data Length: 20
        Authentication Data: a705477bc22dc0d6842c2462125a35dbfa34b0ee
        Mapping Record 1, EID Prefix: [100] 172.16.100.10/32, TTL: 1440, Action: No-
Action, Authoritative
            <snipped>
            EID Prefix AFI: LISP Canonical Address Format (LCAF) (16387)
            EID Prefix: [100] 172.16.100.10
                LCAF: Instance ID: 100, Address: 172.16.100.10
                    LCAF Header: 00000220000a
                    Instance ID: 100
                    Address AFI: IPv4 (1)
                    Address: 172.16.100.10
            Locator Record 1, Local RLOC: 192.168.0.11, Reachable, Priority/Weight:
1/1, Multicast Priority/Weight: 1/1
                 <snipped>
```

Capture 1-9: *EID-to-RLOC Map-Register - Registration Over TCP (Step 1).*

Capture 1-10 verifies that both registration Ack and Map-Notify messages are sent within the same IP packet. Capture 1-11 shows the TCP ACK sent by Edge-xTR-11.

```
Internet Protocol Version 4, Src: 192.168.0.22, Dst: 192.168.0.11
Transmission Control Protocol, Src Port: 4342, Dst Port: 46666, Seq: 16, Ack: 131,
Len: 107
    <snipped>
Locator/ID Separation Protocol (Reliable Transport), Msg: 1, Registration ACK for
[100] 172.16.100.10/32
    Type: Registration ACK (18)
    Length: 31
    Message ID: 1
    EID Prefix: [100] 172.16.100.10/32
    Message End Marker: 0x9facade9 (correct)
Locator/ID Separation Protocol (Reliable Transport), Msg: 2, Mapping Notification for
[100] 172.16.100.10/32
    <snipped>
    Mapping Record 1, EID Prefix: [100] 172.16.100.10/32, TTL: 1440, Action: No-
Action, Not Authoritative
        <snipped>
        EID Prefix AFI: LISP Canonical Address Format (LCAF) (16387)
        EID Prefix: [100] 172.16.100.10
            LCAF: Instance ID: 100, Address: 172.16.100.10
                <snipped>
        Locator Record 1, RLOC: 192.168.0.11, Reachable, Priority/Weight: 1/1,
Multicast Priority/Weight: 1/1
            <snipped>
```

Capture 1-10: *EID-to-RLOC Map-Register - Registration Over TCP (Step 2).*

```
Internet Protocol Version 4, Src: 192.168.0.11, Dst: 192.168.0.22
Transmission Control Protocol, Src Port: 46666, Dst Port: 4342, Seq: 131, Ack: 123,
<snipped>
    Flags: 0x010 (ACK)
    <snipped>
```

Capture 1-11: *EID-to-RLOC Map-Register - Registration Over TCP (Step 3).*

In this phase, MapSrv-22 can respond to Map-request messages on behalf of Edge-xTR-11. The next chapter describes how the EID-to-RLOC information is advertised to Border-PxTR-13. Because the LISP Control-Plane node, in our case MapSrv-22, does not redistribute EID-to-RLOC information in an unsolicited way like BGP Route-Reflector, we need another mechanism for that and that’s where MP-BGP VPNv4 comes into play.

BGP VPNv4 Update from Control Plane to Border

Introduction

The previous chapter describes how Edge-xTR-11 used LISP Map-Register message to advertise EID-to-RLOC information to MapServ-22. It also explained how MapSrv-22, as a role of Mapping Server, stores the information into Mapping Data Base. MapSrv-22 is also *Map-Resolver*. This means that when it receives the LISP *Map-Request* message from the xTR device, it will respond with a *Map-Reply* message. If MapSrv-22 knows the EID-to-RLOC mapping, it places this information into the Map-Reply message. If MapSrv-22 doesn't have mapping information, it instructs requesting xTR to forward traffic to its Proxy-xTR. This, however, is not the case in our example. What we want to do is advertise the EP1 reachability information to Border-PxTR. In order to do that, we need to a) export EID-to-RLOC information from the Mapping Data Base to instance-specific VRF_100 RIB. Then we can advertise it by using BGP and because we want to include virtual network identifier into update we use MP-BGP VPNv4 because there we have Route Target Attribute. The next sections describe the process in detail.

Phase 1: Map Server - RIB Update

LISP Map-Server doesn't install EID-to-RLOC mapping information from the Mapping Database into a RIB by default. To do that we need to export the information from the LISP Mapping Data Base to RIB by using the LISP Instance-specific command *route-export site-registrations*. Example 1-6 illustrates the update process. Example 1-7 shows the RIB entry concerning EP1 IP address 172.16.100.10/32 in VRF 100_NWKT. Due to redistribution, the route is shown as directly connected, via Null0. If you take a look at the timestamps in example 1-6 and compare it to timestamps in example 1-3, you will see that the RIB update happens right after the unreliable EID-to-RLOC registration process.

Complete device configuration can be found in chapter 1 Appendix 1.

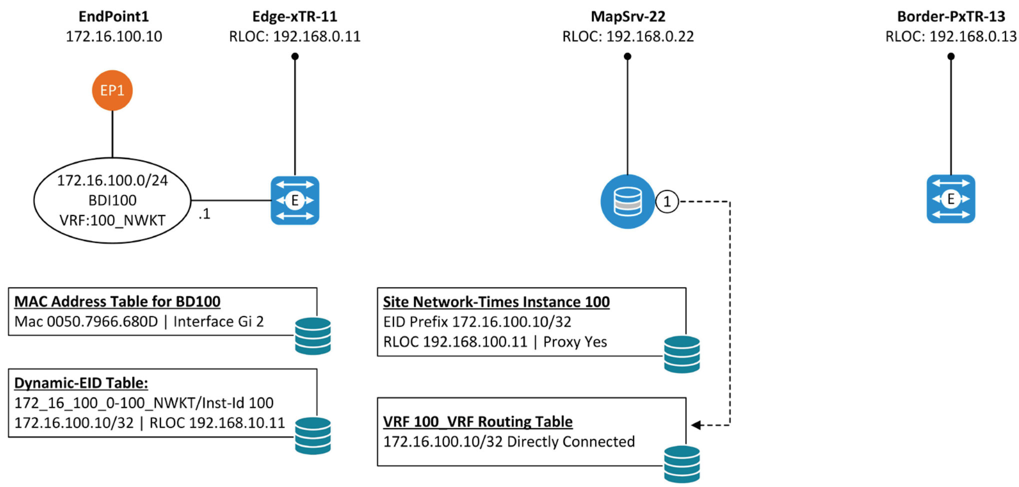


Figure 1-10: *EID-to-RLOC information from LISP to RIB.*

```
00:50:43: LISP: RIB Watch Group default 192.168.0.11/32 , created.

00:50:43: LISP: RIB Watch Group default 192.168.0.11/32 , scheduling RIB update.

00:50:43: LISP-0: MS EID IID 100 prefix 172.16.100.10/32 SVC_IP_IAF_IPv4 site Network-
Times, MS EID prefix export route update/create.

00:50:43: LISP: RIBtable [IPv4:100_NWKT:172.16.100.10/32] created.

00:50:43: LISP: RIBtable [IPv4:100_NWKT:172.16.100.10/32] add source ms-site-reg.
```

Example 1-6: *EID-to-RLOC information from LISP to RIB process.*

```
MapSrv-22#sh ip route vrf 100_NWKT 172.16.100.10

Routing Table: 100_NWKT
Routing entry for 172.16.100.10/32
  Known via "lisp", distance 250, metric 1, type intra area
  Redistributing via bgp 65010
  Advertised by bgp 65010 metric 10 route-map LISP-to-BGP
  Routing Descriptor Blocks:
  * directly connected, via Null0
      Route metric is 1, traffic share count is 1
```

Example 1-7: *RIB Entry About EP1 IP address 172.16.100.10/32*

Phase 2: Map-Server - BGP Table Update

When the LISP route is installed into VRF Specific RIB, we can export the route to the BGP process. LISP routes have to be redistributed to BGP through the route-map where we allow only subnets used within the LISP domain. The reason is that when Border-PxTR-13 receives routing update from SD-WAN Edge device, it will send a LISP Map-Register message to MapSrv-22, which in turn installs the information into the LISP Mapping Data Base. Without redistribution policy, this route is redistributed to BGP and advertised back to Border-PxTR. LISP Instances are used as a Virtual Network Identifier (VN-Id) in LISP Domain. MP-BGP VPNv4 update carries VRF specific Extended Community Route-Target, which is used to import routes in the right BRIB. We also need to create an aggregate route, otherwise we end up advertising all host routes.

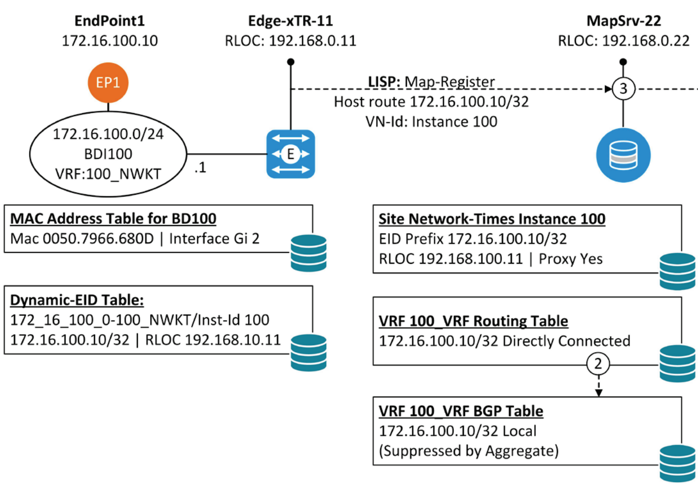


Figure 1-11: *LSIP Route from RIB to BGP Process.*

```
00:50:43: LISP: RIBtable [IPv4:100_NWKT:172.16.100.10/32] route create nh NULL if
Null0 topoid INVALID route-tag 0.
00:50:43: BGP(4): redist event (1) request for 1:100:172.16.100.10/32
00:50:43: BGP(4) route 1:100:172.16.100.10/32 gw-1 0.0.0.0 src_proto (lisp) path-limit
1
00:50:43: BGP(4): route 1:100:172.16.100.10/32 up
00:50:43: BGP(4): sourced route for 1:100:172.16.100.10/32 created
00:50:43: BGP(4): sourced route for 1:100:172.16.100.10/32 path 0x7FDB7F0E5348 id 0 gw
0.0.0.0 created (weight 32768)
00:50:43: BGP(4): redistributed route 1:100:172.16.100.10/32 added gw 0.0.0.0
00:50:43: BGP(4): created aggregate route for 1:100:172.16.100.0/24
00:50:43: BGP(4): Revise route installing 1 of 1 routes for 172.16.100.0/24 ->
0.0.0.0(100_NWKT) to 100_NWKT IP table
00:50:43: BGP: topo 100_NWKT:VPNv4 Unicast:base Remove_fwdroute for
1:100:172.16.100.10/32
00:50:43: BGP(4): 192.168.0.13 NEXT_HOP is set to self for net 1:100:172.16.100.0/24,
00:50:43: BGP(4): (base) 192.168.0.13 send UPDATE (format) 1:100:172.16.100.0/24, next
192.168.0.22, label 20, metric 0, path Local, extended community RT:1:100
```

Example 1-8: *LISP Route from RIB to BGP Process.*

Example 1-9 shows the VRF 100_NWKT BRIB entry concerning IP address 172.16.100.10. The reason why MapSrv-22 doesn't advertise the information to any BGP peer is that we are only advertising an aggregate address 172.16.100.0/24. The Route-Target is taken from the VRF 100_NWKT configuration.

```
MapSrv-22#sh ip bgp vpnv4 vrf 100_NWKT 172.16.100.10
BGP routing table entry for 1:100:172.16.100.10/32, version 4
Paths: (1 available, best #1, table 100_NWKT, Advertisements suppressed by an
aggregate.)
  Not advertised to any peer
  Refresh Epoch 1
  Local
    0.0.0.0 (via default) from 0.0.0.0 (192.168.0.22)
      Origin incomplete, metric 10, localpref 100, weight 32768, valid, sourced, best
      Extended Community: RT:1:100
      mpls labels in/out 16/nolabel
```

Example 1-9: *VRF 100_NWKT BRIB.*

Capture 1-12 shows the BGP update message sent by MapSrv-22 where it lists itself as a next-hop.

```
Internet Protocol Version 4, Src: 192.168.0.22, Dst: 192.168.0.13
Transmission Control Protocol, Src Port: 179, Dst Port: 50606, Seq: 1, Ack: 1, Len:
104
Border Gateway Protocol - UPDATE Message
    <snipped>
    Path attributes
        Path Attribute - MP_REACH_NLRI
            Flags: 0x80, Optional, Non-transitive, Complete
            Type Code: MP_REACH_NLRI (14)
            Length: 32
            Address family identifier (AFI): IPv4 (1)
            Subsequent address family identifier (SAFI): Labeled VPN Unicast (128)
            Next hop network address (12 bytes)
                Next Hop: Empty Label Stack RD=0:0 IPv4=192.168.0.22
            Number of Subnetwork points of attachment (SNPA): 0
            Network layer reachability information (15 bytes)
                BGP Prefix
                    Prefix Length: 112
                    Label Stack: 17 (bottom)
                    Route Distinguisher: 1:100
                    MP Reach NLRI IPv4 prefix: 172.16.100.0
        Path Attribute - ORIGIN: IGP
        Path Attribute - AS_PATH: empty
        Path Attribute - AGGREGATOR: AS: 65010 origin: 192.168.0.22
        Path Attribute - ATOMIC_AGGREGATE
        Path Attribute - MULTI_EXIT_DISC: 0
        Path Attribute - LOCAL_PREF: 100
        Path Attribute - EXTENDED_COMMUNITIES
            Flags: 0xc0, Optional, Transitive, Complete
            Type Code: EXTENDED_COMMUNITIES (16)
            Length: 8
            Carried extended communities: (1 community)
                Route Target: 1:100 [Transitive 2-Octet AS-Specific]
```

Capture 1-12: *BGP Update sent by MapSrv-22.*

Phase 3: Border-PxTR – BRIB, RIB, and Dyn-EID Table Update

When Border-PxTR-13 receives the BGP Update from MapSrv-22, the information is first installed into the Adj-RIB-In and from there into VRF 100_NWKT BRIB based on the Route-Target 1:100. From BRIB route is installed into VRF 100_NWKT RIB. Because Border-PxTR-13 receives BGP updates from the SD-WAN Edge device, it imports routes from the BRIB into LISP using itself as an RLOC. This is done by using the command *route-import database bgp [AS Number] locator-set [RLOC set name]* under the LISP instance.

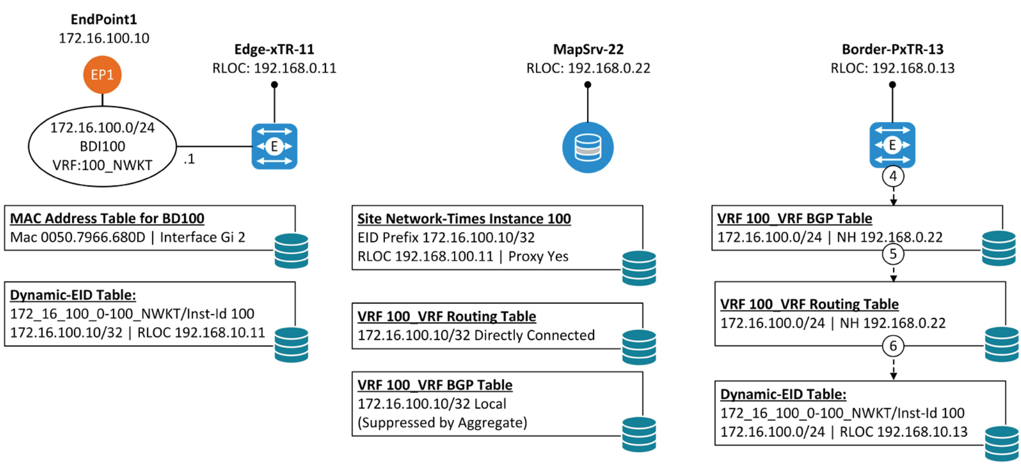


Figure 1-12: *BRIB and RIB Update by Border-PxTR-13.*

Example 1-10 shows that Border-PxTR-13 has import subnet 172.16.100.0/24 NLRI in its BRIB related to VRF 100_NWKT.

```
Border-PxTR-13#sh ip bgp vpnv4 vrf 100_NWKT 172.16.100.0/24
BGP routing table entry for 1:100:172.16.100.0/24, version 3
Paths: (1 available, best #1, table 100_NWKT)
Flag: 0x100
  Not advertised to any peer
  Refresh Epoch 1
  Local, (aggregated by 65010 192.168.0.22)
    192.168.0.22 (metric 3) (via default) from 192.168.0.22 (192.168.0.22)
      Origin IGP, metric 0, localpref 100, valid, internal, atomic-aggregate, best
      Extended Community: RT:1:100
      mpls labels in/out nolabel/17
```

Example 1-10: *Border-PxTR-13 VRF-specific BRIB.*

Example 1-11 shows that routing information towards 172.16.100.0/24 is installed from the BRIB into RIB of VRF 100_NWKT.

```
Border-PxTR-13#sh ip route vrf 100_NWKT 172.16.100.0

Routing Table: 100_NWKT
Routing entry for 172.16.100.0/24
  Known via "bgp 65010", distance 200, metric 0, type internal
  Redistributing via lisp
  Last update from 192.168.0.22 00:24:09 ago
  Routing Descriptor Blocks:
  * 192.168.0.22 (default), from 192.168.0.22, 00:24:09 ago
      Route metric is 0, traffic share count is 1
      AS Hops 0
      MPLS label: 17
      MPLS Flags: MPLS Required
```

Example 1-11: *Border-PxTR-13 VRF-specific RIB.*

Example 1-12 shows the RIB to LISP redistribution process. Remember that routing information can't be installed from one protocol DB to another protocol DB. This means that valid BGP table entry is first encoded into the RIB and from there it is redistributed to LISP DB.

```
00:47:40: LISP-0: RIB Redist Spec IID 100 database bgp 65010, RIB event
172.16.100.0/24 UP.
00:47:40: LISP-0: RtImp IID 100 prefix 172.16.100.0/24 database bgp 65010 created.

00:47:40: LISP-0: Local EID IID 100 prefix 172.16.100.0/24 RLOC 192.168.0.13
pri/wei=1/1, Created locator from cfg-intf (state: active, rlocs: 1/1, sources: route-
import).

00:47:40: LISP-0: Local EID IID 100 prefix 172.16.100.0/24 RLOC 192.168.0.13
pri/wei=1/1, Added numeric RLOC (IPv4 intf RLOC Loopback0) (state: active, rlocs: 1/1,
sources: route-import).

00:47:40: LISP-0: Local EID IID 100 prefix 172.16.100.0/24 RLOC 192.168.0.13
pri/wei=1/1, Created (IPv4 intf RLOC Loopback0) (state: active, rlocs: 1/1, sources:
route-import).
```

Example 1-12: *RIB to LISP distribution process on Border-PxTR-13.*

Example 1-13 that Border-PxTR-13 has installed subnet 172.16.100.0/24 with RLOC 192.168.0.13 (its Loopback 0) into the Dynamic-EID Data Base.

```
Border-PxTR-13#show lisp instance-id 100 ipv4 database
LISP ETR IPv4 Mapping Database for EID-table vrf 100_NWKT (IID 100), LSBs: 0x1
Entries total 1, no-route 0, inactive 0

172.16.100.0/24, route-import, inherited from default locator-set RLOC-SET1
  Locator       Pri/Wgt  Source     State
  192.168.0.13    1/1    cfg-intf   site-self, reachable
```

Example 1-13: *Dynamic-EID Data Base on Border-PxTR-13.*

Phase 4: Border-PxTR – LISP Registration Process about 172.16.100.0/24

Installing routing information from BRIB to LISP Mapping Cache triggers the same registration process as happened when EP1 sends a GARP message in the previous chapter. First, Border-PxT-13 sends an unreliable Map-Register message, which is confirmed by MapSrv-22 using the Map-Notify message. After that, Border-PxTR-13 opens TCP connection and sends the reliable Map-Register message over TCP connection and so on.

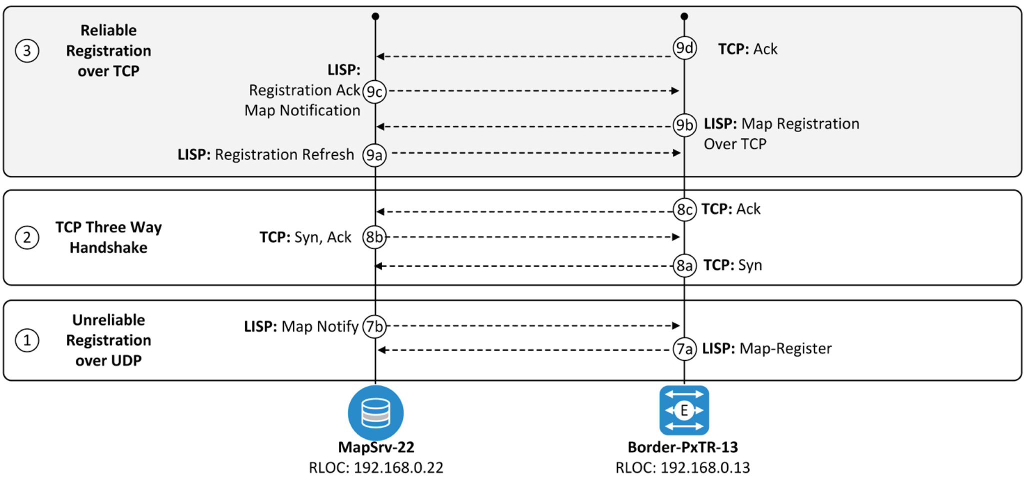


Figure 1-13: *Registration of subnet 172.16.100.0/24 by Border-PxTR-13.*

This means that route to subnet 172.16.100.0/24 is reflected to MapSrv-22. However, this update is ignored due to fact that MapSrv-22 already has a null route to subnet 172.16.100.0/24 in its RIB as can be seen example below.

```
MapSrv-22#sh ip route vrf 100_NWKT | b Gateway
Gateway of last resort is not set

      172.16.0.0/16 is variably subnetted, 3 subnets, 2 masks
B        172.16.100.0/24 [200/0], 00:16:28, Null0
l        172.16.100.10/32 [250/1], 00:16:28, Null0
l        172.16.100.12/32 [250/1], 00:05:45, Null0
```

Example 1-14: *Routing Table of MapSrv-22.*

Example 1-15 shows the process. We can see that MapSrv-22 sends a BGP update about subnet 172.16.100.0/24 to Border-PxTR-13. Then it receives Map-Register about the same route from the Border-PxTR-13. MapSrv-22 replies with Map-Notify message but it does not update its RIB because there is already the same route as the directly connected Null route. When the route is not installed into the RIB, it can't be export to the BGP process.

```
00:50:43: BGP(4): (base) 192.168.0.13 send UPDATE (format) 1:100:172.16.100.0/24, next
192.168.0.22, label 20, metric 0, path Local, extended community RT:1:100
00:50:43: LISP: Processing Map-Register proxy, map-notify, no merge, no security, no
mobile-node, not to-RTR, no fast-map-register, no EID-notify, ID-included, 1 record,
nonce 0xE8960D4F-0xA0194093, key-id 1, auth-data-len 20, hash-function sha1, xTR-ID
0xB5C974B                                                    2-0xFB2412BC-
0x533C8AEE-0x88BA11ED, site-ID unspecified
00:50:43: LISP: Processing Map-Register mapping record for IID 100 172.16.100.0/24
LCAF 2, ttl 1440, action none, authoritative, 1 locator
        192.168.0.13 pri/wei=1/1 LpR
00:50:43: LISP-0: IID 100 Parsing SVC_IP_IAF_IPv4 registration for prefix
172.16.100.0/24.
00:50:43: LISP: Processing IP Map-Register mapping record for IID 100 172.16.100.0/24
LCAF 2, ttl 1440, action none, authoritative, 1 locator
        192.168.0.13 pri/wei=1/1 LpR
00:50:43: LISP-0: IID 100 Processing SVC_IP_IAF_IPv4 registration for prefix
172.16.100.0/24.
00:50:43: LISP: Session VRF default, Local UNSPEC, Peer 192.168.0.13, Role: Passive,
State: Up, Received reliable registration message registration for IID 100  EID
172.16.100.0/24  (RX 0, TX 0).
00:50:43: LISP-0: MS registration IID 100 prefix 172.16.100.0/24 192.168.0.13
SVC_IP_IAF_IPv4 site Network-Times, Created new registration.
00:50:43: LISP-0: MS registration IID 100 prefix 172.16.100.0/24 192.168.0.13
SVC_IP_IAF_IPv4 site Network-Times, Adding locator 192.168.0.13.
00:50:43: LISP-0: MS inst member IID 100 192.168.0.13 Became valid.
00:50:43: LISP-0: MS EID IID 100 prefix 172.16.100.0/24 SVC_IP_IAF_IPv4 site Network-
Times, Scheduling map notifications for prefix 172.16.100.0/24.
00:50:43: LISP-0: MS EID IID 100 prefix 172.16.100.0/24 SVC_IP_IAF_IPv4 site Network-
Times, Scheduling reliable map notification for prefix 172.16.100.0/24 to ETR
192.168.0.13 which registered 172.16.100.0/24.
00:50:43: LISP-0: MS EID IID 100 prefix 172.16.100.0/24 SVC_IP_IAF_IPv4 site Network-
Times, MS EID prefix export route update/create.
00:50:43: LISP: RIBtable [IPv4:100_NWKT:172.16.100.0/24] created.
00:50:43: LISP: RIBtable [IPv4:100_NWKT:172.16.100.0/24] add source ms-site-reg.
00:50:43: LISP-0: MS EID IID 100 prefix 172.16.100.0/24 SVC_IP_IAF_IPv4 site Network-
Times, ALT route update/create.
00:50:43: LISP-0: ALTroute IID 100 prefix 172.16.100.0/24 <-> created.
00:50:43: LISP-0: ALTroute IID 100 prefix 172.16.100.0/24 <-> add source MS-EID.
00:50:43: LISP-0: Building reliable registration message registration-ack for IID 100
EID 172.16.100.0/24 .
00:50:43: LISP-0: Building reliable registration message mapping-notification for IID
100  EID 172.16.100.0/24 .
00:50:43: LISP-0: MS EID IID 100 prefix 172.16.100.10/32 SVC_IP_IAF_IPv4 site Network-
Times, Scheduling reliable map notification for prefix 172.16.100.10/32 to ETR
192.168.0.13 which registered 172.16.100.0/24.
00:50:43: LISP: RIBtable [IPv4:100_NWKT:172.16.100.0/24] route create nh NULL if Null0
topoid INVALID route-tag 0.
00:50:43: LISP: RIBtable [IPv4:100_NWKT:172.16.100.0/24] route create failed to stick.
00:50:43: LISP-0: ALTroute IID 100 prefix 172.16.100.0/24 <MS-EID> RIB route ignore
create, no ALT RIB.
```

Example 1-15: *LISP Debug on MapSrv-22.*

Now we have seen how the EID-to-RLOC mapping information is a) advertised by Edge-xTR-11 to MapSrv-22 by using LISP Map-Register messages, and b) how MapSrv-22 first installs the information into the RIB and from there export it to the BGP process and advertises it to Border-PxTR-13. The next chapter describes the process of how routing information is first advertised to SD-WAN edge device using BGP IPv4. Then it shows how OMP advertises the routing information to SD-WAN Control-Plane node vSmart which in turn forwards update to another SD-WAN edge vEdge2 in BGP EVPN Datacenter site.

Chapter 1 Appendix A: Device config Configuration

Edge-xTR-11

```
Edge-xTR-11#sh run
Building configuration...

Current configuration : 3143 bytes
!
! Last configuration change at 06:37:46 UTC Sun Jun 13 2021
!
version 16.6
service timestamps debug datetime msec
service timestamps log datetime msec
platform qfp utilization monitor load 80
no platform punt-keepalive disable-kernel-core
platform console serial
!
hostname Edge-xTR-11
!
boot-start-marker
boot-end-marker
!
!
vrf definition 100_NWKT
 rd 1:100
 !
 address-family ipv4
  route-target export 1:100
  route-target import 1:100
 exit-address-family
!
!
no aaa new-model
!
!
no login on-success log
!
!
!
!
subscriber templating
!
multilink bundle-name authenticated
!
!
!
crypto pki trustpoint TP-self-signed-2500303463
 enrollment selfsigned
 subject-name cn=IOS-Self-Signed-Certificate-2500303463
 revocation-check none
 rsakeypair TP-self-signed-2500303463
!
!
crypto pki certificate chain TP-self-signed-2500303463
!
!
!
license udi pid CSR1000V sn 9KUZ0RAC3VI
```

```
diagnostic bootup level minimal
spanning-tree extend system-id
!
!
redundancy
!
interface Loopback0
 ip address 192.168.0.11 255.255.255.255
 ip ospf 1 area 0
!
interface LISP0
!
interface LISP0.100
!
interface GigabitEthernet1
 ip unnumbered Loopback0
 ip lisp source-locator Loopback0
 ip ospf network point-to-point
 ip ospf 1 area 0
 negotiation auto
 medium p2p
 no mop enabled
 no mop sysid
!
interface GigabitEthernet2
 no ip address
 negotiation auto
 no mop enabled
 no mop sysid
 service instance 100 ethernet
  encapsulation untagged
  bridge-domain 100
 !
!
interface GigabitEthernet3
 no ip address
 shutdown
 negotiation auto
 no mop enabled
 no mop sysid
!
interface GigabitEthernet4
 no ip address
 shutdown
 negotiation auto
 no mop enabled
 no mop sysid
!
interface BDI100
 mac-address 0001.0001.0001
 vrf forwarding 100_NWKT
 ip address 172.16.100.1 255.255.255.0
 no ip redirects
 ip local-proxy-arp
 ip route-cache same-interface
 lisp mobility 172_16_100_0-100_NWKT
 no mop enabled
 no mop sysid
!
router lisp
 locator-table default
 locator-set RLOC-SET1
```

```
  IPv4-interface Loopback0 priority 1 weight 1
  exit-locator-set
 !
 locator default-set RLOC-SET1
 service ipv4
  encapsulation vxlan
  itr map-resolver 192.168.0.22
  itr
  etr map-server 192.168.0.22 key nwktimes
  etr map-server 192.168.0.22 proxy-reply
  etr
  exit-service-ipv4
 !
 service ethernet
  itr map-resolver 192.168.0.22
  itr
  etr map-server 192.168.0.22 key nwktimes
  etr map-server 192.168.0.22 proxy-reply
  etr
  exit-service-ethernet
 !
 instance-id 100
  remote-rloc-probe on-route-change
  dynamic-eid 172_16_100_0-100_NWKT
   database-mapping 172.16.100.0/24 locator-set RLOC-SET1
   exit-dynamic-eid
  !
  service ipv4
   eid-table vrf 100_NWKT
   map-cache 0.0.0.0/0 map-request
   use-petr 192.168.0.13
   exit-service-ipv4
  !
  exit-instance-id
 !
 exit-router-lisp
!
router ospf 1
!
virtual-service csr_mgmt
!
ip forward-protocol nd
ip http server
ip http authentication local
ip http secure-server
!
control-plane
!
line con 0
 stopbits 1
line vty 0
 login
line vty 1 2
 login
 length 0
line vty 3 4
 login
!
end

Edge-xTR-11#
```

Border-PxTR-13

```
Border-PxTR-13#sh run
Building configuration...

Current configuration : 3252 bytes
!
! Last configuration change at 06:40:03 UTC Sun Jun 13 2021
!
version 16.6
service timestamps debug datetime localtime
service timestamps log datetime msec
platform qfp utilization monitor load 80
no platform punt-keepalive disable-kernel-core
platform console serial
!
hostname Border-PxTR-13
!
boot-start-marker
boot-end-marker
!
!
vrf definition 100_NWKT
 rd 1:100
 !
 address-family ipv4
  route-target export 1:100
  route-target import 1:100
 exit-address-family
!
!
no aaa new-model
!
no login on-success log
!
subscriber templating
!
multilink bundle-name authenticated
!
crypto pki trustpoint TP-self-signed-530020971
 enrollment selfsigned
 subject-name cn=IOS-Self-Signed-Certificate-530020971
 revocation-check none
 rsakeypair TP-self-signed-530020971
!
!
crypto pki certificate chain TP-self-signed-530020971
!
!
license udi pid CSR1000V sn 9L6R8516KGG
diagnostic bootup level minimal
spanning-tree extend system-id
!
redundancy
!
ingress-class-map 2
!
interface Loopback0
 ip address 192.168.0.13 255.255.255.255
 ip ospf 1 area 0
!
interface LISP0
```

```
!
interface LISP0.100
!
interface GigabitEthernet1
 vrf forwarding 100_NWKT
 ip address 172.16.10.13 255.255.255.0
 negotiation auto
 no mop enabled
 no mop sysid
!
interface GigabitEthernet2
 ip unnumbered Loopback0
 ip ospf network point-to-point
 ip ospf 1 area 0
 negotiation auto
 medium p2p
 no mop enabled
 no mop sysid
!
interface GigabitEthernet3
 no ip address
 negotiation auto
 no mop enabled
 no mop sysid
!
interface GigabitEthernet4
 ip unnumbered Loopback0
 ip ospf 1 area 0
 negotiation auto
 medium p2p
 no mop enabled
 no mop sysid
!
router lisp
 locator-table default
 locator-set RLOC-SET1
  IPv4-interface Loopback0 priority 1 weight 1
  exit-locator-set
 !
 locator default-set RLOC-SET1
 service ipv4
  encapsulation vxlan
  itr map-resolver 192.168.0.22
  itr
  etr map-server 192.168.0.22 key nwktimes
  etr map-server 192.168.0.22 proxy-reply
  etr
  exit-service-ipv4
 !
 service ethernet
  itr map-resolver 192.168.0.22
  itr
  etr map-server 192.168.0.22 key nwktimes
  etr map-server 192.168.0.22 proxy-reply
  etr
  exit-service-ethernet
 !
 instance-id 100
  service ipv4
   eid-table vrf 100_NWKT
   map-cache 0.0.0.0/0 map-request
   route-import database bgp 65010 locator-set RLOC-SET1
```

```
   map-cache site-registration
   proxy-etr
   proxy-itr 192.168.0.13
   exit-service-ipv4
  !
  exit-instance-id
 !
 exit-router-lisp
!
router ospf 1
!
router bgp 65010
 bgp log-neighbor-changes
 neighbor 172.16.10.1 remote-as 65100
 neighbor 192.168.0.22 remote-as 65010
 neighbor 192.168.0.22 update-source Loopback0
 !
 address-family vpnv4
  neighbor 192.168.0.22 activate
  neighbor 192.168.0.22 send-community extended
 exit-address-family
 !
 address-family ipv4 vrf 100_NWKT
  neighbor 172.16.10.1 remote-as 65100
  neighbor 172.16.10.1 activate
 exit-address-family
!
!
virtual-service csr_mgmt
!
ip forward-protocol nd
ip http server
ip http authentication local
ip http secure-server
!
control-plane
!
line con 0
 stopbits 1
line vty 0
 login
line vty 1 2
 login
 length 0
line vty 3 4
 login
!
end

Border-PxTR-13#
```

MapSrv-22

```
MapSrv-22#sh run
Building configuration...

Current configuration : 3161 bytes
!
! Last configuration change at 06:37:46 UTC Sun Jun 13 2021
!
version 16.6
service timestamps debug datetime msec
service timestamps log datetime msec
platform qfp utilization monitor load 80
no platform punt-keepalive disable-kernel-core
platform console serial
!
hostname MapSrv-22
!
boot-start-marker
boot-end-marker
!
!
vrf definition 100_NWKT
 rd 1:100
 !
 address-family ipv4
  route-target export 1:100
  route-target import 1:100
 exit-address-family
!
!
no aaa new-model
!
no login on-success log
!
subscriber templating
!
multilink bundle-name authenticated
!
crypto pki trustpoint TP-self-signed-540876456
 enrollment selfsigned
 subject-name cn=IOS-Self-Signed-Certificate-540876456
 revocation-check none
 rsakeypair TP-self-signed-540876456
!
!
crypto pki certificate chain TP-self-signed-540876456
!
license udi pid CSR1000V sn 9YOPWUC9N3C
diagnostic bootup level minimal
spanning-tree extend system-id
!
redundancy
!
interface Loopback0
 ip address 192.168.0.22 255.255.255.255
 ip ospf 1 area 0
!
interface GigabitEthernet1
 ip unnumbered Loopback0
 ip lisp source-locator Loopback0
 ip ospf network point-to-point
```

```
 ip ospf 1 area 0
 negotiation auto
 medium p2p
 no mop enabled
 no mop sysid
!
interface GigabitEthernet2
 no ip address
 shutdown
 negotiation auto
 no mop enabled
 no mop sysid
!
interface GigabitEthernet3
 no ip address
 shutdown
 negotiation auto
 no mop enabled
 no mop sysid
!
interface GigabitEthernet4
 no ip address
 shutdown
 negotiation auto
 no mop enabled
 no mop sysid
!
router lisp
 locator-table default
 service ipv4
  encapsulation vxlan
  map-server
  map-resolver
  exit-service-ipv4
 !
 service ethernet
  map-server
  map-resolver
  exit-service-ethernet
 !
 instance-id 100
  service ipv4
   eid-table vrf 100_NWKT
   route-export site-registrations
   distance site-registrations 250
   exit-service-ipv4
  !
  exit-instance-id
 !
 site Network-Times
  authentication-key nwktimes
  eid-record instance-id 100 172.16.30.0/24 accept-more-specifics
  eid-record instance-id 100 172.16.100.0/24 accept-more-specifics
  exit-site
 !
 exit-router-lisp
!
router ospf 1
!
router bgp 65010
 bgp log-neighbor-changes
 neighbor 192.168.0.13 remote-as 65010
```

```
 neighbor 192.168.0.13 update-source Loopback0
 !
 address-family ipv4
  no neighbor 192.168.0.13 activate
 exit-address-family
 !
 address-family vpnv4
  neighbor 192.168.0.13 activate
  neighbor 192.168.0.13 send-community extended
 exit-address-family
 !
 address-family ipv4 vrf 100_NWKT
  aggregate-address 172.16.100.0 255.255.255.0 summary-only
  redistribute lisp metric 10 route-map LISP-to-BGP
 exit-address-family
!
!
virtual-service csr_mgmt
!
ip forward-protocol nd
ip http server
ip http authentication local
ip http secure-server
!
!
!
ip prefix-list LISP-to-BGP seq 5 permit 172.16.100.0/24 ge 32
!
!
route-map LISP-to-BGP permit 10
 match ip address prefix-list LISP-to-BGP
!
control-plane
!
line con 0
 stopbits 1
line vty 0
 login
line vty 1 2
 login
 length 0
line vty 3 4
 login
!
end

MapSrv-22#
```

Chapter 1 Appendix B: BGP and LISP Debugs

MapSrv-22

```
00:50:43: LISP: Processing received Map-Register(3) message on GigabitEthernet1 from
192.168.0.11:4342 to 192.168.0.22:4342
00:50:43: LISP: Processing Map-Register proxy, map-notify, no merge, no security, no
mobile-node, not to-RTR, no fast-map-register, no EID-notify, ID-included, 1 record,
nonce 0xE297E49D-0xEDA98C98, key-id 1, auth-data-len 20, hash-function sha1, xTR-ID
0xB8F66E3                                                        B-0x101CCD7C-
0x5B9F516E-0x3526FC6E, site-ID unspecified
00:50:43: LISP: Processing Map-Register mapping record for IID 100 172.16.100.10/32
LCAF 2, ttl 1440, action none, authoritative, 1 locator
        192.168.0.11 pri/wei=1/1 LpR
00:50:43: LISP-0: IID 100 Parsing SVC_IP_IAF_IPv4 registration for prefix
172.16.100.10/32.
00:50:43: LISP: Processing IP Map-Register mapping record for IID 100 172.16.100.10/32
LCAF 2, ttl 1440, action none, authoritative, 1 locator
        192.168.0.11 pri/wei=1/1 LpR
00:50:43: LISP-0: IID 100 Processing SVC_IP_IAF_IPv4 registration for prefix
172.16.100.10/32.
00:50:43: LISP-0: MS EID IID 100 prefix 172.16.100.10/32 SVC_IP_IAF_IPv4 site *,
Created.
00:50:43: LISP-0: MS Site EID IID 100 prefix 172.16.100.10/32 SVC_IP_IAF_IPv4 site
Network-Times, Created dynamic site EID prefix entry.
00:50:43: LISP-0: MS registration IID 100 prefix 172.16.100.10/32 192.168.0.11
SVC_IP_IAF_IPv4 site Network-Times, Created new registration.
00:50:43: LISP-0: MS registration IID 100 prefix 172.16.100.10/32 192.168.0.11
SVC_IP_IAF_IPv4 site Network-Times, Adding locator 192.168.0.11.
00:50:43: LISP: RIB Watch Group default 192.168.0.11/32 , created.
00:50:43: LISP: RIB Watch Group default 192.168.0.11/32 , scheduling RIB update.
00:50:43: LISP-0: MS inst member IID 100 192.168.0.11 Became valid.
00:50:43: LISP: Session VRF default, Local UNSPEC, Peer 192.168.0.11, Role: Passive,
State: Unknown, Created (RX 0, TX 0).
00:50:43: LISP: Session VRF default, Local UNSPEC, Peer 192.168.0.11, Role: Passive,
State: Unknown, Scheduled update (RX 0, TX 0).
00:50:43: LISP-0: MS EID IID 100 prefix 172.16.100.10/32 SVC_IP_IAF_IPv4 site Network-
Times, Scheduling map notifications for prefix 172.16.100.10/32.
00:50:43: LISP-0: MS EID IID 100 prefix 172.16.100.10/32 SVC_IP_IAF_IPv4 site Network-
Times, Scheduling unreliable map notification for prefix 172.16.100.10/32 to ETR
192.168.0.11 which registered 172.16.100.10/32.
00:50:43: LISP: RIB Watch Group default 192.168.0.11/32 , installing in RIB.
00:50:43: LISP-0: MS EID IID 100 prefix 172.16.100.10/32 SVC_IP_IAF_IPv4 site Network-
Times, ALT route update/create.
00:50:43: LISP-0: ALTroute IID 100 prefix 172.16.100.10/32 <-> created.
00:50:43: LISP-0: ALTroute IID 100 prefix 172.16.100.10/32 <-> add source MS-EID.
00:50:43: LISP: Session VRF default, Local UNSPEC, Peer 192.168.0.11, Role: Passive,
State: Unknown, Destroyed with OS adaptor (RX 0, TX 0).
00:50:43: LISP-0: Map-Notify 192.168.0.22:4342->192.168.0.11:4342 xTR-ID 0xB8F66E3B-
0x101CCD7C-0x5B9F516E-0x3526FC6E, sending with 1 prefix, nonce 0xE297E49D-0xEDA98C98
00:50:43: LISP-0: ALTroute IID 100 prefix 172.16.100.10/32 <MS-EID> RIB route ignore
create, no ALT RIB.
00:50:43: LISP: Session VRF default, Local UNSPEC, Peer 192.168.0.11, Role: Passive,
State: Unknown, Peer is reachable (RX 0, TX 0).
00:50:43: LISP: Session VRF default, Local UNSPEC, Peer 192.168.0.11, Role: Passive,
State: Unknown, Scheduled update (RX 0, TX 0).
00:50:43: LISP-0: MS EID IID 100 prefix 172.16.100.10/32 SVC_IP_IAF_IPv4 site Network-
Times, MS EID prefix export route update/create.
00:50:43: LISP: RIBtable [IPv4:100_NWKT:172.16.100.10/32] created.
00:50:43: LISP: RIBtable [IPv4:100_NWKT:172.16.100.10/32] add source ms-site-reg.
```

```
00:50:43: LISP-0: MS EID IID 100 prefix 172.16.100.10/32 SVC_IP_IAF_IPv4 site Network-
Times, ALT route update/create.
00:50:43: LISP-0: ALTroute IID 100 prefix 172.16.100.10/32 <MS-EID> update source MS-
EID.
00:50:43: LISP: Session VRF default, Local UNSPEC, Peer 192.168.0.11, Role: Passive,
State: Init, Created with OS adaptor (RX 0, TX 0).
00:50:43: LISP: RIBtable [IPv4:100_NWKT:172.16.100.10/32] route create nh NULL if
Null0 topoid INVALID route-tag 0.
00:50:43: BGP(4): redist event (1) request for 1:100:172.16.100.10/32
00:50:43: BGP(4) route 1:100:172.16.100.10/32 gw-1 0.0.0.0 src_proto (lisp) path-limit
1
00:50:43: BGP(4): route 1:100:172.16.100.10/32 up
00:50:43: BGP(4): sourced route for 1:100:172.16.100.10/32 created
00:50:43: BGP(4): sourced route for 1:100:172.16.100.10/32 path 0x7FDB7F0E5348 id 0 gw
0.0.0.0 created (weight 32768)
00:50:43: BGP(4): redistributed route 1:100:172.16.100.10/32 added gw 0.0.0.0
00:50:43: BGP(4): created aggregate route for 1:100:172.16.100.0/24
00:50:43: BGP(4): Revise route installing 1 of 1 routes for 172.16.100.0/24 ->
0.0.0.0(100_NWKT) to 100_NWKT IP table
00:50:43: BGP: topo 100_NWKT:VPNv4 Unicast:base Remove_fwdroute for
1:100:172.16.100.10/32
00:50:43: BGP(4): 192.168.0.13 NEXT_HOP is set to self for net 1:100:172.16.100.0/24,
00:50:43: BGP(4): (base) 192.168.0.13 send UPDATE (format) 1:100:172.16.100.0/24, next
192.168.0.22, label 20, metric 0, path Local, extended community RT:1:100
00:50:43: LISP: Processing Map-Register proxy, map-notify, no merge, no security, no
mobile-node, not to-RTR, no fast-map-register, no EID-notify, ID-included, 1 record,
nonce 0xE8960D4F-0xA0194093, key-id 1, auth-data-len 20, hash-function sha1, xTR-ID
0xB5C974B                                                  2-0xFB2412BC-
0x533C8AEE-0x88BA11ED, site-ID unspecified
00:50:43: LISP: Processing Map-Register mapping record for IID 100 172.16.100.0/24
LCAF 2, ttl 1440, action none, authoritative, 1 locator
        192.168.0.13 pri/wei=1/1 LpR
00:50:43: LISP-0: IID 100 Parsing SVC_IP_IAF_IPv4 registration for prefix
172.16.100.0/24.
00:50:43: LISP: Processing IP Map-Register mapping record for IID 100 172.16.100.0/24
LCAF 2, ttl 1440, action none, authoritative, 1 locator
        192.168.0.13 pri/wei=1/1 LpR
00:50:43: LISP-0: IID 100 Processing SVC_IP_IAF_IPv4 registration for prefix
172.16.100.0/24.
00:50:43: LISP: Session VRF default, Local UNSPEC, Peer 192.168.0.13, Role: Passive,
State: Up, Received reliable registration message registration for IID 100  EID
172.16.100.0/24  (RX 0, TX 0).
00:50:43: LISP-0: MS registration IID 100 prefix 172.16.100.0/24 192.168.0.13
SVC_IP_IAF_IPv4 site Network-Times, Created new registration.
00:50:43: LISP-0: MS registration IID 100 prefix 172.16.100.0/24 192.168.0.13
SVC_IP_IAF_IPv4 site Network-Times, Adding locator 192.168.0.13.
00:50:43: LISP-0: MS inst member IID 100 192.168.0.13 Became valid.
00:50:43: LISP-0: MS EID IID 100 prefix 172.16.100.0/24 SVC_IP_IAF_IPv4 site Network-
Times, Scheduling map notifications for prefix 172.16.100.0/24.
00:50:43: LISP-0: MS EID IID 100 prefix 172.16.100.0/24 SVC_IP_IAF_IPv4 site Network-
Times, Scheduling reliable map notification for prefix 172.16.100.0/24 to ETR
192.168.0.13 which registered 172.16.100.0/24.
00:50:43: LISP-0: MS EID IID 100 prefix 172.16.100.0/24 SVC_IP_IAF_IPv4 site Network-
Times, MS EID prefix export route update/create.
00:50:43: LISP: RIBtable [IPv4:100_NWKT:172.16.100.0/24] created.
00:50:43: LISP: RIBtable [IPv4:100_NWKT:172.16.100.0/24] add source ms-site-reg.
00:50:43: LISP-0: MS EID IID 100 prefix 172.16.100.0/24 SVC_IP_IAF_IPv4 site Network-
Times, ALT route update/create.
00:50:43: LISP-0: ALTroute IID 100 prefix 172.16.100.0/24 <-> created.
00:50:43: LISP-0: ALTroute IID 100 prefix 172.16.100.0/24 <-> add source MS-EID.
00:50:43: LISP-0: Building reliable registration message registration-ack for IID 100
EID 172.16.100.0/24 .
```

```
00:50:43: LISP-0: Building reliable registration message mapping-notification for IID
100  EID 172.16.100.0/24 .
00:50:43: LISP-0: MS EID IID 100 prefix 172.16.100.10/32 SVC_IP_IAF_IPv4 site Network-
Times, Scheduling reliable map notification for prefix 172.16.100.10/32 to ETR
192.168.0.13 which registered 172.16.100.0/24.
00:50:43: LISP: RIBtable [IPv4:100_NWKT:172.16.100.0/24] route create nh NULL if Null0
topoid INVALID route-tag 0.
00:50:43: LISP: RIBtable [IPv4:100_NWKT:172.16.100.0/24] route create failed to stick.
00:50:43: LISP-0: ALTroute IID 100 prefix 172.16.100.0/24 <MS-EID> RIB route ignore
create, no ALT RIB.
00:50:43: LISP-0: Building reliable registration message mapping-notification for IID
100  EID 172.16.100.10/32 .
00:50:43: LISP: Processing received Encap-Control(8) message on GigabitEthernet1 from
192.168.0.13:4342 to 192.168.0.22:4342
00:50:43: LISP: Processing received Map-Request(1) message on GigabitEthernet1 from
172.16.100.10:4342 to 172.16.100.10:4342
00:50:43: LISP: Received map request for IID 100 172.16.100.10/32, source_eid IID 100
172.16.100.0, ITR-RLOCs: 192.168.0.13, records 1, nonce 0x6F9CFA26-0x427F4D1C,
FromPITR
00:50:43: LISP-0: MS EID IID 100 prefix 172.16.100.10/32 SVC_IP_IAF_IPv4 site Network-
Times, Sending proxy reply to 192.168.0.13.
00:50:47: LISP: OS TCP Session VRF: IPv4 default, Peer: 192.168.0.11:35400, Local:
UNSPEC, TCB: 0x7FDB7F12A878 Accepted connection, local address 192.168.0.22, remote
port 35400 (State: Init).
00:50:47: LISP-0: Session User Type Error-Reporter/3 Peer 192.168.0.11 Role Passive
State Down , IPv4 RLOC default, Init.
00:50:47: LISP-0: Building reliable registration message registration-refresh for IID
0  EID 0.0.0.0/0 , Refresh Rejected: FALSE, Scope: global/0, EID AFI: invalid/0.
00:50:47: LISP: Session VRF default, Local UNSPEC, Peer 192.168.0.11, Role: Passive,
State: Up, Received reliable registration message wlc subscribe for IID 100  EID
0.0.0.0/0 , EID AFI: IPv4/1 (RX 0, TX 0).
00:50:47: LISP-0: MS rr client IID 100 192.168.0.11 Created (AF: SVC_IP_IAF_IPv4,
state: idle, session: Down).
00:50:47: LISP-0: Session User Type reliable-registration-ms/4 Peer 192.168.0.11 Role
Passive State Down MS 0 IID 100 AF IPv4 reliable registration client 192.168.0.11,
Init.
00:50:47: LISP-0: MS WLC client IID 100 AFI IPv4 Peer 192.168.0.11, Created.
00:50:47: LISP-0: MS WLC client IID 100 AFI IPv4 Peer 192.168.0.11, Scheduling full
update.
00:50:47: LISP: Processing Map-Register proxy, map-notify, no merge, no security, no
mobile-node, not to-RTR, no fast-map-register, no EID-notify, ID-included, 1 record,
nonce 0xE297E49D-0xEDA98C98, key-id 1, auth-data-len 20, hash-function sha1, xTR-ID
0xB8F66E3                                                         B-0x101CCD7C-
0x5B9F516E-0x3526FC6E, site-ID unspecified
00:50:47: LISP: Processing Map-Register mapping record for IID 100 172.16.100.10/32
LCAF 2, ttl 1440, action none, authoritative, 1 locator
        192.168.0.11 pri/wei=1/1 LpR
00:50:47: LISP-0: IID 100 Parsing SVC_IP_IAF_IPv4 registration for prefix
172.16.100.10/32.
00:50:47: LISP: Processing IP Map-Register mapping record for IID 100 172.16.100.10/32
LCAF 2, ttl 1440, action none, authoritative, 1 locator
        192.168.0.11 pri/wei=1/1 LpR
00:50:47: LISP-0: IID 100 Processing SVC_IP_IAF_IPv4 registration for prefix
172.16.100.10/32.
00:50:47: LISP: Session VRF default, Local UNSPEC, Peer 192.168.0.11, Role: Passive,
State: Up, Received reliable registration message registration for IID 100  EID
172.16.100.10/32  (RX 0, TX 0).
00:50:47: LISP-0: MS registration IID 100 prefix 172.16.100.10/32 192.168.0.11
SVC_IP_IAF_IPv4 site Network-Times, Found existing registration, updating.
00:50:47: LISP-0: MS EID IID 100 prefix 172.16.100.10/32 SVC_IP_IAF_IPv4 site Network-
Times, Scheduling map notifications for prefix 172.16.100.10/32.
```

```
00:50:47: LISP-0: MS EID IID 100 prefix 172.16.100.10/32 SVC_IP_IAF_IPv4 site Network-
Times, Scheduling reliable map notification for prefix 172.16.100.10/32 to ETR
192.168.0.11 which registered 172.16.100.10/32.
00:50:47: LISP-0: Building reliable registration message registration-ack for IID 100
EID 172.16.100.10/32 .
00:50:47: LISP-0: Building reliable registration message mapping-notification for IID
100  EID 172.16.100.10/32 .
00:50:47: LISP-0: MS EID IID 100 prefix 172.16.100.10/32 SVC_IP_IAF_IPv4 site Network-
Times, Scheduling reliable map notification for prefix 172.16.100.10/32 to ETR
192.168.0.13 which registered 172.16.100.0/24.
00:50:47: LISP-0: Building reliable registration message mapping-notification for IID
100  EID 172.16.100.10/32 .
00:50:47: LISP-0: Session User Type Error-Reporter/3 Peer 192.168.0.11 Role Passive
State Down , IPv4 RLOC default, Deleting.
*Jun 18 08:36:54.435: %PLATFORM-4-ELEMENT_WARNING:  F0: smand:  RP/0: Used Memory
value 92% exceeds warning level 88%
```

Border-PxTR-22

```
00:47:40: BGP(4): 192.168.0.22 rcvd UPDATE w/ attr: nexthop 192.168.0.22, origin i,
localpref 100, metric 0, atomic-aggregate, aggregated by 65010 192.168.0.22, extended
community RT:1:100
00:47:40: BGP(4): 192.168.0.22 rcvd 1:100:172.16.100.0/24, label 20
00:47:40: BGP(4): Revise route installing 1 of 1 routes for 172.16.100.0/24 ->
192.168.0.22(100_NWKT) to 100_NWKT IP table
00:47:40: LISP-0: RIB Redist Spec IID 100 database bgp 65010, RIB event
172.16.100.0/24 UP.
00:47:40: LISP-0: RtImp IID 100 prefix 172.16.100.0/24 database bgp 65010 created.
00:47:40: LISP-0: RtImp IID 100 prefix 172.16.100.0/24 database bgp 65010 updating
cache (up).
00:47:40: LISP-0: Local EID IID 100 prefix 172.16.100.0/24, Setting state to active
(state: inactive, rlocs: 0/0, sources: NONE).
00:47:40: LISP-0: Local EID IID 100 prefix 172.16.100.0/24, found RIB route
172.16.100.0/24 to EID prefix (state: active, rlocs: 0/0, sources: NONE).
00:47:40: LISP-0: IAF IID 100 SVC_IP_IAF_IPv4, found route to local EID prefixes,
switching local RLOCs to reachable.
00:47:40: LISP: RIB Watch Group 100_NWKT 172.16.100.0/24 (no-default), created.
00:47:40: LISP: RIB Watch Group 100_NWKT 172.16.100.0/24 (no-default), scheduling RIB
update.
00:47:40: LISP-0: Local EID IID 100 prefix 172.16.100.0/24, Added source route-import
(state: active, rlocs: 0/0, sources: route-import).
00:47:40: LISP-0: Local EID IID 100 prefix 172.16.100.0/24, Best source invalid ->
route-import (state: active, rlocs: 0/0, sources: route-import).
00:47:40: LISP-0: IAF IID 100 SVC_IP_IAF_IPv4, Recalculating locator status bits from
0x0 to 0x1.
00:47:40: LISP-0: Local EID IID 100 prefix 172.16.100.0/24 RLOC 192.168.0.13
pri/wei=1/1, Created locator from cfg-intf (state: active, rlocs: 1/1, sources: route-
import).
00:47:40: LISP-0: Local EID IID 100 prefix 172.16.100.0/24 RLOC 192.168.0.13
pri/wei=1/1, Added numeric RLOC (IPv4 intf RLOC Loopback0) (state: active, rlocs: 1/1,
sources: route-import).
00:47:40: LISP-0: Local EID IID 100 prefix 172.16.100.0/24 RLOC 192.168.0.13
pri/wei=1/1, Created (IPv4 intf RLOC Loopback0) (state: active, rlocs: 1/1, sources:
route-import).
00:47:40: LISP-0: IPv4 Map Server IID 100 192.168.0.22, Built reliable map register
message. Size = 112.
00:47:40: LISP-0: Remote EID IID 100 prefix 172.16.100.0/24, Covered by local EID
prefix entry (sources: <map-rep|away>, state: forward-native, rlocs: 0, local).
```

```
00:47:40: LISP-0: IAF IID 100 SVC_IP_IAF_IPv4, Processed 1 SMR table prefix fwd
update.
00:47:40: LISP-0: Away Table entry IID 100 prefix 172.16.100.0/24, Deleted.
00:47:40: LISP-0: Remote EID IID 100 prefix 172.16.100.0/24, Change state to deleted
(sources: <>, state: forward-native, rlocs: 0, local).
00:47:40: LISP-0: IAF IID 100 SVC_IP_IAF_IPv4, Reevaluated 1 local EID away entries.
00:47:40: LISP: RIB Watch Group 100_NWKT 172.16.100.0/24 (no-default), installing in
RIB.
00:47:40: LISP-0: IAF IID 100 SVC_IP_IAF_IPv4, Reevaluated 0 more specific local EID
away entries.
00:47:40: LISP-0: IAF IID 100 SVC_IP_IAF_IPv4, Processed 1 SMR table prefix fwd
update.
00:47:40: LISP-0: Local EID IID 100 prefix 172.16.100.0/24, found RIB route
172.16.100.0/24 to EID prefix (state: active, rlocs: 1/1, sources: route-import).
00:47:40: LISP: Session VRF default, Local 192.168.0.13, Peer 192.168.0.22, Role:
Active, State: Up, Received reliable registration message registration-ack for IID 100
EID 172.16.100.0/24  (RX 0, TX 0).
00:47:40: LISP-0: IPv4 Map Server IID 100 192.168.0.22, Processing reliable ACK
message for prefix 172.16.100.0/24.
00:47:40: LISP: Session VRF default, Local 192.168.0.13, Peer 192.168.0.22, Role:
Active, State: Up, Received reliable registration message mapping-notification for IID
100  EID 172.16.100.0/24  (RX 0, TX 0).
00:47:40: LISP-0: Local EID IID 100 prefix 172.16.100.0/24, Received map notify
(state: active, rlocs: 1/1, sources: route-import).
00:47:40: LISP-0: IPv4 Map Server IID 100 192.168.0.22, Processing reliable NOTIFY
message for prefix 172.16.100.0/24.
00:47:40: LISP: Session VRF default, Local 192.168.0.13, Peer 192.168.0.22, Role:
Active, State: Up, Received reliable registration message mapping-notification for IID
100  EID 172.16.100.10/32  (RX 0, TX 0).
00:47:40: LISP-0: Away Table entry IID 100 prefix 172.16.100.10/32, Created.
00:47:40: LISP-0: SMR EID prefix IID 100 prefix 172.16.100.10/32, Created.
00:47:40: LISP-0: Remote EID IID 100 prefix 172.16.100.10/32, Change state to send-
map-request (sources: <away>, state: unknown, rlocs: 0, local).
00:47:40: LISP-0: Remote EID IID 100 prefix 172.16.100.10/32, Starting idle timer
(delay 00:02:30) (sources: <away>, state: send-map-request, rlocs: 0, local).
00:47:40: LISP-0: Remote EID IID 100 prefix 172.16.100.10/32, Change state to reused
(sources: <invalid|away>, state: send-map-request, rlocs: 0, local).
00:47:40: LISP-0: Remote EID IID 100 prefix 172.16.100.10/32, Change state to
incomplete (sources: <away|away>, state: reused, rlocs: 0, local).
00:47:40: LISP-0: Remote EID IID 100 prefix 172.16.100.10/32, [incomplete] Scheduling
map requests delay 00:00:00 min_elapsed 00:00:01 (sources: <away|away>, state:
incomplete, rlocs: 0, local).
00:47:40: LISP-0: IPv4 Map Server IID 100 192.168.0.22, Processing reliable NOTIFY
message for prefix 172.16.100.10/32.
00:47:40: LISP-0: IAF IID 100 SVC_IP_IAF_IPv4, Processed 1 SMR table prefix fwd
update.
00:47:40: LISP-0: Remote EID IID 100 prefix 172.16.100.10/32, Extranet IID passed to
CEF is 16777216 (sources: <away|away>, state: incomplete, rlocs: 0, local).
00:47:41: LISP-0: IID 100 Request processing of remote EID prefix map requests to
IPv4.
00:47:41: LISP: Send map request type remote EID prefix
00:47:41: LISP: Send map request for EID prefix IID 100 172.16.100.10/32
00:47:41: LISP-0: Remote EID IID 100 prefix 172.16.100.10/32, Send map request (1)
(sources: <away|away>, state: incomplete, rlocs: 0, local).
00:47:41: LISP-0: EID-AF IPv4, Sending map-request from 172.16.100.10 to 172.16.100.10
for EID 172.16.100.10/32, ITR-RLOCs 1, nonce 0x6F9CFA26-0x427F4D1C (encap src
192.168.0.13, dst 192.168.0.22), FromPITR.
00:47:41: LISP: Processing received Map-Reply(2) message on GigabitEthernet2 from
192.168.0.22:4342 to 192.168.0.13:4342
00:47:41: LISP: Received map reply nonce 0x6F9CFA26-0x427F4D1C, records 1
00:47:41: LISP: Processing Map-Reply mapping record for IID 100 172.16.100.10/32 LCAF
2, ttl 1440, action none, not authoritative, 1 locator
```

```
        192.168.0.11 pri/wei=1/1 lpR
00:47:41: LISP-0: Map Request IID 100 prefix 172.16.100.10/32 remote EID prefix[LL],
Received reply with rtt 3ms.
00:47:41: LISP: Processing mapping information for EID prefix IID 100 172.16.100.10/32
00:47:41: LISP-0: Remote EID IID 100 prefix 172.16.100.10/32, Change state to reused
(sources: <map-rep|away>, state: incomplete, rlocs: 0, local).
00:47:41: LISP-0: IAF IID 100 SVC_IP_IAF_IPv4, Persistent db: ignore writing request,
disabled.
00:47:41: LISP-0: Remote EID IID 100 prefix 172.16.100.10/32, Change state to complete
(sources: <map-rep|away>, state: reused, rlocs: 0, local).
00:47:41: LISP: RIB Watch Group default 192.168.0.11/32 , created.
00:47:41: LISP: RIB Watch Group default 192.168.0.11/32 , scheduling RIB update.
00:47:41: LISP-0: Remote EID IID 100 prefix 172.16.100.10/32, RLOCs pending rwatch
update, defer fwd update (sources: <map-rep|away>, state: complete, rlocs: 0, local).
00:47:41: LISP-0: Remote EID IID 100 prefix 172.16.100.10/32, 1 RLOCs pending rwatch
update, defer fwd update (sources: <map-rep|away>, state: complete, rlocs: 0, local).
00:47:41: LISP-0: Remote EID IID 100 prefix 172.16.100.10/32, Recalculated RLOC status
bits from 0x0 to 0x1 (sources: <map-rep|away>, state: complete, rlocs: 1, local).
00:47:41: LISP-0: Remote EID IID 100 prefix 172.16.100.10/32, 1 RLOCs pending rwatch
update, defer fwd update (sources: <map-rep|away>, state: complete, rlocs: 1, local).
00:47:41: LISP: RIB Watch Group default 192.168.0.11/32 , installing in RIB.
00:47:41: LISP-0: Remote shrRLOC 192.168.0.11, Reachability notification, up* allow*
remote.
00:47:41: LISP-0: Remote EID IID 100 prefix 172.16.100.10/32, No more RLOCs pending
rwatch update, schedule deferred fwd update (sources: <map-rep|away>, state: complete,
rlocs: 1, local).
00:47:41: LISP-0: Remote EID IID 100 prefix 172.16.100.10/32, Extranet IID passed to
CEF is 16777216 (sources: <map-rep|away>, state: complete, rlocs: 1, local).
00:47:41: LISP-0: IID 100 Request processing of remote EID RLOC map requests to IPv4.
00:47:41: LISP: Send map request type remote EID RLOC
00:47:41: LISP: Send map request for EID prefix IID 100 172.16.100.10/32
00:47:41: LISP-0: Remote EID IID 100 prefix 172.16.100.10/32 [RRLOCset 0x7F6833D84C78
1/1] 192.168.0.11 pri/wei=1/1, Send RLOC probe (sources: <map-rep|away>, state:
complete, rlocs: 1, local).
00:47:41: LISP-0: EID-AF IPv4, Sending probe map-request from 192.168.0.13 to
192.168.0.11 for EID 172.16.100.10/32, ITR-RLOCs 1, nonce 0x1793BCB0-0xDF340702,
FromPITR.
00:47:41: LISP: Processing received Map-Reply(2) message on GigabitEthernet2 from
192.168.0.11:4342 to 192.168.0.13:4342
00:47:41: LISP: Received map reply nonce 0x1793BCB0-0xDF340702, records 1, probe
00:47:41: LISP: Processing Map-Reply mapping record for IID 100 172.16.100.10/32 LCAF
2, ttl 1440, action none, authoritative, 1 locator
        192.168.0.11 pri/wei=1/1 LPR
00:47:41: LISP-0: Map Request IID 100 prefix 172.16.100.10/32 192.168.0.11 remote EID
RLOC[MM], Received reply with rtt 2ms.
00:47:44: LISP: Session VRF default, Local 192.168.0.13, Peer 192.168.0.22, Role:
Active, State: Up, Received reliable registration message mapping-notification for IID
100  EID 172.16.100.10/32  (RX 0, TX 0).
00:47:44: LISP-0: IPv4 Map Server IID 100 192.168.0.22, Processing reliable NOTIFY
message for prefix 172.16.100.10/32.
```

Chapter 2: SD-WAN Control-Plane: OMP

Introduction

Figure 2-1 shows the Control Plane operation when host EP1 using IP 172.16.100.10/32 joins the network. In the previous chapter, we saw how Edge-xTR-11 learned the IP address and registered it to MapSrv-22 by using the LISP Map-register Message where the Instance-Id 100 represents Virtual Network-Id (VN-Id). MapSrv-22, in turn, advertised the NLRI to Border-PxTR-13 as BGP VPNv4 Update message where extended community RT 1:100 (=VN-Id) is used as a kind of VPN identifier (BGP VPNv4 route import/export policy is based on RT value). This chapter first explains how Border-PxTR-13 sends BGP IPv4 Update message to local SD-WAN device vEdge-1. The eBGP peering between Border-PxTR-13 and vEdge1 is VRF based and BGP updates over it don't carry any VN-Id. vEdge-1 imports the routing information from BRIB to RIB. Then it advertises the routing information by using OMP (Overlay Management Protocol) to the SD-WAN centralized Control Plane vSmart over the DTLS tunnel using System-IP as an originator-Id and VPN label 1003 as a VN-Id. vSmart forwards this update to SD-WAN device vEdge-2 located in Datacenter. vEdge-2 installs the route into the routing table. After that, it exports the routing information from the RIB to the BGP process and sends the BGP IPv4 Update message to Border-Leaf-13 over VRF NWKT eBGP peering without VN-Id.

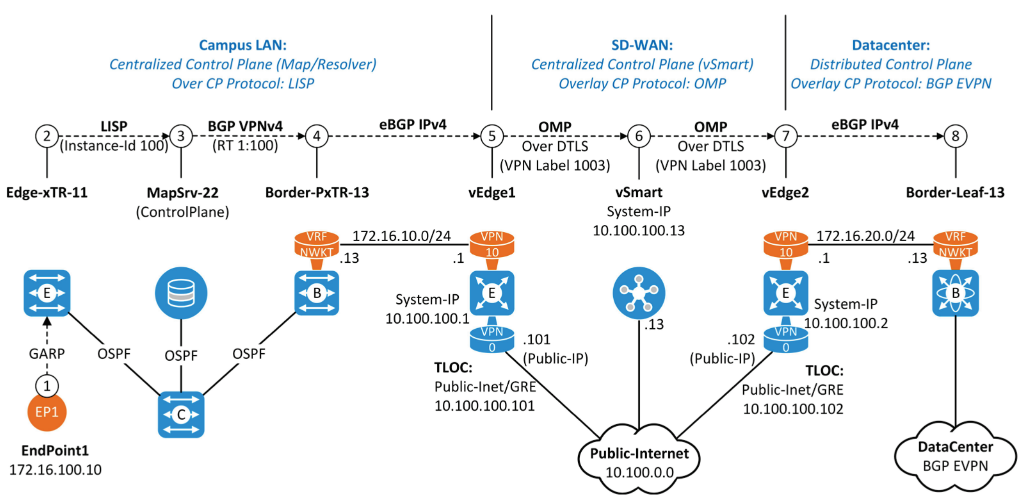


Figure 2-1: *Overall Control-Plane Operation.*

BRIB, RIB, and OMP on Local Edge

When vEdge-1 receives the BGP update about 172.16.100.0/24, it first imports the route into the VPN 10 BGP table and from there to VPN 10 RIB. The next-hop for the route is 172.16.10.13 (Border-PxTR-13). From VPN 10 RIB routing information is exported to the OMP database (Overlay Management Protocol) with Label 1003 that is used as VPN-Id. The OMP table includes several entry-specific attributes, just like BGP. Originator, Site-Id, and Origin Protocol attributes are quite self-explanatory. The Transport Locator (TLOC) describes, as the name implies, the transport network along which you can use to get to the destination network. We only have one transport network, Public-Internet, in our example but in real life, there will be more than just one transport network. TLOC also describes the tunneling mechanism which the originator uses with this particular transport network as well as the originator's System Identifier (Sys-ID).

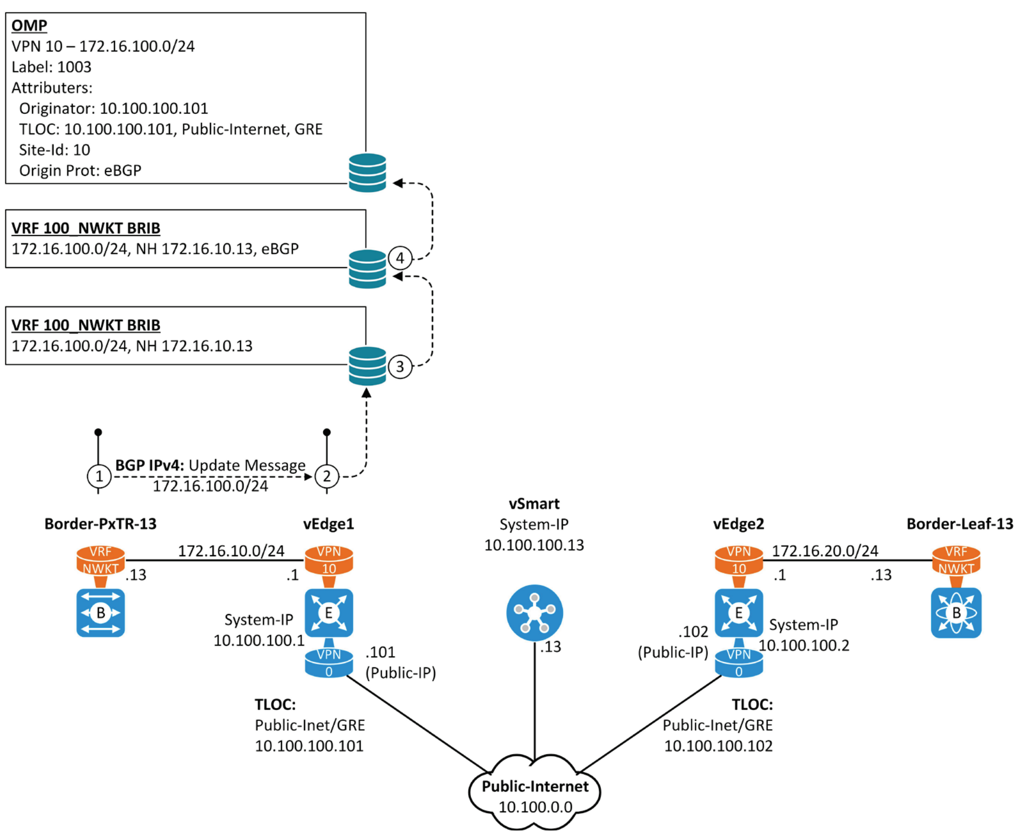


Figure 2-2: *vEdge Local Processes - Import Route into BRIB, RIB, and OMP.*

Capture 2-1 shows the BGP Update sent by Border-PxTR-13. Notice that the aggregate origin is the LISP Control-Plane Node MapSrv-22.

```
Internet Protocol Version 4, Src: 172.16.10.13, Dst: 172.16.10.1
Transmission Control Protocol, Src Port: 179, Dst Port: 43025, Seq: 1, Ack:
1, Len: 61
Border Gateway Protocol - UPDATE Message
    Marker: ffffffffffffffffffffffffffffffff
    Length: 61
    Type: UPDATE Message (2)
    Withdrawn Routes Length: 0
    Total Path Attribute Length: 34
    Path attributes
        Path Attribute - ORIGIN: IGP
        Path Attribute - AS_PATH: 65010
        Path Attribute - AGGREGATOR: AS: 65010 origin: 192.168.0.22
        Path Attribute - ATOMIC_AGGREGATE
        Path Attribute - NEXT_HOP: 172.16.10.13
    Network Layer Reachability Information (NLRI)
        172.16.100.0/24
            NLRI prefix length: 24
            NLRI prefix: 172.16.100.0
```

Capture 2-1: *BGP Update from Border-PxTR-13 to vEdge1.*

Example 2-1 shows the routing information about network 172.16.100.0/24 in the VPN 10 RIB of vEdge-1. The route is selected as the best route and installed into FIB (Forwarding Information Base - Data-Plane) with the next-hop 172.16.10.13 reachable via the interface ge0/2.

```
vEdge-1# show ip routes bgp detail
Codes Proto-sub-type:
  IA -> ospf-intra-area, IE -> ospf-inter-area,
  E1 -> ospf-external1, E2 -> ospf-external2,
  N1 -> ospf-nssa-external1, N2 -> ospf-nssa-external2,
  e -> bgp-external, i -> bgp-internal
Codes Status flags:
  F -> fib, S -> selected, I -> inactive,
  B -> blackhole, R -> recursive, L -> import
""-------------------------------------------
 VPN 10      PREFIX 172.16.100.0/24
--------------------------------------------
 proto            bgp
 proto-sub-type   e
 distance         20
 metric           0
 uptime           0:00:06:45
 nexthop-ifname   ge0/2
 nexthop-addr     172.16.10.13
 status           F,S
```

Example 2-1: *VPN 10 RIB Entry about Network 172.16.100.0/24.*

Example 2-2 shows the OMP table of vEdge-1about 172.16.100.0/24. The route is redistributed to OMP and that is why the peer listed under *received from* section is 0.0.0.0 and the origin attribute is the system-id of vEdge-1. vEdge-1 advertises the information to SD-WAN Control-Plane node vSmart using Label 1003 as VN-Id.

```
vEdge-1# show omp routes received detail

---------------------------------------------------
omp route entries for vpn 10 route 172.16.100.0/24
---------------------------------------------------
            RECEIVED FROM:
peer            0.0.0.0
path-id         37
label           1003
status          C,Red,R
    Attributes:
     originator       10.100.100.101
     type             installed
     tloc             10.100.100.101, public-internet, gre
     overlay-id        1
     site-id          10
     origin-proto     eBGP
     origin-metric    0

            ADVERTISED TO:
peer    10.100.100.13
    Attributes:
     originator       10.100.100.101
     label            1003
     path-id          37
     tloc             10.100.100.101, public-internet, gre
     site-id          10
     overlay-id        1
     origin-proto     eBGP
     origin-metric    0
```

Example 2-2: *OMP Routing Information about Network 172.16.100.0/24.*

Setting Up Data-Path - OMP TLOC Route

Before explaining how the routing information about network 172.16.100.0/24 is propagated to remote vEdge-2 over SD-WAN I will shortly introduce the OMP TLOC routes. In order to send traffic to network 172.16.100.0/24, the remote SD-WAN device vEdge-2 has to know a) the transport network where vEdge-1 is connected (Public-Internet), b) What is the public IP address that vEdge-1 uses for the transport connection (10.100.0.101), c) what is the tunneling protocol (GRE in our case), and d) does vEdge-1 allow tunneling only within one transport network (true = restrict bit set to 1). Based on this information vEdge-2 builds the Overlay Network using GRE (Generic Route Encapsulation) tunneling solution. Note, I'm using GRE only because I want clear text packet captures from data packets send between vEdge-1 and vEdge-2. Before setting up the GRE tunnel, vEdge-2 confirms that vEdge-1 is reachable by using BFD (Bi-Directional Forwarding Detection). Figure 2-3 shows how vEdge-1 sends OMP TLOC information to vSmart, which forwards the information to the remote edge device vEdge-2. edge devices exchange routing information only with vSmart, never between themselves.

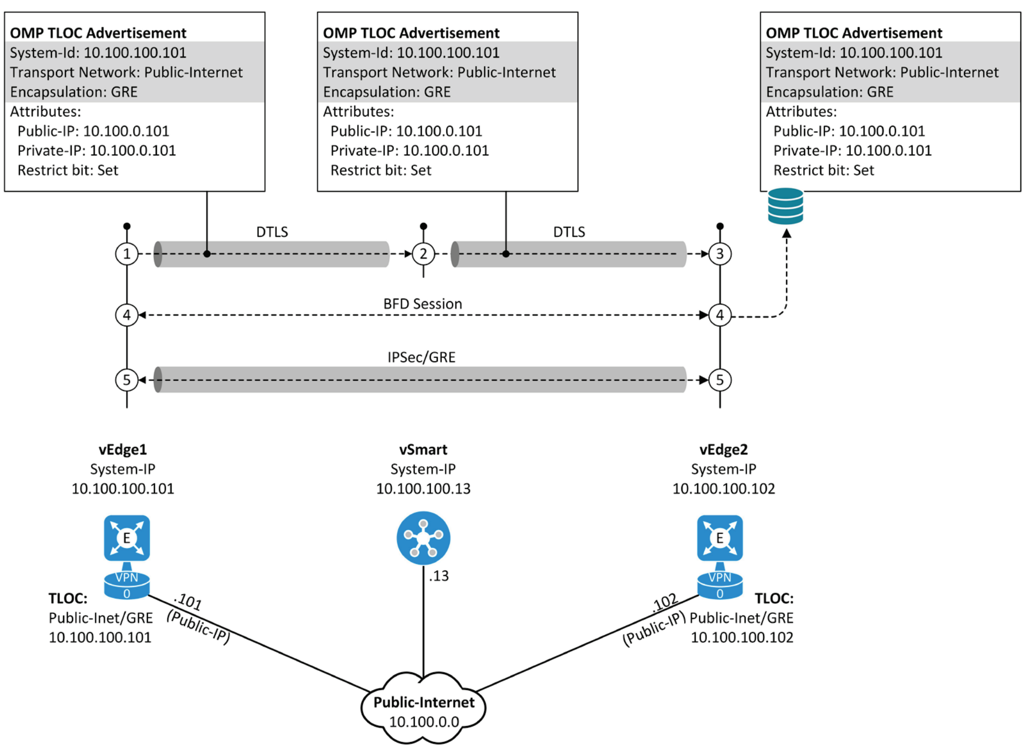


Figure 2-3: *OMP TLOC Process.*

Example 2-3 shows the OMP TLOC route that vEdge-1 advertises to SD-WAN Control-Plane node vSmart (10.100.100.13). vEdge-1 tells that it belongs to site 10 and is connected to the transport network Public-Internet by using IP address 10.100.0.101 (listed as public-ip). It also describes that it uses GRE as a tunnel protocol and accepts tunneling only from edge nodes that are connected to the same transport network (restrict bit set to one).

```
vEdge-1# show omp tlocs | exclude not | nomore
---------------------------------------------------
tloc entries for 10.100.100.101
                 public-internet
                 gre
---------------------------------------------------
            RECEIVED FROM:
peer            0.0.0.0
status          C,Red,R
    Attributes:
     attribute-type    installed
     encap-key         0
     public-ip         10.100.0.101
     public-port       0
     private-ip        10.100.0.101
     private-port      0
     <IPv6 part snipped>
     bfd-status        up
     site-id           10
     preference        0
     weight            1
     version           3
     gen-id             0x80000034
     carrier           default
     restrict          1
     <snipped>
            ADVERTISED TO:
peer    10.100.100.13
    Attributes:
     encap-key         0
     public-ip         10.100.0.101
     public-port       0
     private-ip        10.100.0.101
     private-port      0
     <IPv6 part snipped>
     site-id           10
     preference        0
     weight            1
     version           3
     gen-id            0x80000034
     carrier           default
     restrict          1
     on-demand          0
     <snipped>
```

Example 2-3: *vEdge-1 - OMP TLOC entries.*

Example 2-4 shows OMP TLOC routing information that vSmart has received from vEdge-1 (10.100.100.101).

```
vsmart# show omp tlocs | exclude not | nomore
---------------------------------------------------
tloc entries for 10.100.100.101
                 public-internet
                 gre
---------------------------------------------------
            RECEIVED FROM:
peer            10.100.100.101
status          C,I,R
    Attributes:
     attribute-type    installed
     encap-key         0
     public-ip         10.100.0.101
     public-port       0
     private-ip        10.100.0.101
     private-port      0
     <snipped>
     site-id           10
     <snipped>
     restrict          1
     <snipped>
```

Example 2-4: *vSmart - OMP TLOC entries.*

Example 2-5 shows the OMP TLOC routing information that remote vEdge-2 has learned from vSmart (10.100.100.13).

```
vEdge2# show omp tlocs | exclude not | nomore
---------------------------------------------------
tloc entries for 10.100.100.101
                 public-internet
                 gre
---------------------------------------------------
            RECEIVED FROM:
peer            10.100.100.13
status          C,I,R
    Attributes:
     attribute-type    installed
     encap-key         0
     public-ip         10.100.0.101
     public-port       0
     private-ip        10.100.0.101
     private-port      0
     <snipped>
     bfd-status        up
     site-id           10
     <snipped>
     restrict          1
     <snipped>
```

Example 2-5: *vEdge-2 - OMP TLOC entries.*

Example 2-6 verifies the BFD session between vEdge-1 and vEdge-2.

```
vEdge2# show bfd sessions system-ip 10.100.100.101 | beg 10.

10.100.100.101 10  up public-internet public-internet  10.100.0.102
10.100.0.101  gre
```

Example 2-6: *BFD Session Between vEdge-2 and vEdge-1.*

Example 2-7 shows that GRE tunneling between vEdge-1 and vEdg-2. Note, that statistic is not shown in the example.

```
vEdge2# show tunnel statistics
tunnel stats gre 10.100.0.102 10.100.0.101 0 0
 system-ip      10.100.100.101
 local-color    public-internet
 remote-color   public-internet
 <snipped>
```

Example 2-7: *GRE Tunneling Between vEdge-1 and vEdge-2.*

Overlay Network Control-Plane - OMP Service Route

When vEdge-1 and vEdge-2 know each other's TLOC information, they can install customer routing information into the RIB. Figure 2-4 illustrates the process, where vEdge-1 advertises the OMP Service route about network 172.16.100.0/24 to vSmart over the DTLS tunnel. The VPN-Id VPN 10 is used as VN-Id in Control-Plane. Based on that the vEdge-2 installs the information into the correct VPN RIB. OMP Service route also carries a set of attributes. These attributes describe the VN-Id used in Data-Plane (Label 1003), TLOC information, and Site-Id among the other things. vSmart verifies that the vEdge-1 is reachable before forwarding the OMP Service route to vEdge-2. vEdge-2 also verifies that it can route data to vEdge-1 before installing the information into the RIB. The OMP Service route doesn't carry the Underlay Network routable public IP address. vEdge-2 knows how to reach vEdge-1 based TLOC database.

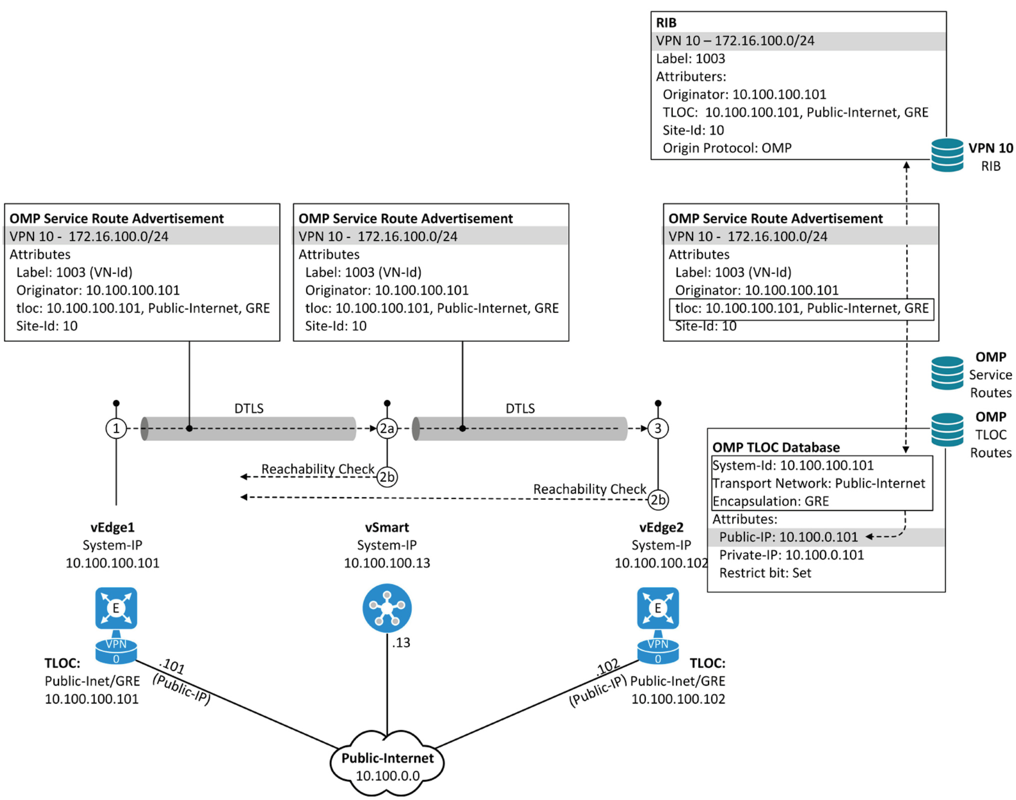


Figure 2-4: *OMP Service Route Propagation.*

Example 2-8 shows that vEdge-1 is redistributed the route from the eBGP into OMP Process. Status bits describe that the route is Chosen (C), Redistributed (R), and the Origin is Reachable (R). vEdge-1 advertises this information to vSmart (10.100.100.13).

```
vEdge1# show omp routes received detail

---------------------------------------------------
omp route entries for vpn 10 route 172.16.100.0/24
---------------------------------------------------
            RECEIVED FROM:
peer            0.0.0.0
path-id         37
label           1003
status          C,Red,R
    Attributes:
     originator       10.100.100.101
     type             installed
     tloc             10.100.100.101, public-internet, gre
     overlay-id        1
     site-id          10
     origin-proto     eBGP
     origin-metric    0
            ADVERTISED TO:
peer    10.100.100.13
    Attributes:
     originator       10.100.100.101
     label            1003
     path-id          37
     tloc             10.100.100.101, public-internet, gre
     site-id          10
     overlay-id        1
     origin-proto     eBGP
     origin-metric    0
vEdge-1#
```

Example 2-8: *OMP Service Route Information on vEdge-1.*

Example 2-9 shows that vSmart has received the OMP Service route from vEdge-1. Status bits describe that the route is Chosen (C) and that the origin is Reachable (R). Example 2-9 also shows that vSmart advertises the information to vEdge-2.

```
vsmart# show omp routes received detail | exclude not\ set | nomore
---------------------------------------------------
omp route entries for vpn 10 route 172.16.100.0/24
---------------------------------------------------
            RECEIVED FROM:
peer            10.100.100.101
path-id         37
label           1003
status          C,R
    Attributes:
     originator       10.100.100.101
     type             installed
     tloc             10.100.100.101, public-internet, gre
     overlay-id        1
     site-id          10
     origin-proto     eBGP
     origin-metric    0
            ADVERTISED TO:
peer    10.100.100.102
    Attributes:
     originator       10.100.100.101
     label            1003
     path-id          6
     tloc             10.100.100.101, public-internet, gre
     site-id          10
     overlay-id        1
     origin-proto     eBGP
     origin-metric    0
```

Example 2-9: *OMP Service Route Information on vSmart.*

Example 2-10 shows that vEdge-2 has received the OMP Service route from vSmart.

```
vEdge2# show omp routes received detail | exclude not | nomore
---------------------------------------------------
omp route entries for vpn 10 route 172.16.100.0/24
---------------------------------------------------
            RECEIVED FROM:
peer            10.100.100.13
path-id         6
label           1003
status          C,I,R
    Attributes:
     originator       10.100.100.101
     type             installed
     tloc             10.100.100.101, public-internet, gre
     overlay-id        1
     site-id          10
     origin-proto     eBGP
     origin-metric    0
vEdge2#
```

Example 2-10: *OMP Service Route Information on vEdge-2.*

Example 2-11 shows VPN 10 RIB o vEdge-2. vEdge-2 does recursive route lookup for *tloc-ip* from the TLOC database in order to find the IP address used in the tunnel header. vEdge-2 sends data packets to VPN 10 network 172.16.100.0/24 over the Public-Internet transport network using GRE encapsulation and VPN label 1003.

```
vEdge2# show ip routes vpn 10 172.16.100.0/24 detail
Codes Proto-sub-type:
  IA -> ospf-intra-area, IE -> ospf-inter-area,
  E1 -> ospf-external1, E2 -> ospf-external2,
  N1 -> ospf-nssa-external1, N2 -> ospf-nssa-external2,
  e -> bgp-external, i -> bgp-internal
Codes Status flags:
  F -> fib, S -> selected, I -> inactive,
  B -> blackhole, R -> recursive, L -> import

""---------------------------------------------
 VPN 10      PREFIX 172.16.100.0/24
---------------------------------------------
 proto           omp
 distance        250
 metric          0
 uptime          0:00:04:42
 tloc-ip         10.100.100.101
 tloc-color      public-internet
 tloc-encap      gre
 nexthop-label   1003
 status          F,S
```

Example 2-11: *VPN 10 RIB on vEdge-2.*

OMP to BGP on Remote Edge

Figure 2-5 illustrates the process of how vEdge-2 redistributes OMP VPN 10 route 172.16.100.0/24 from the RIB to the BGP process to Border-Leaf-13. We are using VRF-based eBGP IPv4 peering between vEdge-2 and Border-Leaf-13, so there is no VN-Id carried within BGP Update message. Border-Leaf-13 verifies that the BGP Update is valid (configured peer with the right ASN, available next-hop, and so on). In case that update is valid, it installs route into VRF-specific RIB.

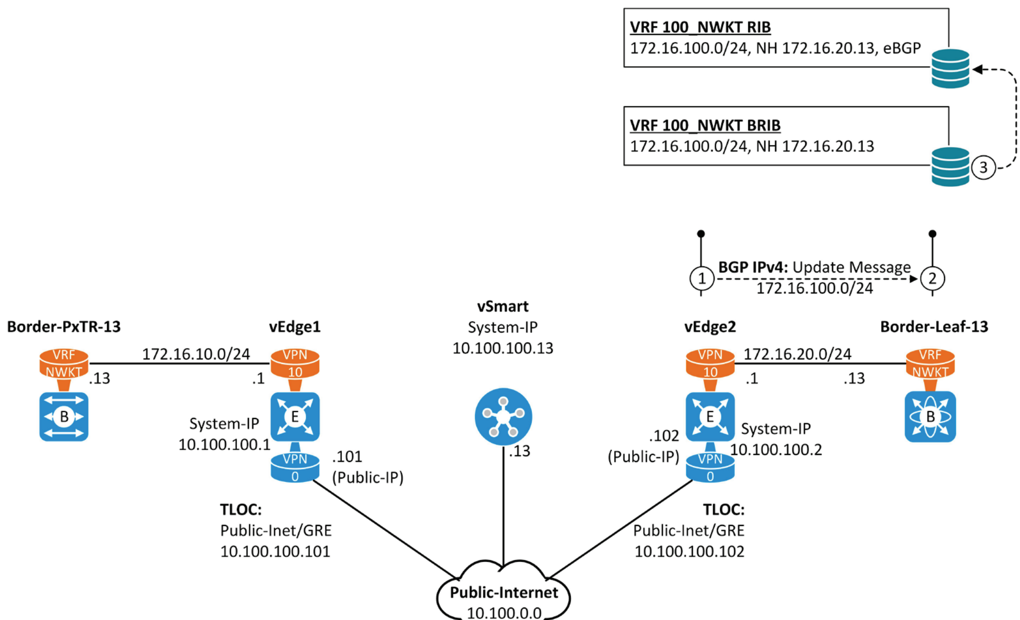


Figure 2-5: *OMP Service Route to BGP on vEdge-2.*

Capture 2-1 shows the BGP Update message sent by vEdge-2 to Border-Leaf-13. Note that the Path Attribute AS_PATH has only AS 65100 that is the VPN 10 specific ASN on vEdge-2. The OMP Service route updates don't carry BGP ASN information and that is why the LISP domain BGP ASN 65010 is not shown. The VPN Site-Id 20 is encoded in Route Origin extended community as AN: 20.

```
Internet Protocol Version 4, Src: 172.16.20.1, Dst: 172.16.20.13
Transmission Control Protocol, Src Port: 40769, Dst Port: 179, Seq: 20, Ack: 1, Len: 66
Border Gateway Protocol - UPDATE Message
    Marker: ffffffffffffffffffffffffffffffff
    Length: 66
    Type: UPDATE Message (2)
    Withdrawn Routes Length: 0
    Total Path Attribute Length: 39
    Path attributes
        Path Attribute - ORIGIN: INCOMPLETE
        Path Attribute - AS_PATH: 65100
```

```
        Path Attribute - NEXT_HOP: 172.16.20.1
        Path Attribute - MULTI_EXIT_DISC: 1000
        Path Attribute - EXTENDED_COMMUNITIES
            Flags: 0xc0, Optional, Transitive, Complete
            Type Code: EXTENDED_COMMUNITIES (16)
            Length: 8
            Carried extended communities: (1 community)
                Route Origin: 0:20 [Transitive 2-Octet AS-Specific]
                    Type: Transitive 2-Octet AS-Specific (0x00)
                    Subtype (AS2): Route Origin (0x03)
                    2-Octet AS: 0
                    4-Octet AN: 20
    Network Layer Reachability Information (NLRI)
        172.16.100.0/24
            NLRI prefix length: 24
            NLRI prefix: 172.16.100.0
```

Capture 2-2: *BGP Update from vEdge-2 to Border-Leaf-13.*

Example 2-13 shows the VRF NWKT BGP table on Border-Leaf-13.

```
Border-leaf-13# show ip bgp vrf NWKT
BGP routing table information for VRF NWKT, address family IPv4 Unicast
BGP table version is 9, Local Router ID is 172.16.30.1
Status: s-suppressed, x-deleted, S-stale, d-dampened, h-history, *-valid, >-best
Path type: i-internal, e-external, c-confed, l-local, a-aggregate, r-redist, I-i
njected
Origin codes: i - IGP, e - EGP, ? - incomplete, | - multipath, & - backup, 2 - b
est2

   Network            Next Hop            Metric     LocPrf     Weight Path
*>e172.16.100.0/24    172.16.20.1           1000                     0 65100 ?
```

Example 2-12: *VRF NWKT BRIB on Border-Leaf-13.*

Example 2-13 verifies that the route is also installed into VRF NWKT RIB with the next-hop 172.16.20.1 (vEdge-2).

```
Border-leaf-13# show ip route 172.16.100.0/24 vrf NWKT
IP Route Table for VRF "NWKT"
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
'%<string>' in via output denotes VRF <string>

172.16.100.0/24, ubest/mbest: 1/0
    *via 172.16.20.1, [20/1000], 00:11:45, bgp-65030, external, tag 65100
```

Example 2-13: *VRF NWKT BRIB on Border-Leaf-13.*

This chapter describes how the routing information is sent from LISP domain edge device Border-PxTR-13 to local SD-WAN device vEdge-1. Besides we saw the process of how the routing information ends up into RIB of remote SD-WAN edge device vEdge-2 and finally how it is sent to Datacenter border leaf Border-Leaf-13. The next chapter explains how the route to network 172.16.100.0/24 ends up in the RIB of Leaf-11.

Chapter 2 Appendix A: Device config Configuration

vEdge-1

```
vEdge-1# show run
system
 host-name               vEdge-1
 system-ip               10.100.100.101
 site-id                 10
 admin-tech-on-failure
 no route-consistency-check
 sp-organization-name    nwkt
 organization-name       nwkt
 vbond 10.100.0.11
 aaa
  auth-order local radius tacacs
  usergroup basic
   task system read write
   task interface read write
  !
  usergroup netadmin
  !
  usergroup operator
   task system read
   task interface read
   task policy read
   task routing read
   task security read
  !
  user admin
   password
$6$siwKBQ==$wT2lUa9BSreDPI6gB8sl4E6PAJoVXgMbgv/whJ8F1C6sWdRazdxorYYTLrL6syiG6qnLABTnrE
96HJiKF6QRq1
  !
  user ciscotacro
   description CiscoTACReadOnly
   group       operator
   status      enabled
  !
  user ciscotacrw
   description CiscoTACReadWrite
   group       netadmin
   status      enabled
  !
 !
 logging
  disk
   enable
  !
 !
 ntp
  server 10.100.0.14
   version 4
  exit
 !
!
omp
 no shutdown
 graceful-restart
 advertise bgp
```

```
 advertise connected
 advertise static
!
security
 ipsec
  authentication-type sha1-hmac ah-sha1-hmac
 !
!
vpn 0
 interface ge0/0
  description "Internet Transport"
  ip address 10.100.0.101/24
  tunnel-interface
   encapsulation gre
   color public-internet restrict
   allow-service all
   no allow-service bgp
   allow-service dhcp
   allow-service dns
   allow-service icmp
   no allow-service sshd
   no allow-service netconf
   no allow-service ntp
   no allow-service ospf
   no allow-service stun
   allow-service https
  !
  no shutdown
 !
 interface ge0/1
  description "MPLS Transport"
  ip address 10.200.0.101/24
  tunnel-interface
   encapsulation gre
   color mpls restrict
   max-control-connections 0
   no control-connections
   allow-service all
   no allow-service bgp
   allow-service dhcp
   allow-service dns
   allow-service icmp
   no allow-service sshd
   no allow-service netconf
   no allow-service ntp
   no allow-service ospf
   no allow-service stun
   allow-service https
  !
  no shutdown
 !
 ip route 0.0.0.0/0 10.100.0.1
 ip route 0.0.0.0/0 10.200.0.1
!
vpn 10
 router
  bgp 65100
   address-family ipv4-unicast
    redistribute omp
   !
   neighbor 172.16.10.13
    no shutdown
```

```
    remote-as 65010
    address-family ipv4-unicast
    !
   !
  !
 !
 interface ge0/2
  ip address 172.16.10.1/24
  no shutdown
 !
!
vpn 512
 interface eth0
  description Management
  ip dhcp-client
  no shutdown
 !
 ip route 0.0.0.0/0 192.168.10.1
!
```

vEdge-2

```
vEdge2# show run
system
 host-name               vEdge2
 system-ip               10.100.100.102
 site-id                 20
 admin-tech-on-failure
 no route-consistency-check
 sp-organization-name    nwkt
 organization-name       nwkt
 vbond 10.100.0.11
 aaa
  auth-order local radius tacacs
  usergroup basic
   task system read write
   task interface read write
  !
  usergroup netadmin
  !
  usergroup operator
   task system read
   task interface read
   task policy read
   task routing read
   task security read
  !
  user admin
   password
$6$siwKBQ==$wT2lUa9BSreDPI6gB8sl4E6PAJoVXgMbgv/whJ8F1C6sWdRazdxorYYTLrL6syiG6qnLABTnrE
96HJiKF6QRq1
  !
  user ciscotacro
   description CiscoTACReadOnly
   group       operator
   status      enabled
  !
  user ciscotacrw
```

```
   description CiscoTACReadWrite
   group       netadmin
   status      enabled
  !
 !
 logging
  disk
   enable
  !
 !
 ntp
  server 10.100.0.14
   version 4
  exit
 !
!
omp
 no shutdown
 graceful-restart
 advertise bgp
 advertise connected
 advertise static
!
security
 ipsec
  authentication-type sha1-hmac ah-sha1-hmac
 !
!
vpn 0
 interface ge0/0
  description "Internet Transport"
  ip address 10.100.0.102/24
  tunnel-interface
   encapsulation gre
   color public-internet restrict
   allow-service all
   no allow-service bgp
   allow-service dhcp
   allow-service dns
   allow-service icmp
   no allow-service sshd
   no allow-service netconf
   no allow-service ntp
   no allow-service ospf
   no allow-service stun
   allow-service https
  !
  no shutdown
 !
 interface ge0/1
  description "MPLS Transport"
  ip address 10.200.0.102/24
  tunnel-interface
   encapsulation gre
   color mpls restrict
   max-control-connections 0
   no control-connections
   allow-service all
   no allow-service bgp
   allow-service dhcp
   allow-service dns
   allow-service icmp
```

```
   no allow-service sshd
   no allow-service netconf
   no allow-service ntp
   no allow-service ospf
   no allow-service stun
   allow-service https
  !
  no shutdown
 !
 ip route 0.0.0.0/0 10.100.0.1
 ip route 0.0.0.0/0 10.200.0.1
!
vpn 10
 router
  bgp 65100
   address-family ipv4-unicast
    redistribute omp
   !
   neighbor 172.16.20.13
    no shutdown
    remote-as 65030
    address-family ipv4-unicast
    !
   !
  !
 !
 interface ge0/2
  ip address 172.16.20.1/24
  no shutdown
 !
!
vpn 512
 interface eth0
  description Management
  ip dhcp-client
  no shutdown
 !
 ip route 0.0.0.0/0 192.168.10.1
```

vSmart

```
vsmart# show run
system
 host-name             vsmart
 system-ip             10.100.100.13
 site-id               100
 admin-tech-on-failure
 organization-name     nwkt
 vbond 10.100.0.11
 aaa
  auth-order local radius tacacs
  usergroup basic
   task system read write
   task interface read write
  !
  usergroup netadmin
  !
  usergroup operator
   task system read
   task interface read
```

```
   task policy read
   task routing read
   task security read
  !
  usergroup tenantadmin
  !
  user admin
   password
$6$jKzSSqC2GCJveJV4$VxMCv59Qv2J.lDd2luqXXJ9dUuv3izVKXPEbE3b43AAry3n6ptI7DqunO0y0TzxaUV
RGAUZ7E/ySEiWdyt8/60
  !
  user ciscotacro
   description CiscoTACReadOnly
   group       operator
   status      enabled
  !
  user ciscotacrw
   description CiscoTACReadWrite
   group       netadmin
   status      enabled
  !
 !
 logging
  disk
   enable
  !
 !
 ntp
  server 10.100.0.14
   version 4
  exit
 !
!
omp
 no shutdown
 graceful-restart
!
vpn 0
 interface eth0
  ip address 10.100.0.13/24
  ipv6 dhcp-client
  tunnel-interface
   allow-service all
   allow-service dhcp
   allow-service dns
   allow-service icmp
   no allow-service sshd
   allow-service netconf
   no allow-service ntp
   no allow-service stun
  !
  no shutdown
 !
 ip route 0.0.0.0/0 10.100.0.1
!
vpn 512
 interface eth1
  ip dhcp-client
  no shutdown
 !
!
```

Chapter 3: DC Control-Plane: BGP EVPN - IP Prefix Route

Introduction

Figure 3-1 shows the Overlay Network Control-Plane interaction. Edge-xTR-11 registers reachability information (IP address and location) of EP1 to MapSrv-22. MapSrv-22 stores the information into the Mapping database and then installs it to RIB. Then MapSrv-22 exports the information to the BGP process and sends BGP Update using VPNv4 address format to Border-PxTR-13. Border-PxTR-13 imports NLRI into BRIB and RIB. Then it sends BGP Update to the local SD-WAN edge device vEdge-1. After importing the received information into BRIB and RIB, vEdge-1 exports the information to the OMP process and advertises it over a DTLS connection to vSmart (SD-WAN Control-Plane node). vSmart, in turn, advertises information to remote SD-WAN device vEdge-2. After importing the received information into the RIB, vEdge-2 exports the information to the BGP process and sends BGP Update to Border-Leaf-13. Border-Leaf-13 installs the information into BRIB and RIB. Next, Border-Leaf-13 sends BGP Update message using EVPN route type 5 (IP Prefix Route) to its iBGP peer Spine-1 (BGP Route-Reflector) using auto-generated Route-Target 65030:10077. Spine-1 forwards the BGP Update to Leaf-1, which imports the information into L3VNI used with VRF NWKT and installs the route into the VRF NWKT RIB.

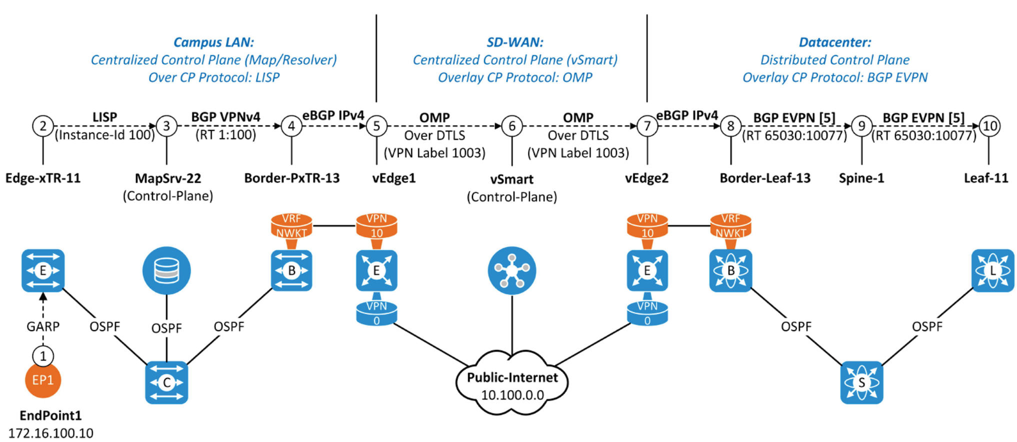


Figure 3-1: *Overall Control-Plane Operation.*

BGP Process on Border-Leaf-13

When Border-Leaf-13 receives BGP Update from vEdge-1 it first installs the route into the BGP Adj-Rib-In table. After the BGP validation process (valid next-hop address, correct BGP ASN, and so on) it installs the route into BGP Loc-RIB with auto-generated Route-Target extended community value 65030:10077 (ASN:L3VNI). Then Border-Leaf-13 installs the route into the VRF NWKT specific RIB. After installing route into the RIB, Border-leaf-13 exports the information from there to BRIB as EVPN Route. It adds the BGP Route Distinguisher (RD) which is auto derived from Router-Id (192.168.100.13) and VRF-Id (3). The destination IPv4 Prefix 172.16.100.0/24 is changed to EVPN format [5]:[0]:[0]:[24]:[172.16.100.0]/224 where [5] defines EVPN Route-Type, 5 = IP prefix Route. Following [0]:[0] defines the ESI Value and Ethernet Tag-Id which are not used in our case. Next, there is the subnet mask and IP address. The last notation /224 is the total address length. The received Label value is the L3 Virtual Network Id that is used in Data-Plane. The SoO (Site of Origin) Extended Community value is used for BGP loop prevention and the ENCAP (Encapsulation) value 8 describes that VXLAN encapsulation has to be used on Data-Plane. The Router MAC address is needed because VXLAN uses MAC-in-UDP encapsulation and there has to be some MAC information carried within the VXLAN tunnel header.

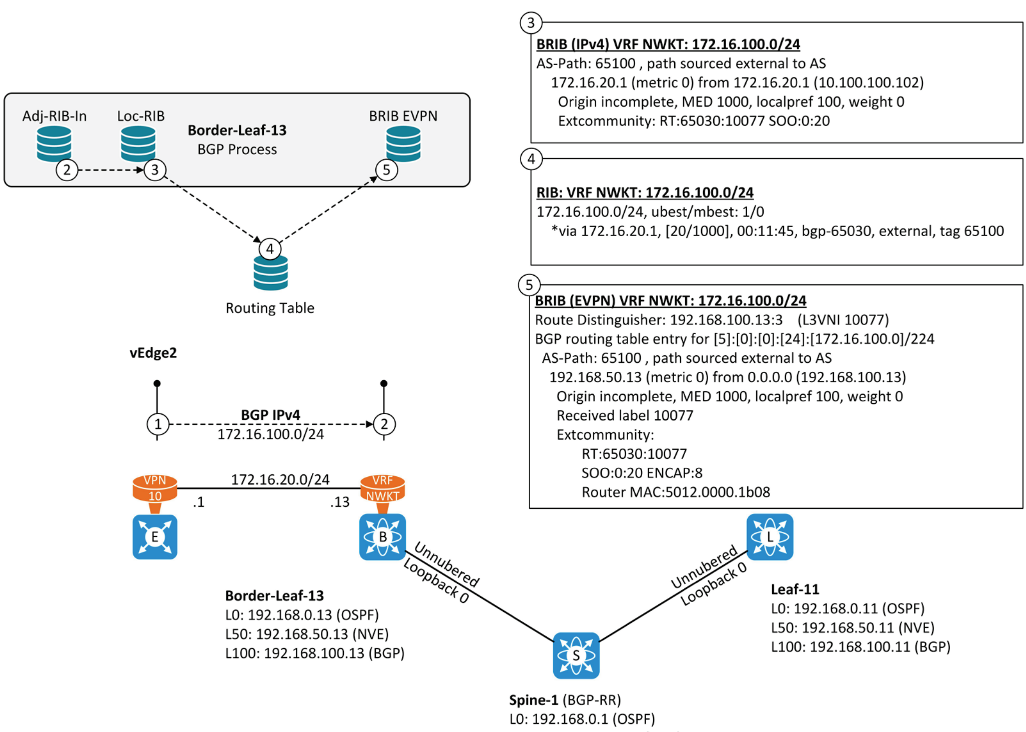


Figure 3-2: *Border-Leaf-13 Control-Plane Processes.*

Example 3-1 shows the BGP IPv4 information about network 172.16.100.0/24. The route is learned from vEdge-2 which Border-Leaf-13 peers using its link address 172.16.20.1. vEdge-2 uses its OMP System-Id 10.100.100.102 as RID. The IPv4 Unicast route is not advertised to any peer.

```
Border-leaf-13# show ip bgp vrf NWKT 172.16.100.0
BGP routing table information for VRF NWKT, address family IPv4 Unicast
BGP routing table entry for 172.16.100.0/24, version 5
<snipped>
  Advertised path-id 1, VPN AF advertised path-id 1
  Path type: external, path is valid, is best path, no labeled nexthop, in rib
  AS-Path: 65100 , path sourced external to AS
    172.16.20.1 (metric 0) from 172.16.20.1 (10.100.100.102)
      Origin incomplete, MED 1000, localpref 100, weight 0
      Extcommunity: RT:65030:10077 SOO:0:20

  VRF advertise information:
  Path-id 1 not advertised to any peer

  VPN AF advertise information:
  Path-id 1 not advertised to any peer
```

Example 3-1: *VRF NWKR Specific BGP IPv4 Entry on Border-Leaf-13.*

Example 3-2 verifies that Border-Leaf-13 has been installed the route from its BRIB into the RIB.

```
Border-leaf-13# sh ip route 172.16.100.0/24 vrf NWKT
IP Route Table for VRF "NWKT"
'*' denotes best ucast next-hop
<snipped>

172.16.100.0/24, ubest/mbest: 1/0
    *via 172.16.20.1, [20/1000], 00:11:45, bgp-65030, external, tag 65100
```

Example 3-2: *VRF NWKR Specific RIB on Border-Leaf-13.*

Example 3-3 shows the BGP EVPN entry about the network 172.16.100.0/24.

```
Border-leaf-13# sh bgp l2vpn evpn 172.16.100.0 vrf NWKT
Route Distinguisher: 192.168.100.13:3    (L3VNI 10077)
BGP routing table entry for [5]:[0]:[0]:[24]:[172.16.100.0]/224, version 5
Paths: (1 available, best #1)
Flags: (0x000002) (high32 00000000) on xmit-list, is not in l2rib/evpn

  Advertised path-id 1
  Path type: local, path is valid, is best path, no labeled nexthop
  Gateway IP: 0.0.0.0
  AS-Path: 65100 , path sourced external to AS
    192.168.50.13 (metric 0) from 0.0.0.0 (192.168.100.13)
      Origin incomplete, MED 1000, localpref 100, weight 0
      Received label 10077
      Extcommunity: RT:65030:10077 SOO:0:20 ENCAP:8 Router MAC:5012.0000.1b08

  Path-id 1 advertised to peers:
    192.168.100.1
```

Example 3-3: *VRF NWKR Specific BGP EVPN L3VNI Entry on Border-Leaf-13.*

BGP Process on Leaf-11

Border-Leaf-13 sends BGP Update that carries EVPN NLRI about network 172.16.100.0/24 to its iBGP peer Spine-1. Spine-1 is a BGP Route-Reflector and forwards the BGP Update to Leaf-11. When Leaf-11 gets the BGP Update it installs the information into the BGP Adj-RIB-In. It imports the information into the BGP Loc-RIB after BGP update validation. Leaf-11 changes the original RD value 192.168.100.13:3 to local RD value 192.168.100.11:3 during the import process. The next-hop IP address 192.168.50.13 is bind to the NVE1 interface on Border-Leaf-13. It is used in the outer IP header as a destination IP address when sending VXLAN encapsulated data towards network 172.16.100.0/24. As mentioned before, the Label value is L3VNI which describes VRF/Tenant in the VXLAN tunnel header. The router MAC is used in the inner Ethernet header as the destination MAC address. L3VNI (segid 10077) and encapsulation type (8 = VXLAN) are imported together with the next-hop IP address 192.168.50.13 into the VRF NWKT RIB.

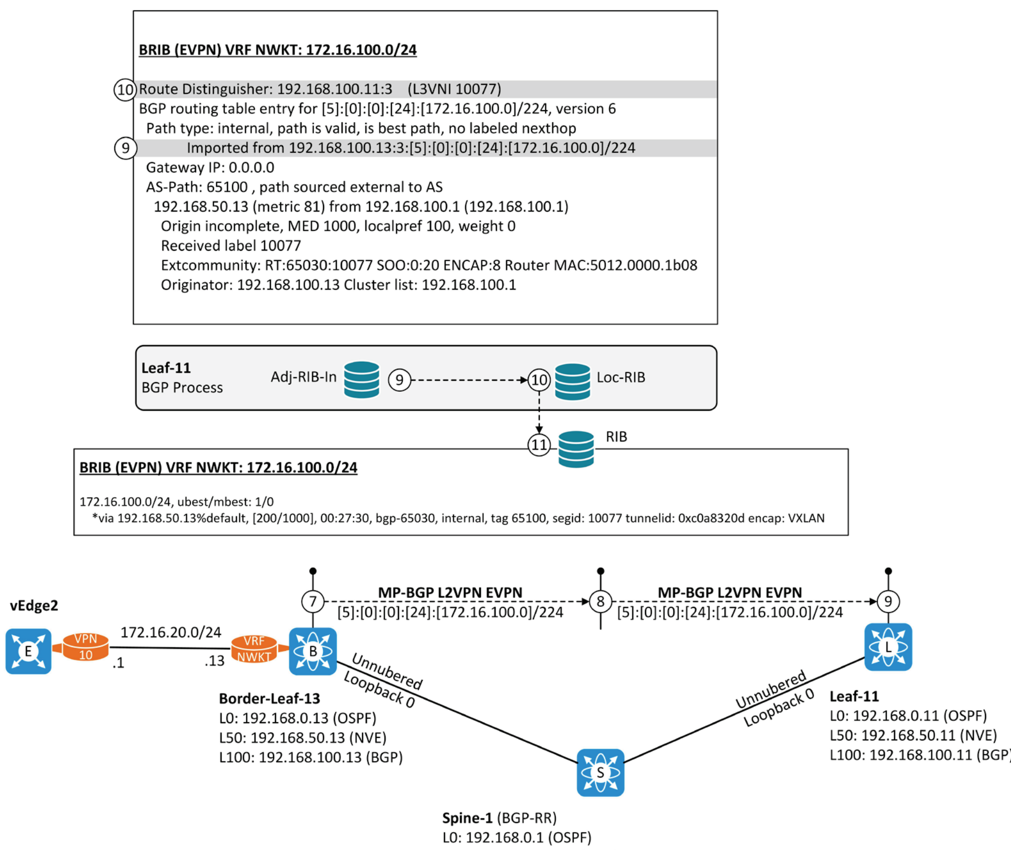


Figure 3-3: *Leaf-11 Control-Plane Processes.*

Capture 3-1 shows the packet capture about the BGP Update packet that Border-Leaf-13 sent to Spine-1. The BGP Path Attribute MP_REACH-NLRI includes EVPN NLRI route type IP Prefix Route (type 5). The Route Distinguisher is set to 192.168.50.13:3 (NVE1 IP:VRF-Id). It is used in order to Spine-1 can differentiate possible overlapping IP Prefixes/IP addresses between different VRFs. The ESI (Ethernet Segment Id) value is note set because we are not using ESI multihoming. Both AS_PATH and MED attributes are copied from the IPv4 BGP Update received from vEdge-2. The BGP Extended Community attribute Route-Target 65030:10077 (ASN:L3VNI) is used for BGP import policy. Note that the unknown subtype 0x03 defines Router MAC.

```
Internet Protocol Version 4, Src: 192.168.100.13, Dst: 192.168.100.1
Transmission Control Protocol, Src Port: 18200, Dst Port: 179, Seq: 39, Ack: 20, Len:
134
Border Gateway Protocol - UPDATE Message
  <snipped>
     Path attributes
        Path Attribute - MP_REACH_NLRI
            Flags: 0x90, Optional, Extended-Length, Non-transitive, Complete
            Type Code: MP_REACH_NLRI (14)
            Length: 45
            Address family identifier (AFI): Layer-2 VPN (25)
            Subsequent address family identifier (SAFI): EVPN (70)
            Next hop network address (4 bytes): NEXT_HOP: 192.168.50.13
            Number of Subnetwork points of attachment (SNPA): 0
            Network layer reachability information (36 bytes)
                EVPN NLRI: IP Prefix route
                    Route Type: IP Prefix route (5)
                    Length: 34
                    Route Distinguisher: 0001c0a8640d0003 (192.168.100.13:3)
                    ESI: 00:00:00:00:00:00:00:00:00:00
                    Ethernet Tag ID: 0
                    IP prefix length: 24
                    IPv4 address: 172.16.100.0
                    IPv4 Gateway address: 0.0.0.0
                    MPLS Label Stack: 629 (bottom)
        Path Attribute - ORIGIN: INCOMPLETE
        Path Attribute - AS_PATH: 65100
        Path Attribute - MULTI_EXIT_DISC: 1000
        Path Attribute - LOCAL_PREF: 100
        Path Attribute - EXTENDED_COMMUNITIES
            Flags: 0xc0, Optional, Transitive, Complete
            Type Code: EXTENDED_COMMUNITIES (16)
            Length: 32
            Carried extended communities: (4 communities)
                Route Target: 65030:10077 [Transitive 2-Octet AS-Specific]
                    <snipped>
                Route Origin: 0:20 [Transitive 2-Octet AS-Specific]
                    <snipped>
                Encapsulation: VXLAN Encapsulation [Transitive Opaque]
                    <snipped>
                    Tunnel type: VXLAN Encapsulation (8)
                Unknown subtype 0x03: 0x5012 0x0000 0x1b08 [Transitive EVPN]
                    <snipped>
                    Raw Value: 0x5012 0x0000 0x1b08
```

Capture 3-1: *BGP Update Sent from Border-Leaf-13 to Spine-1.*

Example 3-4 shows both BGP Adj-RIB-In and Loc-RIB. They are almost the same excluding Route Distinguisher values and Imported -statements. The Adj-RIB-In entry is imported into BGP Loc-RIB and to the RIB and that is why there is an h to 2 destination statement. The information is not installed into L2RIB because this is an IP Prefix Route, not a MAC Advertisement route used to describe the host MAC address. Note that the IP address 192.168.50.13 is used as the next-hop but BGP peering is established with Spine-1 (192.168.100.13).

```
Leaf-11# show bgp l2vpn evpn 172.16.100.0
BGP routing table information for VRF default, address family L2VPN EVPN
>>>> COMMENT >>>> This entry is in BGP Adj-RIB-In >>>>>
Route Distinguisher: 192.168.100.13:3
BGP routing table entry for [5]:[0]:[0]:[24]:[172.16.100.0]/224, version 5
Paths: (1 available, best #1)
Flags: (0x000002) (high32 00000000) on xmit-list, is not in l2rib/evpn, is not in HW

  Advertised path-id 1
  Path type: internal, path is valid, is best path, no labeled nexthop
             Imported to 2 destination(s)
             Imported paths list: NWKT L3-10077
  Gateway IP: 0.0.0.0
  AS-Path: 65100 , path sourced external to AS
    192.168.50.13 (metric 81) from 192.168.100.1 (192.168.100.1)
      Origin incomplete, MED 1000, localpref 100, weight 0
      Received label 10077
      Extcommunity: RT:65030:10077 SOO:0:20 ENCAP:8 Router MAC:5012.0000.1b08
      Originator: 192.168.100.13 Cluster list: 192.168.100.1

  Path-id 1 not advertised to any peer
>>>> COMMENT >>>> This entry is in BGP BGP LOC-RIB >>>>>
Route Distinguisher: 192.168.100.11:3    (L3VNI 10077)
BGP routing table entry for [5]:[0]:[0]:[24]:[172.16.100.0]/224, version 6
Paths: (1 available, best #1)
Flags: (0x000002) (high32 00000000) on xmit-list, is not in l2rib/evpn, is not in HW

  Advertised path-id 1
  Path type: internal, path is valid, is best path, no labeled nexthop
             Imported from 192.168.100.13:3:[5]:[0]:[0]:[24]:[172.16.100.0]/224
  Gateway IP: 0.0.0.0
  AS-Path: 65100 , path sourced external to AS
    192.168.50.13 (metric 81) from 192.168.100.1 (192.168.100.1)
      Origin incomplete, MED 1000, localpref 100, weight 0
      Received label 10077
      Extcommunity: RT:65030:10077 SOO:0:20 ENCAP:8 Router MAC:5012.0000.1b08
      Originator: 192.168.100.13 Cluster list: 192.168.100.1

  Path-id 1 not advertised to any peer
```

Example 3-4: *BGP Adj-RIB-In and Loc-RIB tables on Border-Leaf-13.*

Example 3-5 shows how the route is installed into the RIB.

```
Leaf-11# show ip route 172.16.100.0 detail vrf NWKT
IP Route Table for VRF "NWKT"
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
'%<string>' in via output denotes VRF <string>

172.16.100.0/24, ubest/mbest: 1/0
    Extended Community: 0x1a 1c 00 03 00 00 00 00 00 14 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00
    *via 192.168.50.13%default, [200/1000], 00:27:30, bgp-65030, internal, tag 65100,
segid: 10077 tunnelid: 0xc0a8320d encap: VXLAN

         BGP-EVPN: VNI=10077 (EVPN)
         client-specific data: 2
         recursive next hop: 192.168.50.13/32%default
         extended route information: BGP origin AS 65100 BGP peer AS 65100
```

Example 3-5: *VRF NWKR Specific BGP EVPN L3VNI Entry on Border-Leaf-13.*

So far, we have seen how routing information concerning the host EP1 IP address on the LISP domain is propagated from the LISP domain as the aggregate route over the SD-WAN domain all the way to the BGP EVPN domain. From the Control-Plane perspective, we are halfway. In order to have IP connectivity between hosts EP1 in the LISP domain and EP2 in the BGP EVPN domain, we need to advertise also the IP reachability information concerning EP2 all the way down to Edge- 11. The next chapter discusses how the routing about IP address 172.16.30.10.2 is advertised as EVPN MAC Route from Leaf-11 to vEdge-2.

Chapter 3 Appendix A: Device config Configuration

Border-Leaf-13

```
Border-leaf-13# sh run

!Command: show running-config
!No configuration change since last restart
!Time: Thu Jul 29 12:10:27 2021

version 9.3(5) Bios:version
hostname Border-leaf-13
vdc Border-leaf-13 id 1
  limit-resource vlan minimum 16 maximum 4094
  limit-resource vrf minimum 2 maximum 4096
  limit-resource port-channel minimum 0 maximum 511
  limit-resource u4route-mem minimum 248 maximum 248
  limit-resource u6route-mem minimum 96 maximum 96
  limit-resource m4route-mem minimum 58 maximum 58
  limit-resource m6route-mem minimum 8 maximum 8
  limit-resource vni_bd minimum 4096 maximum 4096

nv overlay evpn
feature ospf
feature bgp
feature pim
feature fabric forwarding
feature interface-vlan
feature vn-segment-vlan-based
feature nv overlay

fabric forwarding anycast-gateway-mac 0001.0001.0001
ip pim rp-address 192.168.238.238 group-list 224.0.0.0/4
ip pim ssm range 232.0.0.0/8
vlan 1,10,77
vlan 10
  name VLAN10-mapped-to-VNI10000
  vn-segment 10000
vlan 77
  name NWKT
  vn-segment 10077

vrf context NWKT
  vni 10077
  rd auto
  address-family ipv4 unicast
    route-target both auto
    route-target both auto evpn
vrf context management
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 Vlan1
```

```
interface Vlan10
  no shutdown
  vrf member NWKT
  ip address 172.16.30.1/24
  fabric forwarding mode anycast-gateway

interface Vlan77
  no shutdown
  mtu 9216
  vrf member NWKT
  ip forward

interface nve1
  no shutdown
  host-reachability protocol bgp
  source-interface loopback50
  member vni 10000
    mcast-group 238.0.0.10
  member vni 10077 associate-vrf

interface Ethernet1/1
  no switchport
  vrf member NWKT
  ip address 172.16.20.13/24
  no shutdown

interface Ethernet1/2
  no switchport
  medium p2p
  ip unnumbered loopback0
  ip ospf network point-to-point
  ip router ospf UNDERLAY-NET area 0.0.0.0
  ip pim sparse-mode
  no shutdown



interface mgmt0
  vrf member management

interface loopback0
  ip address 192.168.0.13/32
  ip router ospf UNDERLAY-NET area 0.0.0.0
  ip pim sparse-mode

interface loopback50
  ip address 192.168.50.13/32
  ip router ospf UNDERLAY-NET area 0.0.0.0
  ip pim sparse-mode

interface loopback100
  ip address 192.168.100.13/32
  ip router ospf UNDERLAY-NET area 0.0.0.0
  ip pim sparse-mode
```

```
icam monitor scale

line console
line vty
boot nxos bootflash:/nxos.9.3.5.bin sup-1
router ospf UNDERLAY-NET
  router-id 192.168.0.13
router bgp 65030
  router-id 192.168.100.13
  address-family ipv4 unicast
  address-family l2vpn evpn
  neighbor 192.168.100.1
    remote-as 65030
    update-source loopback100
    address-family l2vpn evpn
      send-community extended
  vrf NWKT
    address-family ipv4 unicast
      advertise l2vpn evpn
      aggregate-address 172.16.30.0/24 summary-only
    neighbor 172.16.20.1
      remote-as 65100
      address-family ipv4 unicast
evpn
  vni 10000 l2
    rd auto
    route-target import auto
    route-target export auto
```

Spine-1

```
Spine-1# sh run

!Command: show running-config
!No configuration change since last restart
!Time: Thu Jul 29 12:11:42 2021

version 9.3(5) Bios:version
hostname Spine-1
vdc Spine-1 id 1
  limit-resource vlan minimum 16 maximum 4094
  limit-resource vrf minimum 2 maximum 4096
  limit-resource port-channel minimum 0 maximum 511
  limit-resource u4route-mem minimum 248 maximum 248
  limit-resource u6route-mem minimum 96 maximum 96
  limit-resource m4route-mem minimum 58 maximum 58
  limit-resource m6route-mem minimum 8 maximum 8
  limit-resource vni_bd minimum 4096 maximum 4096

nv overlay evpn
feature ospf
feature bgp
feature pim
feature fabric forwarding
```

```
feature interface-vlan
feature vn-segment-vlan-based
feature nv overlay

username                    admin                    password                    5
$5$MEIGJJ$mivgNqGgIyMgZlXe85B1NaKxJTHL173JBUg0/agjoC.
 role network-admin
ip domain-lookup
copp profile strict
snmp-server       user       admin       network-admin       auth       md5
0xfc84745eecde487041c3b4552b5495d7
 priv 0xfc84745eecde487041c3b4552b5495d7 localizedkey
rmon event 1 log trap public description FATAL(1) owner PMON@FATAL
rmon event 2 log trap public description CRITICAL(2) owner PMON@CRITICAL
rmon event 3 log trap public description ERROR(3) owner PMON@ERROR
rmon event 4 log trap public description WARNING(4) owner PMON@WARNING
rmon event 5 log trap public description INFORMATION(5) owner PMON@INFO

ip pim rp-address 192.168.238.238 group-list 224.0.0.0/4
ip pim ssm range 232.0.0.0/8
ip pim anycast-rp 192.168.238.238 192.168.238.11
vlan 1

vrf context management

interface Vlan1

interface Ethernet1/1
  no switchport
  medium p2p
  ip unnumbered loopback0
  ip ospf network point-to-point
  ip router ospf UNDERLAY-NET area 0.0.0.0
  ip pim sparse-mode
  no shutdown

interface Ethernet1/2
  no switchport
  medium p2p
  ip unnumbered loopback0
  ip ospf network point-to-point
  ip router ospf UNDERLAY-NET area 0.0.0.0
  ip pim sparse-mode
  no shutdown

interface Ethernet1/3
  no switchport
  medium p2p
  ip unnumbered loopback0
  ip ospf network point-to-point
  ip router ospf UNDERLAY-NET area 0.0.0.0
  ip pim sparse-mode
  no shutdown

interface mgmt0
```

```
  vrf member management

interface loopback0
  ip address 192.168.0.1/32
  ip router ospf UNDERLAY-NET area 0.0.0.0
  ip pim sparse-mode

interface loopback50
  ip address 192.168.50.1/32
  ip router ospf UNDERLAY-NET area 0.0.0.0
  ip pim sparse-mode

interface loopback100
  ip address 192.168.100.1/32
  ip router ospf UNDERLAY-NET area 0.0.0.0
  ip pim sparse-mode

interface loopback238
  description ** Anycast-RP address **
  ip address 192.168.238.238/32
  ip router ospf UNDERLAY-NET area 0.0.0.0
  ip pim sparse-mode

interface loopback511
  description ** Unique Address for Anycast-RP **
  ip address 192.168.238.11/32
  ip router ospf UNDERLAY-NET area 0.0.0.0
  ip pim sparse-mode
icam monitor scale

line console
line vty
boot nxos bootflash:/nxos.9.3.5.bin sup-1
router ospf UNDERLAY-NET
  router-id 192.168.0.1
router bgp 65030
  router-id 192.168.100.1
  address-family ipv4 unicast
  address-family l2vpn evpn
  neighbor 192.168.100.11
    remote-as 65030
    update-source loopback100
    address-family l2vpn evpn
      send-community extended
      route-reflector-client
  neighbor 192.168.100.12
    remote-as 65030
    update-source loopback100
    address-family l2vpn evpn
      send-community extended
      route-reflector-client
  neighbor 192.168.100.13
    remote-as 65030
    update-source loopback100
    address-family l2vpn evpn
```

```
      send-community extended
      route-reflector-client
```

Leaf-11

```
Leaf-11# sh run

!Command: show running-config
!No configuration change since last restart
!Time: Thu Jul 29 12:12:27 2021

version 9.3(5) Bios:version
hostname Leaf-11
vdc Leaf-11 id 1
  limit-resource vlan minimum 16 maximum 4094
  limit-resource vrf minimum 2 maximum 4096
  limit-resource port-channel minimum 0 maximum 511
  limit-resource u4route-mem minimum 248 maximum 248
  limit-resource u6route-mem minimum 96 maximum 96
  limit-resource m4route-mem minimum 58 maximum 58
  limit-resource m6route-mem minimum 8 maximum 8
  limit-resource vni_bd minimum 4096 maximum 4096

nv overlay evpn
feature ospf
feature bgp
feature pim
feature fabric forwarding
feature interface-vlan
feature vn-segment-vlan-based
feature nv overlay

username                  admin                  password                  5
$5$GINJDF$TTUDQXpzOy5wSjAcSZPHoAQQm2epkKbYi8kqRGTdjf2
 role network-admin
ip domain-lookup
copp profile strict
snmp-server       user       admin       network-admin       auth       md5
0x8f513ee9fa080d6ad6d20b5f962c6cc1
 priv 0x8f513ee9fa080d6ad6d20b5f962c6cc1 localizedkey
rmon event 1 log trap public description FATAL(1) owner PMON@FATAL
rmon event 2 log trap public description CRITICAL(2) owner PMON@CRITICAL
rmon event 3 log trap public description ERROR(3) owner PMON@ERROR
rmon event 4 log trap public description WARNING(4) owner PMON@WARNING
rmon event 5 log trap public description INFORMATION(5) owner PMON@INFO

fabric forwarding anycast-gateway-mac 0001.0001.0001
ip pim rp-address 192.168.238.238 group-list 224.0.0.0/4
ip pim ssm range 232.0.0.0/8
vlan 1,10,77
vlan 10
  name VLAN10-mapped-to-VNI10000
  vn-segment 10000
vlan 77
  name NWKT
```

```
  vn-segment 10077

vrf context NWKT
  vni 10077
  rd auto
  address-family ipv4 unicast
    route-target both auto
    route-target both auto evpn
vrf context management
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 Vlan1

interface Vlan10
  no shutdown
  vrf member NWKT
  ip address 172.16.30.1/24
  fabric forwarding mode anycast-gateway

interface Vlan77
  no shutdown
  mtu 9216
  vrf member NWKT
  ip forward

interface nve1
  no shutdown
  host-reachability protocol bgp
  source-interface loopback50
  member vni 10000
    mcast-group 238.0.0.10
  member vni 10077 associate-vrf

interface Ethernet1/1
  no switchport
  medium p2p
  ip unnumbered loopback0
  ip ospf network point-to-point
  ip router ospf UNDERLAY-NET area 0.0.0.0
  ip pim sparse-mode
  no shutdown

interface Ethernet1/2
  switchport access vlan 10



interface mgmt0
  vrf member management

interface loopback0
  ip address 192.168.0.11/32
```

```
  ip router ospf UNDERLAY-NET area 0.0.0.0
  ip pim sparse-mode

interface loopback50
  ip address 192.168.50.11/32
  ip router ospf UNDERLAY-NET area 0.0.0.0
  ip pim sparse-mode

interface loopback100
  ip address 192.168.100.11/32
  ip router ospf UNDERLAY-NET area 0.0.0.0
  ip pim sparse-mode
icam monitor scale

line console
line vty
boot nxos bootflash:/nxos.9.3.5.bin sup-1
router ospf UNDERLAY-NET
  router-id 192.168.0.11
router bgp 65030
  router-id 192.168.100.11
  address-family ipv4 unicast
  address-family l2vpn evpn
  neighbor 192.168.100.1
    remote-as 65030
    update-source loopback100
    address-family l2vpn evpn
      send-community extended
  vrf NWKT
    address-family ipv4 unicast
      advertise l2vpn evpn
evpn
  vni 10000 l2
    rd auto
    route-target import auto
    route-target export auto
```

Chapter 4: DC Control-Plane: BGP EVPN - MAC Address

Introduction

We have seen in previous chapters how the IP address 172.16.100.10 assigned to EP1 is advertised within the LISP domain and advertised as an aggregate route all the way down to Leaf-11 in the BGP EVPN domain. This chapter first explains how the EP3 's IP address 172.16.30.3 is first advertised by Leaf-11 as BGP EVPN MAC Advertisement Route (Route-Type 2) via Spine-1 to Border-Leaf-13. Next, you will learn how Border-Leaf-13 advertises the aggregate route 172.16.30.0/24 to SD-WAN edge device vEdge-2. The last section briefly shows how the routing information is propagated over the SD-WAN. The BGP EVPN NLRI MAC Advertisement Route carries to MPLS Labels which identify L2VN (10000) and L3VN (10077). In our example, VLAN 10 is part of the VRF NWKT and it is attached to L2VN 10000. L3VNI for VRF NWKT is 10077.

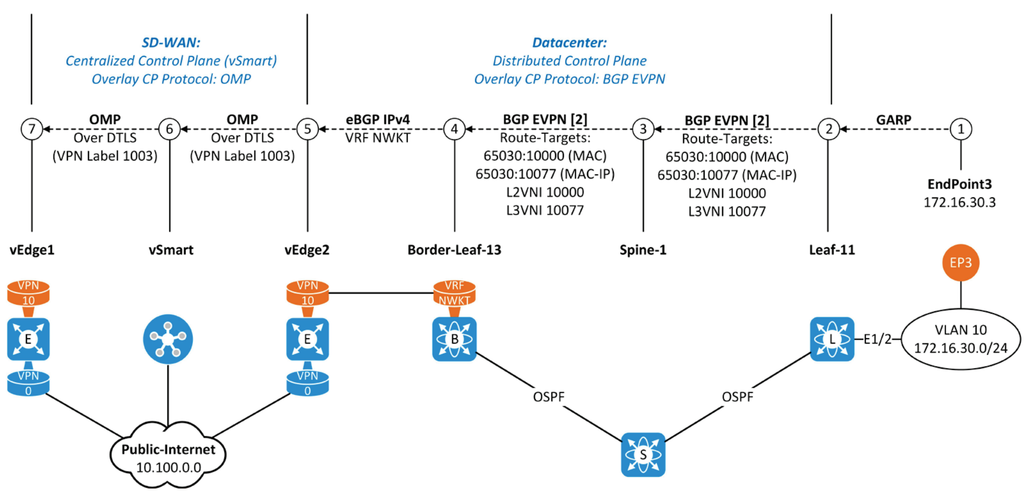


Figure 4-1: *Overall Control-Plane Operation: BGP EVPN to OMP to LISP.*

Local Leaf Processes: MAC Address

When EP3 comes up, it verifies that no one else is using its IP address 172.16.30.3. It does this by sending a GARP message by using the Broadcast MAC address as a destination. Leaf-11 learns both MAC and IP addresses from the ingress GARP message. It stores the MAC address into the VLAN 10 specific MAC table (A1). Besides, Leaf-11 installs the IP address into the ARP Cache (A2). The MAC address information is programmed from the MAC table to EVPN Instance 10000 L2RIB (B1). Also, the IP information is exported from the ARP table into the L2RIB as MAC-IP entry (B2). The reason why we have these two L2 tables is simple, we can't export information straight from either MAC or ARP tables into the BGP process. Next, Leaf-11 installs MAC and IP information into BGP Loc-RIB as two separate entries, one for the MAC (C1) and the other for the MAC-IP (C2) binding information. Leaf-11 builds Spine-1 specific Adj-RIB-Out entries with a Next-Hop Path-Attribute and then sends the BGP Update message to Spine-1. Spine-1 is unaware of any customer EVPN Instances, meaning it doesn't have any Route-Target import/export policy, so it just forwards the BGP Update message without installing the information into its local tables. It, however, adds BGP Originator-Id and Cluster-Id Path-Attributes to the BGP Update message.

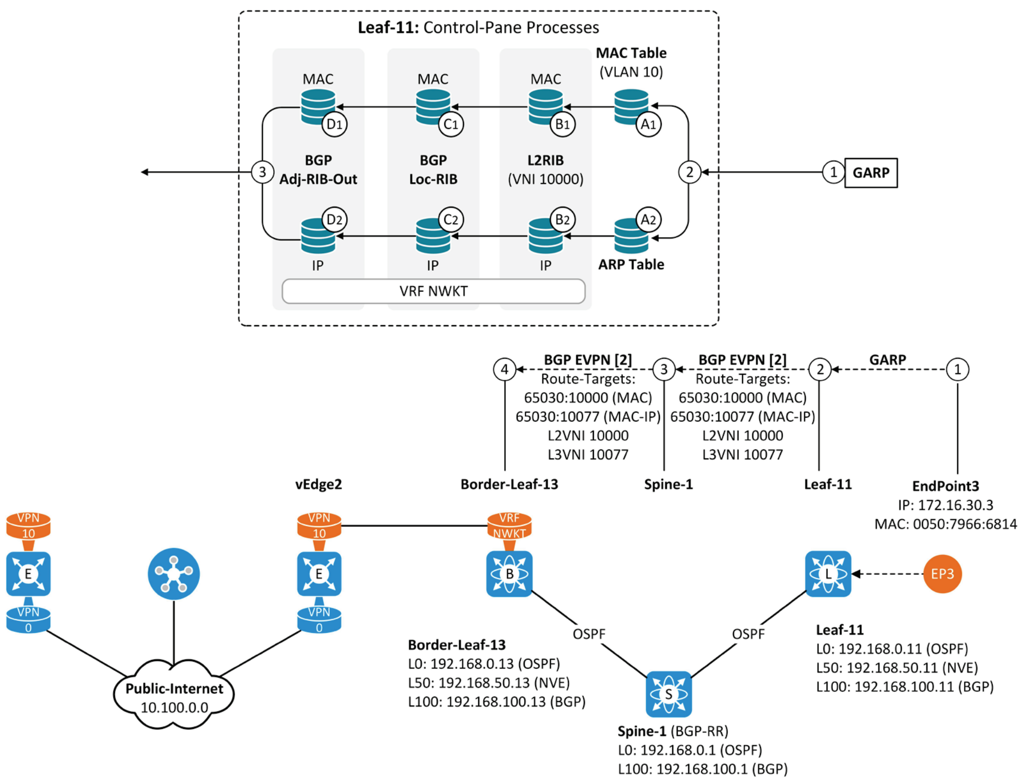


Figure 4-2: Local Leaf *Leaf-11 Processes.*

Phase A1: MAC Address Table

Example 4-1 shows that Leaf-11 has installed the MAC address 0050.7966.6814 (EP3) into the VLAN 10 MAC table with port E1/2.

```
Leaf-11# show mac address-table vlan 10
Legend:
        * - primary entry, G - Gateway MAC, (R) - Routed MAC, O - Overlay MAC
        age - seconds since last seen,+ - primary entry using vPC Peer-Link,
        (T) - True, (F) - False, C - ControlPlane MAC, ~ - vsan
   VLAN     MAC Address      Type      age     Secure NTFY Ports
---------+-----------------+--------+---------+------+----+------------------
*   10     0050.7966.6814   dynamic  0         F      F    Eth1/2
G   10     5010.0000.1b08   static   -         F      F    sup-eth1(R)
```

Example 4-1: *Leaf-11 VLAN 10 MAC Address Table.*

I have added example 4-2 just to show that MAC tables can also be viewed per L2VNI.

```
Leaf-11# show mac address-table vni 10000
Legend:
        * - primary entry, G - Gateway MAC, (R) - Routed MAC, O - Overlay MAC
        age - seconds since last seen,+ - primary entry using vPC Peer-Link,
        (T) - True, (F) - False, C - ControlPlane MAC, ~ - vsan
   VLAN     MAC Address      Type      age     Secure NTFY Ports
---------+-----------------+--------+---------+------+----+------------------
*   10     0050.7966.6814   dynamic  0         F      F    Eth1/2
```

Example 4-2: *Leaf-11 VNI 10000 MAC Address Table.*

Phase A2: ARP Table

Example 4-3 shows the VRF NWKTs ARP table where we have an IP address 172.16.30.3 bind to MAC address 0050.7966.6814.

```
Leaf-11# show ip arp vrf NWKT

Flags: * - Adjacencies learnt on non-active FHRP router
       + - Adjacencies synced via CFSoE
       # - Adjacencies Throttled for Glean
       CP - Added via L2RIB, Control plane Adjacencies
       PS - Added via L2RIB, Peer Sync
       RO - Re-Originated Peer Sync Entry
       D - Static Adjacencies attached to down interface

IP ARP Table for context NWKT
Total number of entries: 1
Address         Age       MAC Address     Interface       Flags
172.16.30.3     00:16:49  0050.7966.6814  Vlan10
```

Example 4-3: *Leaf-11 ARP Table.*

Phase B1: L2RIB MAC

Example 4-4 illustrates the L2RIB concerning MAC entry. Topology 10 points to L2VN 10000, and the L-Flag indicates that this is a local route. The Next-Hop is interface E1/2.

```
Leaf-11# show l2route evpn mac evi 10

Flags -(Rmac):Router MAC (Stt):Static (L):Local (R):Remote (V):vPC link
(Dup):Duplicate (Spl):Split (Rcv):Recv (AD):Auto-Delete (D):Del Pending
(S):Stale (C):Clear, (Ps):Peer Sync (O):Re-Originated (Nho):NH-Override
(Pf):Permanently-Frozen, (Orp): Orphan

Topology    Mac Address    Prod   Flags         Seq No     Next-Hops
----------- -------------- ------ ------------- ---------- ---------------------
10          0050.7966.6814 Local  L,            0          Eth1/2
```

Example 4-4: *Leaf-11 L2RIB MAC.*

Phase B2: L2RIB MAC-IP

Example 4-5 shows the MAC-IP entry in L2RIB. It is installed into the L2RIB by the Host Mobility Manager (HMM).

```
Leaf-11# show l2route mac-ip topology 10
Flags -(Rmac):Router MAC (Stt):Static (L):Local (R):Remote (V):vPC link
(Dup):Duplicate (Spl):Split (Rcv):Recv(D):Del Pending (S):Stale (C):Clear
(Ps):Peer Sync (Ro):Re-Originated (Orp):Orphan
Topology    Mac Address    Host IP      Prod   Flags  Seq No   Next-Hops
----------- -------------- ------------ ------ ------ -------- -----------------
10          0050.7966.6814 172.16.30.3   HMM    L,    0        Local
```

Example 4-5: *Leaf-11 L2RIB MAC-IP.*

Phase C1-2, D1-2: BGP Table

Example 4-6 shows BGP L2VPN EVPN entries about MAC and MAC-IP in Leaf-11's BGP table. The Route-Distinguisher (RD) is a part of the address, and it makes it possible to use overlapping MAC and IP addresses within VRF/Tenant. The IP address part (192.168.100.11) in BGP Loc-RIB is the BGP RID of the local device. The latter part comes from the base value 32676 plus VLAN Id 10. The Next-Hop is the IP attached to Loopback 50 that, in turn, is bind to NVE Interface (tunnel interface). The address [2]:[0]:[0]:[48]:[0050.7966.6814]:[0]:[0.0.0.0]/216 tells that this MAC Advertisement Route (Route Type [2]), following [0]:[0] means that we are not using ESI Multihoming and that there is no Ethernet Tag-Id. the [48] is the length of the following MAC address [0050.7966.6814]. The last part [0]:[0.0.0.0] means that this is MAC-Only. The /216 is the address bit count.

The Extended Community Received Label 10000 is used in Data-Plane as L2VN Identifier, which means that all Intra-VN (switched) traffic within VLAN 10 carries this label. The Route-Target Extended Community 65030:10000 is auto-generated from BGP ASN and L2VNI. The Encapsulation value eight means VXLAN encapsulation will be used in Data-Plane.

The second entry [2]:[0]:[0]:[48]:[0050.7966.6814]:[32]:[172.16.30.3]/272 describes the MAC-IP binding information. It is identical to the first entry with three additional Extended Communities. The first one is an additional Label 10077 used in Data-Plane for Inter-VN (routed) traffic between different VNs. The second one, the additional Route-Target 65030:10077 that is used for importing routes into BGP EVPN L3VNI. It is needed in order to do routing between subnets/VNs. The last one is the Route MAC address that is needed because VXLAN is MAC-in-IP/UDP tunneling mechanism and the inner destination MAC address in routed traffic uses Router MAC address (host MAC addresses are not available in routed traffic). I will go through the Inter-VN routing process later in the Data-Plane chapter.

```
Leaf-11# show bgp l2vpn evpn 0050.7966.6814
BGP routing table information for VRF default, address family L2VPN EVPN
Route Distinguisher: 192.168.100.11:32777    (L2VNI 10000)
BGP routing table entry for [2]:[0]:[0]:[48]:[0050.7966.6814]:[0]:[0.0.0.0]/216, version
7
Paths: (1 available, best #1)
Flags: (0x000102) (high32 00000000) on xmit-list, is not in l2rib/evpn

  Advertised path-id 1
  Path type: local, path is valid, is best path, no labeled nexthop
  AS-Path: NONE, path locally originated
    192.168.50.11 (metric 0) from 0.0.0.0 (192.168.100.11)
      Origin IGP, MED not set, localpref 100, weight 32768
      Received label 10000
      Extcommunity: RT:65030:10000 ENCAP:8

  Path-id 1 advertised to peers:
    192.168.100.1
BGP routing table entry for [2]:[0]:[0]:[48]:[0050.7966.6814]:[32]:[172.16.30.3]/272,
version 6
Paths: (1 available, best #1)
Flags: (0x000102) (high32 00000000) on xmit-list, is not in l2rib/evpn

  Advertised path-id 1
  Path type: local, path is valid, is best path, no labeled nexthop
  AS-Path: NONE, path locally originated
    192.168.50.11 (metric 0) from 0.0.0.0 (192.168.100.11)
      Origin IGP, MED not set, localpref 100, weight 32768
      Received label 10000 10077
      Extcommunity: RT:65030:10000 RT:65030:10077 ENCAP:8 Router MAC:5010.0000.1b08

  Path-id 1 advertised to peers:
    192.168.100.1
```

Example 4-6: *Leaf-11 BGP Tables.*

Capture 4-1 shows how Leaf-11 encodes information into the BGP Update message. Note that the L2VN Identifier is encoded into MPLS 1 field and the L3VN Identifier is encoded into the MPLS 2 field. In order to get the VNI from the MPLS fields, we need to convert Binary to HEX and HEX to Decimal. So, the Binary number 0010 0111 0001 0000 (last four digits not shown in capture) in HEX representation is 2710, which is 10 000 converted to Decimal. The Decimal value 625 for MPLS Label 1 comes from the Binary value 0010 0111 0001 (last four digits excluded) that in HEX format is 0271 (=Decimal format 625). The same conversion process also applies to MPLS Label 2. The last Extended Community with an Unknown subtype is the Router MAC address.

```
Internet Protocol Version 4, Src: 192.168.100.11, Dst: 192.168.100.1
Transmission Control Protocol, Src Port: 40654, Dst Port: 179, Seq: 20, Ack: 20, Len:
127
Border Gateway Protocol - UPDATE Message
    Marker: ffffffffffffffffffffffffffffffff
    Length: 127
    Type: UPDATE Message (2)
    Withdrawn Routes Length: 0
    Total Path Attribute Length: 104
    Path attributes
        Path Attribute - MP_REACH_NLRI
            Flags: 0x90, Optional, Extended-Length, Non-transitive, Complete
            Type Code: MP_REACH_NLRI (14)
            Length: 51
            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 (42 bytes)
                EVPN NLRI: MAC Advertisement Route
                    Route Type: MAC Advertisement Route (2)
                    Length: 40
                    Route Distinguisher: 0001c0a8640b8009 (192.168.100.11:32777)
                    ESI: 00:00:00:00:00:00:00:00:00:00
                    Ethernet Tag ID: 0
                    MAC Address Length: 48
                    MAC Address: Private_66:68:14 (00:50:79:66:68:14)
                    IP Address Length: 32
                    IPv4 address: 172.16.30.3
                    0000 0000 0010 0111 0001 .... = MPLS Label 1: 625
                    0000 0000 0010 0111 0101 .... = MPLS Label 2: 629
        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: 32
            Carried extended communities: (4 communities)
                Route Target: 65030:10000 [Transitive 2-Octet AS-Specific]
                Route Target: 65030:10077 [Transitive 2-Octet AS-Specific]
                Encapsulation: VXLAN Encapsulation [Transitive Opaque]
                Unknown subtype 0x03: 0x5010 0x0000 0x1b08 [Transitive EVPN]
```

Capture 4-1: *Leaf-11 BGP Update to Spine-1.*

Remote Leaf Processes: MAC Address

Border-Leaf-13 receives the BGP update carrying both MAC and MAC-IP EVPN NLRI from Spine-1. It installs the received information into the BGP Adj-RIB-In, and from there it imports L2VNI 10000 information (MAC and MAC-IP address) into Loc-RIB based on RT 65030:10000. It also imports L3VNI information (IP) based on the RT 65030:10077 into Loc-RIB. The MAC address information is installed first into L2RIB and from there to VLAN 10 MAC address table. The MAC address information is used for Intra-VN switching. In addition, Border-Leaf-13 installs IP/MAC binding into local ARP Cache from the L2VNI 10000 MAC-IP entry. Based on L3VNI RT 65030:10077, Border-Leaf-13 imports routing information also into L3RIB as a host route. From there, the route is exported to vEdge-2 specific Adj-RIB-Out. The Route-Target is taken from the VRF NWKT configuration. Then Border-Leaf-13 builds the BGP update message and sends it to vEdge-2. Note that we are using route aggregation in Border-Leaf-13, so the NLRI is subnet 172.16.30.0/24.

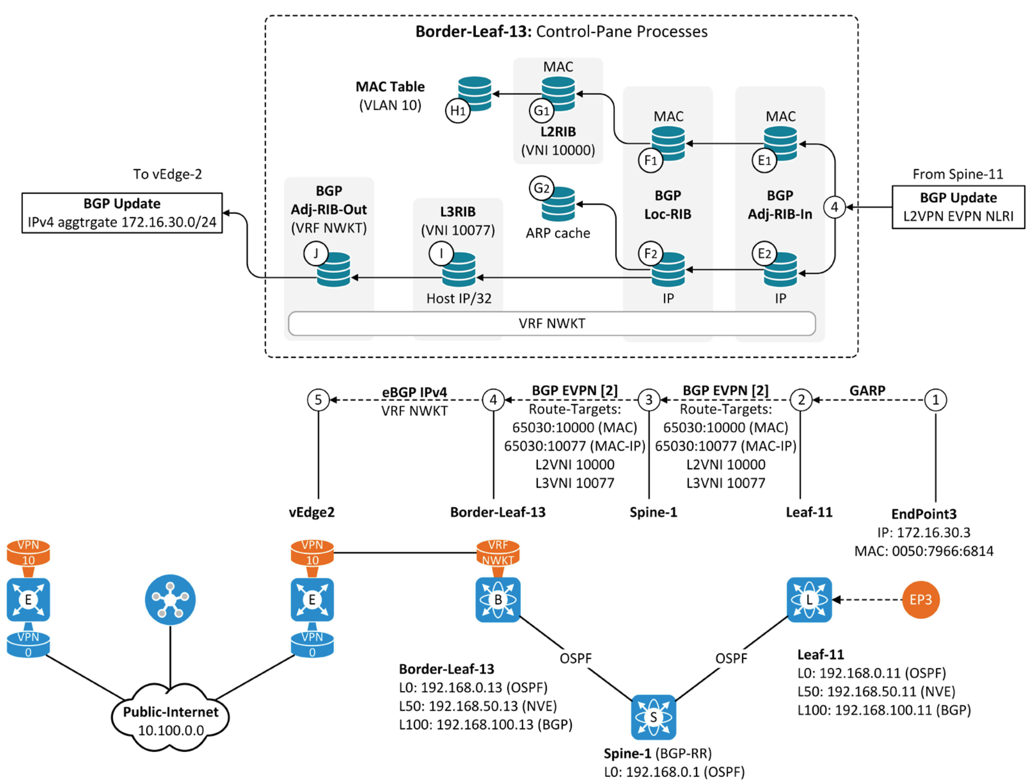


Figure 4-3: *Remote Leaf Border-Leaf-13 Processes.*

Phase E1-2, F1-2: BGP Table

Example 4-7 shows the Border-Leaf-13's BGP table EVPN entries about IP address 172.16. 30.3. Note, by using the IP address instead of the MAC address in the show command, we get only L2VN MAC-IP entries, and MAC-only entries are excluded from the output. The first entry describes unmodified EVPN NLRI in BGP Adj-RIB-In. It has the original RD 192.168.100.11:32777 set by Leaf-11. The second entry shows the L2VNI 10000 specific BGP Loc-RIB information. The import policy is based on Route-Target 65030:10000. This information is later installed into L2RIB from where it is programmed into the VLAN 10 MAC address table. The RD is changed from 192.18.100.11:32777 to 192.168.100.13:32777 during the import process. In addition, MAC-IP binding information is exported to the local ARP Cache. The third highlighted entry shows how the information is installed into BGP Loc-RIB as the L3VNI entry. The Import is based on the RT 65030:10077 and the information is also used when the route is installed into the VRF NWKT RIB.

```
Border-leaf-13# show bgp l2vpn evpn 172.16.30.3
BGP routing table information for VRF default, address family L2VPN EVPN
Route Distinguisher: 192.168.100.11:32777
BGP routing table entry for [2]:[0]:[0]:[48]:[0050.7966.6814]:[32]:[172.16.30.3]/272,
version 5
Paths: (1 available, best #1)
Flags: (0x000202) (high32 00000000) on xmit-list, is not in l2rib/evpn, is not in HW

  Advertised path-id 1
  Path type: internal, path is valid, is best path, no labeled nexthop
             Imported to 3 destination(s)
             Imported paths list: NWKT L3-10077 L2-10000
  AS-Path: NONE, path sourced internal to AS
    192.168.50.11 (metric 81) from 192.168.100.1 (192.168.100.1)
      Origin IGP, MED not set, localpref 100, weight 0
      Received label 10000 10077
      Extcommunity: RT:65030:10000 RT:65030:10077 ENCAP:8 Router MAC:5010.0000.1b08
      Originator: 192.168.100.11 Cluster list: 192.168.100.1

  Path-id 1 not advertised to any peer

Route Distinguisher: 192.168.100.13:32777    (L2VNI 10000)
BGP routing table entry for [2]:[0]:[0]:[48]:[0050.7966.6814]:[32]:[172.16.30.3]/272,
version 6
Paths: (1 available, best #1)
Flags: (0x000212) (high32 00000000) on xmit-list, is in l2rib/evpn, is not in HW

  Advertised path-id 1
  Path type: internal, path is valid, is best path, no labeled nexthop, in rib
             Imported from
192.168.100.11:32777:[2]:[0]:[0]:[48]:[0050.7966.6814]:[32]:[172.16.30.3]/272
  AS-Path: NONE, path sourced internal to AS
    192.168.50.11 (metric 81) from 192.168.100.1 (192.168.100.1)
      Origin IGP, MED not set, localpref 100, weight 0
      Received label 10000 10077
      Extcommunity: RT:65030:10000 RT:65030:10077 ENCAP:8 Router MAC:5010.0000.1b08
      Originator: 192.168.100.11 Cluster list: 192.168.100.1

  Path-id 1 not advertised to any peer
```

```
Route Distinguisher: 192.168.100.13:3    (L3VNI 10077)
BGP routing table entry for [2]:[0]:[0]:[48]:[0050.7966.6814]:[32]:[172.16.30.3]/272,
version 7
Paths: (1 available, best #1)
Flags: (0x000202) (high32 00000000) on xmit-list, is not in l2rib/evpn, is not in HW

  Advertised path-id 1
  Path type: internal, path is valid, is best path, no labeled nexthop
             Imported from
192.168.100.11:32777:[2]:[0]:[0]:[48]:[0050.7966.6814]:[32]:[172.16.30.3]/272
  AS-Path: NONE, path sourced internal to AS
    192.168.50.11 (metric 81) from 192.168.100.1 (192.168.100.1)
      Origin IGP, MED not set, localpref 100, weight 0
      Received label 10000 10077
      Extcommunity: RT:65030:10000 RT:65030:10077 ENCAP:8 Router MAC:5010.0000.1b08
      Originator: 192.168.100.11 Cluster list: 192.168.100.1

  Path-id 1 not advertised to any peer
```
Example 4-7: Border-*Leaf-13 BGP Tables.*

Phase I: L3RIB

Example 4-8 verifies that route is exported from the BGP Loc-RIB into the VRF NWKT's RIB.

```
Border-leaf-13# show ip route 172.16.30.3 vrf NWKT
IP Route Table for VRF "NWKT"
<snipped>

172.16.30.3/32, ubest/mbest: 1/0
    *via 192.168.50.11%default, [200/0], 01:32:33, bgp-65030, internal, tag 65030,
segid: 10077 tunnelid: 0xc0a8320b encap: VXLAN
```
Example 4-8: Border-*Leaf-13 RIB.*

Phase G1: L2RIB

Example 4-9 shows that route is exported from the BGP Loc-RIB into L2RIB. Topology Id points to VLAN 10 that is attached to VN10000. Information is produced by BGP and the next-hop is the tunnel interface NVE1 of Leaf-11.

```
Border-leaf-13# show l2route evpn mac evi 10

Flags -(Rmac):Router MAC (Stt):Static (L):Local (R):Remote (V):vPC link
(Dup):Duplicate (Spl):Split (Rcv):Recv (AD):Auto-Delete (D):Del Pending
(S):Stale (C):Clear, (Ps):Peer Sync (O):Re-Originated (Nho):NH-Override
(Pf):Permanently-Frozen, (Orp): Orphan

Topology    Mac Address    Prod   Flags    Seq No  Next-Hops
----------- -------------- ------ -------- ------- ----------------------------
10          0050.7966.6814 BGP    SplRcv   0       192.168.50.11 (Label: 10000)
```
Example 4-9: Border-*Leaf-13 L2RIB.*

Phase H1: MAC Address Table

Example 4-10 shows that the MAC address 0050.7966.6814 is exported from the L2RIB into the MAC address table. The Port describes the next-hop address and that the frames switched to the network will be sent via the NVE1 tunnel interface.

```
Border-leaf-13# show mac address-table vlan 10
Legend:
        * - primary entry, G - Gateway MAC, (R) - Routed MAC, O - Overlay MAC
        age - seconds since last seen,+ - primary entry using vPC Peer-Link,
        (T) - True, (F) - False, C - ControlPlane MAC, ~ - vsan
   VLAN     MAC Address      Type      age     Secure NTFY Ports
---------+-----------------+--------+---------+------+----+------------------
C   10     0050.7966.6814   dynamic  0         F      F    nve1(192.168.50.11)
G   10     5012.0000.1b08   static   -         F      F    sup-eth1(R)
```

Example 4-10: *Border-Leaf-13 MAC address table.*

In order to Border-Leaf -13 be able to switch frames to destination MAC address, it has to know what tunnel encapsulation is used with the next-hop device 192.168.50.11 (Leaf-11). This information is stored in the Recursive Next-Hop Database (RNH DB). Example 4-11 shows that the encapsulation type is VLAN and the VN for egress frames is 10000. The Peer-MAC is not used when switching frames because the inner header (payload) has sender and receiver MAC addresses. Anyway, MAC address information is needed in the case where one of the locally connected hosts wants to send data to a remote host in the same L2 domain and that is not our case because Border-Leaf-13 doesn't have any local hosts.

```
Border-leaf-13# show nve internal bgp rnh database vni 10000
--------------------------------------------
Total peer-vni msgs recvd from bgp: 1
Peer add requests: 2
Peer update requests: 0
Peer delete requests: 0
Peer add/update requests: 2
Peer add ignored (peer exists): 0
Peer update ignored (invalid opc): 0
Peer delete ignored (invalid opc): 0
Peer add/update ignored (malloc error): 0
Peer add/update ignored (vni not cp): 0
Peer delete ignored (vni not cp): 0
--------------------------------------------
Showing BGP RNH Database, size : 2 vni 10000

Flag codes: 0 - ISSU Done/ISSU N/A       1 - ADD_ISSU_PENDING
            2 - DEL_ISSU_PENDING         3 - UPD_ISSU_PENDING


VNI     Peer-IP        Peer-MAC       Tunnel-ID  Encap  (A /S ) Flags PT  Egress VNI
10000   192.168.50.11  0000.0000.0000 0x0        vxlan  (1 /0 ) 0     FAB 10000
```

Example 4-11: *Border-Leaf-13 Recursive Next-Hop Database for VNI 10000.*

Capture 4-2 shows the BGP Update message, which Border-Leaf-13 sent to the local SD-WAN edge device vEdge-2. Instead of advertising individual host IP address 172.16.30.3/32, we are using an aggregate route 172.16.30.0/24.

```
Internet Protocol Version 4, Src: 172.16.20.13, Dst: 172.16.20.1
Transmission Control Protocol, Src Port: 37301, Dst Port: 179, Seq: 1, Ack: 1, Len: 61
Border Gateway Protocol - UPDATE Message
    Marker: ffffffffffffffffffffffffffffffff
    Length: 61
    Type: UPDATE Message (2)
    Withdrawn Routes Length: 0
    Total Path Attribute Length: 34
    Path attributes
        Path Attribute - ORIGIN: IGP
        Path Attribute - AS_PATH: 65030
        Path Attribute - NEXT_HOP: 172.16.20.13
        Path Attribute - AGGREGATOR: AS: 65030 origin: 172.16.30.1
        Path Attribute - ATOMIC_AGGREGATE
    Network Layer Reachability Information (NLRI)
        172.16.30.0/24
            NLRI prefix length: 24
            NLRI prefix: 172.16.30.0
```

Capture 4-2: *Border-Leaf-13 BGP Update to vEdge.*

SD-WAN OMP Processes

This section briefly recaps the OMP operation. You can go back to chapter two (SD-WAN Control-Plane: OMP) to the check details. The local SD-WAN device vEdge-2 installs the information received from the BGP Update in its BGP Table and from there into the RIB. Then it exports the route to the OMP process and sends the routing update to the centralized SD-WAN Control-Plane node vSmart over the DTLS tunnel. vSmart forwards an unmodified update to vEdge-1.

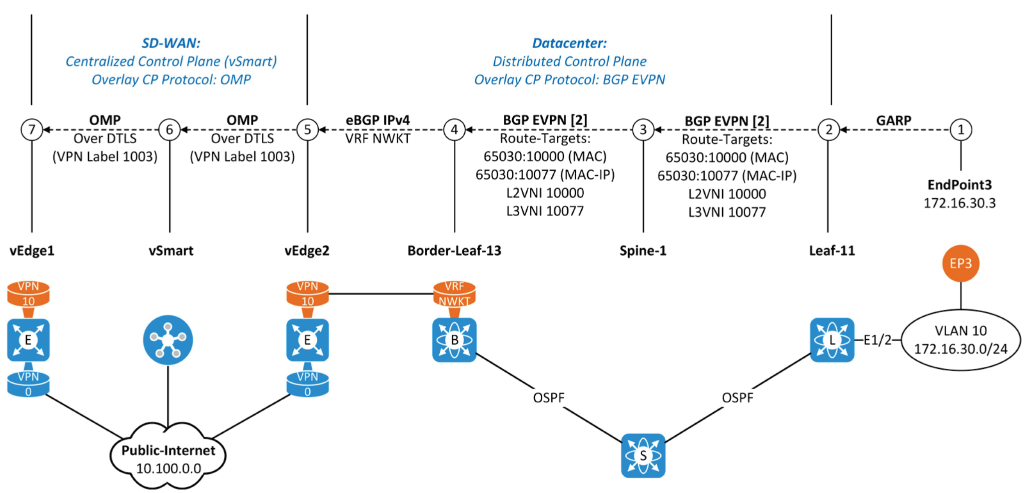


Figure 4-4: *SD-WAN Processes.*

Example 4-12 shows the routes that vEdge-2 has learned from BGP routing updates. The next-hop is Border-Leaf-13.

```
vEdge2# show ip routes bgp detail
<snipped>

""--------------------------------------------
 VPN 10      PREFIX 172.16.30.0/24
--------------------------------------------
 proto            bgp
 proto-sub-type   e
 distance         20
 metric           0
 uptime           0:00:45:18
 nexthop-ifname   ge0/2
 nexthop-addr     172.16.20.13
 status           F,S
```

Example 4-12: *vEdge-2's RIB.*

Example 4-13 shows the OMP route to 172.16.30.0/24 (VPN 10). Packets sent over SD-WAN are using Label 1003 as VN-Id. The tloc attribute (Transport Locator) describes the vEdge-2's System-Id (10.100.100.2), transport network (public-internet), and tunnel encapsulation (GRE). vEdge-2 sends this information to vSmart (10.100.100.13).

```
vEdge2# show omp routes received detail | exclude not\ set | nomore

---------------------------------------------------
omp route entries for vpn 10 route 172.16.30.0/24
---------------------------------------------------
            RECEIVED FROM:
peer            0.0.0.0
path-id         37
label           1003
status          C,Red,R
    Attributes:
     originator       10.100.100.102
     type             installed
     tloc             10.100.100.102, public-internet, gre
     overlay-id        1
     site-id          20
     origin-proto     eBGP
     origin-metric    0
            ADVERTISED TO:
peer    10.100.100.13
    Attributes:
     originator       10.100.100.102
     label            1003
     path-id          37
     tloc             10.100.100.102, public-internet, gre
     site-id          20
     overlay-id        1
     origin-proto     eBGP
     origin-metric    0
```

Example 4-13: *vEdge-2's OMP Table.*

Example 4-14 verifies that the remote SD-WAN device vEdge-1 has received the OMP update from vSmart.

```
vEdge-1# show omp routes received detail | exclude not\ set | nomore

---------------------------------------------------
omp route entries for vpn 10 route 172.16.30.0/24
---------------------------------------------------
            RECEIVED FROM:
peer            10.100.100.13
path-id         4
label           1003
status          C,I,R
    Attributes:
     originator       10.100.100.102
     type             installed
     tloc             10.100.100.102, public-internet, gre
     overlay-id        1
     site-id          20
     origin-proto     eBGP
     origin-metric    0
```

Example 4-14: *vEdge-1's OMP Table.*

Example 4-15 shows that the remote SD-WAN device vEdge-1 has installed the routing information from the OMP table into the RIB.

```
vEdge-1# show ip routes vpn 10 172.16.30.0/24 detail
Codes Proto-sub-type:
  IA -> ospf-intra-area, IE -> ospf-inter-area,
  E1 -> ospf-external1, E2 -> ospf-external2,
  N1 -> ospf-nssa-external1, N2 -> ospf-nssa-external2,
  e -> bgp-external, i -> bgp-internal
Codes Status flags:
  F -> fib, S -> selected, I -> inactive,
  B -> blackhole, R -> recursive, L -> import

""--------------------------------------------
 VPN 10      PREFIX 172.16.30.0/24
--------------------------------------------
 proto           omp
 distance        250
 metric          0
 uptime          0:00:52:57
 tloc-ip         10.100.100.102
 tloc-color      public-internet
 tloc-encap      gre
 nexthop-label   1003
 status          F,S
```

Example 4-15: *vEdge-1's RIB.*

Example 4-16 shows the TLOC Database, which vEdge-1 uses for resolving the tunnel destination IP address for packets to network 172.16.30.0/24 over the Public-Internet transport network.

```
vEdge1# show omp tlocs | exclude not | nomore
---------------------------------------------------
tloc entries for 10.100.100.102
                 public-internet
                 gre
---------------------------------------------------
            RECEIVED FROM:
peer            10.100.100.13
status          C,I,R
    Attributes:
     attribute-type    installed
     encap-key         0
     public-ip         10.100.0.102
     public-port       0
     private-ip        10.100.0.102
     private-port      0
     <snipped>
     bfd-status        up
     site-id           10
     <snipped>
     restrict          1
     <snipped>
```

Example 4-16: *vEdge-1's TLOC Table.*

This chapter explained how the routing information originated by Leaf-11 is advertised from the BGP EVPN domain to local SD-WAN. It also briefly recaps how the information is advertised over the SD-WAN to the remote device vEdge-1. In the next chapter, I will explain the Control-Plane processes of how the information is advertised to LISP domain Control-Plane node MapSrv-22.

Chapter 5: LISP Control-Plane: External Registration

Introduction

This chapter introduces how Border-PxTR-13 registers the external IP prefix 172.16.30.0/24 received as a BGP update from vEdge-1 to MapSrv-22 using LISP Map-register messages. Chapter 2 explains the LISP RLOC-to-EID mapping process in detail so this chapter just briefly recaps the operation. Figure 5-1 illustrates the overall process. vEdge-1 sends a BGP Update message where it describes the NLRI for prefix 172.16.30.0/24. Border-PxTR-13 first imports the information into the LISP processes. Next, it sends a LISP Map-Register message to MapSrv-22. In addition to IP prefix information, the Map-Register message carries Locator Record information that describes the destination IP address used in the outer IP header (tunnel header) when devices route IP packets towards the advertised subnet.

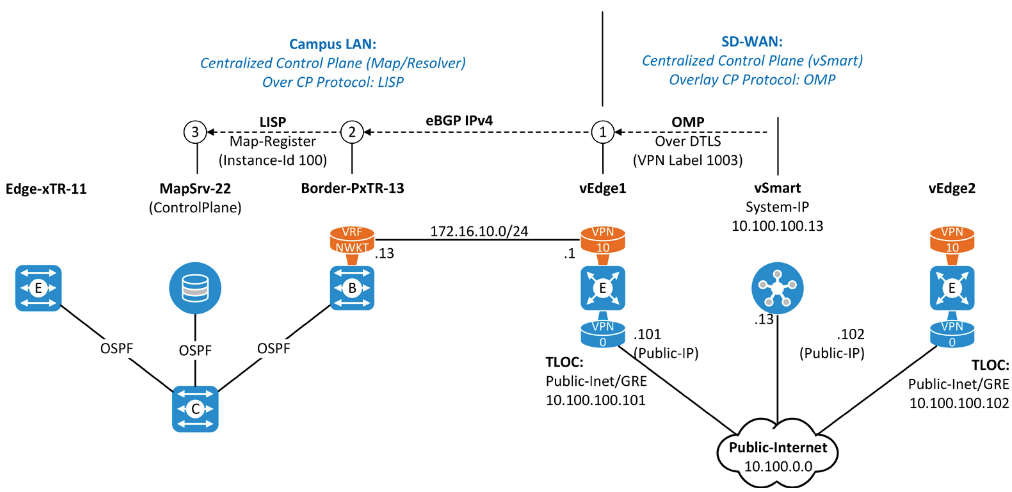


Figure 5-1: *Overall Control-Plane Operation: OMP to LISP.*

Phases 1-2: OMP to BGP in vEdge-1

Capture 5-1 shows the BGP Update message from vEdge-1 to Border-PxTR-13. The Route-Origin BGP extended community attribute carries the Site-Id of advertising device vEdge-1. BGP uses this attribute for preventing other edge devices from advertising the route back to SD-WAN edge devices.

```
Internet Protocol Version 4, Src: 172.16.10.1, Dst: 172.16.10.13
Transmission Control Protocol, Src Port: 51909, Dst Port: 179, Seq: 39, Ack: 39, Len:
66
Border Gateway Protocol - UPDATE Message
    Marker: ffffffffffffffffffffffffffffffff
    Length: 66
    Type: UPDATE Message (2)
    Withdrawn Routes Length: 0
    Total Path Attribute Length: 39
    Path attributes
        Path Attribute - ORIGIN: INCOMPLETE
        Path Attribute - AS_PATH: 65100
        Path Attribute - NEXT_HOP: 172.16.10.1
        Path Attribute - MULTI_EXIT_DISC: 1000
        Path Attribute - EXTENDED_COMMUNITIES
            Flags: 0xc0, Optional, Transitive, Complete
            Type Code: EXTENDED_COMMUNITIES (16)
            Length: 8
            Carried extended communities: (1 community)
                Route Origin: 0:10 [Transitive 2-Octet AS-Specific]
                    Type: Transitive 2-Octet AS-Specific (0x00)
                    Subtype (AS2): Route Origin (0x03)
                    2-Octet AS: 0
                    4-Octet AN: 10
    Network Layer Reachability Information (NLRI)
        172.16.30.0/24
            NLRI prefix length: 24
            NLRI prefix: 172.16.30.0
```

Capture 5-1: *BGP Update from vEdge-1 to Border-PxTR-13.*

Phase 3: LISP Map-Register to MapSrv-22 from Border-PxTR-13

Capture 5-1 shows the LISP Map-Register message from Border-PxTR-13 to MapSrv-22. This message is the first one and that is why it uses unreliable transport protocol UDP. The mapping process follows the basic procedures where Border-PxTR-13 starts a TCP three-way handshake process to open the TCP connection with MapSrv-22. I have explained the complete Map-Register process in chapter 2. If we compare the LISP EID-to-RLOC registration processes to the operation of OMP, we can notice lots of similarities. The later highlighted section in capture 5-2 shows the Locator Record 1. It describes the public IP address used in the tunneling header as a destination IP address. In other words, the Locator Record section describes the location just like we describe the device location in OMP TLOC Routes (Public IP used by device). The first highlighted section Mapping

Record, in turn, describes the IP Prefix and its LISP Instance-Id (Virtual Network) just like OMP Service Route. The main difference is that LISP sends information within one message while OMP uses separate messages.

```
Internet Protocol Version 4, Src: 192.168.0.13, Dst: 192.168.0.22
User Datagram Protocol, Src Port: 4342, Dst Port: 4342
Locator/ID Separation Protocol
    0011 .... .... .... .... .... = Type: Map-Register (3)
    .... 1... .... .... .... .... = P bit (Proxy-Map-Reply): Set
    .... .0.. .... .... .... .... = S bit (LISP-SEC capable): Not set
    .... ..1. .... .... .... .... = I bit (xTR-ID present): Set
    .... ...0 .... .... .... .... = R bit (Built for an RTR): Not set
    .... .... 0000 0000 0000 000. = Reserved bits: 0x0000
    .... .... .... .... .... ...1 = M bit (Want-Map-Notify): Set
    Record Count: 1
    Nonce: 0x54faec8822e5bc8b
    Key ID: 0x0001
    Authentication Data Length: 20
    Authentication Data: 25292a07c6b2deac2811fdb92b9ff9e9d9e57b74
    Mapping Record 1, EID Prefix: [100] 172.16.30.0/24, TTL: 1440, Action: No-Action, 
Authoritative
        Record TTL: 1440
        Locator Count: 1
        EID Mask Length: 24
        000. .... .... .... = Action: No-Action (0)
        ...1 .... .... .... = Authoritative bit: Set
        .... .000 0000 0000 = Reserved: 0x000
        0000 .... .... .... = Reserved: 0x0
        .... 0000 0000 0000 = Mapping Version: 0
        EID Prefix AFI: LISP Canonical Address Format (LCAF) (16387)
        EID Prefix: [100] 172.16.30.0
            LCAF: Instance ID: 100, Address: 172.16.30.0
                LCAF Header: 00000220000a
                Instance ID: 100
                Address AFI: IPv4 (1)
                Address: 172.16.30.0
        Locator  Record  1,  Local  RLOC: 192.168.0.13,  Reachable,  Priority/Weight:  1/1, 
Multicast Priority/Weight: 1/1
            Priority: 1
            Weight: 1
            Multicast Priority: 1
            Multicast Weight: 1
            Flags: 0x0005
            AFI: IPv4 (1)
            Locator: 192.168.0.13
    xTR-ID: efd8cace5906016ad7c49f5f5cc352db
    Site-ID: 0000000000000000
```

Capture 5-2: *LISP Map-Register Message from Border-PxTR-13 to MapSrv-22.*

Example 5-1 shows the VRF 100_NWKT specific BRIB entry about subnet 172.16.30.0/24. Not that the Route-Origin extended BGP community is encoded as Site-of-Origin (SoO).

```
Border-PxTR-13#sh bgp vrf 100_NWKT 172.16.30.0
BGP routing table entry for 1:100:172.16.30.0/24, version 3
Paths: (1 available, best #1, table 100_NWKT)
  Advertised to update-groups:
     2
  Refresh Epoch 1
  65100
    172.16.10.1 (via vrf 100_NWKT) from 172.16.10.1 (10.100.100.101)
      Origin incomplete, metric 1000, localpref 100, valid, external, best
      Extended Community: SoO:0:10 RT:1:100
      mpls labels in/out 17/nolabel
      rx pathid: 0, tx pathid: 0x0
```

Example 5-1: *Border-PxTR-13's BGP Table.*

Example 5-2 verifies that the route is imported into the RIB of Border-PxTR-13.

```
Border-PxTR-13#sh ip route vrf 100_NWKT 172.16.30.0

Routing Table: 100_NWKT
Routing entry for 172.16.30.0/24
  Known via "bgp 65010", distance 20, metric 1000
  Tag 65100, type external
  Redistributing via lisp
  Last update from 172.16.10.1 00:07:16 ago
  Routing Descriptor Blocks:
  * 172.16.10.1, from 172.16.10.1, 00:07:16 ago
      Route metric is 1000, traffic share count is 1
      AS Hops 1
      Route tag 65100
      MPLS label: none
```

Example 5-2: *Border-PxTR-13's RIB.*

Example 5-3 shows that Border-PxTR-13 has installed the IP prefix 172.16.30.0/24 to its LISP process. Remember that the locator-ser RLOC-SET1 defines the Locator IP address. You can find the RLOC-SET1 configuration from Appendix A of chapter 1.

```
Border-PxTR-13#show lisp instance-id 100 ipv4 database 172.16.30.0/24
LISP ETR IPv4 Mapping Database for EID-table vrf 100_NWKT (IID 100), LSBs: 0x1
Entries total 2, no-route 0, inactive 0

172.16.30.0/24, route-import, inherited from default locator-set RLOC-SET1
  Locator       Pri/Wgt  Source     State
  192.168.0.13    1/1    cfg-intf   site-self, reachable
```

Example 5-3: *Border-PxTR-13's LISP Mapping Database.*

Example 5-4 shows that MapSrv-22 has installed the information into its LISP Mapping Database.

```
MapSrv-22#show lisp site 172.16.30.0/24 instance-id 100
LISP Site Registration Information

Site name: Network-Times
Allowed configured locators: any
Requested EID-prefix:

  EID-prefix: 172.16.30.0/24 instance-id 100
    First registered:     00:48:51
    Last registered:      00:48:32
    Routing table tag:    0
    Origin:               Configuration, accepting more specifics
    Merge active:         No
    Proxy reply:          Yes
    TTL:                  1d00h
    State:                complete
    Registration errors:
      Authentication failures:   0
      Allowed locators mismatch: 0
    ETR 192.168.0.13, last registered 00:48:32, proxy-reply, map-notify
                  TTL  1d00h,  no  merge,  hash-function  sha1,  nonce  0x54FAEC88-
0x22E5BC8B
                  state complete, no security-capability
                  xTR-ID 0xEFD8CACE-0x5906016A-0xD7C49F5F-0x5CC352DB
                  site-ID unspecified
                  sourced by reliable transport
      Locator       Local  State      Pri/Wgt  Scope
      192.168.0.13  yes    up           1/1    IPv4 none
```

Example 5-4: *MapSrv-22's LISP Mapping Database.*

In order to advertise internal IP Prefixes to Border-PxTR-13, MapSrv-22 exports all LISP Mapping information to the BGP process. This means that also the external IP prefix 172.16.30.0/24 will be exported to the BGP process. For this reason, we are using route maps to permit only the LISP domain internal networks to be advertised to Border-PxTR-13. We can see from the example 5-5 that subnet 172.16.30.0/24 is not advertised to any peer. You can find MapSrv-22's configuration from the Appendix A of chapter one.

```
MapSrv-22#sh ip bgp vpnv4 vrf 100_NWKT 172.16.30.0/24
BGP routing table entry for 1:100:172.16.30.0/24, version 5
Paths: (1 available, best #1, table 100_NWKT)
  Not advertised to any peer
  Refresh Epoch 1
  65100
    192.168.0.13 (metric 3) (via default) from 192.168.0.13 (192.168.0.13)
      Origin incomplete, metric 1000, localpref 100, valid, internal, best
      Extended Community: SoO:0:10 RT:1:100
      mpls labels in/out nolabel/17
      rx pathid: 0, tx pathid: 0x0
```

Example 5-5: *MapSrv-22's BGP Table.*

Summary

Figure 5-2 summarizes the End-to-End Control-Plane operation. Edge-xTR-11 learns the IP address of EP1 from the ingress GARP message (1). It sends the host-specific EID-to-RLOC information to MapSrv-22 using LISP Map-Register messages (2). LISP Mapping Server publishes mapping information only when separately asked with LISP Map-Request messages. That is why MapSrv-22 exports the information to the BGP process and advertises the aggregate 172.16.30.0/24 to Border-PxTR-13 as VPNv4 NLRI (3). The VPNv4 NLRI is needed because we need to differentiate routes belonging to separate VRFs/Tenants. Border-PxTR-13 advertises the NLRI as IPv4 Unicast route to vEdge-1 (4). It then advertises the information to vSmart as OMP Service (5), which in turn, forwards information to vEdge-2 (6). vEdge-2 exports the route to the BGP process and sends a BGP Update message to Border-Leaf-13 as IPv4 Unicast NLRI (7). Border-Leaf-13 sends the BGP Update message to its iBGP peer Spine-1 as L2VPN EVPN NLRI (8), which forwards the message retaining the Next-Hop Path Attribute installed by Border-Leaf-13 to Leaf-11 (9). As the last step, Leaf-11 installs the information to its BGP table as an L3 entry. The process is the almost same concerning EP3's IP address propagation. The first difference is that Border-PxTR-13 doesn't use BGP for advertising NLRI to MapSrv-22. Instead, it imports BGP routes to LISP and sends LISP Map-Register messages to MapSrv-22. Besides, MapSrv-22 doesn't forward information to Edge-xTR-11 automatically.

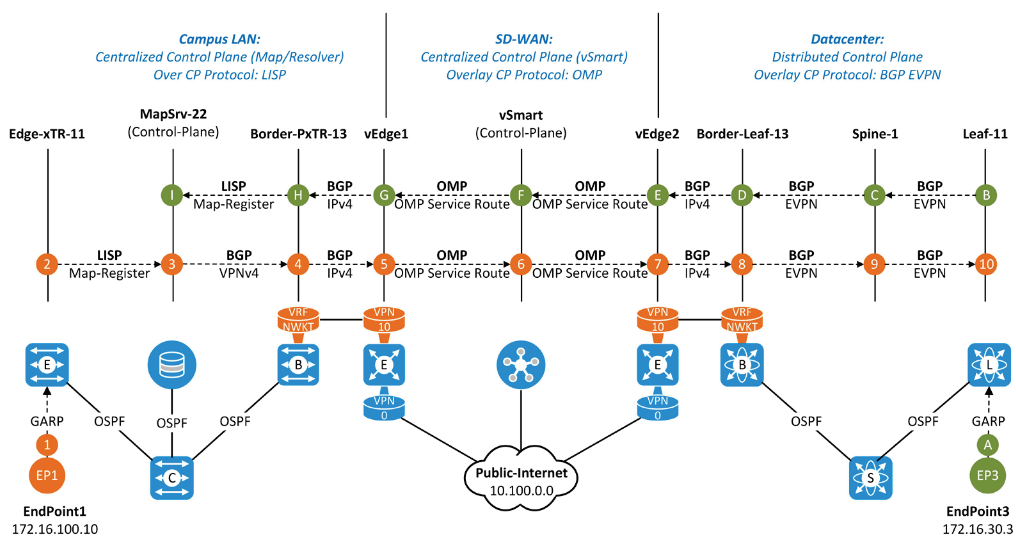


Figure 5-2: *End-to-End Control-Plane Protocols.*

The Next chapter goes through the whole Data-Plane operation when EP3 in Datacenter pings EP1 in Campus Fabric.

Chapter 6: Complete Data-Plane Operation

Introduction

This chapter introduces Data-Plane operation and explains how the data packets from EP3 (IP 172.16.30.3) in Datacenter Fabric are forwarded via SD-WAN to EP1 (IP 172.16.100.10) in Campus Fabric. (1) EndPoint3 sends the ICMP Request packet to its gateway switch Leaf-11. Leaf-11 makes routing decision based on the VRF NWKT routing table. Before forwarding the packet, Leaf-11 adds a VXLAN header where it uses L3VNI 10077. It also sets the outer IP header where it uses the Border-Leaf-13 tunnel interface's IP address 192.168.50.13 as a destination. Spine-1 routes the packet to Border-Leaf-13 based on the outer IP address. Border-Leaf-13 notices that the destination IP address of the received IP packet belongs to its's NVE1 tunnel interface. It removes the outer IP header and based UDP destination port it notices that this is VXLAN encapsulated packet. It knows that L3VNI 10077 belongs to VRF NWKT. It strips off the VXLAN header and routes the packet to vEdge-2. The ingress interface towards DC in vEdge-2 belongs to VPN 10. vEdge-2 consults its routing table. Based on it, vEdge-2 constructs tunnel headers and sends ICMP Request to vEdge-1 via Public-Internet using MPLS Label 1003 as a VPN identifier. Routers in Internet routes packet based on the outer destination IP address. When vEdge-1 receives the packet, it notices that the destination IP address is its' Public IP address. It first removes the outer IP header. Then it checks the tunnel header. Based on the Label value 1003, it knows that packet belongs to VPN 10. It consults the VPN 10 RIB and routes the packet to Border-PxTR-13. The ingress interface on Border-PxTR-13 belongs to VRF 100_NWKT that belongs to LISP Instance 100. It checks the Instance 100 specific LISP mapping in order to know how it should route the packet. The LISP mapping Database does not contain the information because this is the first packet to destination 172.16.100.10. Border-PxTR-13 sends a LISP Map-Request message to MapSrv-22, which replies with a LISP Map-Reply message, where it describes the RLOC of Edge-xTR-11 that has registered the IP address 172.16.100.10. I have excluded the Map-Request/Reply processes from figure 6-1 to keep the figure simple. Border-Leaf-13 encapsulates the ICMP Request packet with a tunnel header. It sets the Instance-Id 100 on the VXLAN header and adds the outer IP header where it uses the Edge-xTR-11's IP address 192.168.0.13 as a destination address. Core-1 routes the packet to Edge-xTR-11 based on the outer IP header destination address. Edge-xTR-11 processes the ingress IP packet because the destination IP address belongs to it. Based on the destination UDP port 4789, it knows that the following header is a VXLAN header. Edge-xTR-11 knows that the LISP Instance-Id 100 is bind to BD 100. Because Edge-xTR-11 has an L3 interface in BD 100, it resolves the MAC address for the IP address 172.16.100.10 from the ARP table and the egress interface for the MAC from the MAC address table. EP1 processes the ICMP Request packet and sends the ICMP Reply to EP3.

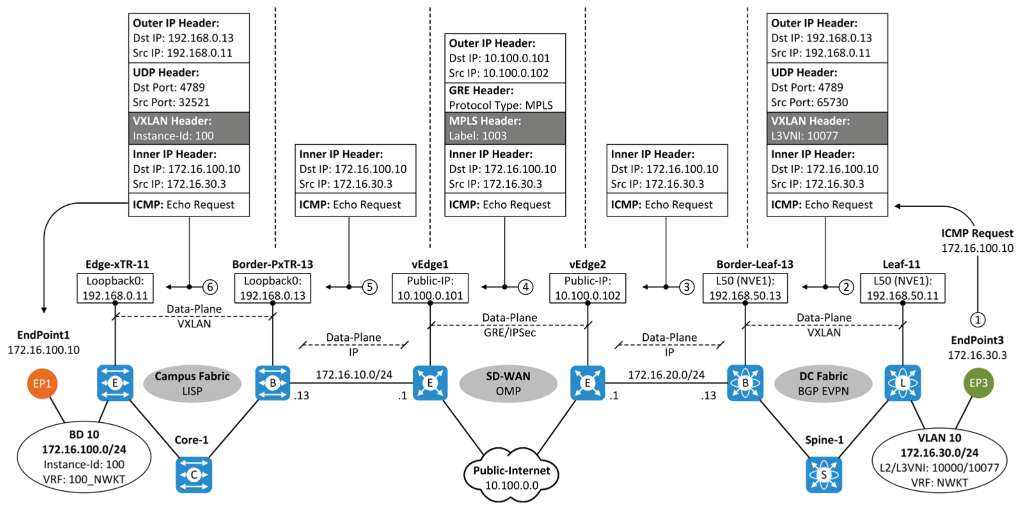


Figure 6-1: *End-to-End Data-Plane Operation.*

Next sections explain the processes in greater detail.

DC Fabric Internal Data-Plane: VXLAN

EP3 starts pinging EP1. Because the destination IP address is not within its subnet 172.16.30.0/24, EP3 sends ICMP Request messages to its' gateway device Leaf-11. The source IP address belongs to VRF NWKT so Leaf-11 does routing lookup from the VRF NWKT's RIB. Example 6-2 shows that the next-hop IP address is 172.16.50.13. It also shows that the encapsulation is VXLAN and the VNI is 10077. Leaf-11 has received this information from the Border-Leaf-13 via Spine-1 as a BGP EVPN Route-Type 5 advertisement (IP Prefix Route). In order to route the packet to the destination IP address 172.16.50.13, Leaf-11 does recursive route lookup from the global RIB to resolve a real next-hop and sends ICMP request packets to Spine-1 that makes its routing decision based on the outer IP header destination address.

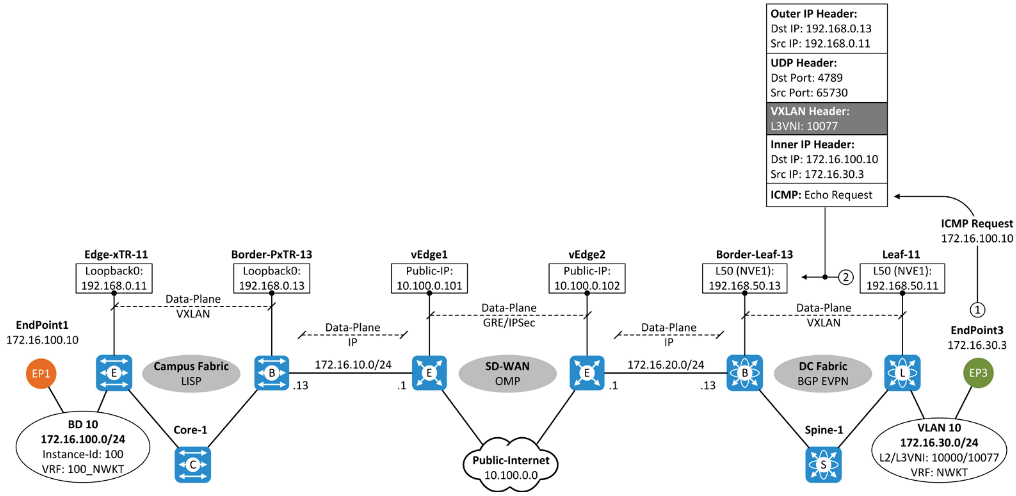


Figure 6-2: *Leaf-11 Data-Plane Processes.*

```
EP3> ping 172.16.100.10

84 bytes from 172.16.100.10 icmp_seq=1 ttl=57 time=41.347 ms
84 bytes from 172.16.100.10 icmp_seq=2 ttl=57 time=18.682 ms
84 bytes from 172.16.100.10 icmp_seq=3 ttl=57 time=20.703 ms
84 bytes from 172.16.100.10 icmp_seq=4 ttl=57 time=24.209 ms
84 bytes from 172.16.100.10 icmp_seq=5 ttl=57 time=19.733 ms
```

Example 6-1: *Ping from EP3 to EP1.*

```
Leaf-11# sh ip route 172.16.100.0 vrf NWKT
IP Route Table for VRF "NWKT"
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
'%<string>' in via output denotes VRF <string>

172.16.100.0/24, ubest/mbest: 1/0
  *via 192.168.50.13%default, [200/1000], 00:45:06, bgp-65030, internal, tag 65100,
segid: 10077 tunnelid: 0xc0a8320d encap: VXLAN
```

Example 6-2: *VRF NWKT's RIB on Leaf-11.*

```
Leaf-11# sh ip route 192.168.50.13
IP Route Table for VRF "default"
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
'%<string>' in via output denotes VRF <string>

192.168.50.13/32, ubest/mbest: 1/0
    *via 192.168.0.1, Eth1/1, [110/81], 03:04:41, ospf-UNDERLAY-NET, intra
```

Example 6-3: *Global RIB on Leaf-11.*

Capture 6-1 shows the packet that Leaf-11 has sent towards Border-Leaf-13. Note that the destination MAC address in the inner MAC header is the Router MAC of Border-Leaf-13 address. The capture 3-1 in chapter 3 shows that Border-Leaf-13 includes its MAC address into BGP Update as Extended Community Path Attribute, and that is how Leaf-11 learns the Router MAC. As I mentioned before, the inner MAC is required because VXLAN is MAC-in-UDP/IP encapsulation method.

```
Ethernet II, Src: 50:11:00:00:1b:08, Dst: 50:10:00:00:1b:08
Internet Protocol Version 4, Src: 192.168.50.13, Dst: 192.168.50.11
User Datagram Protocol, Src Port: 65370, Dst Port: 4789
Virtual eXtensible Local Area Network
    Flags: 0x0800, VXLAN Network ID (VNI)
        0... .... .... .... = GBP Extension: Not defined
        .... .... .0.. .... = Don't Learn: False
        .... 1... .... .... = VXLAN Network ID (VNI): True
        .... .... .... 0... = Policy Applied: False
        .000 .000 0.00 .000 = Reserved(R): 0x0000
    Group Policy ID: 0
    VXLAN Network Identifier (VNI): 10077
    Reserved: 0
Ethernet II, Src: 50:12:00:00:1b:08, Dst: 50:10:00:00:1b:08
Internet Protocol Version 4, Src: 172.16.100.10, Dst: 172.16.30.3
Internet Control Message Protocol
```

Capture 6-1: *Packet Sent by Leaf-11 to Border-Leaf-13.*

DC Fabric to SD-WAN Data-Plane: IP

When Border-leaf-13 receives the encapsulated ICMP request from Spine-1, it removes the outer IP header and UDP header. It knows that the following header is the VXLAN header based on the UDP destination port number. The VNI-Id 10077 identifies L3VNI for VRF NWKT 100 and Border-Leaf-13 routes packet on based VRF NWKT's RIB.

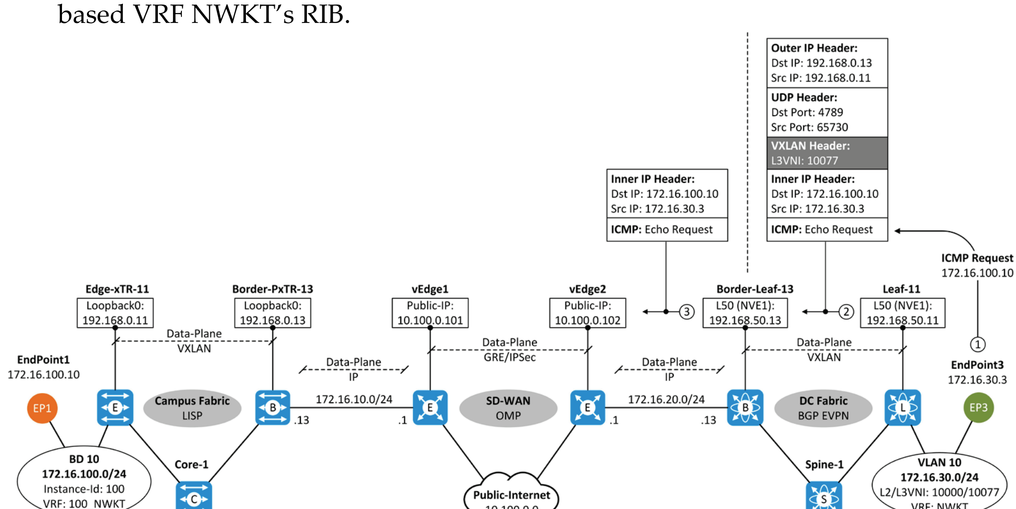


Figure 6-3: *Border-Leaf-13 Data-Plane Processes.*

Example 6-4 shows the VRF NWKT's RIB on Border-Leaf-13 concerning the destination network 172.16.100.0/24. The next-hop IP address is the vEdge-2's DC-link address. Border-leaf-13 has learned this information from the BGP Update message that vEdge-2 has sent to it.

```
Border-leaf-13# sh ip route 172.16.100.0 vrf NWKT
IP Route Table for VRF "NWKT"
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
'%<string>' in via output denotes VRF <string>

172.16.100.0/24, ubest/mbest: 1/0
    *via 172.16.20.1, [20/1000], 00:47:53, bgp-65030, external, tag 65100
```

Example 6-4: *VRF NWKT's RIB on Border-Leaf-13.*

Capture 6-2 shows the packet that Border-Leaf-13 has sent towards SD-WAN edge device vEdge-2.

```
Ethernet  II,  Src:  50:12:00:00:1b:08  (50:12:00:00:1b:08),  Dst:  50:00:00:0b:00:03
(50:00:00:0b:00:03)
Internet Protocol Version 4, Src: 172.16.30.3, Dst: 172.16.100.10
Internet Control Message Protocol
    Type: 8 (Echo (ping) request)
    Code: 0
    Checksum: 0x006f [correct]
    [Checksum Status: Good]
    Identifier (BE): 8091 (0x1f9b)
    Identifier (LE): 39711 (0x9b1f)
    Sequence number (BE): 2 (0x0002)
    Sequence number (LE): 512 (0x0200)
    [Response frame: 10]
    Data (56 bytes)
```

Capture 6-2: *Packet Sent by Border-Leaf-13 to vEdge-2.*

SD-WAN Internal Data-Plane: GRE

The Datacentre SD-WAN edge device vEdge-2 receives the ICMP request from Border-Leaf-13. The ingress interface belongs to VPN 10. VPN 10 RIB (example 6-5) shows that the destination TLOC IP address is 10.100.100.101 (vEdge-1's System-id), the transport network is Public-Internet, and the VPN label is 1003. The RIB entry also shows that the route is learned as an OMP Service route update. However, the RIB entry does describe the destination IP address for the outer IP header used in Data-Plane. To resolve it, vEdge-2 has to consult its OMP TLOC table (example 6-6).

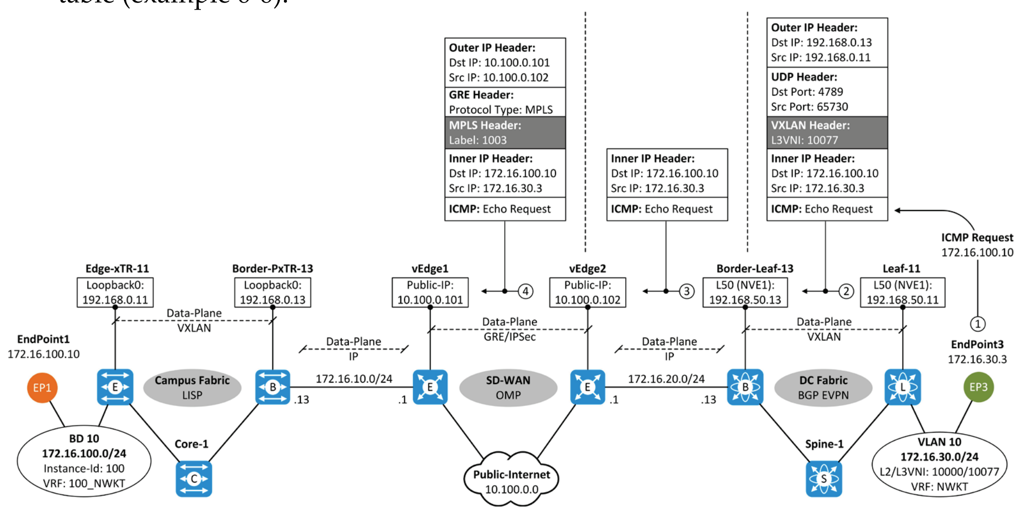


Figure 6-4: *vEdge-2 Data-Plane Processes.*

```
Edge2# show ip routes vpn 10 172.16.100.0/24 detail
Codes Proto-sub-type:
  IA -> ospf-intra-area, IE -> ospf-inter-area,
  E1 -> ospf-external1, E2 -> ospf-external2,
  N1 -> ospf-nssa-external1, N2 -> ospf-nssa-external2,
  e -> bgp-external, i -> bgp-internal
Codes Status flags:
  F -> fib, S -> selected, I -> inactive,
  B -> blackhole, R -> recursive, L -> import

""---------------------------------------------
 VPN 10      PREFIX 172.16.100.0/24
--------------------------------------------
 proto           omp
 distance        250
 metric          0
 uptime          0:00:52:18
 tloc-ip         10.100.100.101
 tloc-color      public-internet
 tloc-encap      gre
 nexthop-label   1003
 status          F,S
```

Example 6-5: *The RIB (OMP Service Routes) of VPN 10 on vEdge-2.*

The partial OMP TLOC database output in example 6-6 shows the Public IP address by which vEdge-1 is connected to the Public-Internet. vEdge-2 uses this IP address in the outer IP header as a destination address when sending packets via Public-Internet transport network towards destination subnet 172.16.100.0/24. vEdge-2 has received OMP TLOC routing information from the SD-WAN Control-Plane node vSmart.

```
vEdge2# show omp tlocs ip 10.100.100.101 color public-internet | exclude not | nomore

---------------------------------------------------
tloc entries for 10.100.100.101
                 public-internet
                 gre
---------------------------------------------------
            RECEIVED FROM:
peer            10.100.100.13
status          C,I,R
    Attributes:
     attribute-type    installed
     encap-key         0
     public-ip         10.100.0.101
     public-port       0
     private-ip        10.100.0.101
     private-port      0
     <snipped>
     bfd-status        up
     site-id           10
     preference        0
     weight            1
     version           3
    gen-id             0x80000039
     carrier           default
     restrict          1
     on-demand          0
     groups            [ 0 ]
     bandwidth         0
     qos-group         default-group
vEdge2#
```

Example 6-6: *OMP TLOC Information on vEdge-2.*

Capture 6-3 shows that vEdge-2 uses MPLS Label 1003 as VPN identifier when sending a GRE encapsulated packet via Public-Internet to vEdge-1.

```
Ethernet  II,  Src:  50:00:00:0b:00:01  (50:00:00:0b:00:01),  Dst:  50:00:00:02:00:01
(50:00:00:02:00:01)
Internet Protocol Version 4, Src: 10.100.0.102, Dst: 10.100.0.101
Generic Routing Encapsulation (MPLS label switched packet)
    Flags and Version: 0x0000
    Protocol Type: MPLS label switched packet (0x8847)
MultiProtocol Label Switching Header, Label: 1003, Exp: 0, S: 1, TTL: 62
Internet Protocol Version 4, Src: 172.16.30.3, Dst: 172.16.100.10
Internet Control Message Protocol
    Type: 8 (Echo (ping) request)
```

Capture 6-3: *Packet Sent by vEdge-2 to vEdge-1.*

SD-WAN to Campus Fabric Data-Plane: IP

vEdge-1 receives the encapsulated ICMP Request from the Public-Internet. The destination IP address is its Public IP address, so it processes the packet. It first removes the outer IP header. Then it checks the MPLS Label value to resolve to which VPN this packet belongs. MPLS Label 1003 defines the VPN 10 on vEdge-1. It does a routing lookup from the VPN 10 RIB. Example 6-7 shows that the next-hop is 172.16.10.13 and that the egress interface is ge0/2. vEdge-2 has learned this information from BGP Update that Border-PxTR-13 has previously sent to it.

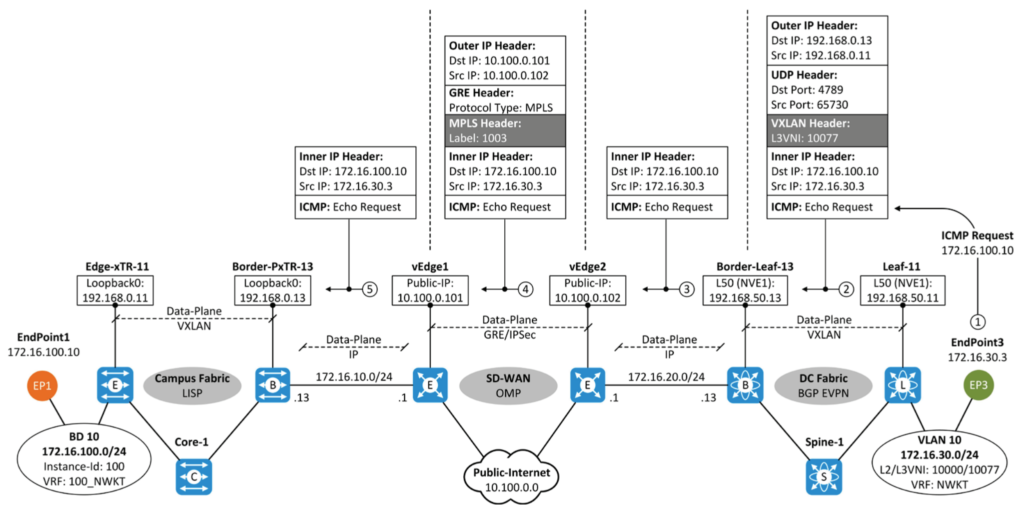


Figure 6-5: *vEdge-1 Data-Plane Processes.*

```
vEdge-1# show ip routes vpn 10 172.16.100.0/24 detail
<snipped>
""---------------------------------------------
 VPN 10      PREFIX 172.16.100.0/24
-------------------------------------------
 proto              bgp
 proto-sub-type     e
 distance           20
 metric             0
 uptime             0:00:54:42
 nexthop-ifname     ge0/2
 nexthop-addr       172.16.10.13
 status             F,S
```

Example 6-7: *The RIB of VPN 10 on vEdge-1.*

Capture 6-4 shows the packet that vEdge-1 has sent to Border-PxTR-13.

```
Ethernet II, Src: 50:00:00:02:00:03, Dst: 50:00:00:0a:00:00
Internet Protocol Version 4, Src: 172.16.30.3, Dst: 172.16.100.10
Internet Control Message Protocol
    Type: 8 (Echo (ping) request)
```

Capture 6-4: *Packet Sent by vEdge-1 to Border-PxTR-13.*

Campus Fabric Internal Data-Plane: VXLAN

Border-PxTR-13 receives the ICMP Request message from vEdge-1. The ingress interface belongs to VRF 100_NWKT, which in turn is mapped to LISP Instance 100. Example 6-8 shows the VRF 100_NWKT RIB where the next-hop for packets towards network 172.16.100.0/24 is 192.168.0.1 (MapSrv-22). Example 6-9, in turn, shows that this information is also programmed to Forwarding Information Base (FIB). However, this is an aggregate route that MapSrv-22 has sent to Border-Leaf-13. When Border-PxTR-13 receives the first packet, it routes it to MapSrv-22, which doesn't have routing information in its RIB even though the mapping information is stored in its EID-to-RLOC mapping database. As a result, MapSrv-22 drops the packet. However, the ingress packet triggers the LISP process on Border-Leaf-13. It checks the local Map-Cache to find out EID-to-RLOC mapping information. The result is a miss. It then sends a Map-Request message to MapSrv-22 where it asks what is the RLOC IP address that it should use when sending packets to 172.16.100.10. MapSrv-22 has the information because Edge-xTR-11 has already registered the mapping information. MapSrv-22 sends a Map-Reply message to Border-PxTR-13 where it tells that Edge-xTR-11 (192.168.0.11) is the Egress Tunnel Router (ETR). Border-PxTR-13 installs the information to FIB (example 6-11). At this phase, it has one FIB entry for 172.16.100.0/24 and the other one for 172.16.100.10/32. When the second ICMP Request message arrives, Border-Leaf-13 forwards it correctly to Edge-xTR-11. Border-PxTR-13 adds an outer IP header on top of the packet where it uses the Edge-xTR-11 IP address as a destination. It also adds a UDP header with the destination port 4789. It uses the LISP Instance-Id 100 as a VPN identifier in the VXLAN header. Then it sends the encapsulated packet to Core-1 that routes the packet based on the outer destination IP address.

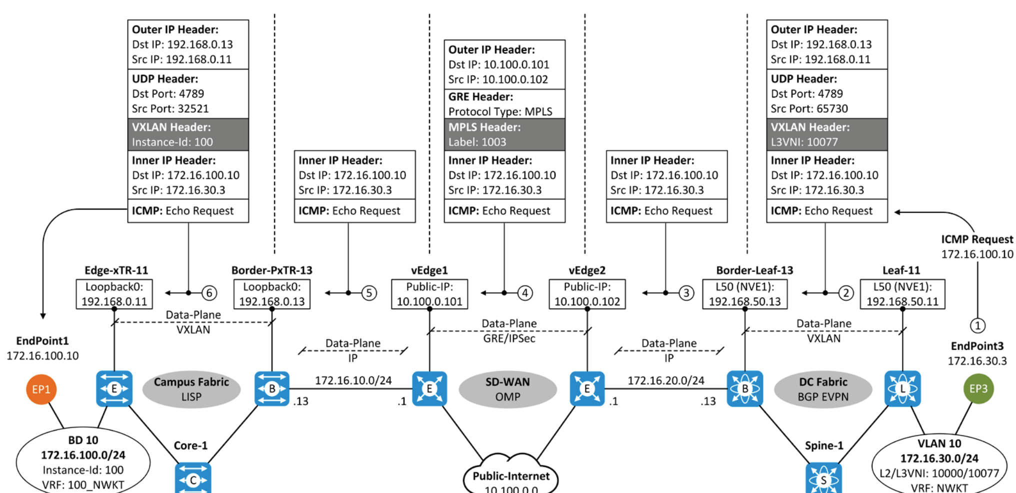


Figure 6-6: *Border-PxTR-13 Data-Plane Processes.*

Example 6-8 shows the routing information to network 172.16.100.0/24 that Border-PxTR-13 has learned from MapSrv-22 via BGP Update message. We can see that the next is MapSrv-22.

```
Border-PxTR-13#sh ip route vrf 100_NWKT 172.16.100.10

Routing Table: 100_NWKT
Routing entry for 172.16.100.0/24
  Known via "bgp 65010", distance 200, metric 0, type internal
  Redistributing via lisp
  Last update from 192.168.0.22 00:57:00 ago
  Routing Descriptor Blocks:
  * 192.168.0.22 (default), from 192.168.0.22, 00:57:00 ago
      Route metric is 0, traffic share count is 1
      AS Hops 0
      MPLS label: 17
      MPLS Flags: MPLS Required
```

Example 6-8: *VRF 100_NWKT's RIB on Border-PxTR-13.*

Border-PxTR-13 has also installed the information to FIB.

```
Border-PxTR-13#sh ip cef vrf 100_NWKT 172.16.100.0 detail
172.16.100.0/24, epoch 2, flags [cover dependents, subtree context, rib defined all
labels]
  SC owned,sourced: LISP local EID -
  LISP EID attributes: localEID Yes, c-dynEID No, d-dynEID No
  SC owned,sourced: LISP generalised SMR - [disabled, not inheriting, 0x7F38D0EFFDA8
locks: 1]
  Covered dependent prefixes: 1
    notify cover updated: 1
  2 IPL sources [no flags]
  recursive via 192.168.0.22 label 17
    nexthop 192.168.0.1 GigabitEthernet2 unusable: no label
```

Example 6-9: *FIB Entry about Network 172.16.100.0/24 on Bordr-PxTR-13.*

Example 6-10 shows the EID-to-RLOC Mapping-Cache of Border-PxTR-13 after MapSrv-22 is replied to it. Note that Border-PxTR-13 verifies that the RLOC is reachable by sending probes to it.

```
Border-PxTR-13#show lisp eid-table vrf 100_NWKT ipv4 map-cache 172.16.100.10
LISP IPv4 Mapping Cache for EID-table vrf 100_NWKT (IID 100), 2 entries

172.16.100.10/32, uptime: 01:00:44, expires: 22:59:15, via map-reply, self, complete
  Sources: map-reply
  State: complete, last modified: 01:00:44, map-source: 192.168.0.11
  Idle, Packets out: 10(980 bytes) (~ 00:57:55 ago)
  Configured as EID address space
  Locator       Uptime    State      Pri/Wgt     Encap-IID
  192.168.0.11  01:00:44  up           1/1          -
    Last up-down state change:        01:00:44, state change count: 1
    Last route reachability change:   01:00:44, state change count: 1
    Last priority / weight change:    never/never
    RLOC-probing loc-status algorithm:
      Last RLOC-probe sent:           01:00:44 (rtt 3ms)
```

Example 6-10: *EID-to-RLOC Mapping-Cache on Border-PxTR-13.*

Example 6-11 shows that Border-PxTR-13 has installed the host entry to its FIB.

```
Border-PxTR-13#sh ip cef vrf 100_NWKT 172.16.100.10 detail
172.16.100.10/32, epoch 2, flags [subtree context, check lisp eligibility]
  SC inherited: LISP local EID -
  SC owned,sourced: LISP remote EID - locator status bits 0x00000001
  LISP remote EID: 10 packets 980 bytes fwd action encap, cfg as EID space
  LISP EID attributes: localEID Yes, c-dynEID No, d-dynEID No
  SC owned,sourced: LISP generalised SMR - [enabled, not inheriting, 0x7F38D0EFFD90
locks: 1]
  LISP source path list
    nexthop 192.168.0.11 LISP0.100
  3 IPL sources [no flags]
  nexthop 192.168.0.11 LISP0.100
```

Example 6-11: *FIB Entry about Host IP Address 172.16.100.10/32 on Bordr-PxTR-13.*

Capture 6-5 shows that Border-PxTR-13 has sent the second ICMP Request message to Edge-xTR-11. It verifies that Border-PxTR-13 has used the LISP Instance-Id 100 as VXLAN Network Identifier.

```
Ethernet II, Src: 50:00:00:0e:00:00 (50:00:00:0e:00:00), Dst: Broadcast
(ff:ff:ff:ff:ff:ff)
Internet Protocol Version 4, Src: 192.168.0.13, Dst: 192.168.0.11
User Datagram Protocol, Src Port: 32521, Dst Port: 4789
Virtual eXtensible Local Area Network
    Flags: 0x0840, Don't Learn, VXLAN Network ID (VNI)
        0... .... .... .... = GBP Extension: Not defined
        .... .... .1.. .... = Don't Learn: True
        .... 1... .... .... = VXLAN Network ID (VNI): True
        .... .... .... 0... = Policy Applied: False
        .000 .000 0.00 .000 = Reserved(R): 0x0000
    Group Policy ID: 0
    VXLAN Network Identifier (VNI): 100
    Reserved: 0
Ethernet II, Src: 50:00:00:0a:00:01 (50:00:00:0a:00:01), Dst: ba:25:cd:f4:ad:38
(ba:25:cd:f4:ad:38)
Internet Protocol Version 4, Src: 172.16.30.3, Dst: 172.16.100.10
Internet Control Message Protocol
    Type: 8 (Echo (ping) request)
    Code: 0
    Checksum: 0xfe6c [correct]
    [Checksum Status: Good]
    Identifier (BE): 8603 (0x219b)
    Identifier (LE): 39713 (0x9b21)
    Sequence number (BE): 4 (0x0004)
    Sequence number (LE): 1024 (0x0400)
    [Response frame: 60]
    Data (56 bytes)
```

Capture 6-5: *Packet Sent by Border-Leaf-13 to Edge-xTR-11.*

Example 6-12 shows the EID-to-RLOC Mapping database on MapSrv-22. Border-Leaf-13 has registered subnets 172.16.30.0/24 and 172.16.100.0/24 to MapSrv-22. Edge-xTR-11, in turn, has registered the host-specific entry.

```
MapSrv-22#show lisp site name Network-Times instance-id 100
Site name: Network-Times
Allowed configured locators: any
Allowed EID-prefixes:

  EID-prefix: 172.16.30.0/24 instance-id 100
    First registered:     00:03:21
    Last registered:      00:03:19
    Routing table tag:    0
    Origin:               Configuration, accepting more specifics
    Merge active:         No
    Proxy reply:          Yes
    TTL:                  1d00h
    State:                complete
    Registration errors:
      Authentication failures:   0
      Allowed locators mismatch: 0
    ETR 192.168.0.13, last registered 00:03:19, proxy-reply, map-notify
                      TTL  1d00h,  no  merge,  hash-function  sha1,  nonce  0x54FAEC88-
0x22E5BC8B
                      state complete, no security-capability
                      xTR-ID 0xEFD8CACE-0x5906016A-0xD7C49F5F-0x5CC352DB
                      site-ID unspecified
                      sourced by reliable transport
      Locator       Local  State      Pri/Wgt  Scope
      192.168.0.13  yes    up           1/1    IPv4 none

  EID-prefix: 172.16.100.0/24 instance-id 100
    First registered:     00:01:54
    Last registered:      00:01:54
    Routing table tag:    0
    Origin:               Configuration, accepting more specifics
    Merge active:         No
    Proxy reply:          Yes
    TTL:                  1d00h
    State:                complete
    Registration errors:
      Authentication failures:   0
      Allowed locators mismatch: 0
    ETR 192.168.0.13, last registered 00:01:54, proxy-reply, map-notify
                      TTL  1d00h,  no  merge,  hash-function  sha1,  nonce  0x54FAEC88-
0x22E5BC8B
                      state complete, no security-capability
                      xTR-ID 0xEFD8CACE-0x5906016A-0xD7C49F5F-0x5CC352DB
                      site-ID unspecified
                      sourced by reliable transport
      Locator       Local  State      Pri/Wgt  Scope
      192.168.0.13  yes    up           1/1    IPv4 none

  EID-prefix: 172.16.100.10/32 instance-id 100
    First registered:     00:01:54
    Last registered:      00:01:49
    Routing table tag:    0
    Origin:               Dynamic, more specific of 172.16.100.0/24
    Merge active:         No
    Proxy reply:          Yes
    TTL:                  1d00h
    State:                complete
    Registration errors:
      Authentication failures:   0
      Allowed locators mismatch: 0
    ETR 192.168.0.11, last registered 00:01:49, proxy-reply, map-notify
```

```
                      TTL  1d00h,  no  merge,  hash-function  sha1,  nonce  0x2A38F7C7-
0x79599A97
                      state complete, no security-capability
                      xTR-ID 0xB29EE3E2-0x6BDDF328-0x6524BF4B-0x9769F92A
                      site-ID unspecified
                      sourced by reliable transport
      Locator       Local  State      Pri/Wgt  Scope
      192.168.0.11  yes    up           1/1    IPv4 none
```

Example 6-13: *EID-to-RLOC Mapping Table on MapSrv-22.*

Campus Fabric Internal Data-Plane (Edge-11): Switched

When Edge-xTR-11 receives the encapsulated ICMP Request from Spine-1, it notices that the destination IP address belongs to it. It removes the outer IP header and based on the destination UDP port, it knows that the next header is VXLAN header. It knows that the packet belongs to the VRF 100_NWKT based on the LISP Instance-Id 100 encoded to the VXLAN header. It does ARP lookup to find the IP/MAC address mapping information and then checks from the MAC address table to find out the egress interface where EP1 is connected.

```
Edge-xTR-11#show bridge-domain 100
Bridge-domain 100 (2 ports in all)
State: UP                    Mac learning: Enabled
Aging-Timer: 300 second(s)
    BDI100  (up)
    GigabitEthernet2 service instance 100
   AED MAC address    Policy  Tag       Age  Pseudoport
   -   0001.0001.0001 to_bdi  static    0    BDI100
   0   0050.7966.680D forward dynamic   265  GigabitEthernet2.EFP100
```

Example 6-15: *MAC Address Table of Bridge Domain 100.*

Now we have seen the Control-Plane processes how devices using different Control-Plane protocols advertises routing information between themself. We have also seen how devices forward packets across domains that don't use the same encapsulation methods. The next chapter discusses the similarities of Control-Plane operation between Campus Fabric with LISP, SD-WAN with OMP, and Datacenter Fabric with BGP EVPN.

Chapter 7: LISP, BGP EVPN, and OMP Comparison

IP reachability

Every Overlay Network solution requires IP reachability between edge devices via Underlay Network. This section explains the basic routing solution in Underlay Network from Campus Fabric, SD-WAN, and Datacenter Fabric perspectives. Figure 7-1 illustrates the IP reachability requirements for Campus Fabric, SD-WAN, and Datacenter Fabric.

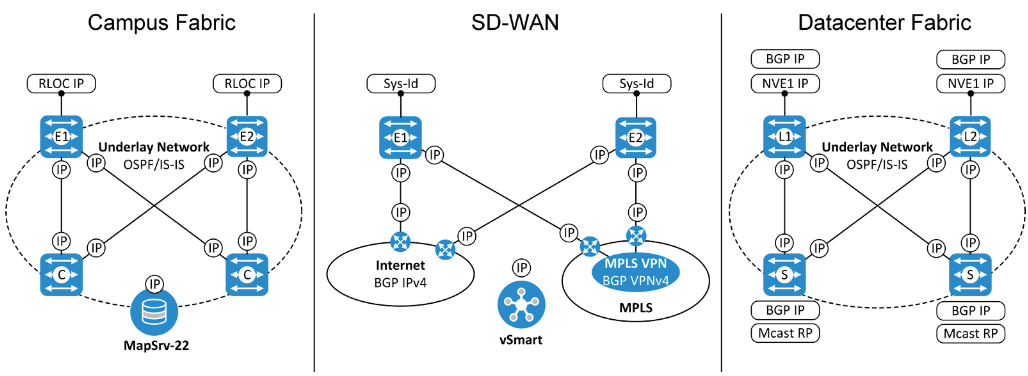


Figure 7-1: *IP Reachability Requirements.*

Campus Fabric

LISP Control-Plane uses Remote Locator (RLOC) IP address in LISP Map-Registers message as a destination address when it registers its local hosts to Mapping Server. Edge devices use RLOC addresses as the destination IP address in the outer IP header for VXLAN encapsulated data packets. Core routers route tunneled data packets based on it. These are the reasons why RLOC addresses have to be routable in Underlay Network. When an edge device sends data packets from one of its locally connected hosts to a remote host behind another edge device, it does recursive routing lookup for the destination RLOC address to find out the real next-hop. Also, the Mapping Server has to have IP reachability between all edge devices.

Datacenter Fabric

Leaf switches use the IP address of the Network Virtual Edge (NVE) interface as BGP next-hop in EVPN NLRI update. Other Leaf switches use it as a destination IP address in the outer IP header for VXLAN encapsulated data packets, just like edge devices in Campus Fabric use RLOC IP address. That is why NVE interface IP addresses have to be routable in Underlay Network. In addition, IP addresses which Leaf and Spine are using for iBGP peering, as well as IP addresses assigned to Multicast-RP, have to be routable.

SD-WAN

Devices in Campus and Datacenter fabrics belong to the trusted network, and that is why the Underlay Network routing architecture is quite simple. Also, we don't have to use data encryption on LAN or DC. That is not the case with SD-WAN. Public Internet and mobile networks are pretty far from the secure network, while MPLS VPN/private APNs are secure but routing/label switching are administrated by third parties. That does not affect the IP reachability requirements. Devices connected to the same transport network have to have an IP connection between themselves. SD-WAN edge devices don't use their System-Id in Data-Plane, which is why it is not routable in Underlay Network. Edge devices identify themselves with their System-Id in OMP TLOC (Transport Locator) route advertisement, which they send over each Transport Network. The TLOC describes the device location, just like RLOC and NVE. However, it also defines the color (Transport Network), Data-Plane encapsulation (IPSec/GRE), and Public-IP address (Transport Network uplink IP). Edge devices publish customer networks using OMP Service route advertisements. These updates describe that the next-hop is System-Id X and it is reachable via Transport Network Y. Remote edge device then checks the Public-IP address for Y/X combination from TLOC table. Besides, vSmart has to have IP reachability with all edge devices and other Control-Plane components.

Overlay Network

Edge devices build an Overlay Network between themselves. This simply means that sending device encapsulates the original data with the IP address that the remote device uses for network virtualization.

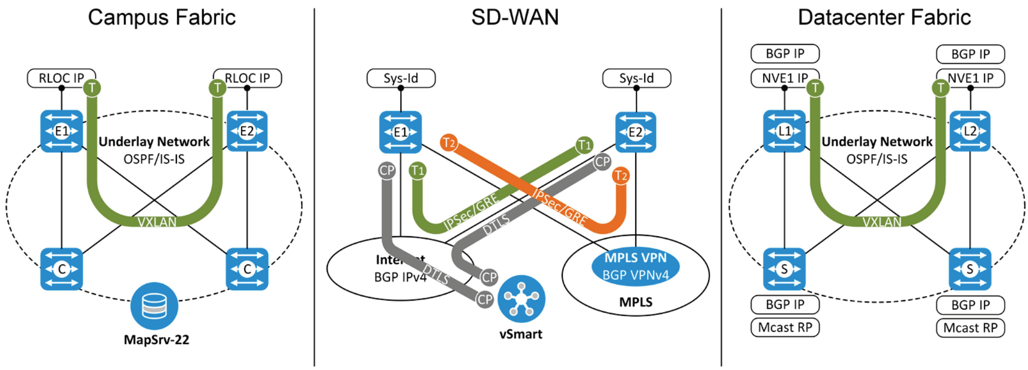


Figure 7-2: *Virtual Network Tunneling.*

Campus Fabric

Edge-xTRs in the LISP domain use their RLOC IP addresses as a source and destination in the outer IP header. Core devises route encapsulated data based on the destination IP address in the tunnel header. Campus Fabric uses VXLAN tunneling where the Virtual Network Identifier (VNI) is LISP Instance-Id. Edge-xTR devices use flow-based load-balancing for data packets over equal costs core uplinks.

Datacenter Fabric

VTEP/Leaf switches in the BGP EVPN domain use their NVE interface IP addresses as a source and destination in the outer IP header. Core devises route encapsulated data based on the destination IP address in the tunnel header. Datacenter Fabric solution with BGP EVPN Control-Plane also uses VXLAN tunneling where the L2VNI identifies Layer 2 segments while L3VNI identifies routing domain (VRF/Tenant). VTEP/Leaf switches use flow-based load-balancing for data packets over equal cost core uplinks.

SD-WAN

Edge devices build IPSec/GRE tunnels for data packets between their Transport Network uplink IP addresses. Edge devices can establish tunnels with other edge devices which are directly connected to the same Transport Network. Besides, edge devices can build tunnels over different Transport Networks if there is an IP connection and the tunneling policy allows it (restrict bit not set in TLOC updates). Edge devises exchange routing updates with vSmart over DTLS tunnel. All three solutions use the same Control-Plane peering model. Edge-xTR registers and requests EID-to-RLOC mapping information to and from Mapping Server. Leaf switches exchange BGP EVPN updates with Spine switches. SD-WAN edge devices exchange routing information with vSmart.

Final words

Networks virtualization, no matter which solution we are using, tends to be quite complex. However, they all serve the same need, getting data securely and resilient from point A to B. How this goal is achieved varies between solutions. LISP uses quite a simple EID-to-RLOC mapping system to describe endpoint/prefix reachability information. BGP EVPN has its route types for hosts, prefixes, L2 Multicast, L3 Tenant routed Multicast, ESI Multi-homing, and Ingress/headend replication. OMP, in turn, uses TLOC and Service routes. Each solution uses some tunneling solution where virtual networks are identified in a solution-specific way. LISP uses Instance-Id, BGP EVPN uses L2VNI/L3VNI, and OMP uses VPN/Label binding. Connecting different solutions might feel like a nightmare but in the end, it is not so complex. I have used good old BGP IPv4 Unicast with VRF Lite solution throughout this book.

Cheers - Toni Pasanen

In case you want to learn more about LISP, OMP, or BGP EVPN…
These books are available at Leanpub.com (pdf) and Amazon (Kindle, paperback)