Toronto Metropolitan University

Computer Networks Program
Faculty of Engineering & Architectural Science

EVPN-VXLAN Implementation and interoperability between

Case Study Project report

Authors: Aria Armani, Hamidreza Naghizadeh, Neda Ghafarzadeh

Project Supervisor: Professor Ahmad Enaya

Date: 7/31/2024

Summary

The document elaborates on the following key aspects:

- VXLAN Basics: Introduction to VXLAN, its header structure, and the encapsulation process.
- MP-BGP and EVPN: Explanation of MP-BGP for EVPN, its configuration, and the different EVPN route types.
- Interoperability Challenges: Discussion of challenges and solutions for implementing VXLAN in a mixed-vendor environment, focusing on differences in feature sets between Cisco and Arista devices.
- Scenario-based Implementation: Step-by-step implementation of various scenarios demonstrating the practical application of VXLAN, including iBGP Neighborship, EVPN configuration, and the handling of Layer 3 routing using VXLAN Route Type 5.
- Technical Configurations: Detailed configurations for Cisco and Arista devices, highlighting specific features such as Cisco's VPC and Arista's MLAG, and the setup of Virtual Tunnel Endpoints (VTEPs).

Certification of Authorship

I certify that I am the sole author of this case study. This is a true and original version, including any necessary final revisions approved by my examiners. I grant Toronto Metropolitan University permission to lend this case study to other institutions or individuals for scholarly research purposes. Additionally, I authorize Toronto Metropolitan University to reproduce this case study, either in full or in part, by photocopying or other means, upon request by other institutions or individuals for scholarly research. I acknowledge that my dissertation may be made available electronically to the public.

Table of Contents

Table of Figures

Introduction

The document provides an in-depth exploration of the implementation and interoperability of EVPN-VXLAN across multi-vendor environments, specifically focusing on Cisco and Arista devices. VXLAN (Virtual Extensible LAN) is a network virtualization technology that enables the creation of a virtual Layer 2 network over a Layer 3 infrastructure, addressing the limitations of traditional VLANs in large-scale data centers. This technology enhances network scalability, flexibility, and isolation. The document details the foundational configurations, the role of Multi-Protocol BGP (MP-BGP) in EVPN, and the specific challenges encountered in a multi-vendor setup. It also covers various scenarios demonstrating the practical implementation of VXLAN, highlighting differences in feature sets and configurations between Cisco and Arista devices.

Purpose and Importance of VXLAN Interoperability Testing

Since data centers often utilize multi-vendor products, especially for connecting different data centers, operating VXLAN in a multi-vendor environment is a concern. There are several use cases for using multi-vendors in data centers:

- Moving from legacy environments to new environments
- Connecting data centers
- Organizations that depend on multi-vendor models
- Cost optimization (some brands offer more competitive pricing)
- Vendor redundancy (minimizing the risk of dependency on a single vendor)
- Best-of-breed solutions (combining the best features and capabilities from different vendors to build a more robust and efficient network)
- Etc.

To address these concerns, we decided to test multi-vendor inter-operability in VXLAN with EVPN. This testing provides opportunities for companies (buyers) to combine the best features and capabilities from different vendors to build a more robust and efficient network. It also allows companies (sellers) to optimize the features of their products to achieve better interoperability. Previous works, such as EANTC (EANTC-InteropTest2023-TestReport.pdf), have been published in this area.

The following section addresses the operation of VXLAN between different brands (Cisco and Arista). We chose Cisco and Arista because of the resource that we had in TMU LAB environment.

Topology Overview

The network design consists of eight devices, forming a spine-and-leaf topology with full mesh connectivity, except for some specific parts. The devices include two spines and six-leaf switches:

Spines:

- Spine 1: Arista

- Spine 2: Cisco

Leafs:

- Single-homing Leafs:

- Leaf 1: Cisco (connected to Cisco Spine)

- Leaf 2: Arista (connected to Arista Spine)

Multi-homing Leafs:

- Pair 1: Two Cisco switches

- Pair 2: Two Arista switches

Each leaf switch is connected to a server. In single-homing scenarios, each server connects directly to its respective leaf switch. In multi-homing scenarios, each pair of leaf switches connects to a common switch, which then connects to a server, ensuring redundancy and load balancing.

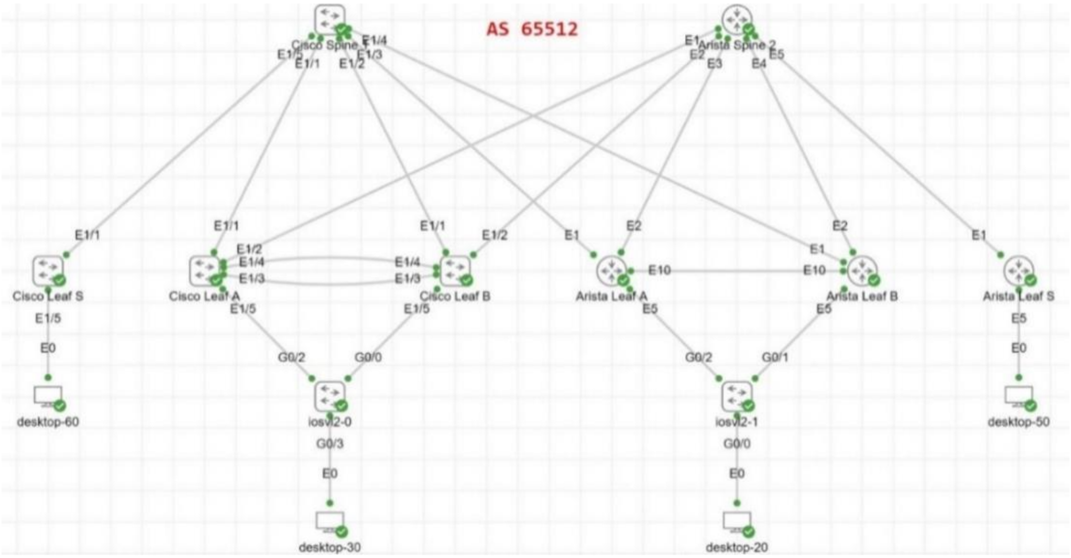


Figure 1: Topology Overview

Connectivity Details

- Cisco Leafs:

- Two links: one for Peer link and one for Keepalive.

- Arista Leafs:

- One link: one for Peer link and Keepalive.

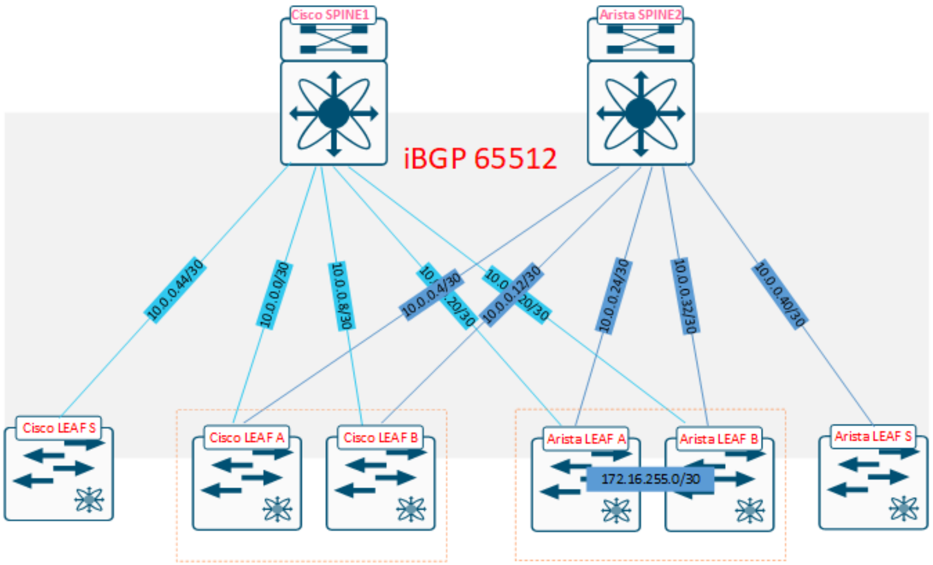


Figure 2: Connectivity Details

IP Plan

Physical Links			
Host	IP Range	Subnet Mask	Description
Cisco LEAF A -Cisco SPINE 1	10.0.0.0	/30	Inter Link
Cisco LEAF A -Arista SPINE 2	10.0.0.4	/30	Inter Link
Cisco LEAF B -Cisco SPINE 1	10.0.0.8	/30	Inter Link
Cisco LEAF B -Arista SPINE 2	10.0.0.12	/30	Inter Link
Cisco LEAF A -Cisco LEAF B	10.0.0.16	/30	Inter Link
Arista LEAF A -Cisco SPINE 1	10.0.0.20	/30	Inter Link
Arista LEAF A -Arista SPINE 2	10.0.0.24	/30	Inter Link
Arista LEAF B -Cisco SPINE 1	10.0.0.28	/30	Inter Link
Arista LEAF B -Arista SPINE 2	10.0.0.32	/30	Inter Link
Arista LEAF A -Arista LEAF B	10.0.0.36	/30	Inter Link
Stand Alone Arista SPINE2	10.0.0.40	/30	Inter Link
Stand Alone Cisco SPINE1	10.0.0.44	/30	Inter Link
Loop Back Interface			
Host	IP Range	Subnet Mask	Description
Cisco LEAF A Lo 1	10.10.10.1	/32	BGP RID/Source
Cisco LEAF A Lo 2	**10.20.20.1**	/32	VTEP Source
Cisco LEAF B Lo 1	10.10.10.2	/32	BGP RID/Source
Cisco LEAF B Lo 2	**10.20.20.1**	/32	VTEP Source

Arista LEAF A Lo 1	10.10.10.3	/32	BGP RID/Source
Arista LEAF A Lo 2	**10.20.20.2**	/32	VTEP Source
Arista LEAF B Lo 1	10.10.10.4	/32	BGP RID/Source
Arista LEAF B Lo 2	**10.20.20.2**	/32	VTEP Source
Cisco Spine 1 Lo1	10.10.10.5	/32	BGP RID/Source
Cisco Spine 1 Lo 2	**10.20.20.3**	/32	VTEP Source
Arista Spine 2 Lo1	10.10.10.6	/32	BGP RID/Source
Arista Spine 2 Lo2	10.20.20.3	/32	VTEP Source
Arista LEAF S Lo1	10.10.10.7	/32	BGP RID/Source
Arista LEAF S Lo2	**10.20.20.4**	/32	VTEP Source
Cisco LEAF S Lo1	10.10.10.8	/32	BGP RID/Source
Cisco LEAF S Lo2	**20.20.20.5**	/32	VTEP Source

BGP			
Host	BGP AS Number	Update source	Network
Cisco LEAF A	65514	Lo 1	Lo 2
Cisco LEAF B	65514	Lo 1	Lo 2
Arista LEAF A	65514	Lo 1	Lo 2
Arista LEAF B	65514	Lo 1	Lo 2
Cisco LEAF S	65514	Lo 1	Lo 2
Arista LEAF S	65514	Lo 1	Lo 2
Cisco Spine 1	65514	Lo 1	Lo 2
Cisco Spine 2	65514	Lo 1	Lo 2

1. Basic Configuration and Implementation

1.1 Introduction to VXLAN

VXLAN (Virtual Extensible LAN) is a network virtualization technology designed to overcome the limitations of traditional VLANs in large-scale data centers. It enables the creation of a virtual Layer 2 network over a Layer 3 infrastructure, facilitating network scalability, flexibility, and isolation.

Key Features:

Scalability: Supports up to 16 million unique network identifiers (VNIs).

Flexibility: Extends Layer 2 networks over Layer 3.

Mobility: Allows VM movement without IP changes.

Multi-Tenancy: Provides isolation for multiple tenants.

1.1.1 VXLAN Header

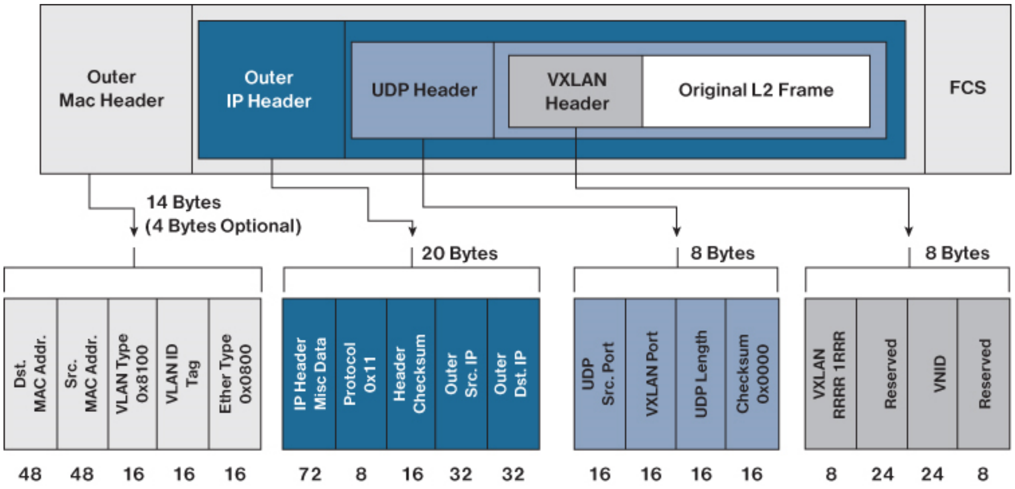


Figure 3: VXLAN Header (https://learningnetwork.cisco.com)

The VXLAN header encapsulates the original Ethernet frame for transport over an IP network. It includes:

Flags (8 bits): The "I" bit (most significant bit) is set to 1 to indicate a valid VNI; the remaining 7 bits are reserved.

VXLAN Network Identifier (VNI) (24 bits): Uniquely identifies the VXLAN segment, supporting up to 16 million segments.

Reserved Fields: Set to 0, reserved for future use.

UDP Header: UDP is used for transport, with the destination port typically being 4789.

Outer IP Header: Contains source and destination IP addresses of VXLAN tunnel endpoints (VTEPs).

Outer Ethernet Header: Allows transmission over the physical Ethernet network.

1.1.2 Encapsulation Process

Encapsulation: Adds VXLAN, UDP, IP, and Ethernet headers to the original Ethernet frame.

Transmission: Sends the encapsulated packet across the IP network.

Decapsulation: Strips off the outer headers at the destination VTEP, delivering the original Ethernet frame. [1]

1.2 MP-BGP and Multi-Protocol BGP for EVPN

MP-BGP (Multi-Protocol Border Gateway Protocol): An extension of BGP that supports multiple network layer protocols, enabling the exchange of diverse routing information within a single BGP session. Multi-Protocol BGP for EVPN: Uses MP-BGP to exchange Layer 2 and Layer 3 reachability information for Ethernet VPNs (EVPNs). It facilitates scalable and flexible multi-tenant Layer 2 and Layer 3 VPN services over a common IP infrastructure.

AFI/SAFI for EVPN:

AFI (Address Family Identifier): 25 (Layer 2 VPN address family).

SAFI (Subsequent Address Family Identifier): 70 (EVPN address family).

By using MP-BGP with these AFI/SAFI values, EVPN can effectively advertise and manage routing information, supporting virtualized networks. [2],[3]

1.3 Ethernet VPN (EVPN)

Ethernet VPN (EVPN) is an advanced network technology that leverages BGP for the control plane to provide scalable and flexible Layer 2 and Layer 3 VPN services. EVPN allows for the extension of Layer 2 networks over an IP/MPLS backbone, making it ideal for modern data center interconnects and multi-tenant cloud environments. It provides features such as MAC address learning over the control plane, efficient multicast handling, and integrated routing and bridging.[4]

1.4 EVPN Route Types

- Type-1 Route (Ethernet A-D Route): Advertises reachability of multi-homed Ethernet segments for fast convergence and split-horizon filtering.
- Type-2 Route (MAC/IP Advertisement Route): Advertises MAC addresses or MAC and IP bindings, supporting ARP suppression and reducing flooding.
- Type-3 Route (Inclusive Multicast Route): Advertises VNI membership to enable dynamic discovery of remote VTEPs and BUM traffic forwarding.

- Type-4 Route (Ethernet Segment Route): Discovers VTEPs on the same Ethernet segment and supports multi-homing and DF election.
- Type-5 Route (IP Prefix Route): Advertises IP prefixes for Layer 3 VPNs, optimizing IP prefix advertisement and reducing churn.[5]

1.5 BGP Implementation

BGP (Border Gateway Protocol) is utilized for both underlay and overlay networks, specifically using iBGP (Internal BGP) within a single Autonomous System (AS). The AS number assigned for this project is 65512.

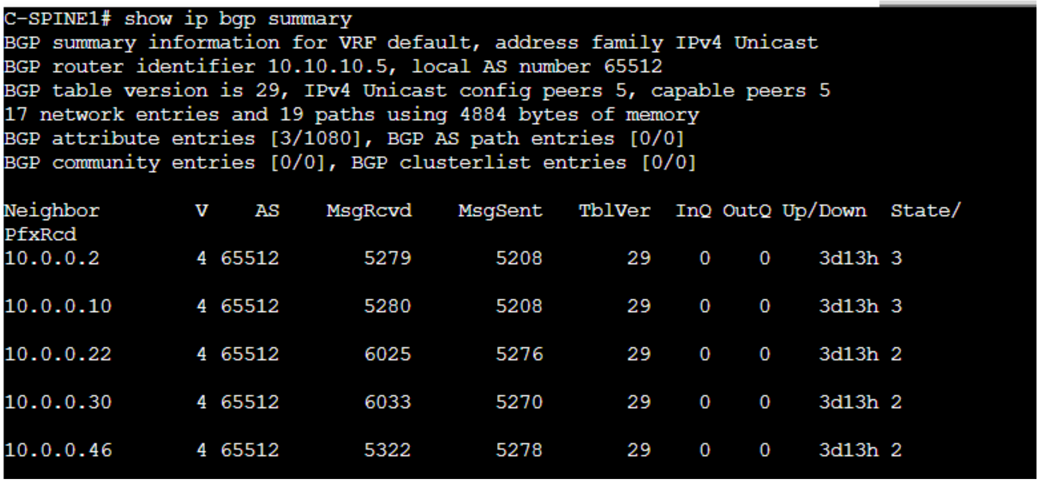
- Underlay Network:

- BGP is configured to establish routes between the spine and leaf switches, ensuring connectivity within the network.

- Overlay Network:

- EVPN (Ethernet VPN) is implemented over BGP to facilitate VXLAN (Virtual Extensible LAN) tunnels, enabling scalable and flexible Layer 2 and Layer 3 network services.[6]

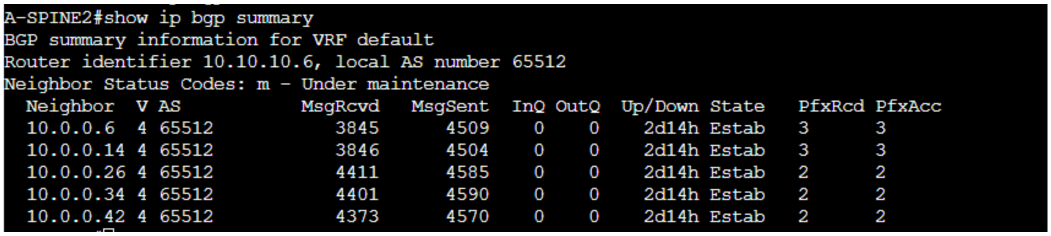
Show BGP neighborship Cisco Spine



```
C-SPINE1# show ip bgp summary
BGP summary information for VRF default, address family IPv4 Unicast
BGP router identifier 10.10.10.5, local AS number 65512
BGP table version is 29, IPv4 Unicast config peers 5, capable peers 5
17 network entries and 19 paths using 4884 bytes of memory
BGP attribute entries [3/1080], BGP AS path entries [0/0]
BGP community entries [0/0], BGP clusterlist entries [0/0]

Neighbor        V    AS    MsgRcvd    MsgSent   TblVer  InQ OutQ Up/Down  State/
PfxRcd
10.0.0.2        4 65512       5279       5208       29    0    0    3d13h 3

10.0.0.10       4 65512       5280       5208       29    0    0    3d13h 3

10.0.0.22       4 65512       6025       5276       29    0    0    3d13h 2

10.0.0.30       4 65512       6033       5270       29    0    0    3d13h 2

10.0.0.46       4 65512       5322       5278       29    0    0    3d13h 2
```

Figure 4: Show BGP neighborship Cisco Spine

Show BGP neighborship Arista Spine



```
A-SPINE2#show ip bgp summary
BGP summary information for VRF default
Router identifier 10.10.10.6, local AS number 65512
Neighbor Status Codes: m - Under maintenance
  Neighbor  V AS           MsgRcvd   MsgSent  InQ OutQ  Up/Down State    PfxRcd PfxAcc
  10.0.0.6  4 65512           3845      4509    0    0    2d14h Estab    3      3
  10.0.0.14 4 65512           3846      4504    0    0    2d14h Estab    3      3
  10.0.0.26 4 65512           4411      4585    0    0    2d14h Estab    2      2
  10.0.0.34 4 65512           4401      4590    0    0    2d14h Estab    2      2
  10.0.0.42 4 65512           4373      4570    0    0    2d14h Estab    2      2
```

Figure 5: Show BGP neighborship Arista Spine

1.6 Features and Configurations

- Cisco Leafs:

- VPC (Virtual Port Channel): This feature allows multiple physical links between the switches to appear as a single logical link, providing redundancy and load balancing. It enhances fault tolerance and simplifies network topology.

- Arista Leafs:

- MLAG (Multi-chassis Link Aggregation): Similar to Cisco's VPC, MLAG allows the creation of logical link aggregation groups across two switches. This ensures that even if one switch fails, the traffic continues to flow through the other, maintaining network stability and performance.

Scientific Explanation of Features:

Virtual Port Channel (VPC)

Virtual Port Channel (VPC) is a Cisco-specific technology that allows multiple physical links between two network devices to combine into a single logical link. This configuration improves bandwidth utilization, provides redundancy, and ensures high availability. VPC is crucial in scenarios where link or device failure could otherwise result in network outages. It allows for active-active link utilization, reducing the chances of single points of failure.

Multi-chassis Link Aggregation (MLAG)

Multi-chassis Link Aggregation (MLAG) is an Arista technology similar to Cisco's VPC. It enables the creation of link aggregation groups that span multiple switches, ensuring continuous network service even if one switch fails. MLAG is pivotal for maintaining high availability and load balancing in data center networks, providing an efficient mechanism to manage traffic and ensure network reliability.

Virtual Tunnel End Point (VTEP)

A VTEP is a crucial component in a VXLAN architecture, responsible for encapsulating and de-encapsulating packets. It serves as the termination point for VXLAN tunnels and uses loopback addresses as the source and destination IPs for tunnel endpoints. In Cisco devices, VTEPs are dynamically discovered via BGP, while in Arista devices, they must be manually configured.

2. Scenario 1: VXLAN Route Types 1 and 2 Implementation

2.1 Overview

In this scenario, we delve into the implementation and validation of VXLAN Route Types 1 and 2. This involves configuring iBGP Neighborship, establishing EVPN (Ethernet VPN) over BGP, and creating VXLAN tunnels using VTEP (VXLAN Tunnel Endpoints). The focus is on ensuring seamless interoperability between Cisco and Arista devices within a spine-leaf topology.

2.2 Configuration Steps

1. iBGP Neighborship Configuration

After completing the basic setup, the next step involves building iBGP Neighborship between all devices within the same AS (65512). This ensures that all devices can communicate over the same BGP instance. Each spine and leaf switch are configured with two loopback interfaces, which are advertised through BGP.

- Loopback Interfaces: These interfaces are crucial as they serve as the source IPs for BGP peering and VXLAN tunnels.

- BGP Configuration: Only the networks of the loopback interfaces are advertised, allowing for efficient routing and minimizing unnecessary traffic.

2. EVPN Configuration

With the iBGP neighborhoods established, the EVPN configuration is implemented within the BGP framework. This setup involves creating VXLAN tunnels with loopback interfaces serving as the source for the VTEP addresses.

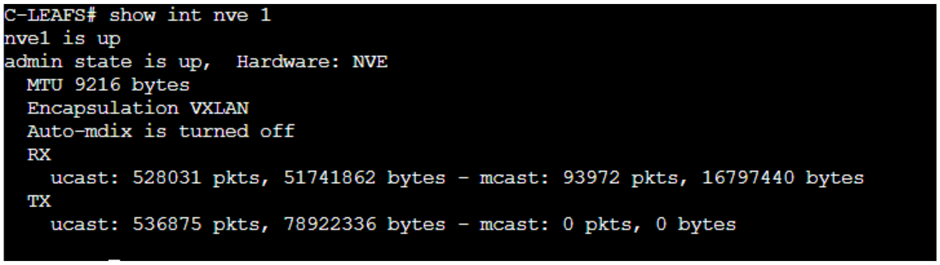
- VXLAN Interfaces:

- Cisco: Interface NVE-1 is created for VXLAN.

- Arista: Interface VXLAN-1 is created for VXLAN.

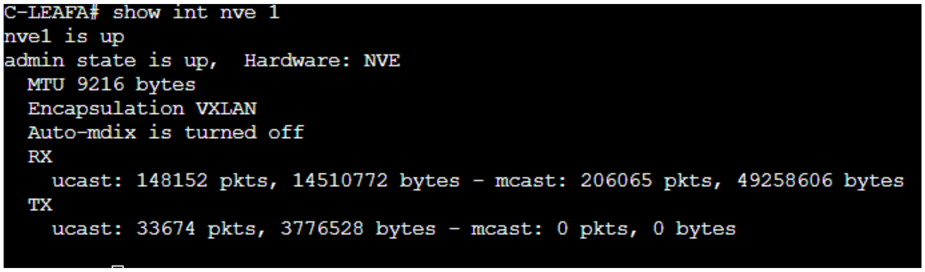
In Cisco, this command is used to display information about Network Virtualization Edge (NVE) interfaces. NVEs are responsible for encapsulating and decapsulating packets for VXLAN (Virtual Extensible LAN) tunnels.

The output typically includes details such as the NVE interface status, associated Virtual Network Identifiers (VNIs), tunnel statistics, and operational parameters. This helps in monitoring and troubleshooting VXLAN tunnels.



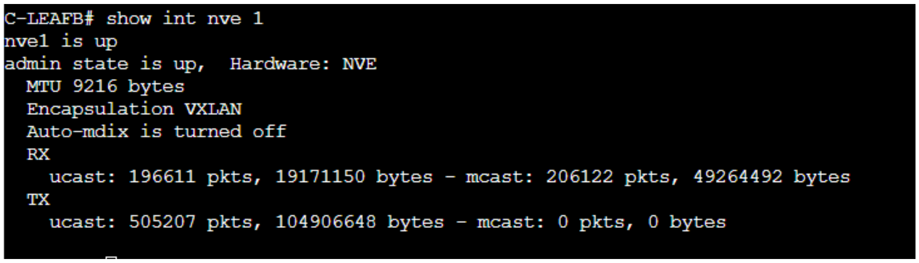
```
C-LEAFS# show int nve 1
nve1 is up
admin state is up,  Hardware: NVE
  MTU 9216 bytes
  Encapsulation VXLAN
  Auto-mdix is turned off
  RX
    ucast: 528031 pkts, 51741862 bytes - mcast: 93972 pkts, 16797440 bytes
  TX
    ucast: 536875 pkts, 78922336 bytes - mcast: 0 pkts, 0 bytes
```

Figure 6: (C-LeAFS) Interface nve 1



```
C-LEAFA# show int nve 1
nve1 is up
admin state is up,  Hardware: NVE
  MTU 9216 bytes
  Encapsulation VXLAN
  Auto-mdix is turned off
  RX
    ucast: 148152 pkts, 14510772 bytes - mcast: 206065 pkts, 49258606 bytes
  TX
    ucast: 33674 pkts, 3776528 bytes - mcast: 0 pkts, 0 bytes
```

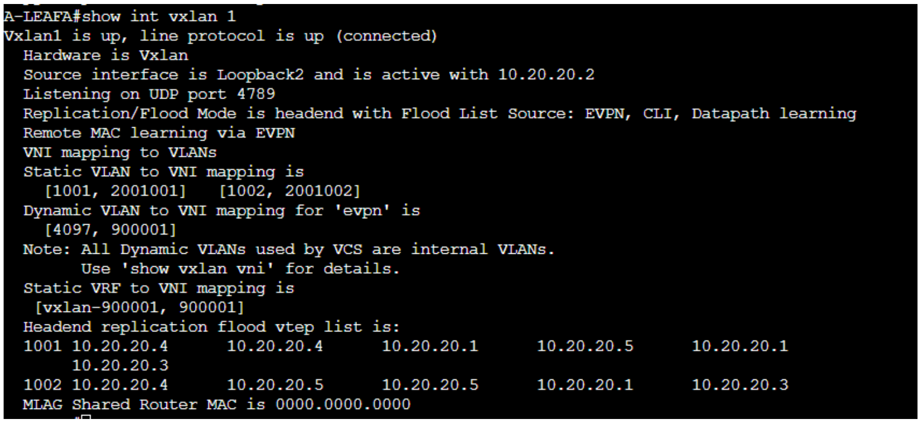
Figure 7: (C-LEAFA) Interface nve 1



```
C-LEAFB# show int nve 1
nve1 is up
admin state is up,  Hardware: NVE
  MTU 9216 bytes
  Encapsulation VXLAN
  Auto-mdix is turned off
  RX
    ucast: 196611 pkts, 19171150 bytes - mcast: 206122 pkts, 49264492 bytes
  TX
    ucast: 505207 pkts, 104906648 bytes - mcast: 0 pkts, 0 bytes
```

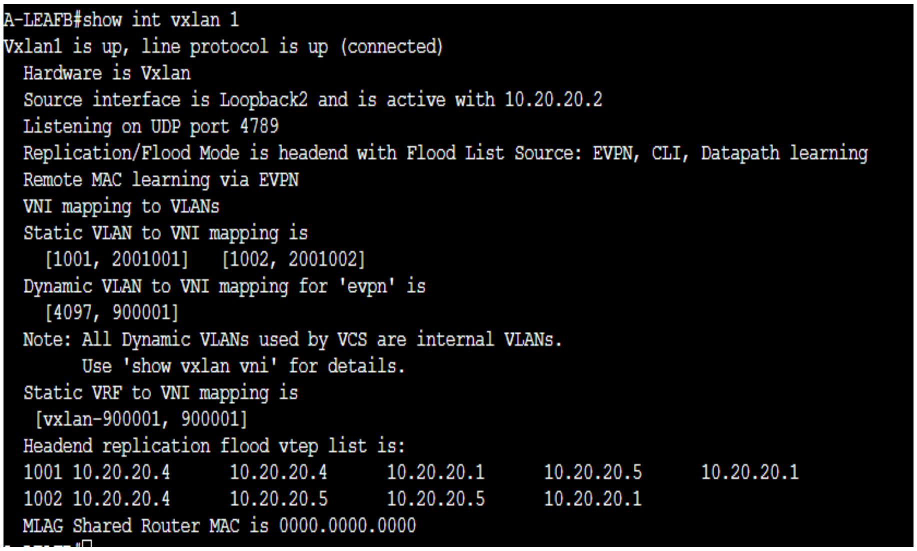
Figure 8: (C-LEAFB) Interface nve 1

In Arista devices, this command is used to display information about VXLAN interfaces. VXLAN interfaces are used to create overlay networks over existing Layer 3 infrastructure, enabling the extension of Layer 2 networks across geographically dispersed locations. The output provides details about the VXLAN interface, such as the interface status, associated VNIs, tunnel endpoints, and packet statistics. This information is crucial for ensuring the proper functioning of VXLAN-based networks.



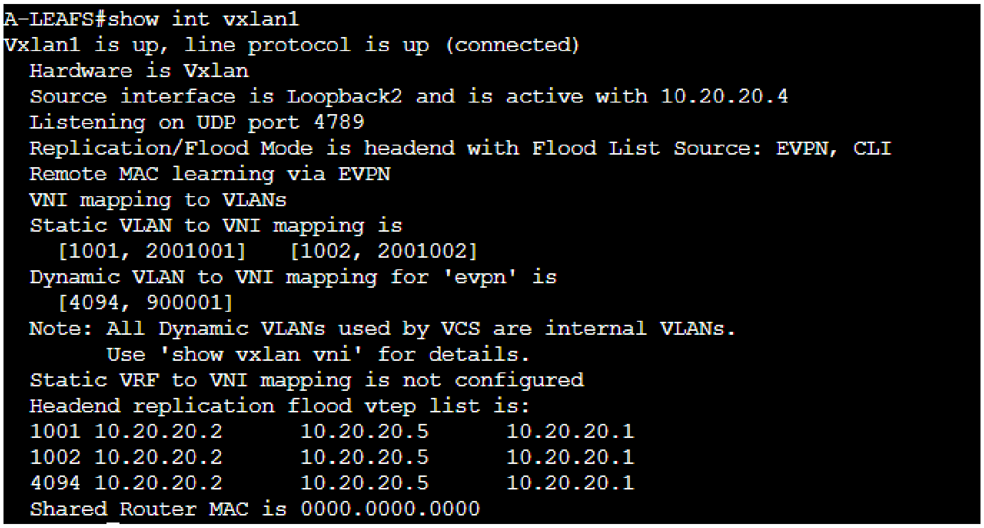
```
A-LEAFA#show int vxlan 1
Vxlan1 is up, line protocol is up (connected)
  Hardware is Vxlan
  Source interface is Loopback2 and is active with 10.20.20.2
  Listening on UDP port 4789
  Replication/Flood Mode is headend with Flood List Source: EVPN, CLI, Datapath learning
  Remote MAC learning via EVPN
  VNI mapping to VLANs
  Static VLAN to VNI mapping is
    [1001, 2001001]   [1002, 2001002]
  Dynamic VLAN to VNI mapping for 'evpn' is
    [4097, 900001]
  Note: All Dynamic VLANs used by VCS are internal VLANs.
        Use 'show vxlan vni' for details.
  Static VRF to VNI mapping is
   [vxlan-900001, 900001]
  Headend replication flood vtep list is:
  1001 10.20.20.4      10.20.20.4      10.20.20.1      10.20.20.5      10.20.20.1
       10.20.20.3
  1002 10.20.20.4      10.20.20.5      10.20.20.5      10.20.20.1      10.20.20.3
  MLAG Shared Router MAC is 0000.0000.0000
```

Figure 9: (A-LEAFA) Interface VXLAN 1



```
A-LEAFB#show int vxlan 1
Vxlan1 is up, line protocol is up (connected)
  Hardware is Vxlan
  Source interface is Loopback2 and is active with 10.20.20.2
  Listening on UDP port 4789
  Replication/Flood Mode is headend with Flood List Source: EVPN, CLI, Datapath learning
  Remote MAC learning via EVPN
  VNI mapping to VLANs
  Static VLAN to VNI mapping is
    [1001, 2001001]   [1002, 2001002]
  Dynamic VLAN to VNI mapping for 'evpn' is
    [4097, 900001]
  Note: All Dynamic VLANs used by VCS are internal VLANs.
        Use 'show vxlan vni' for details.
  Static VRF to VNI mapping is
   [vxlan-900001, 900001]
  Headend replication flood vtep list is:
  1001 10.20.20.4      10.20.20.4      10.20.20.1      10.20.20.5      10.20.20.1
  1002 10.20.20.4      10.20.20.5      10.20.20.5      10.20.20.1
  MLAG Shared Router MAC is 0000.0000.0000
```

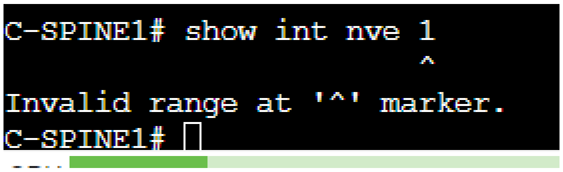
Figure 10: (A-LEAFB) Interface VXLAN 1



```
A-LEAFS#show int vxlan1
Vxlan1 is up, line protocol is up (connected)
  Hardware is Vxlan
  Source interface is Loopback2 and is active with 10.20.20.4
  Listening on UDP port 4789
  Replication/Flood Mode is headend with Flood List Source: EVPN, CLI
  Remote MAC learning via EVPN
  VNI mapping to VLANs
  Static VLAN to VNI mapping is
    [1001, 2001001]    [1002, 2001002]
  Dynamic VLAN to VNI mapping for 'evpn' is
    [4094, 900001]
  Note: All Dynamic VLANs used by VCS are internal VLANs.
        Use 'show vxlan vni' for details.
  Static VRF to VNI mapping is not configured
  Headend replication flood vtep list is:
  1001 10.20.20.2      10.20.20.5      10.20.20.1
  1002 10.20.20.2      10.20.20.5      10.20.20.1
  4094 10.20.20.2      10.20.20.5      10.20.20.1
  Shared Router MAC is 0000.0000.0000
```

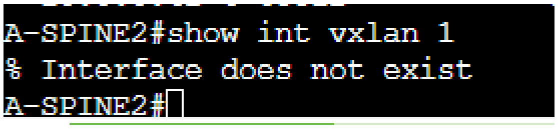
Figure 11: (A-LEAFS) Interface VXLAN 1

Cisco and Arista spine switches do not have VXLAN interfaces because their primary role is to act as transit devices. VXLAN interfaces are typically configured on leaf switches where VXLAN tunnels originate and terminate. This simplifies spine switch configuration and focuses them on high-speed packet forwarding.



```
C-SPINE1# show int nve 1
                        ^
Invalid range at '^' marker.
C-SPINE1# 
```

Figure 12: (C-SPINE1) Interface nve 1



```
A-SPINE2#show int vxlan 1
% Interface does not exist
A-SPINE2#
```

Figure 13: (A-SPINE2) Interface nve 1

For Leafs:

```
interface nve1
  no shutdown
  source-interface loopback2
  host-reachability protocol bgp
  member vni 2001001
    mcast-group 239.0.0.1
  member vni 2001002
    mcast-group 239.0.0.1
```

3. VLAN and VNI Configuration

Two VLANs are configured for this scenario:

- VLAN 1001

- VLAN 1002

Correspondingly, two VNIs (VXLAN Network Identifiers) are created:

- VNI 2001001: Mapped to VLAN 1001

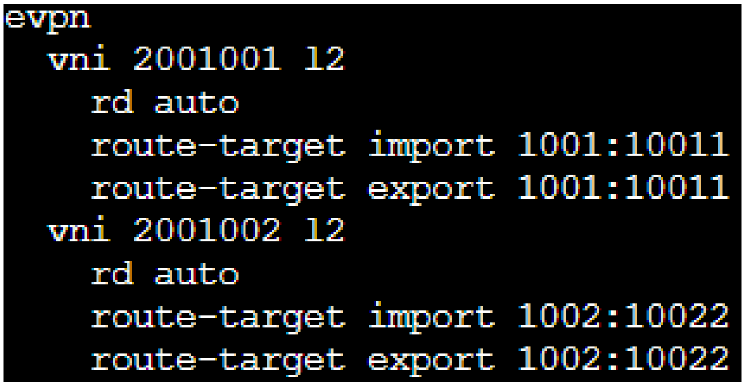
- VNI 2001002: Mapped to VLAN 1002

Evpn configuration:

EVPN (Ethernet VPN) configuration in BGP is essential for enabling VXLAN tunnels. This configuration allows the exchange of MAC address reachability information between the leaf switches, facilitating Layer 2 and Layer 3 connectivity over the VXLAN network.

Cisco:

The image displays a Cisco leaf configuration for setting up EVPN with VXLAN, defining two Layer 2 VNIs, 2001001 and 2001002. Each VNI is configured with an automatic Route Distinguisher (RD) and specific route targets for importing and exporting routes. For VNI 2001001, the import and export route targets are set to 1001:10011, while for VNI 2001002, they are set to 1002:10022. This configuration enables network segmentation and controls the import and export of routes, facilitating efficient traffic management within the overlay network.

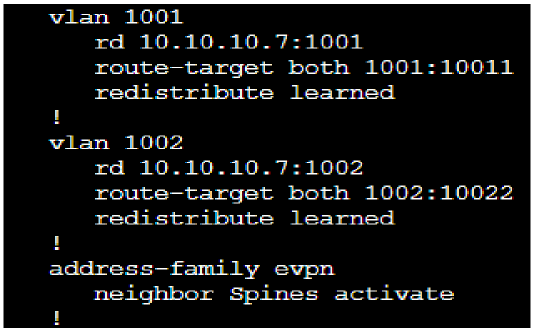


```
evpn
  vni 2001001 l2
    rd auto
    route-target import 1001:10011
    route-target export 1001:10011
  vni 2001002 l2
    rd auto
    route-target import 1002:10022
    route-target export 1002:10022
```

Figure 14: Cisco configuration for EVPN

Arista:

The image shows an Arista configuration for setting up EVPN with VXLAN for VLANs 1001 and 1002. For VLAN 1001, it defines the RD as 10.10.10.7:1001, sets the route target for both import and export to 1001:10011, and redistributes learned routes. Similarly, for VLAN 1002, it defines the RD as 10.10.10.7:1002, sets the route target for both import and export to 1002:10022, and redistributes learned routes. Additionally, the configuration activates the EVPN address family and enables the neighbor "Spines".

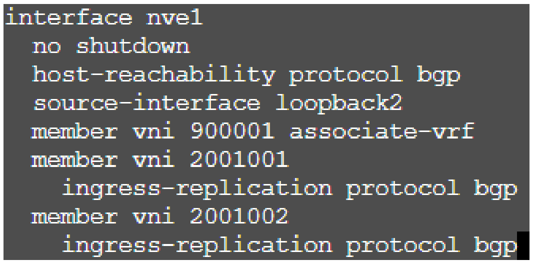


```
   vlan 1001
      rd 10.10.10.7:1001
      route-target both 1001:10011
      redistribute learned
   !
   vlan 1002
      rd 10.10.10.7:1002
      route-target both 1002:10022
      redistribute learned
   !
   address-family evpn
      neighbor Spines activate
   !
```

Figure 15: Arista configuration for EVPN

4. VLAN Extensive Configuration

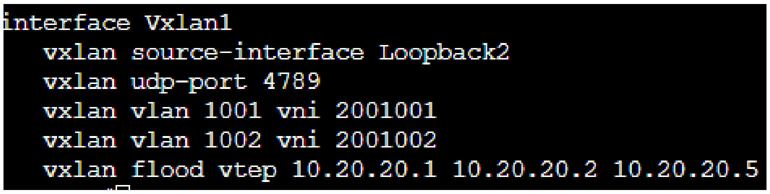
The VLAN's extensive configuration includes encapsulation using a tag to ensure proper communication between VNIs. This encapsulation allows the VXLAN packets to be properly routed across the network.



```
interface nve1
  no shutdown
  host-reachability protocol bgp
  source-interface loopback2
  member vni 900001 associate-vrf
  member vni 2001001
    ingress-replication protocol bgp
  member vni 2001002
    ingress-replication protocol bgp
```

Figure 16: Interface nve1 Configuration

Arista:

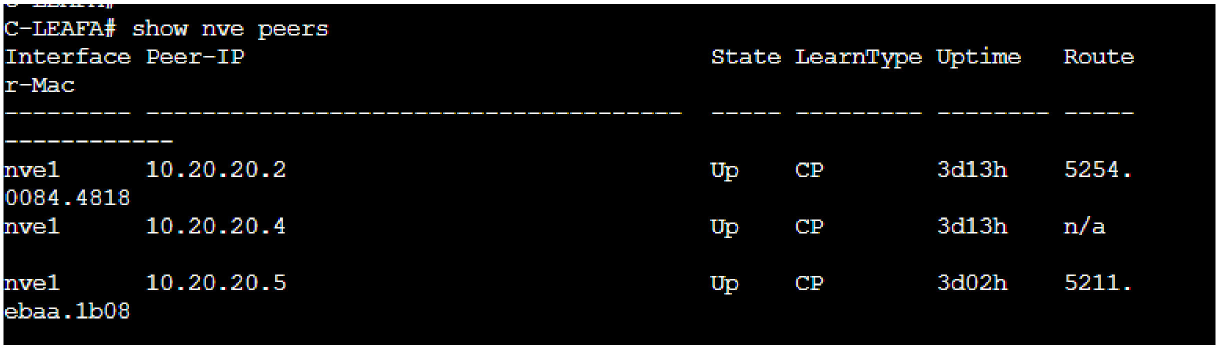


```
interface Vxlan1
   vxlan source-interface Loopback2
   vxlan udp-port 4789
   vxlan vlan 1001 vni 2001001
   vxlan vlan 1002 vni 2001002
   vxlan flood vtep 10.20.20.1 10.20.20.2 10.20.20.5
```

Figure 17: Interface VXLAN1 Configuration

VTEP Pears:

A VTEP is a crucial component in a VXLAN architecture, responsible for encapsulating and de-encapsulating packets. It serves as the termination point for VXLAN tunnels and uses loopback addresses as the source and destination IPs for tunnel endpoints. In Cisco devices, VTEPs are dynamically discovered via BGP, while in Arista devices, they must be manually configured.



```
C-LEAFA# show nve peers
Interface Peer-IP                                 State LearnType Uptime   Route
r-Mac
--------- --------------------------------------  ----- --------- -------- -----
------------
nve1      10.20.20.2                              Up    CP        3d13h    5254.
0084.4818
nve1      10.20.20.4                              Up    CP        3d13h    n/a

nve1      10.20.20.5                              Up    CP        3d02h    5211.
ebaa.1b08
```

Figure 18: Cisco C-LEAFA nve peers

In the Arista configuration, VTEP (Virtual Tunnel End Point) peers are manually specified using their IP addresses. The number 0084.4818 refers to a MAC address format, which is part of the EVPN MAC address learning process.

2.3 Validation and Testing

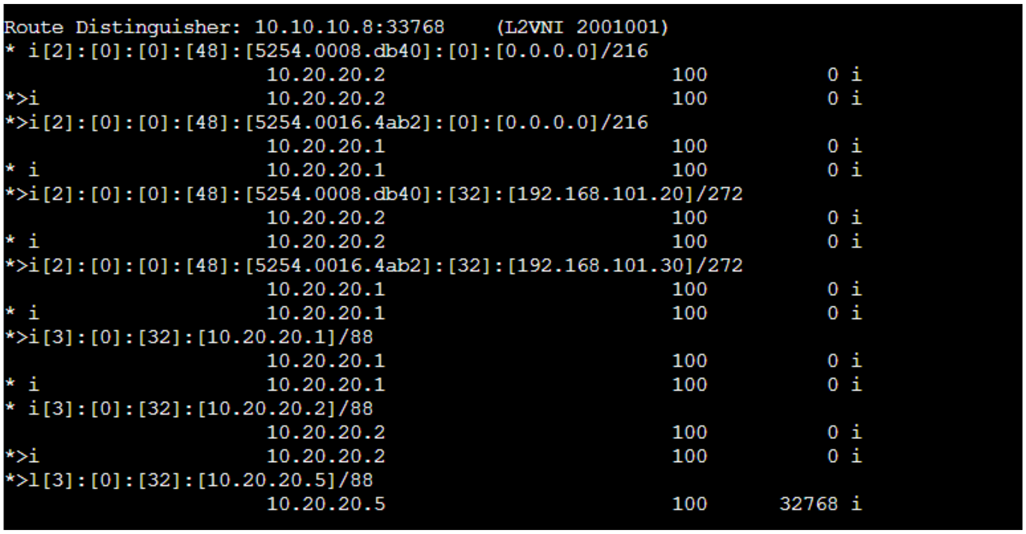
To verify the configuration and ensure that Route Types 1 and 2 are functioning correctly, the following commands are used:

- Cisco:

```
show bgp l2vpn evpn
```

- This command displays information about the EVPN routes, including Route Type 2 details and VNI numbers.

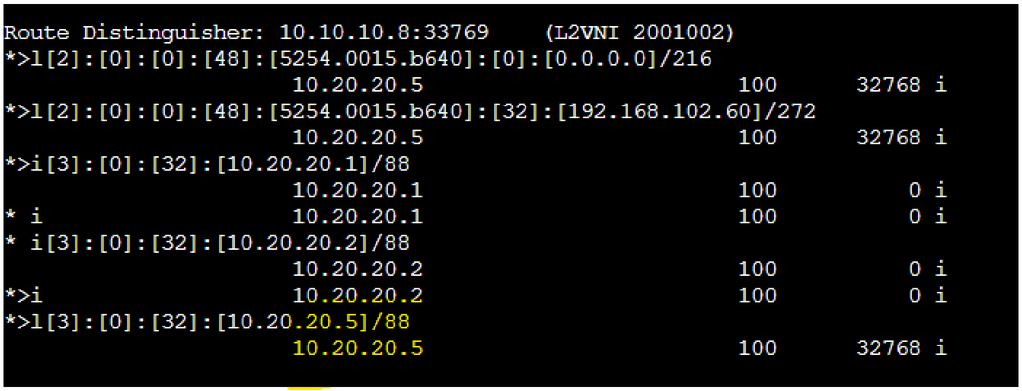
The IP address 192.168.101.30 with MAC address 4ab2 is seen in the Cisco output because it is learned from a vPC (Virtual Port Channel). This means the MAC address is associated with the vPC member links. In a vPC setup, multiple links are combined into a single logical link, providing redundancy and load balancing. Because of that we see two next hops with same IP for that:



```
Route Distinguisher: 10.10.10.8:33768    (L2VNI 2001001)
* i[2]:[0]:[0]:[48]:[5254.0008.db40]:[0]:[0.0.0.0]/216
                      10.20.20.2                          100          0 i
*>i                   10.20.20.2                          100          0 i
*>i[2]:[0]:[0]:[48]:[5254.0016.4ab2]:[0]:[0.0.0.0]/216
                      10.20.20.1                          100          0 i
* i                   10.20.20.1                          100          0 i
*>i[2]:[0]:[0]:[48]:[5254.0008.db40]:[32]:[192.168.101.20]/272
                      10.20.20.2                          100          0 i
* i                   10.20.20.2                          100          0 i
*>i[2]:[0]:[0]:[48]:[5254.0016.4ab2]:[32]:[192.168.101.30]/272
                      10.20.20.1                          100          0 i
* i                   10.20.20.1                          100          0 i
*>i[3]:[0]:[32]:[10.20.20.1]/88
                      10.20.20.1                          100          0 i
* i                   10.20.20.1                          100          0 i
* i[3]:[0]:[32]:[10.20.20.2]/88
                      10.20.20.2                          100          0 i
*>i                   10.20.20.2                          100          0 i
*>l[3]:[0]:[32]:[10.20.20.5]/88
                      10.20.20.5                          100      32768 i
```

Figure 19: Cisco BGP L2VPN VNI 2001001

The IP address 10.20.20.5 is seen as a single endpoint because it is not part of a vPC setup. Unlike vPC, which aggregates multiple physical links, a single link or endpoint will show up as an individual IP address.



```
Route Distinguisher: 10.10.10.8:33769    (L2VNI 2001002)
*>l[2]:[0]:[0]:[48]:[5254.0015.b640]:[0]:[0.0.0.0]/216
                      10.20.20.5                          100      32768 i
*>l[2]:[0]:[0]:[48]:[5254.0015.b640]:[32]:[192.168.102.60]/272
                      10.20.20.5                          100      32768 i
*>i[3]:[0]:[32]:[10.20.20.1]/88
                      10.20.20.1                          100          0 i
* i                   10.20.20.1                          100          0 i
* i[3]:[0]:[32]:[10.20.20.2]/88
                      10.20.20.2                          100          0 i
*>i                   10.20.20.2                          100          0 i
*>l[3]:[0]:[32]:[10.20.20.5]/88
                      10.20.20.5                          100      32768 i
```

Figure 20: Cisco BGP L2VPN VNI 2001002

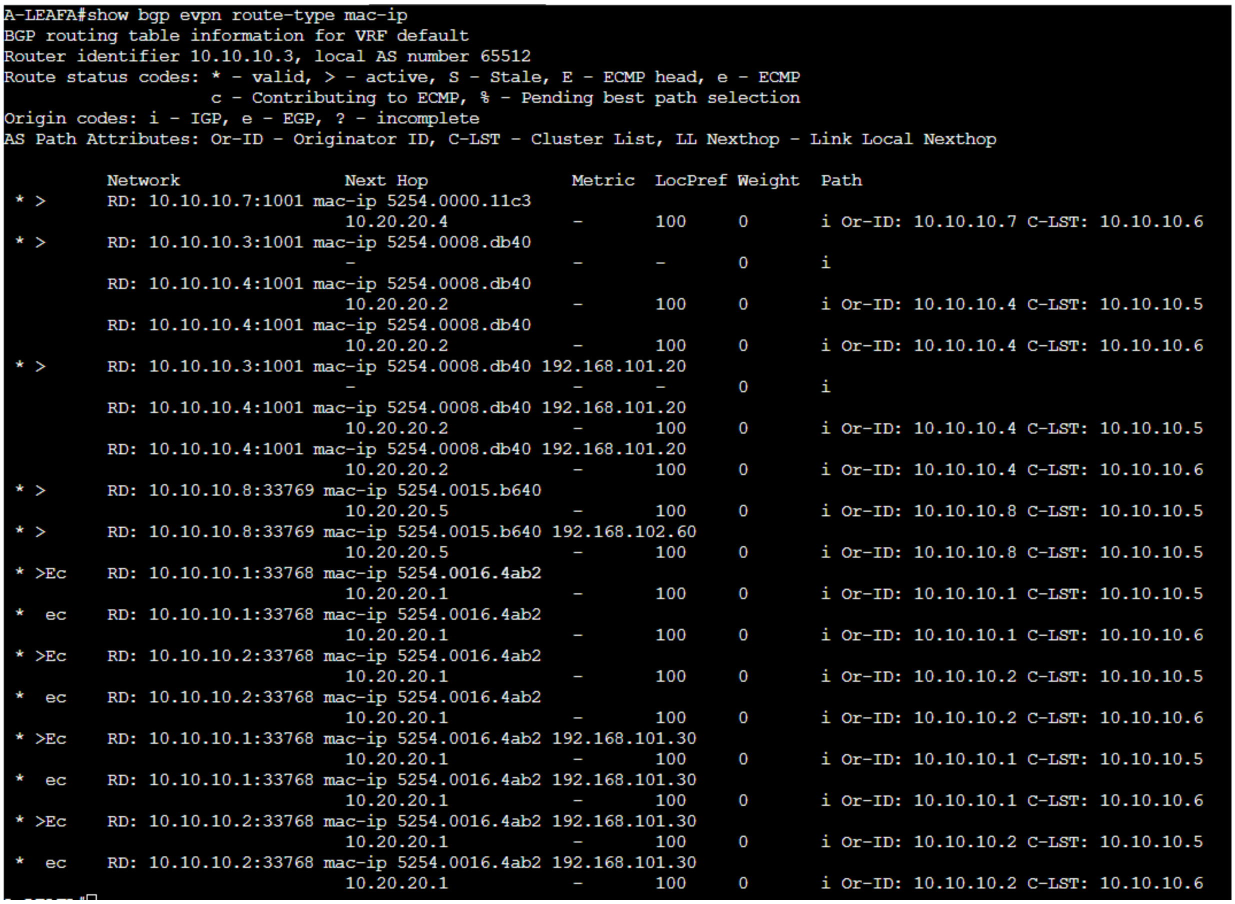
Arista:

show bgp evpn route-type mac-ip

- This command provides similar details for Arista devices, confirming the MAC-IP mappings and VNI configurations.

Both commands confirmed that the configurations were correct and the scenario was functioning as expected.

The MAC address 4ab2 coming from 10.10.10.1 is learned through the BGP EVPN control plane. This shows the MAC address and the associated IP address are reachable via the VTEP with IP 10.10.10.1.

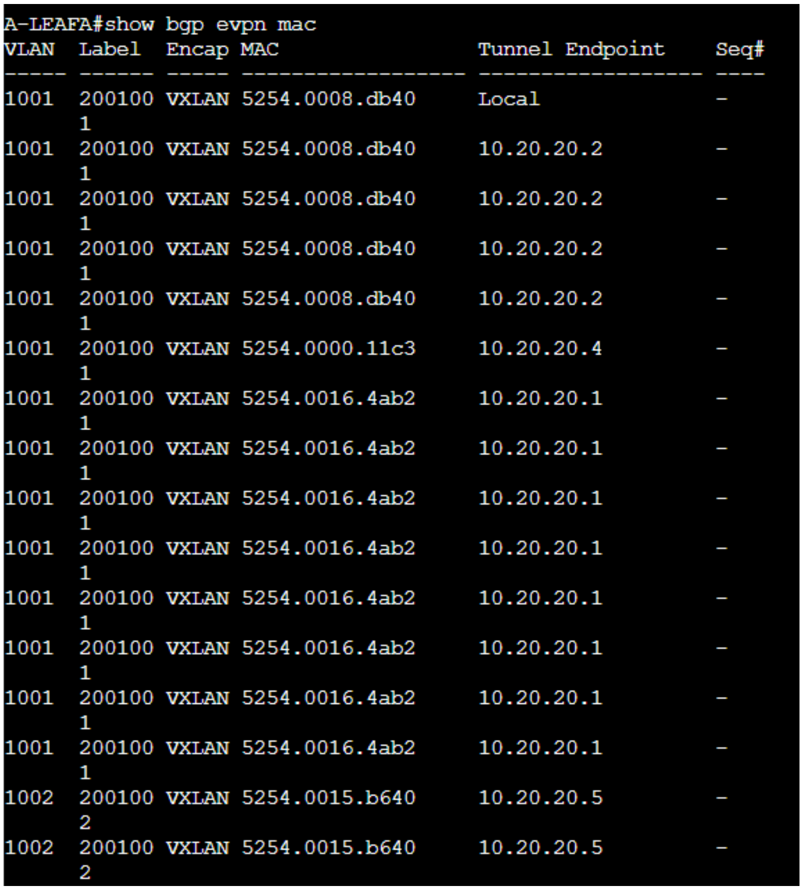


```
A-LEAFA#show bgp evpn route-type mac-ip
BGP routing table information for VRF default
Router identifier 10.10.10.3, local AS number 65512
Route status codes: * - valid, > - active, S - Stale, E - ECMP head, e - ECMP
                    c - Contributing to ECMP, % - Pending best path selection
Origin codes: i - IGP, e - EGP, ? - incomplete
AS Path Attributes: Or-ID - Originator ID, C-LST - Cluster List, LL Nexthop - Link Local Nexthop

          Network                Next Hop              Metric  LocPref Weight  Path
 * >      RD: 10.10.10.7:1001 mac-ip 5254.0000.11c3
                                 10.20.20.4            -       100     0       i Or-ID: 10.10.10.7 C-LST: 10.10.10.6
 * >      RD: 10.10.10.3:1001 mac-ip 5254.0008.db40
                                 -                     -       -       0       i
          RD: 10.10.10.4:1001 mac-ip 5254.0008.db40
                                 10.20.20.2            -       100     0       i Or-ID: 10.10.10.4 C-LST: 10.10.10.5
          RD: 10.10.10.4:1001 mac-ip 5254.0008.db40
                                 10.20.20.2            -       100     0       i Or-ID: 10.10.10.4 C-LST: 10.10.10.6
 * >      RD: 10.10.10.3:1001 mac-ip 5254.0008.db40 192.168.101.20
                                 -                     -       -       0       i
          RD: 10.10.10.4:1001 mac-ip 5254.0008.db40 192.168.101.20
                                 10.20.20.2            -       100     0       i Or-ID: 10.10.10.4 C-LST: 10.10.10.5
          RD: 10.10.10.4:1001 mac-ip 5254.0008.db40 192.168.101.20
                                 10.20.20.2            -       100     0       i Or-ID: 10.10.10.4 C-LST: 10.10.10.6
 * >      RD: 10.10.10.8:33769 mac-ip 5254.0015.b640
                                 10.20.20.5            -       100     0       i Or-ID: 10.10.10.8 C-LST: 10.10.10.5
 * >      RD: 10.10.10.8:33769 mac-ip 5254.0015.b640 192.168.102.60
                                 10.20.20.5            -       100     0       i Or-ID: 10.10.10.8 C-LST: 10.10.10.5
 * >Ec    RD: 10.10.10.1:33768 mac-ip 5254.0016.4ab2
                                 10.20.20.1            -       100     0       i Or-ID: 10.10.10.1 C-LST: 10.10.10.5
 *  ec    RD: 10.10.10.1:33768 mac-ip 5254.0016.4ab2
                                 10.20.20.1            -       100     0       i Or-ID: 10.10.10.1 C-LST: 10.10.10.6
 * >Ec    RD: 10.10.10.2:33768 mac-ip 5254.0016.4ab2
                                 10.20.20.1            -       100     0       i Or-ID: 10.10.10.2 C-LST: 10.10.10.5
 *  ec    RD: 10.10.10.2:33768 mac-ip 5254.0016.4ab2
                                 10.20.20.1            -       100     0       i Or-ID: 10.10.10.2 C-LST: 10.10.10.6
 * >Ec    RD: 10.10.10.1:33768 mac-ip 5254.0016.4ab2 192.168.101.30
                                 10.20.20.1            -       100     0       i Or-ID: 10.10.10.1 C-LST: 10.10.10.5
 *  ec    RD: 10.10.10.1:33768 mac-ip 5254.0016.4ab2 192.168.101.30
                                 10.20.20.1            -       100     0       i Or-ID: 10.10.10.1 C-LST: 10.10.10.6
 * >Ec    RD: 10.10.10.2:33768 mac-ip 5254.0016.4ab2 192.168.101.30
                                 10.20.20.1            -       100     0       i Or-ID: 10.10.10.2 C-LST: 10.10.10.5
 *  ec    RD: 10.10.10.2:33768 mac-ip 5254.0016.4ab2 192.168.101.30
                                 10.20.20.1            -       100     0       i Or-ID: 10.10.10.2 C-LST: 10.10.10.6
```

Figure 21: Arista EVPN Route-Type MAC-IP

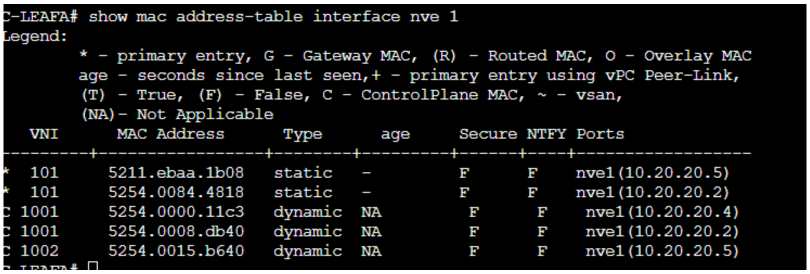
By comparing the output of EVPN route-type 2 commands between Cisco and Arista we can find:

In both cases, the route for mac ending with b640 is advertised by the peer with the same VTEP IP 10.20.20.5. Mac db40 is local for Arista mLAG and 4ab2 is local for Cisco vPC.[7]



```
A-LEAFA#show bgp evpn mac
VLAN  Label  Encap MAC                Tunnel Endpoint    Seq#
----- ------ ----- ------------------ ------------------ ----
1001  200100 VXLAN 5254.0008.db40     Local              -
      1
1001  200100 VXLAN 5254.0008.db40     10.20.20.2         -
      1
1001  200100 VXLAN 5254.0008.db40     10.20.20.2         -
      1
1001  200100 VXLAN 5254.0008.db40     10.20.20.2         -
      1
1001  200100 VXLAN 5254.0008.db40     10.20.20.2         -
      1
1001  200100 VXLAN 5254.0000.11c3     10.20.20.4         -
      1
1001  200100 VXLAN 5254.0016.4ab2     10.20.20.1         -
      1
1001  200100 VXLAN 5254.0016.4ab2     10.20.20.1         -
      1
1001  200100 VXLAN 5254.0016.4ab2     10.20.20.1         -
      1
1001  200100 VXLAN 5254.0016.4ab2     10.20.20.1         -
      1
1001  200100 VXLAN 5254.0016.4ab2     10.20.20.1         -
      1
1001  200100 VXLAN 5254.0016.4ab2     10.20.20.1         -
      1
1001  200100 VXLAN 5254.0016.4ab2     10.20.20.1         -
      1
1001  200100 VXLAN 5254.0016.4ab2     10.20.20.1         -
      1
1002  200100 VXLAN 5254.0015.b640     10.20.20.5         -
      2
1002  200100 VXLAN 5254.0015.b640     10.20.20.5         -
      2
```

Figure 22: Arista Leaf A BGP EVPN Mac-Address Table



```
C-LEAFA# show mac address-table interface nve 1
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,
        (NA)- Not Applicable
   VNI      MAC Address      Type      age     Secure NTFY Ports
---------+-----------------+--------+---------+------+----+------------------
*  101     5211.ebaa.1b08   static   -         F      F    nve1(10.20.20.5)
*  101     5254.0084.4818   static   -         F      F    nve1(10.20.20.2)
C 1001     5254.0000.11c3   dynamic  NA         F      F    nve1(10.20.20.4)
C 1001     5254.0008.db40   dynamic  NA         F      F    nve1(10.20.20.2)
C 1002     5254.0015.b640   dynamic  NA         F      F    nve1(10.20.20.5)
C-LEAFA#
```

Figure 23: Cisco Leaf A Interface nve 1 Mac-Address Table

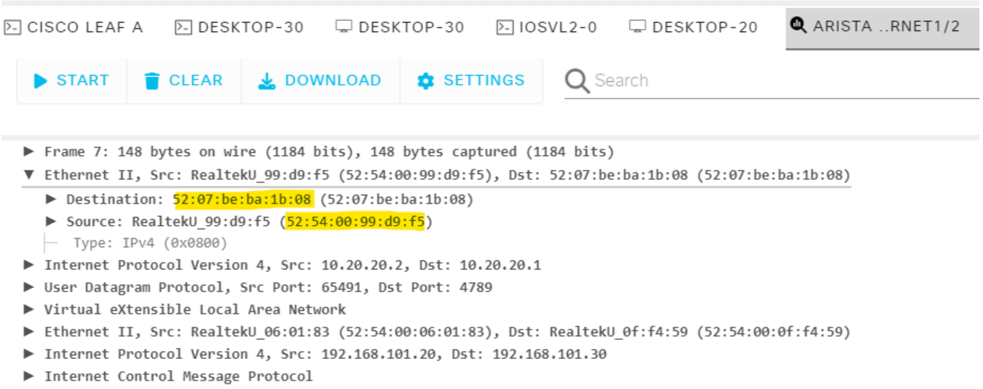
2.4 Packet Capture Deep-Dive

Ping Initiation Details:

- Source IP Address: 192.168.101.20
- Destination IP Address: 192.168.101.30
- Protocol: ICMP

Capture Details:

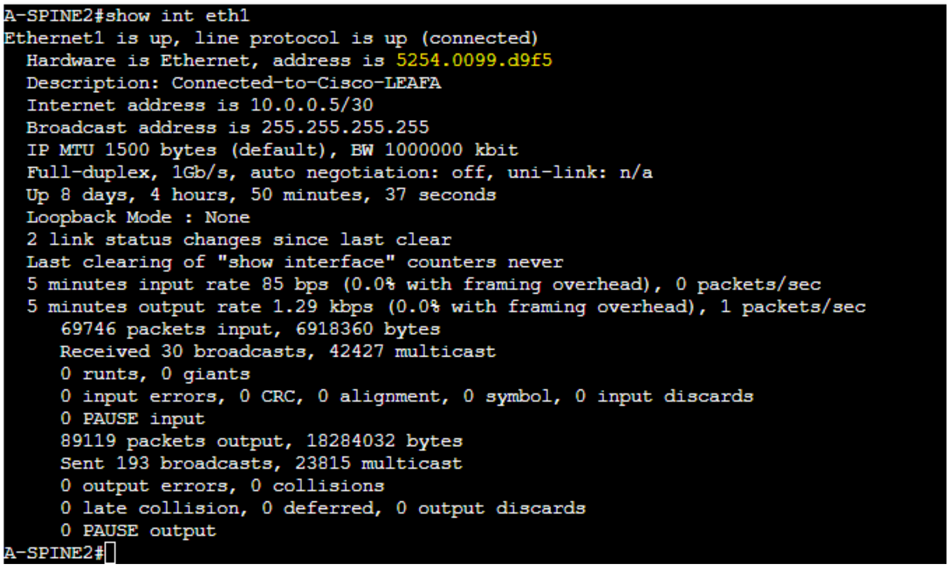
The capture link was initiated between Cisco LEAF A and Arista Spine 2 (Eth1/2 - Eth1). The source and destination MAC addresses of the two endpoints were observed in the Ethernet II layer of the packet capture.



```
▶ Frame 7: 148 bytes on wire (1184 bits), 148 bytes captured (1184 bits)
▼ Ethernet II, Src: RealtekU_99:d9:f5 (52:54:00:99:d9:f5), Dst: 52:07:be:ba:1b:08 (52:07:be:ba:1b:08)
  ▶ Destination: 52:07:be:ba:1b:08 (52:07:be:ba:1b:08)
  ▶ Source: RealtekU_99:d9:f5 (52:54:00:99:d9:f5)
    Type: IPv4 (0x0800)
▶ Internet Protocol Version 4, Src: 10.20.20.2, Dst: 10.20.20.1
▶ User Datagram Protocol, Src Port: 65491, Dst Port: 4789
▶ Virtual eXtensible Local Area Network
▶ Ethernet II, Src: RealtekU_06:01:83 (52:54:00:06:01:83), Dst: RealtekU_0f:f4:59 (52:54:00:0f:f4:59)
▶ Internet Protocol Version 4, Src: 192.168.101.20, Dst: 192.168.101.30
▶ Internet Control Message Protocol
```

Figure 24: Arista Spine Packet Capture

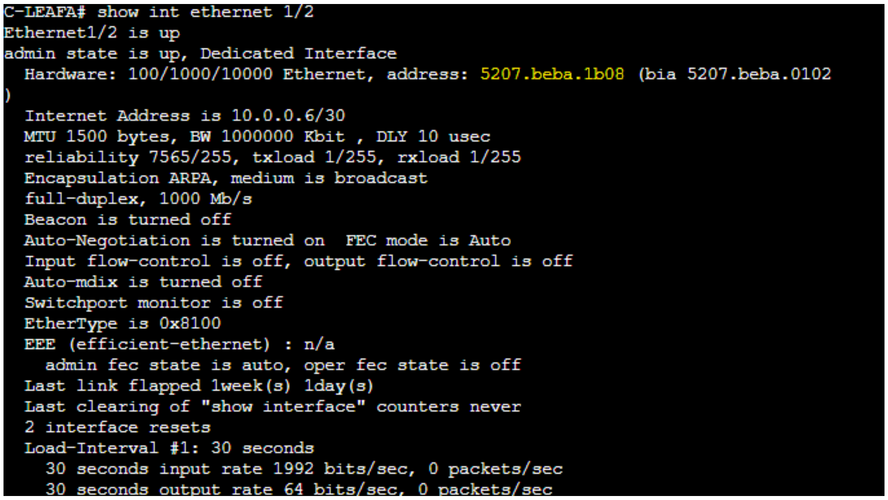
The source MAC address ending with d9 is the MAC address of Eth1 on Arista Spine 2.



```
A-SPINE2#show int eth1
Ethernet1 is up, line protocol is up (connected)
  Hardware is Ethernet, address is 5254.0099.d9f5
  Description: Connected-to-Cisco-LEAFA
  Internet address is 10.0.0.5/30
  Broadcast address is 255.255.255.255
  IP MTU 1500 bytes (default), BW 1000000 kbit
  Full-duplex, 1Gb/s, auto negotiation: off, uni-link: n/a
  Up 8 days, 4 hours, 50 minutes, 37 seconds
  Loopback Mode : None
  2 link status changes since last clear
  Last clearing of "show interface" counters never
  5 minutes input rate 85 bps (0.0% with framing overhead), 0 packets/sec
  5 minutes output rate 1.29 kbps (0.0% with framing overhead), 1 packets/sec
     69746 packets input, 6918360 bytes
     Received 30 broadcasts, 42427 multicast
     0 runts, 0 giants
     0 input errors, 0 CRC, 0 alignment, 0 symbol, 0 input discards
     0 PAUSE input
     89119 packets output, 18284032 bytes
     Sent 193 broadcasts, 23815 multicast
     0 output errors, 0 collisions
     0 late collision, 0 deferred, 0 output discards
     0 PAUSE output
A-SPINE2#
```

Figure 25: Arista Spine Interface Eth1

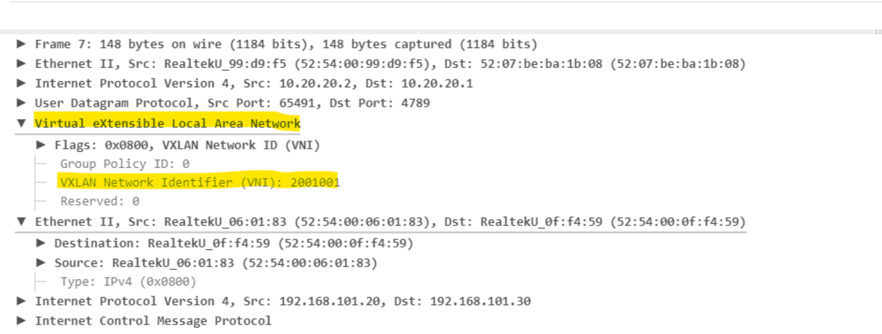
The source MAC address ending with d9 is the MAC address of Eth1 on Arista Spine 2. The destination MAC address ending with 1b:08 is the MAC address of Eth1/2 on Cisco LEAF A.



```
C-LEAFA# show int ethernet 1/2
Ethernet1/2 is up
admin state is up, Dedicated Interface
  Hardware: 100/1000/10000 Ethernet, address: 5207.beba.1b08 (bia 5207.beba.0102
)
  Internet Address is 10.0.0.6/30
  MTU 1500 bytes, BW 1000000 Kbit , DLY 10 usec
  reliability 7565/255, txload 1/255, rxload 1/255
  Encapsulation ARPA, medium is broadcast
  full-duplex, 1000 Mb/s
  Beacon is turned off
  Auto-Negotiation is turned on  FEC mode is Auto
  Input flow-control is off, output flow-control is off
  Auto-mdix is turned off
  Switchport monitor is off
  EtherType is 0x8100
  EEE (efficient-ethernet) : n/a
    admin fec state is auto, oper fec state is off
  Last link flapped 1week(s) 1day(s)
  Last clearing of "show interface" counters never
  2 interface resets
  Load-Interval #1: 30 seconds
    30 seconds input rate 1992 bits/sec, 0 packets/sec
    30 seconds output rate 64 bits/sec, 0 packets/sec
```

Figure 26: Cisco Leaf-A Interface Eth1/2

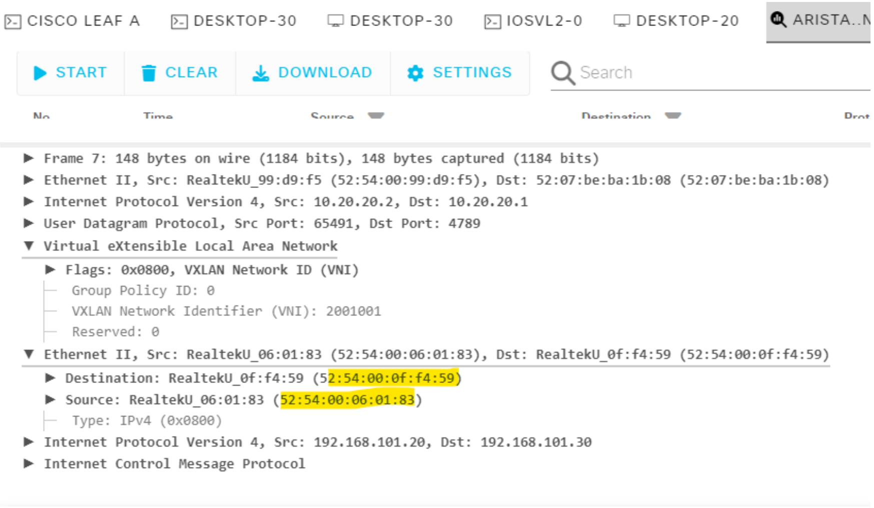
In Packet Capture I found this frame regarding VNI 2001001:



```
▶ Frame 7: 148 bytes on wire (1184 bits), 148 bytes captured (1184 bits)
▶ Ethernet II, Src: RealtekU_99:d9:f5 (52:54:00:99:d9:f5), Dst: 52:07:be:ba:1b:08 (52:07:be:ba:1b:08)
▶ Internet Protocol Version 4, Src: 10.20.20.2, Dst: 10.20.20.1
▶ User Datagram Protocol, Src Port: 65491, Dst Port: 4789
▼ Virtual eXtensible Local Area Network
   ▶ Flags: 0x0800, VXLAN Network ID (VNI)
   ├─ Group Policy ID: 0
   ├─ VXLAN Network Identifier (VNI): 2001001
   └─ Reserved: 0
▼ Ethernet II, Src: RealtekU_06:01:83 (52:54:00:06:01:83), Dst: RealtekU_0f:f4:59 (52:54:00:0f:f4:59)
   ▶ Destination: RealtekU_0f:f4:59 (52:54:00:0f:f4:59)
   ▶ Source: RealtekU_06:01:83 (52:54:00:06:01:83)
   └─ Type: IPv4 (0x0800)
▶ Internet Protocol Version 4, Src: 192.168.101.20, Dst: 192.168.101.30
▶ Internet Control Message Protocol
```

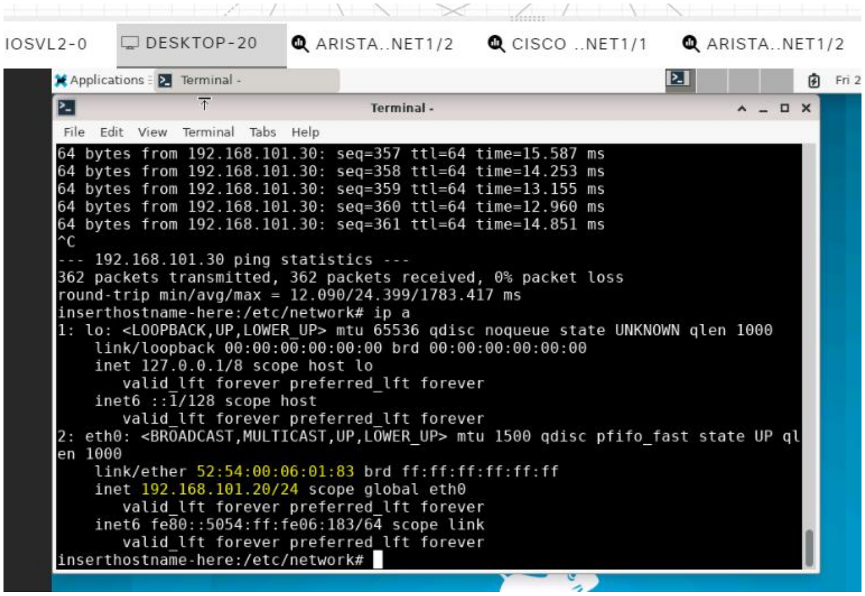
Figure 27: VXLAN Frame Header

Under VXLAN there is an Ethernet II frame with Desktop-20 MAC as a source and Desktop-30 as a destination:



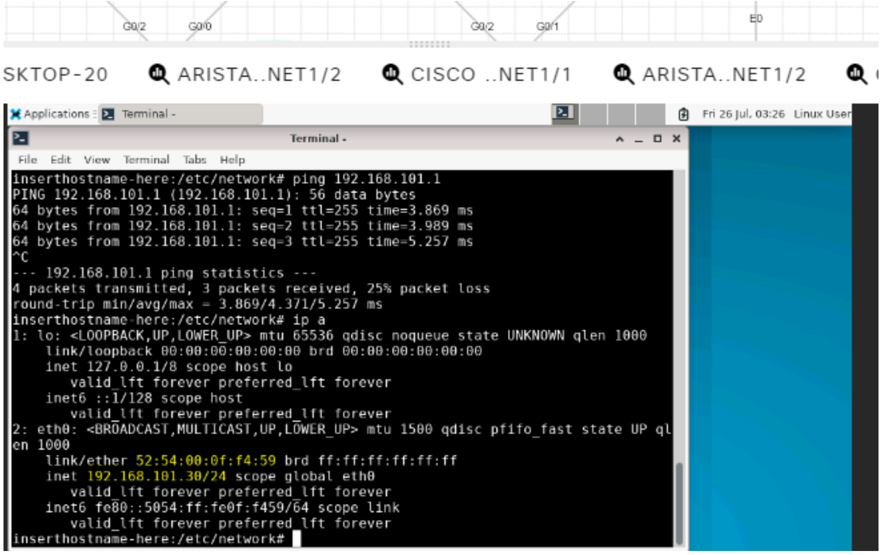
```
▶ Frame 7: 148 bytes on wire (1184 bits), 148 bytes captured (1184 bits)
▶ Ethernet II, Src: RealtekU_99:d9:f5 (52:54:00:99:d9:f5), Dst: 52:07:be:ba:1b:08 (52:07:be:ba:1b:08)
▶ Internet Protocol Version 4, Src: 10.20.20.2, Dst: 10.20.20.1
▶ User Datagram Protocol, Src Port: 65491, Dst Port: 4789
▼ Virtual eXtensible Local Area Network
   ▶ Flags: 0x0800, VXLAN Network ID (VNI)
   ├─ Group Policy ID: 0
   ├─ VXLAN Network Identifier (VNI): 2001001
   └─ Reserved: 0
▼ Ethernet II, Src: RealtekU_06:01:83 (52:54:00:06:01:83), Dst: RealtekU_0f:f4:59 (52:54:00:0f:f4:59)
   ▶ Destination: RealtekU_0f:f4:59 (52:54:00:0f:f4:59)
   ▶ Source: RealtekU_06:01:83 (52:54:00:06:01:83)
   └─ Type: IPv4 (0x0800)
▶ Internet Protocol Version 4, Src: 192.168.101.20, Dst: 192.168.101.30
▶ Internet Control Message Protocol
```

Figure 28: VXLAN Frame Inside



```
64 bytes from 192.168.101.30: seq=357 ttl=64 time=15.587 ms
64 bytes from 192.168.101.30: seq=358 ttl=64 time=14.253 ms
64 bytes from 192.168.101.30: seq=359 ttl=64 time=13.155 ms
64 bytes from 192.168.101.30: seq=360 ttl=64 time=12.960 ms
64 bytes from 192.168.101.30: seq=361 ttl=64 time=14.851 ms
^C
--- 192.168.101.30 ping statistics ---
362 packets transmitted, 362 packets received, 0% packet loss
round-trip min/avg/max = 12.090/24.399/1783.417 ms
inserthostname-here:/etc/network# ip a
1: lo: <LOOPBACK,UP,LOWER_UP> mtu 65536 qdisc noqueue state UNKNOWN qlen 1000
    link/loopback 00:00:00:00:00:00 brd 00:00:00:00:00:00
    inet 127.0.0.1/8 scope host lo
       valid_lft forever preferred_lft forever
    inet6 ::1/128 scope host
       valid_lft forever preferred_lft forever
2: eth0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc pfifo_fast state UP ql
en 1000
    link/ether 52:54:00:06:01:83 brd ff:ff:ff:ff:ff:ff
    inet 192.168.101.20/24 scope global eth0
       valid_lft forever preferred_lft forever
    inet6 fe80::5054:ff:fe06:183/64 scope link
       valid_lft forever preferred_lft forever
inserthostname-here:/etc/network#
```

Figure 29: Desktop-20 IP a



```
inserthostname-here:/etc/network# ping 192.168.101.1
PING 192.168.101.1 (192.168.101.1): 56 data bytes
64 bytes from 192.168.101.1: seq=1 ttl=255 time=3.869 ms
64 bytes from 192.168.101.1: seq=2 ttl=255 time=3.989 ms
64 bytes from 192.168.101.1: seq=3 ttl=255 time=5.257 ms
^C
--- 192.168.101.1 ping statistics ---
4 packets transmitted, 3 packets received, 25% packet loss
round-trip min/avg/max = 3.869/4.371/5.257 ms
inserthostname-here:/etc/network# ip a
1: lo: <LOOPBACK,UP,LOWER_UP> mtu 65536 qdisc noqueue state UNKNOWN qlen 1000
    link/loopback 00:00:00:00:00:00 brd 00:00:00:00:00:00
    inet 127.0.0.1/8 scope host lo
       valid_lft forever preferred_lft forever
    inet6 ::1/128 scope host
       valid_lft forever preferred_lft forever
2: eth0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc pfifo_fast state UP ql
en 1000
    link/ether 52:54:00:0f:f4:59 brd ff:ff:ff:ff:ff:ff
    inet 192.168.101.30/24 scope global eth0
       valid_lft forever preferred_lft forever
    inet6 fe80::5054:ff:fe0f:f459/64 scope link
       valid_lft forever preferred_lft forever
inserthostname-here:/etc/network#
```

Figure 30: Desktop-30 IP a

From this packet capture, we can determine that the outer source and destination MAC and IP addresses correspond to the point-to-point links between the LEAF and SPINES. The inner source and destination addresses contain the information of the hosts.

2.5 Challenges and Differences Between Vendors

During the implementation of Scenario 1, several challenges and differences between Cisco and Arista devices were encountered:

1. Feature Implementation:

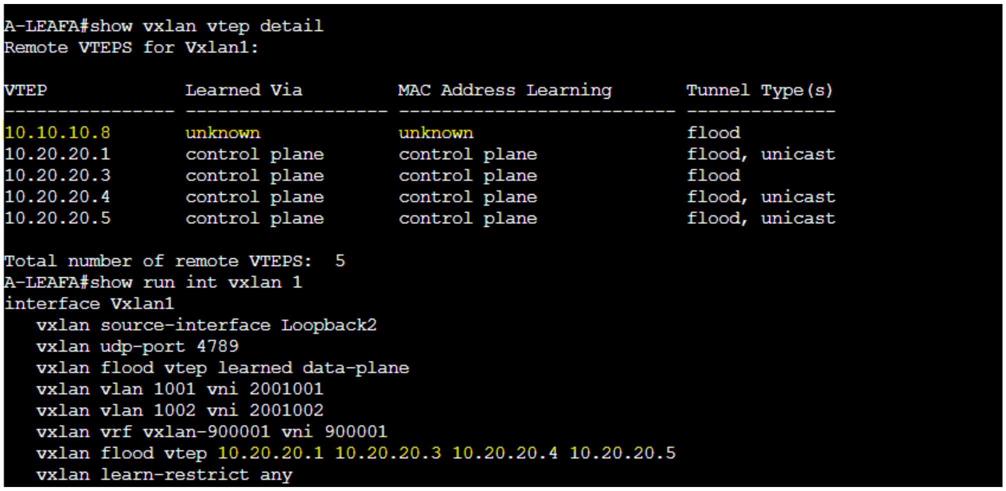
- Cisco: Utilized the VPC feature, which required two links (one for peer link and another for keep-alive link) to ensure redundancy and fault tolerance.

- Arista: Used the MLAG feature, which only required one link, simplifying the configuration process.

2. VTEP Configuration:

- Cisco: Automatically discovers VTEP addresses based on BGP advertisements, reducing manual configuration efforts.

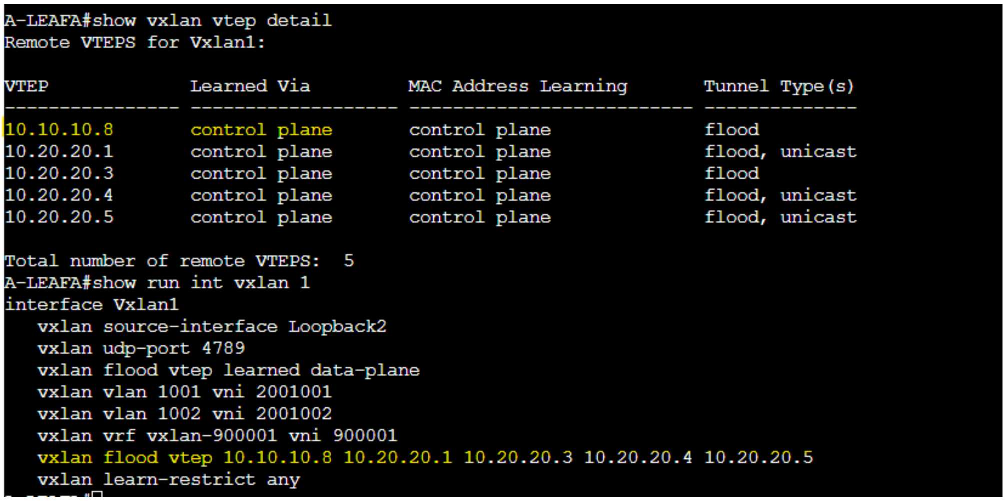
- Arista: Requires manual configuration of VTEP addresses, adding complexity to the setup process.



```
A-LEAFA#show vxlan vtep detail
Remote VTEPS for Vxlan1:

VTEP             Learned Via         MAC Address Learning       Tunnel Type(s)
---------------- ------------------- -------------------------- --------------
10.10.10.8       unknown             unknown                    flood
10.20.20.1       control plane       control plane              flood, unicast
10.20.20.3       control plane       control plane              flood
10.20.20.4       control plane       control plane              flood, unicast
10.20.20.5       control plane       control plane              flood, unicast

Total number of remote VTEPS:  5
A-LEAFA#show run int vxlan 1
interface Vxlan1
   vxlan source-interface Loopback2
   vxlan udp-port 4789
   vxlan flood vtep learned data-plane
   vxlan vlan 1001 vni 2001001
   vxlan vlan 1002 vni 2001002
   vxlan vrf vxlan-900001 vni 900001
   vxlan flood vtep 10.20.20.1 10.20.20.3 10.20.20.4 10.20.20.5
   vxlan learn-restrict any
```

Figure 31: Arista Leaf-A VXLAN VTEP Automatically

When VTEP addresses are not manually configured, they may not be discovered automatically, leading to incomplete VXLAN tunnel setups.

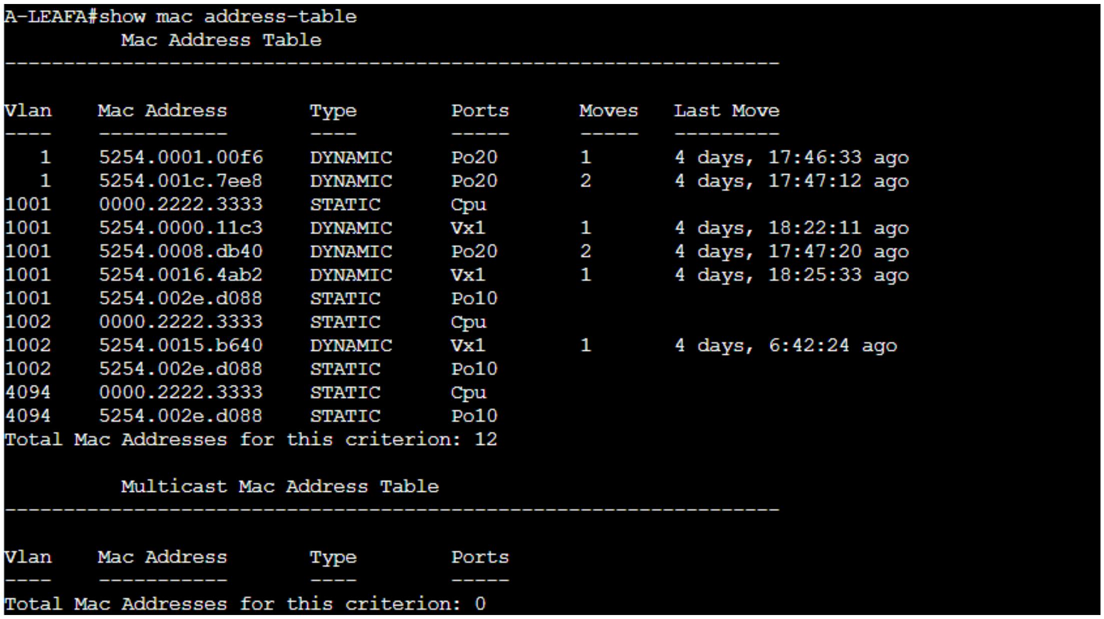


```
A-LEAFA#show vxlan vtep detail
Remote VTEPS for Vxlan1:

VTEP             Learned Via         MAC Address Learning       Tunnel Type(s)
---------------- ------------------- -------------------------- --------------
10.10.10.8       control plane       control plane              flood
10.20.20.1       control plane       control plane              flood, unicast
10.20.20.3       control plane       control plane              flood
10.20.20.4       control plane       control plane              flood, unicast
10.20.20.5       control plane       control plane              flood, unicast

Total number of remote VTEPS:  5
A-LEAFA#show run int vxlan 1
interface Vxlan1
   vxlan source-interface Loopback2
   vxlan udp-port 4789
   vxlan flood vtep learned data-plane
   vxlan vlan 1001 vni 2001001
   vxlan vlan 1002 vni 2001002
   vxlan vrf vxlan-900001 vni 900001
   vxlan flood vtep 10.10.10.8 10.20.20.1 10.20.20.3 10.20.20.4 10.20.20.5
   vxlan learn-restrict any
```

Figure 32: Arista Leaf-A VXLAN VTEP Manually

Manually configuring all VTEPs ensures that the devices know all the endpoints for encapsulation and de-encapsulation, facilitating proper VXLAN operation.

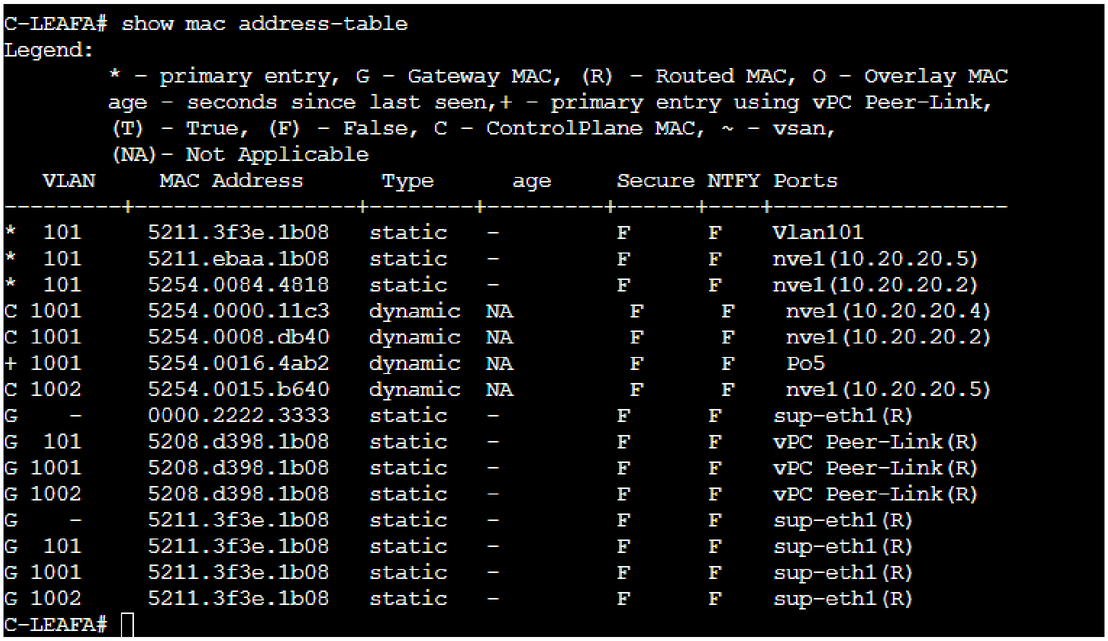
Arista Mac address table:



```
A-LEAFA#show mac address-table
          Mac Address Table
------------------------------------------------------------------

Vlan    Mac Address       Type        Ports      Moves   Last Move
----    -----------       ----        -----      -----   ---------
   1    5254.0001.00f6    DYNAMIC     Po20       1       4 days, 17:46:33 ago
   1    5254.001c.7ee8    DYNAMIC     Po20       2       4 days, 17:47:12 ago
1001    0000.2222.3333    STATIC      Cpu
1001    5254.0000.11c3    DYNAMIC     Vx1        1       4 days, 18:22:11 ago
1001    5254.0008.db40    DYNAMIC     Po20       2       4 days, 17:47:20 ago
1001    5254.0016.4ab2    DYNAMIC     Vx1        1       4 days, 18:25:33 ago
1001    5254.002e.d088    STATIC      Po10
1002    0000.2222.3333    STATIC      Cpu
1002    5254.0015.b640    DYNAMIC     Vx1        1       4 days, 6:42:24 ago
1002    5254.002e.d088    STATIC      Po10
4094    0000.2222.3333    STATIC      Cpu
4094    5254.002e.d088    STATIC      Po10
Total Mac Addresses for this criterion: 12

          Multicast Mac Address Table
------------------------------------------------------------------

Vlan    Mac Address       Type        Ports
----    -----------       ----        -----
Total Mac Addresses for this criterion: 0
```

Figure 33: Arista Leaf-A Mac Address Table

Cisco MAC address table:



```
C-LEAFA# show mac address-table
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,
        (NA)- Not Applicable
   VLAN     MAC Address      Type      age       Secure NTFY Ports
---------+-----------------+--------+---------+------+----+------------------
*  101     5211.3f3e.1b08   static   -         F      F    Vlan101
*  101     5211.ebaa.1b08   static   -         F      F    nve1(10.20.20.5)
*  101     5254.0084.4818   static   -         F      F    nve1(10.20.20.2)
C 1001     5254.0000.11c3   dynamic  NA         F      F    nve1(10.20.20.4)
C 1001     5254.0008.db40   dynamic  NA         F      F    nve1(10.20.20.2)
+ 1001     5254.0016.4ab2   dynamic  NA         F      F    Po5
C 1002     5254.0015.b640   dynamic  NA         F      F    nve1(10.20.20.5)
G    -     0000.2222.3333   static   -         F      F    sup-eth1(R)
G  101     5208.d398.1b08   static   -         F      F    vPC Peer-Link(R)
G 1001     5208.d398.1b08   static   -         F      F    vPC Peer-Link(R)
G 1002     5208.d398.1b08   static   -         F      F    vPC Peer-Link(R)
G    -     5211.3f3e.1b08   static   -         F      F    sup-eth1(R)
G  101     5211.3f3e.1b08   static   -         F      F    sup-eth1(R)
G 1001     5211.3f3e.1b08   static   -         F      F    sup-eth1(R)
G 1002     5211.3f3e.1b08   static   -         F      F    sup-eth1(R)
C-LEAFA#
```

Figure 34: Cisco Leaf-A Mac Address Table

2.6 Conclusion

Scenario 1 successfully demonstrated the implementation of VXLAN Route Types 1 and 2 across a mixed vendor environment using Cisco and Arista devices. The differences in feature implementation and VTEP configuration were effectively managed, ensuring seamless interoperability. This scenario lays the groundwork for more complex configurations and scenarios in subsequent sections.

3. Scenario 2: VXLAN Route Type 5 Implementation

3.1 Overview

Scenario 2 builds on the configuration from Scenario 1, focusing on the implementation of VXLAN Route Type 5, which involves Layer 3 routing. This scenario demonstrates the need for Layer 3 routing in the network to facilitate efficient communication between devices connected to different VLANs across the network, specifically using symmetric routing. Additionally, we will provide an explanation of asymmetric routing for a comprehensive understanding.

Layer 3 Routing Importance

Layer 3 routing is essential for several reasons:

- Avoiding Layer 2 Issues: Problems such as spanning tree protocol (STP) complexities and broadcast storms are mitigated by using Layer 3 routing.

- Efficient Communication: Ensures efficient data transfer between devices on different VLANs without the overhead of Layer 2 bridging.[8]

3.2 Configuration Steps

1. Continuation from Scenario 1

We continue from the configuration established in Scenario 1, where we set up connectivity between VLANs over the entire network using VXLAN.

2. Network Scenario

Consider a scenario with two servers, each running ESXi as a hypervisor:

- Server 1: Connected to an Arista switch (single homing) in VLAN 1002.

- Server 2: Connected to a Cisco switch (single homing) in VLAN 1001.

These servers need to communicate and migrate virtual machines across the network using Layer 3 routing to avoid Layer 2 network issues.

3.3 VXLAN Routing (Route Type 5)

- Asymmetric Routing

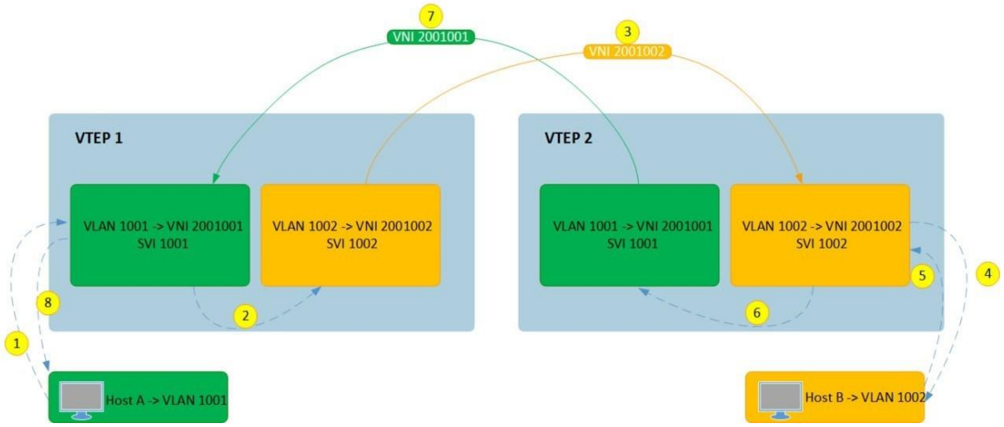


Figure 35: Asymmetric Routing Figure

- Symmetric routing

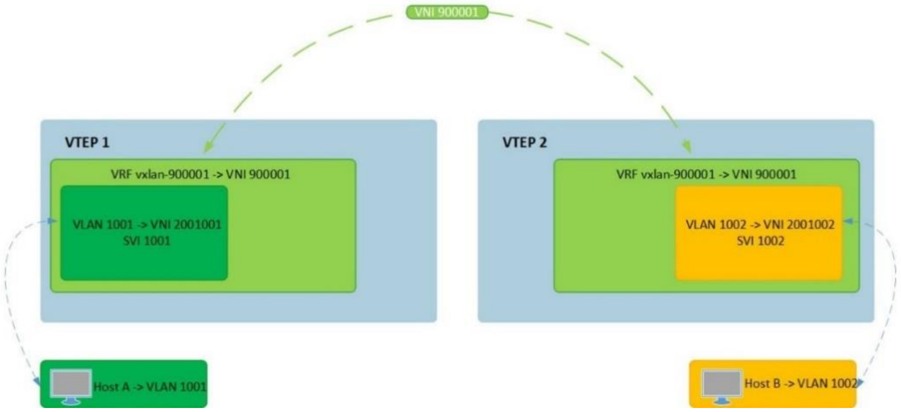


Figure 36: Symmetric routing Figure

In this project, symmetric routing is chosen for VXLAN routing.

3.4 Symmetric Routing Configuration

Ensures that both ingress and egress routing are performed on the same VTEP, simplifying the routing process and improving efficiency.

1. Creating VRF:

- Example VRF: (VXLAN-900001)

- The VRF configuration redistributes both VLANs (1001 and 1002) within the network.

2. Adding VLANs to VRF:

- Each VLAN interface is configured to be a member of the VRF.

- Example configuration for Cisco:

Commands:

```
interface Vlan1001
  vrf member VXLAN-900001
interface Vlan1002
  vrf member VXLAN-900001
```

3. VNI Membership Configuration:

- Under the interface `NVE-1` for Cisco or `VXLAN-1` for Arista, VNIs are assigned to the VRF.

- Example configuration for Cisco:

Commands:

```
interface NVE-1
  member vni 2001001 associate-vrf
  member vni 2001002 associate-vrf
```

3.5 Asymmetric Routing Configuration

In asymmetric routing, the routing decision is made at the ingress VTEP (the switch closest to the source of the traffic). This routing method is more straightforward in terms of initial packet processing but can lead to suboptimal routing paths and increased processing requirements at intermediate nodes.

1. Routing at the Ingress VTEP:

- The ingress VTEP performs the routing decision, encapsulating the packet and sending it to the appropriate egress VTEP based on the destination IP address.

2. No Need for VRF Redistribution:

- Unlike symmetric routing, there is no need to create VRFs and redistribute VLANs. The ingress VTEP handles routing without requiring intermediate nodes to perform additional lookups.

3. Efficiency Considerations:

- While simpler to configure, asymmetric routing can result in higher latencies and potential routing inefficiencies, as each packet must be processed multiple times along its path.

3.6 VXLAN Routing Symmetric vs. Asymmetric

Symmetric Routing

In symmetric routing, both ingress and egress traffic for a flow are handled by the same VTEP. This ensures that the routing decision is consistent regardless of the packet direction. Symmetric routing requires the creation of VRFs and redistribution of VLANs within the VRFs. This approach simplifies troubleshooting and provides a more predictable routing path.

Asymmetric Routing

In asymmetric routing, the ingress VTEP makes the routing decision and encapsulates the packet before sending it to the egress VTEP. The egress VTEP decapsulates the packet and forwards it to the destination. This method can lead to suboptimal routing paths and higher latencies due to multiple processing stages. Asymmetric routing does not require the creation of VRFs, making the initial configuration simpler but potentially leading to more complex traffic patterns.

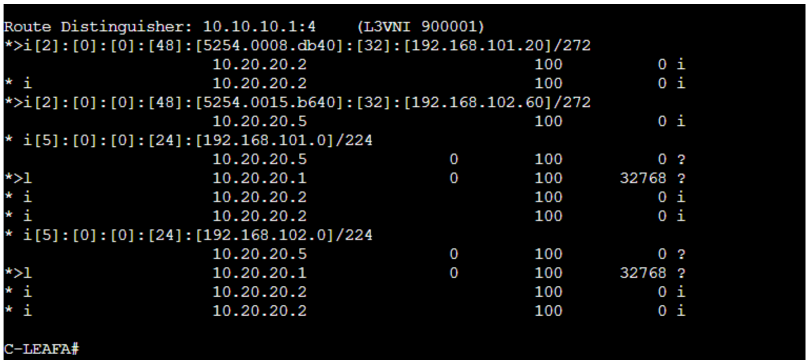
3.7 Verification and Testing

To confirm the configuration and ensure the successful implementation of Scenario 2, the following commands are used:

- Cisco:

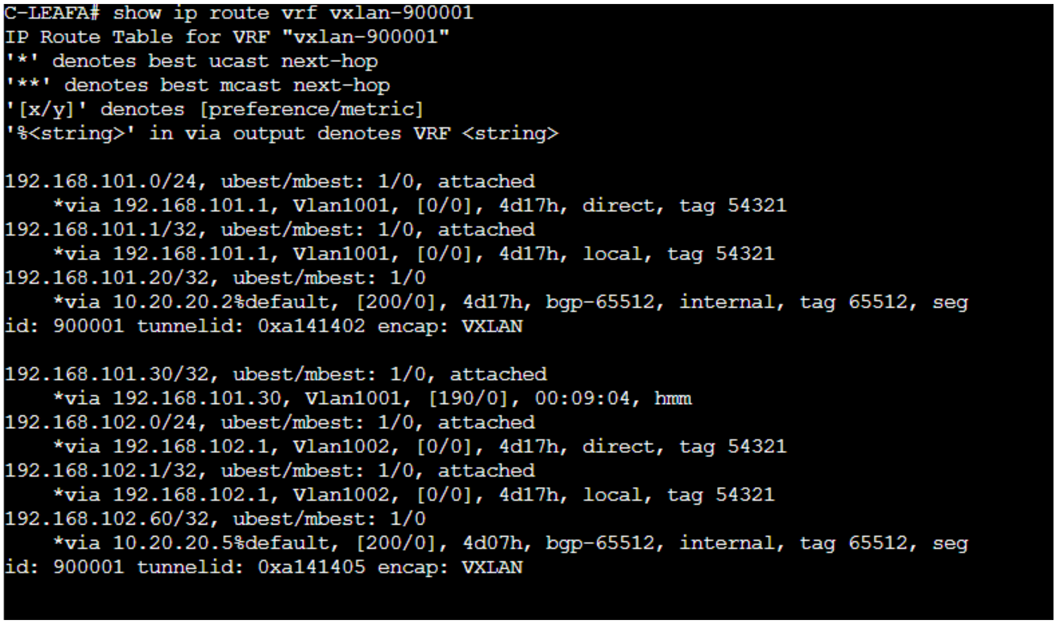
```
show bgp l3vpn vrf VXLAN-900001
show bgp vpnv4 unicast vrf VXLAN-900001
```

- Provides detailed routing information within the VRF and verifies the configuration.



```
Route Distinguisher: 10.10.10.1:4    (L3VNI 900001)
*>i[2]:[0]:[0]:[48]:[5254.0008.db40]:[32]:[192.168.101.20]/272
                      10.20.20.2                        100          0 i
* i                   10.20.20.2                        100          0 i
*>i[2]:[0]:[0]:[48]:[5254.0015.b640]:[32]:[192.168.102.60]/272
                      10.20.20.5                        100          0 i
* i[5]:[0]:[0]:[24]:[192.168.101.0]/224
                      10.20.20.5               0        100          0 ?
*>l                   10.20.20.1               0        100      32768 ?
* i                   10.20.20.2                        100          0 i
* i                   10.20.20.2                        100          0 i
* i[5]:[0]:[0]:[24]:[192.168.102.0]/224
                      10.20.20.5               0        100          0 ?
*>l                   10.20.20.1               0        100      32768 ?
* i                   10.20.20.2                        100          0 i
* i                   10.20.20.2                        100          0 i

C-LEAFA#
```

Figure 37: Cisco BGP L2VPN VNI 900001



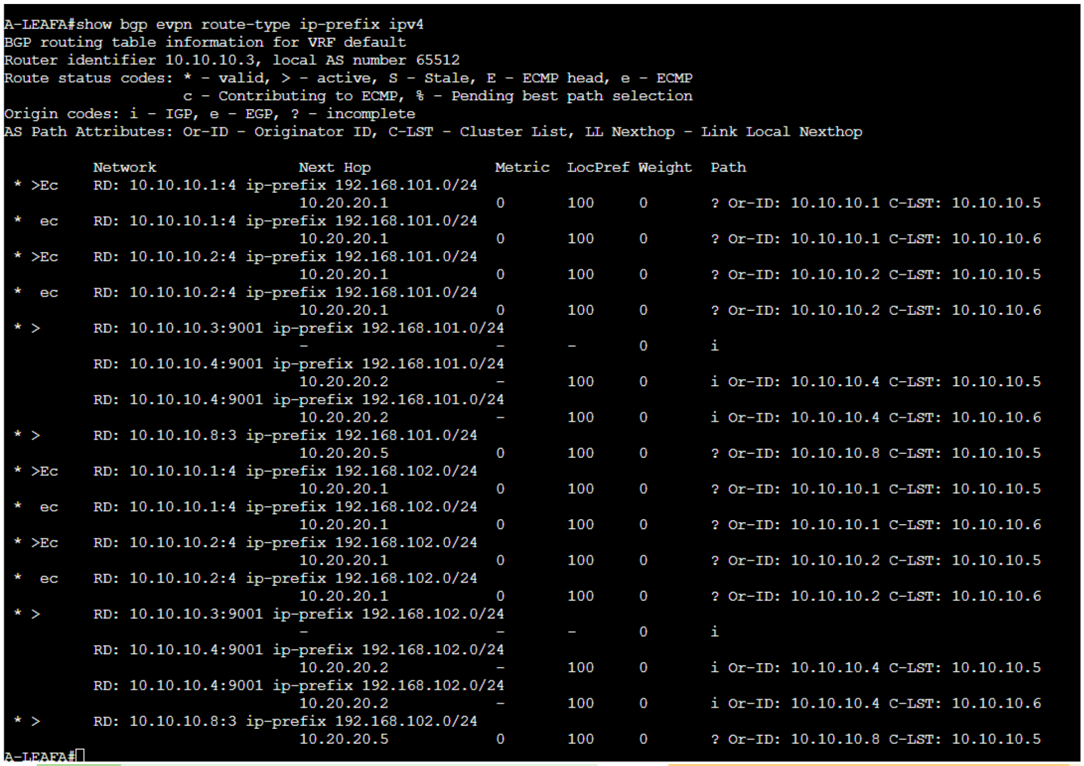
```
C-LEAFA# show ip route vrf vxlan-900001
IP Route Table for VRF "vxlan-900001"
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
'%<string>' in via output denotes VRF <string>

192.168.101.0/24, ubest/mbest: 1/0, attached
    *via 192.168.101.1, Vlan1001, [0/0], 4d17h, direct, tag 54321
192.168.101.1/32, ubest/mbest: 1/0, attached
    *via 192.168.101.1, Vlan1001, [0/0], 4d17h, local, tag 54321
192.168.101.20/32, ubest/mbest: 1/0
    *via 10.20.20.2%default, [200/0], 4d17h, bgp-65512, internal, tag 65512, seg
id: 900001 tunnelid: 0xa141402 encap: VXLAN

192.168.101.30/32, ubest/mbest: 1/0, attached
    *via 192.168.101.30, Vlan1001, [190/0], 00:09:04, hmm
192.168.102.0/24, ubest/mbest: 1/0, attached
    *via 192.168.102.1, Vlan1002, [0/0], 4d17h, direct, tag 54321
192.168.102.1/32, ubest/mbest: 1/0, attached
    *via 192.168.102.1, Vlan1002, [0/0], 4d17h, local, tag 54321
192.168.102.60/32, ubest/mbest: 1/0
    *via 10.20.20.5%default, [200/0], 4d07h, bgp-65512, internal, tag 65512, seg
id: 900001 tunnelid: 0xa141405 encap: VXLAN
```

Figure 38: Cisco Leaf-A VRF Routing Table

- Arista:

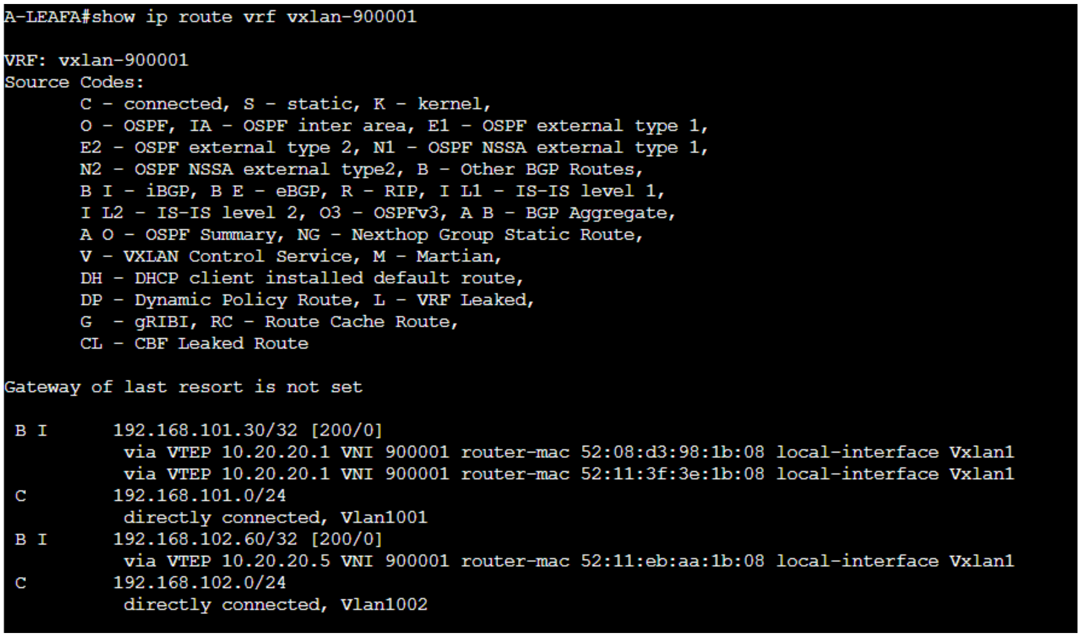
show bgp evpn route-type ip-prefix ipv4

- Displays IP address information and validates the routing.



```
A-LEAFA#show bgp evpn route-type ip-prefix ipv4
BGP routing table information for VRF default
Router identifier 10.10.10.3, local AS number 65512
Route status codes: * - valid, > - active, S - Stale, E - ECMP head, e - ECMP
                    c - Contributing to ECMP, % - Pending best path selection
Origin codes: i - IGP, e - EGP, ? - incomplete
AS Path Attributes: Or-ID - Originator ID, C-LST - Cluster List, LL Nexthop - Link Local Nexthop

          Network                Next Hop              Metric  LocPref Weight  Path
 * >Ec    RD: 10.10.10.1:4 ip-prefix 192.168.101.0/24
                                 10.20.20.1            0       100     0       ? Or-ID: 10.10.10.1 C-LST: 10.10.10.5
 *  ec    RD: 10.10.10.1:4 ip-prefix 192.168.101.0/24
                                 10.20.20.1            0       100     0       ? Or-ID: 10.10.10.1 C-LST: 10.10.10.6
 * >Ec    RD: 10.10.10.2:4 ip-prefix 192.168.101.0/24
                                 10.20.20.1            0       100     0       ? Or-ID: 10.10.10.2 C-LST: 10.10.10.5
 *  ec    RD: 10.10.10.2:4 ip-prefix 192.168.101.0/24
                                 10.20.20.1            0       100     0       ? Or-ID: 10.10.10.2 C-LST: 10.10.10.6
 * >      RD: 10.10.10.3:9001 ip-prefix 192.168.101.0/24
                                 -                     -       -       0       i
          RD: 10.10.10.4:9001 ip-prefix 192.168.101.0/24
                                 10.20.20.2            -       100     0       i Or-ID: 10.10.10.4 C-LST: 10.10.10.5
          RD: 10.10.10.4:9001 ip-prefix 192.168.101.0/24
                                 10.20.20.2            -       100     0       i Or-ID: 10.10.10.4 C-LST: 10.10.10.6
 * >      RD: 10.10.10.8:3 ip-prefix 192.168.101.0/24
                                 10.20.20.5            0       100     0       ? Or-ID: 10.10.10.8 C-LST: 10.10.10.5
 * >Ec    RD: 10.10.10.1:4 ip-prefix 192.168.102.0/24
                                 10.20.20.1            0       100     0       ? Or-ID: 10.10.10.1 C-LST: 10.10.10.5
 *  ec    RD: 10.10.10.1:4 ip-prefix 192.168.102.0/24
                                 10.20.20.1            0       100     0       ? Or-ID: 10.10.10.1 C-LST: 10.10.10.6
 * >Ec    RD: 10.10.10.2:4 ip-prefix 192.168.102.0/24
                                 10.20.20.1            0       100     0       ? Or-ID: 10.10.10.2 C-LST: 10.10.10.5
 *  ec    RD: 10.10.10.2:4 ip-prefix 192.168.102.0/24
                                 10.20.20.1            0       100     0       ? Or-ID: 10.10.10.2 C-LST: 10.10.10.6
 * >      RD: 10.10.10.3:9001 ip-prefix 192.168.102.0/24
                                 -                     -       -       0       i
          RD: 10.10.10.4:9001 ip-prefix 192.168.102.0/24
                                 10.20.20.2            -       100     0       i Or-ID: 10.10.10.4 C-LST: 10.10.10.5
          RD: 10.10.10.4:9001 ip-prefix 192.168.102.0/24
                                 10.20.20.2            -       100     0       i Or-ID: 10.10.10.4 C-LST: 10.10.10.6
 * >      RD: 10.10.10.8:3 ip-prefix 192.168.102.0/24
                                 10.20.20.5            0       100     0       ? Or-ID: 10.10.10.8 C-LST: 10.10.10.5
A-LEAFA#
```

Figure 39: Arista EVPN Route-Type IP-Prefix IPV4



```
A-LEAFA#show ip route vrf vxlan-900001

VRF: vxlan-900001
Source Codes:
       C - connected, S - static, K - kernel,
       O - OSPF, IA - OSPF inter area, E1 - OSPF external type 1,
       E2 - OSPF external type 2, N1 - OSPF NSSA external type 1,
       N2 - OSPF NSSA external type2, B - Other BGP Routes,
       B I - iBGP, B E - eBGP, R - RIP, I L1 - IS-IS level 1,
       I L2 - IS-IS level 2, O3 - OSPFv3, A B - BGP Aggregate,
       A O - OSPF Summary, NG - Nexthop Group Static Route,
       V - VXLAN Control Service, M - Martian,
       DH - DHCP client installed default route,
       DP - Dynamic Policy Route, L - VRF Leaked,
       G  - gRIBI, RC - Route Cache Route,
       CL - CBF Leaked Route

Gateway of last resort is not set

 B I      192.168.101.30/32 [200/0]
           via VTEP 10.20.20.1 VNI 900001 router-mac 52:08:d3:98:1b:08 local-interface Vxlan1
           via VTEP 10.20.20.1 VNI 900001 router-mac 52:11:3f:3e:1b:08 local-interface Vxlan1
 C        192.168.101.0/24
           directly connected, Vlan1001
 B I      192.168.102.60/32 [200/0]
           via VTEP 10.20.20.5 VNI 900001 router-mac 52:11:eb:aa:1b:08 local-interface Vxlan1
 C        192.168.102.0/24
           directly connected, Vlan1002
```

Figure 40: Arista Leaf-A VRF Routing Table

3.8 Differences Between Scenario 1 and Scenario 2

Route Type:

- Scenario 1: Displayed MAC addresses (Layer 2 focus).

- Scenario 2: Displays both IP and MAC addresses, incorporating SVIs (Switched Virtual Interfaces) for Layer 3 routing.

3.9 Explanation of Key Concepts

Virtual Routing and Forwarding (VRF)

VRF is a technology that allows multiple instances of a routing table to coexist within the same router. This enables network segmentation and isolation, providing each VRF with its own unique IP routing table.

VXLAN Network Identifier (VNI)

VNI is a 24-bit identifier used in VXLAN to distinguish different Layer 2 segments, allowing for the creation of isolated virtual networks over a shared physical infrastructure.

Virtual Tunnel End Point (VTEP)

A VTEP is responsible for encapsulating and de-encapsulating VXLAN packets. It uses loopback addresses as the source and destination IPs for tunnel endpoints. In Cisco devices, VTEPs are dynamically discovered via BGP, while in Arista devices, they must be manually configured.

Spanning Tree Protocol (STP)

STP is a Layer 2 network protocol that prevents loops by creating a spanning tree within a network of connected Ethernet switches. While it helps avoid broadcast storms, it can also lead to suboptimal routing paths and increased complexity.

3.10 Conclusion

Scenario 2 successfully demonstrates the implementation of VXLAN Route Type 5, leveraging symmetric routing to enable efficient Layer 3 communication between devices on different VLANs. This scenario highlights the importance of Layer 3 routing in avoiding Layer 2 network issues and ensures seamless interoperability between Cisco and Arista devices. The explanation of asymmetric routing provides additional context, emphasizing the differences in routing approaches and their implications.[9]

4. Scenario 3: Implementing ESI Technology for Multi-Homing

4.1 Overview

Scenario 3 explores the implementation of Ethernet Segment Interface (ESI) technology for multi-homing in a mixed vendor environment involving Cisco and Arista switches. Initially, Virtual Port Channel (VPC) and Multi-chassis Link Aggregation (MLAG) technologies were configured between Cisco and Arista switches, respectively. However, due to compatibility issues with Cisco NXOS version 10 and above, the multi-homing configuration with ESI was successfully implemented only between Arista switches.[10]

4.2 Multi-Chassis Link Aggregation (MLAG)

Definition:

MLAG allows the aggregation of multiple physical links from different switches into a single logical link, providing redundancy and load balancing.

Key Features:

- Redundancy: Reduces single points of failure.

- Load Balancing: Optimizes bandwidth usage.

- Scalability: Increases bandwidth by adding more links.

- Physical Location Requirement: Requires switches to be in the same physical location.

Use Cases:

- Data centers for enhanced link reliability.

- Aggregating uplinks from access to distribution/core switches.[11]

4.3 Ethernet Segment Identifier (ESI) Technology

Definition:

ESI is used in EVPN to manage multi-homed Ethernet segments, ensuring redundancy and load balancing without physical location constraints.

Key Features:

- Redundancy: Supports multiple physical connections.

- Load Balancing: Optimizes network performance.

- EVPN Integration: Scalable Layer 2 and Layer 3 VPN solution.

- No Physical Location Restriction: Allows geographically dispersed deployments.

Use Cases:

- Data center interconnects.

- Reliable multi-homed connections in large networks.[12]

4.4 Comparison

Redundancy and Load Balancing:

- MLAG: Link aggregation across chassis, requires the same location.

- ESI: Multi-homing at Ethernet segment level, no location restriction.

Scalability:

- MLAG: Scales by adding links within the same location.

- ESI: Scales through EVPN, supports dispersed networks.

Use Case Flexibility:

- MLAG: Ideal for data center environments.

- ESI: Suitable for large-scale, flexible deployments.

4.5 ESI Technology and Multi-Vendor Challenges

Ethernet Segment Interface (ESI) is a technology used in Ethernet VPNs (EVPNs) for multi-homing. It provides redundancy and load balancing by allowing multiple links from a single device to be treated as one logical interface, enhancing network resilience and performance.

- Issue with Cisco NXOS 10 and Above:

Cisco NXOS version 10 and later do not support ESI. This limitation arises due to architectural changes in the NXOS software. ESI is defined in [RFC 7432] (https://tools.ietf.org/html/rfc7432), which outlines the standards for EVPN, including multi-homing. The lack of support in NXOS 10+ creates a challenge for multi-vendor environments where seamless ESI integration is required.[13]

4.6 Configuring ESI on Arista Switches

Due to the incompatibility with Cisco NXOS, ESI was implemented between two Arista switches. The existing MLAG configuration was removed to allow for ESI setup.

1. Removing MLAG Configuration:

The previously configured MLAG was disabled to facilitate ESI setup.

```
no mlag configuration
no interface Port-Channel10
```

2. Configuring Port Channel for ESI:

A port channel was created between the Arista leaf switches to establish a logical link.

```
interface Port-Channel20
   description ESI Port Channel
   switchport mode trunk
   switchport trunk allowed vlan all
```

3. ESI Configuration:

ESI was configured on the interfaces, allowing both switches to recognize the logical link.

```
A-LEAFB#show run int po 5
interface Port-Channel5
   switchport trunk allowed vlan 1001-1002
   switchport mode trunk
   !
   evpn ethernet-segment
      identifier 0000:0000:0000:0000:2011
      route-target import 00:1c:73:cd:69:3d
   lacp system-id 0000.0000.2011
A-LEAFB#show run int eth5
interface Ethernet5
   channel-group 5 mode active
A-LEAFB#
```

4.7 ESI Modes

- Active-Active Mode:

Both links are active simultaneously, providing higher bandwidth and redundancy. This mode maximizes link utilization and ensures continuous service even if one link fails.

- Active-Standby Mode:

One link is active while the other is on standby. If the active link fails, the standby link takes over. This mode provides redundancy but does not utilize both links simultaneously, which may be less efficient in terms of bandwidth utilization.

For this scenario, the Active-Active mode was chosen to maximize bandwidth and redundancy.

4.8 Validation and Testing

The successful implementation of ESI was validated by checking the EVPN route types. Specifically, Route Types 1 and 4 were verified to confirm the correct configuration of ESI.

- Route Type 1 (Auto-Discovery): Identifies and advertises the presence of ESI-capable devices.

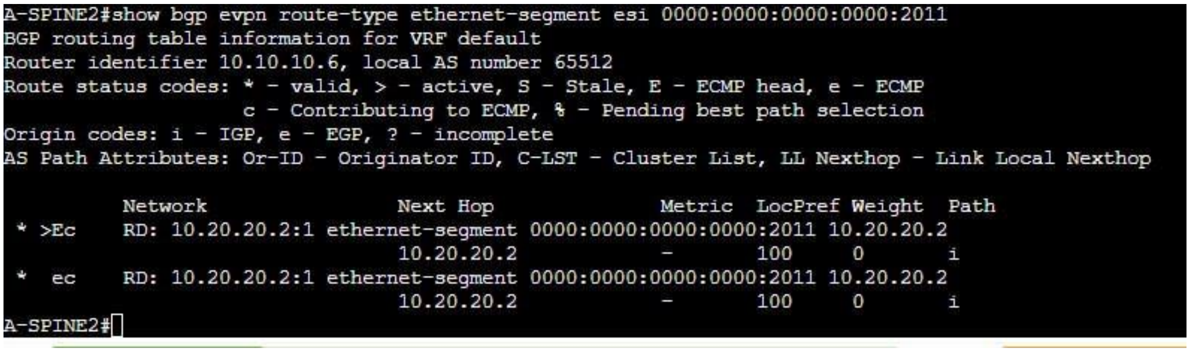
- Route Type 4 (Ethernet Segment): Represents the Ethernet segment, indicating successful ESI configuration.

4.9 Validation Commands:

show bgp evpn route-type Auto-Discovery

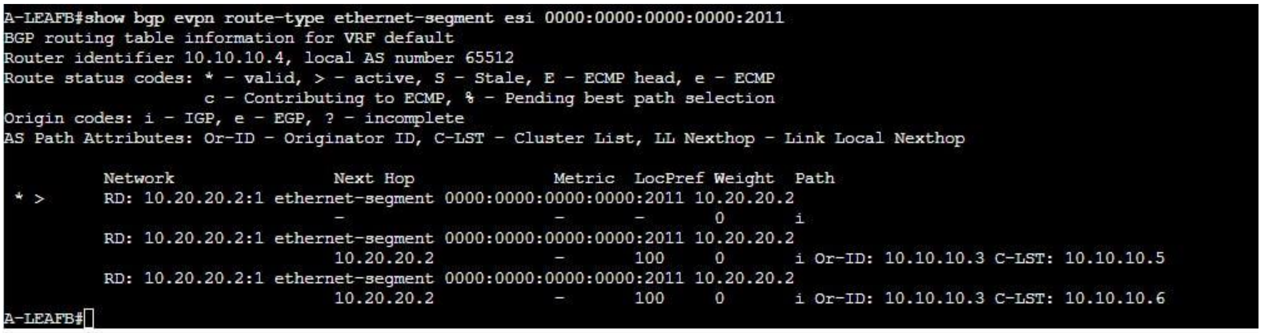
show bgp evpn route-type Ethernet-Segment

The presence of these routes confirmed that ESI was correctly configured and operational between the Arista switches.



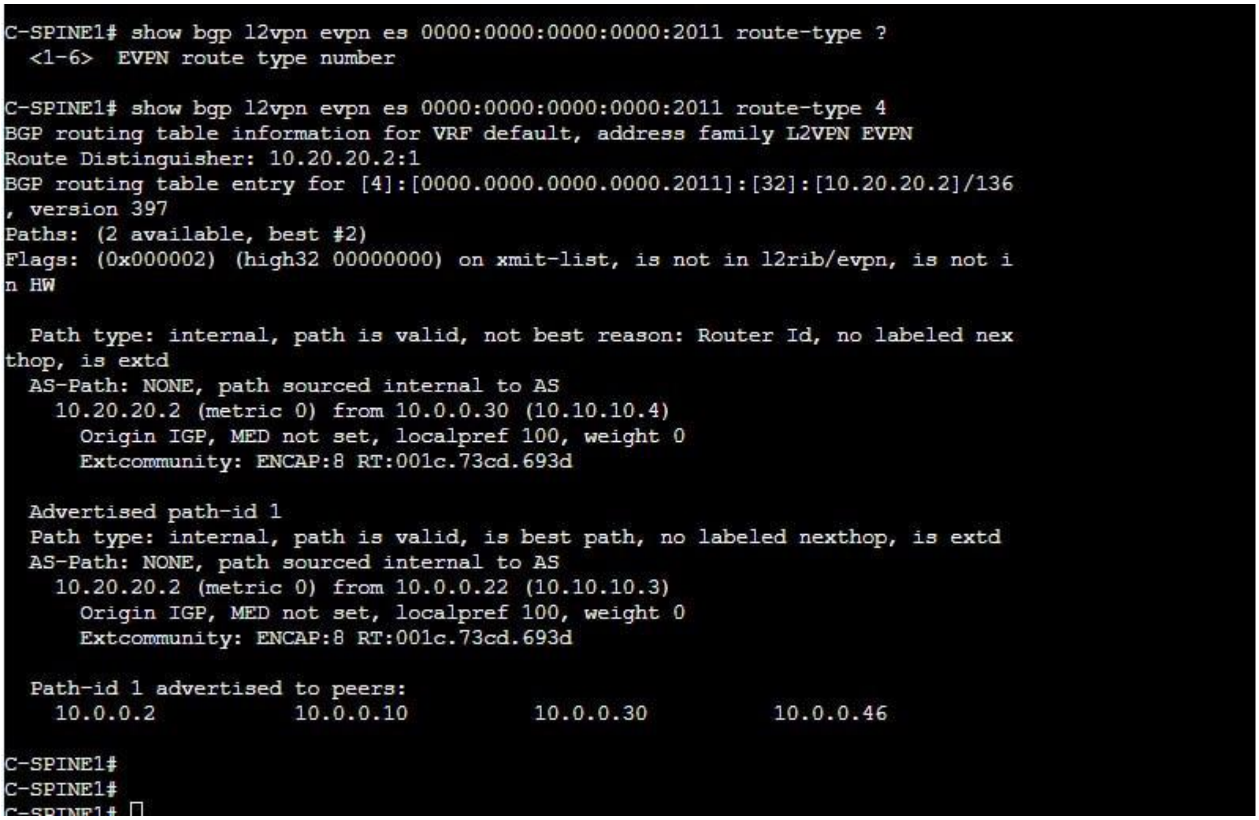
```
A-SPINE2#show bgp evpn route-type ethernet-segment esi 0000:0000:0000:0000:2011
BGP routing table information for VRF default
Router identifier 10.10.10.6, local AS number 65512
Route status codes: * - valid, > - active, S - Stale, E - ECMP head, e - ECMP
                    c - Contributing to ECMP, % - Pending best path selection
Origin codes: i - IGP, e - EGP, ? - incomplete
AS Path Attributes: Or-ID - Originator ID, C-LST - Cluster List, LL Nexthop - Link Local Nexthop

          Network                Next Hop              Metric  LocPref Weight  Path
 * >Ec    RD: 10.20.20.2:1 ethernet-segment 0000:0000:0000:0000:2011 10.20.20.2
                                 10.20.20.2            -       100     0       i
 *  ec    RD: 10.20.20.2:1 ethernet-segment 0000:0000:0000:0000:2011 10.20.20.2
                                 10.20.20.2            -       100     0       i
A-SPINE2#
```

Figure 41: Arista Spine EVPN Route-Type Ethernet-Segment



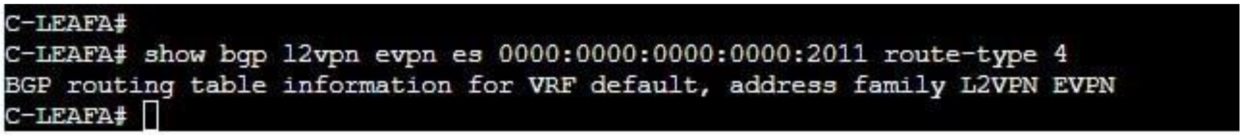
```
A-LEAFB#show bgp evpn route-type ethernet-segment esi 0000:0000:0000:0000:2011
BGP routing table information for VRF default
Router identifier 10.10.10.4, local AS number 65512
Route status codes: * - valid, > - active, S - Stale, E - ECMP head, e - ECMP
                    c - Contributing to ECMP, % - Pending best path selection
Origin codes: i - IGP, e - EGP, ? - incomplete
AS Path Attributes: Or-ID - Originator ID, C-LST - Cluster List, LL Nexthop - Link Local Nexthop

          Network                Next Hop              Metric  LocPref Weight  Path
 * >      RD: 10.20.20.2:1 ethernet-segment 0000:0000:0000:0000:2011 10.20.20.2
                                 -                     -       -       0       i
          RD: 10.20.20.2:1 ethernet-segment 0000:0000:0000:0000:2011 10.20.20.2
                                 10.20.20.2            -       100     0       i Or-ID: 10.10.10.3 C-LST: 10.10.10.5
          RD: 10.20.20.2:1 ethernet-segment 0000:0000:0000:0000:2011 10.20.20.2
                                 10.20.20.2            -       100     0       i Or-ID: 10.10.10.3 C-LST: 10.10.10.6
A-LEAFB#
```

Figure 42: Arista Leaf-B EVPN Route-Type Ethernet-Segment



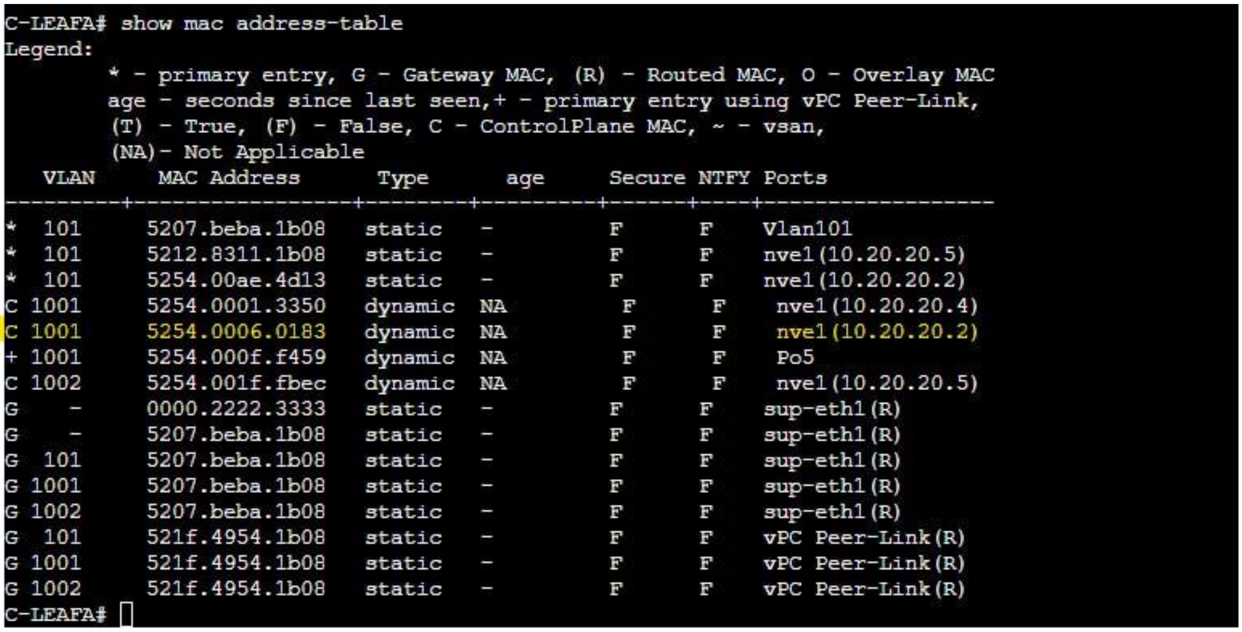
```
C-SPINE1# show bgp l2vpn evpn es 0000:0000:0000:0000:2011 route-type ?
  <1-6>  EVPN route type number

C-SPINE1# show bgp l2vpn evpn es 0000:0000:0000:0000:2011 route-type 4
BGP routing table information for VRF default, address family L2VPN EVPN
Route Distinguisher: 10.20.20.2:1
BGP routing table entry for [4]:[0000.0000.0000.0000.2011]:[32]:[10.20.20.2]/136
, version 397
Paths: (2 available, best #2)
Flags: (0x000002) (high32 00000000) on xmit-list, is not in l2rib/evpn, is not i
n HW

  Path type: internal, path is valid, not best reason: Router Id, no labeled nex
thop, is extd
  AS-Path: NONE, path sourced internal to AS
    10.20.20.2 (metric 0) from 10.0.0.30 (10.10.10.4)
      Origin IGP, MED not set, localpref 100, weight 0
      Extcommunity: ENCAP:8 RT:001c.73cd.693d

  Advertised path-id 1
  Path type: internal, path is valid, is best path, no labeled nexthop, is extd
  AS-Path: NONE, path sourced internal to AS
    10.20.20.2 (metric 0) from 10.0.0.22 (10.10.10.3)
      Origin IGP, MED not set, localpref 100, weight 0
      Extcommunity: ENCAP:8 RT:001c.73cd.693d

  Path-id 1 advertised to peers:
    10.0.0.2           10.0.0.10          10.0.0.30          10.0.0.46

C-SPINE1#
C-SPINE1#
C-SPINE1#
```

Figure 43: Cisco Spine BGP L2VPN Route-Type 4



```
C-LEAFA#
C-LEAFA# show bgp l2vpn evpn es 0000:0000:0000:0000:2011 route-type 4
BGP routing table information for VRF default, address family L2VPN EVPN
C-LEAFA#
```

Figure 44: Cisco Leaf-A BGP L2VPN Route-Type 4



```
C-LEAFA# show mac address-table
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,
        (NA)- Not Applicable
   VLAN     MAC Address      Type      age     Secure NTFY Ports
---------+-----------------+--------+---------+------+----+------------------
*  101     5207.beba.1b08   static   -         F      F    Vlan101
*  101     5212.8311.1b08   static   -         F      F    nve1(10.20.20.5)
*  101     5254.00ae.4d13   static   -         F      F    nve1(10.20.20.2)
C 1001     5254.0001.3350   dynamic  NA         F      F    nve1(10.20.20.4)
C 1001     5254.0006.0183   dynamic  NA         F      F    nve1(10.20.20.2)
+ 1001     5254.000f.f459   dynamic  NA         F      F    Po5
C 1002     5254.001f.fbec   dynamic  NA         F      F    nve1(10.20.20.5)
G    -     0000.2222.3333   static   -         F      F    sup-eth1(R)
G    -     5207.beba.1b08   static   -         F      F    sup-eth1(R)
G  101     5207.beba.1b08   static   -         F      F    sup-eth1(R)
G 1001     5207.beba.1b08   static   -         F      F    sup-eth1(R)
G 1002     5207.beba.1b08   static   -         F      F    sup-eth1(R)
G  101     521f.4954.1b08   static   -         F      F    vPC Peer-Link(R)
G 1001     521f.4954.1b08   static   -         F      F    vPC Peer-Link(R)
G 1002     521f.4954.1b08   static   -         F      F    vPC Peer-Link(R)
C-LEAFA#
```

Figure 45: Cisco Leaf-A Mac-Address Table

4.10 Conclusion

Scenario 3 successfully demonstrates the implementation of ESI technology for multi-homing between Arista switches. The challenge of compatibility with Cisco NXOS 10 and above was addressed by implementing ESI solely on Arista devices. The configuration was validated through EVPN route types, confirming ESI's correct setup and operational status. This scenario highlights the importance of understanding vendor-specific limitations and effectively adapting network configurations to achieve redundancy and load balancing in a mixed vendor environment.[14],[15]

Conclusions

The successful implementation of VXLAN across Cisco and Arista devices demonstrates the feasibility and benefits of using VXLAN for network virtualization in large-scale data centers. The scenarios covered in the document highlight the practical challenges and solutions for achieving seamless interoperability in a multi-vendor environment. By leveraging the strengths of Cisco and Arista technologies, the document provides a comprehensive guide for network engineers to implement scalable, flexible, and efficient network virtualization solutions. The insights gained from these implementations lay the groundwork for more advanced configurations and the continued evolution of data center networks.

References

[1] (n.d.). Retrieved from https://tools.ietf.org/html/rfc7348

[2] (n.d.). Retrieved from https://www.arista.com/en/products/multi-chassis-link-aggregation-mlag

[3] (n.d.). Retrieved from https://www.cisco.com/c/en/us/td/docs/dcn/nx-os/nexus9000/104x/configuration/vxlan/cisco-nexus-9000-series-nx-os-vxlan-configuration-guide-release-104x/m-interoperability-with-mvpn-multi-homing-using-esi.html

[4] (2015, February). Retrieved from https://tools.ietf.org/html/rfc7432

[5] (2015, February). Retrieved from https://tools.ietf.org/html/rfc7432

[6] (2015 , February). Retrieved from https://datatracker.ietf.org/doc/html/rfc7432

[7] (2018, March). Retrieved from https://tools.ietf.org/html/rfc8365 Mahalingam, M. (2014, August).

[8] *RFC 7348*. Retrieved from https://tools.ietf.org/html/rfc7348

[9] *RFC 4271*. (2006 , January). Retrieved from https://tools.ietf.org/html/rfc4271

[10] *RFC 4760*. (2007, January). Retrieved from https://tools.ietf.org/html/rfc4760

[11] *RFC 7432*. (n.d.). Retrieved from MPLS-Based Ethernet VPN: https://tools.ietf.org/html/rfc7432

[12] *RFC 7432*. (n.d.). Retrieved from BGP MPLS-Based Ethernet VPN: https://tools.ietf.org/html/rfc7432

[13] *RFC 7432*. (2015, February). Retrieved from https://tools.ietf.org/html/rfc7432

[14] *RFC 7752*. (2016, March). Retrieved from https://tools.ietf.org/html/rfc7752

[15] *RFC 8365*. (2018, March). Retrieved from https://datatracker.ietf.org/doc/html/rfc8365