Introduction to VXLAN:

The Future Path of Your Data Center

Richard Licon
Principal Technical Marketing Engineer

BRKDCN-1621

Cisco Webex App

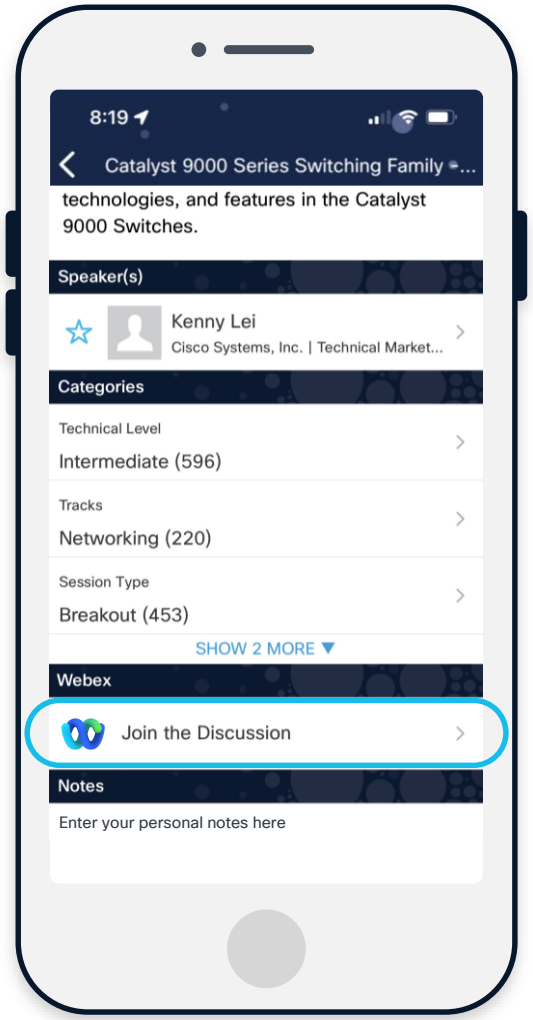
Questions?

Use Cisco Webex App to chat
with the speaker after the session

How

1. Find this session in the Cisco Live Mobile App
2. Click “Join the Discussion”
3. Install the Webex App or go directly to the Webex space
4. Enter messages/questions in the Webex space

Webex spaces will be moderated by the speaker until June 13, 2025.



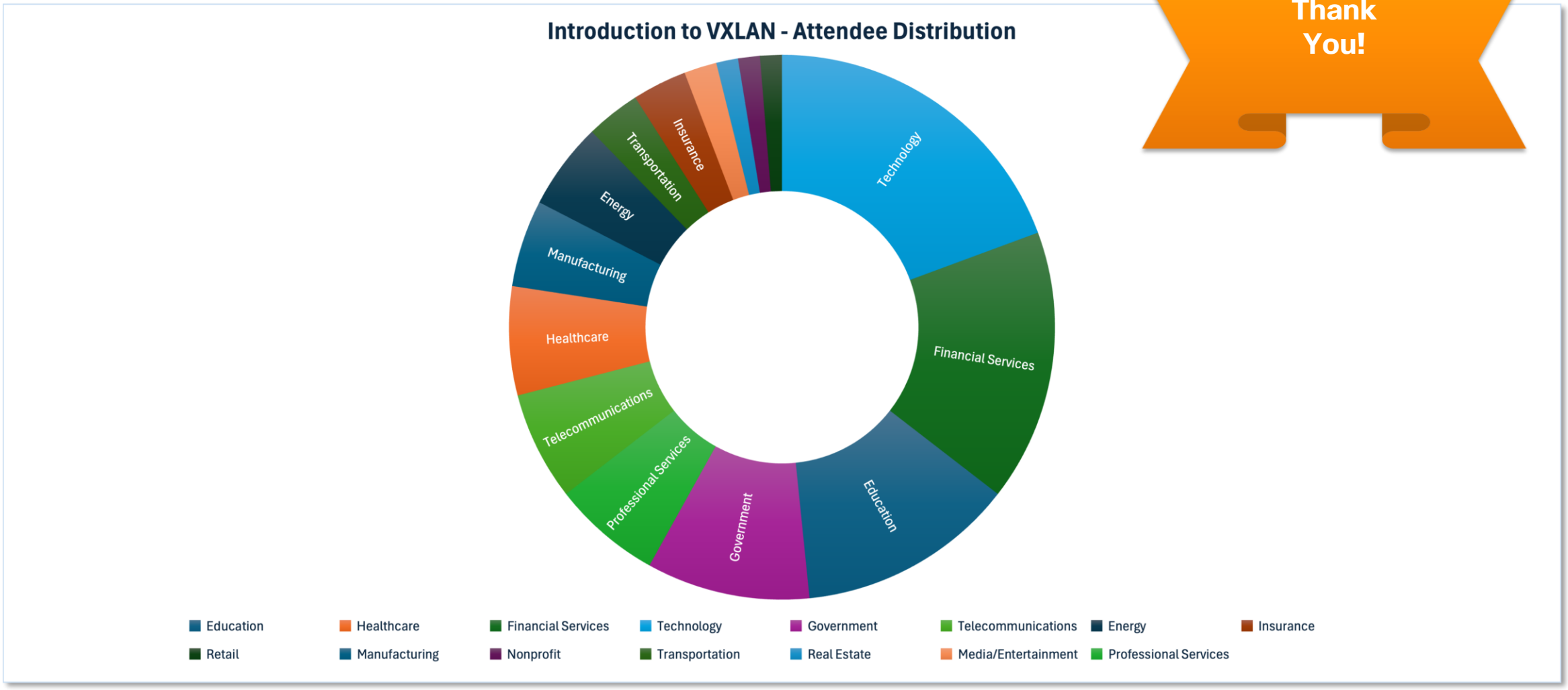
https://ciscolive.ciscoevents.com/
ciscolivebot/**#BRKDCN-1621**

Session Abstract

Learn more about VXLAN and delve deeper into it. This session will cover what VXLAN is, how it works and the problems it solves. The need for flexibility in the data center opens the door to network overlays. Several standards have been proposed and implemented to enable virtualized networks to be layered over a physical network. The data center network needs to be flexible enough to support workloads that can move from one host to another, and for new services to be deployed rapidly. VXLAN is one such technology that provides massively scalable virtual network overlays on top of existing IP infrastructures.

This session will include 101s and packet walks as we cover the insides of VXLAN. No prior knowledge of VXLAN is required. A basic understanding of unicast and multicast routing protocol is needed, along with an understanding of MP BGP.

Attendee Distribution List (Verticals)



Thank You!

Agenda

01 Evolution of the Data Center

02 The Need for network overlays

03 What is VXLAN?

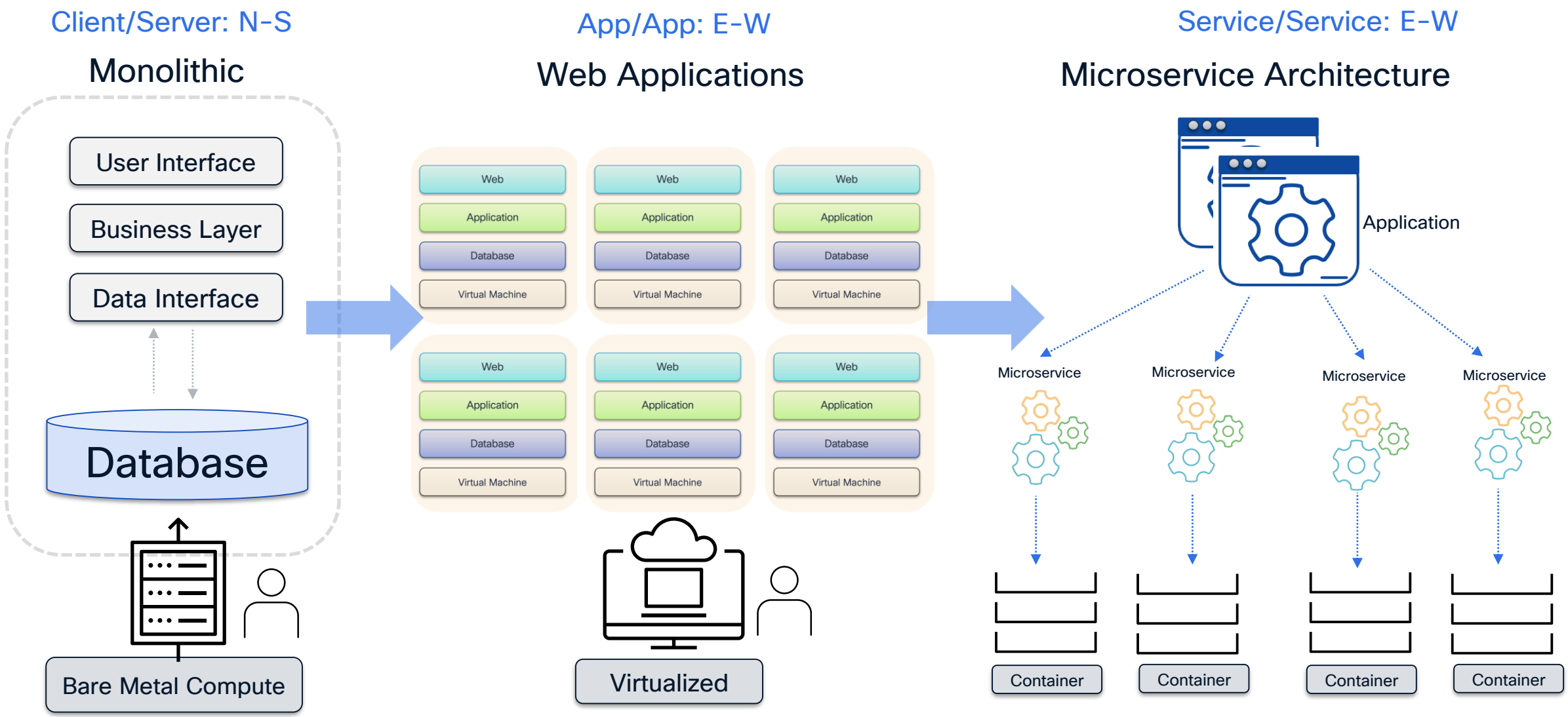
04 The Need for a Control Plane

05 Packet Walks

06 Conclusion

Evolution of the Application Architecture

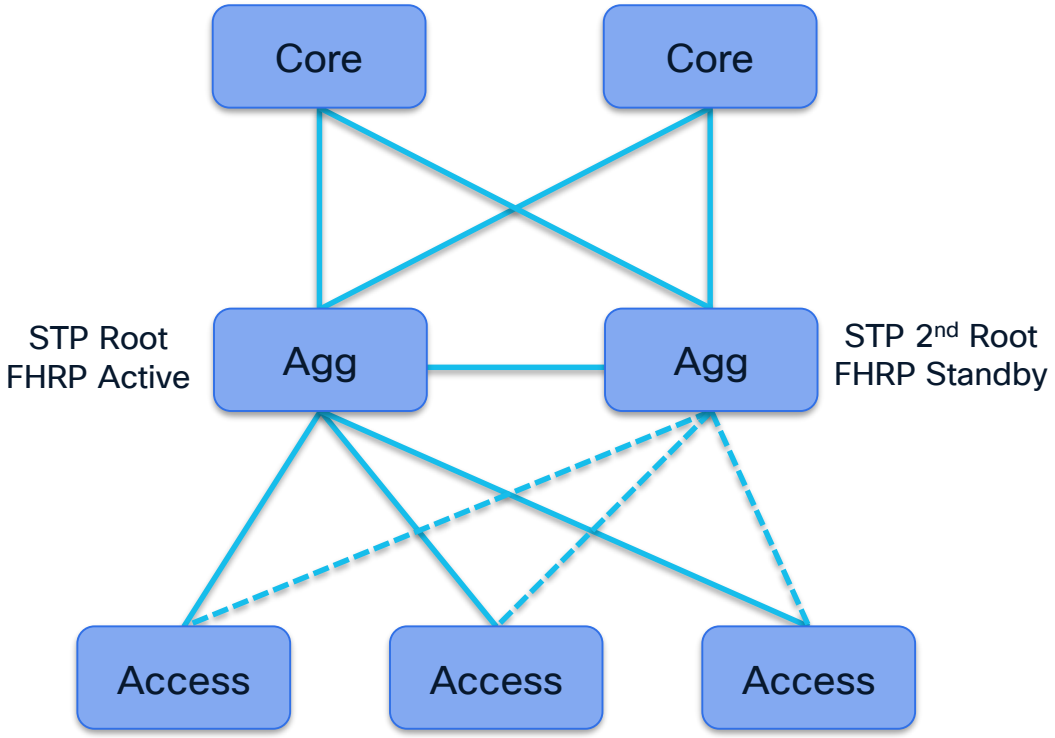
Evolution of the Application Architectures



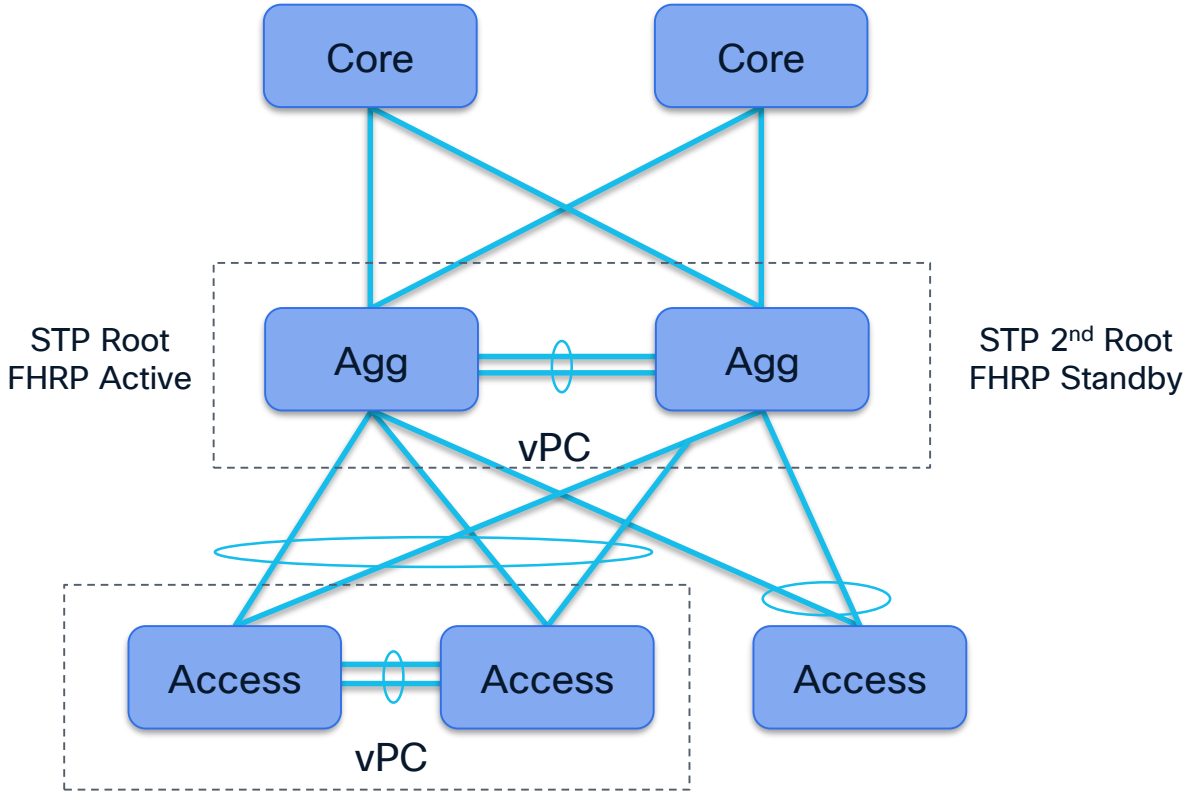
Evolution of the Data Center Network

Evolution of the Data Center Network

Classic Spanning-Tree



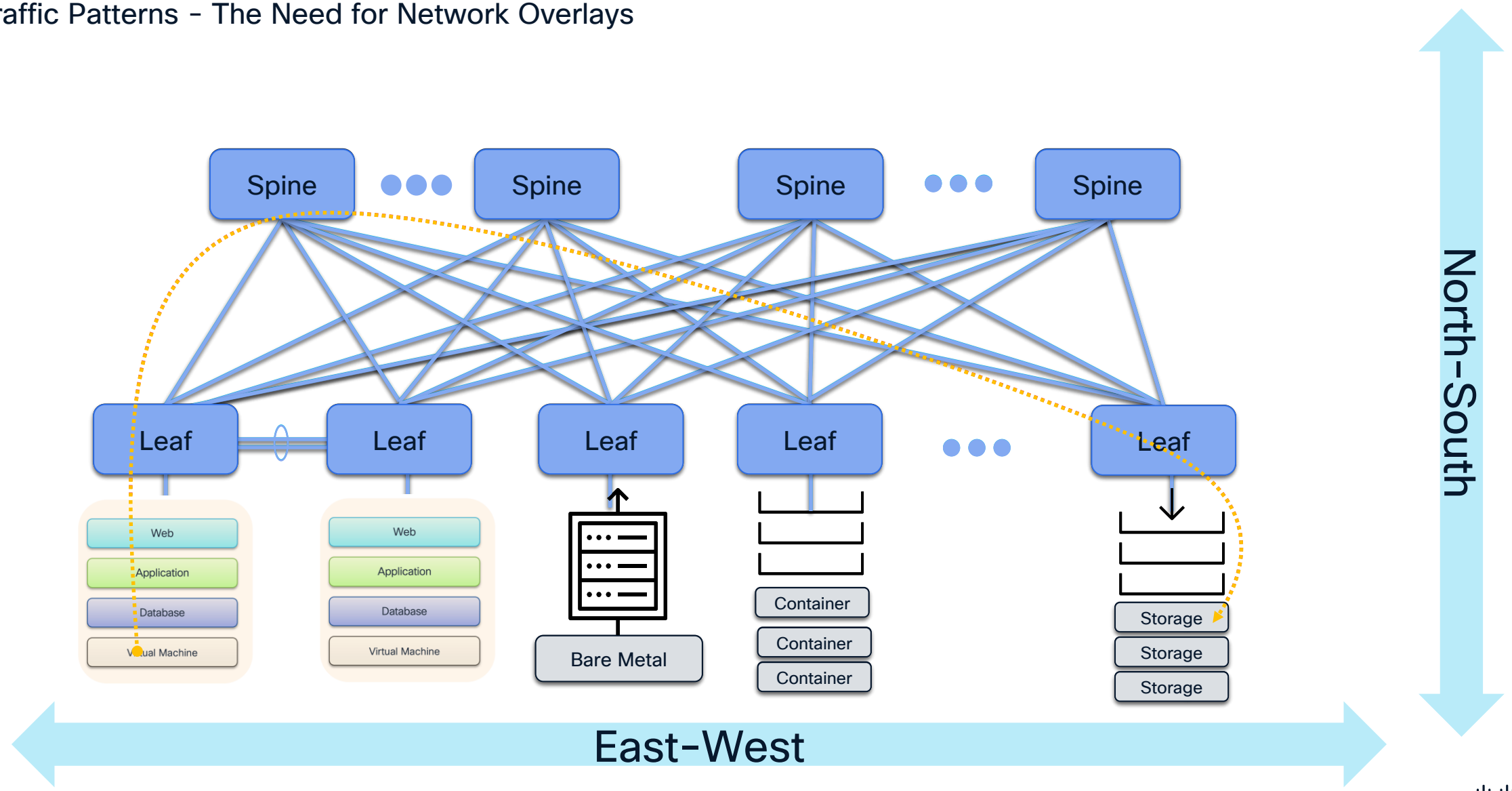
vPC and Spanning-Tree



Evolution of the DC Network Architecture: Clos Networks

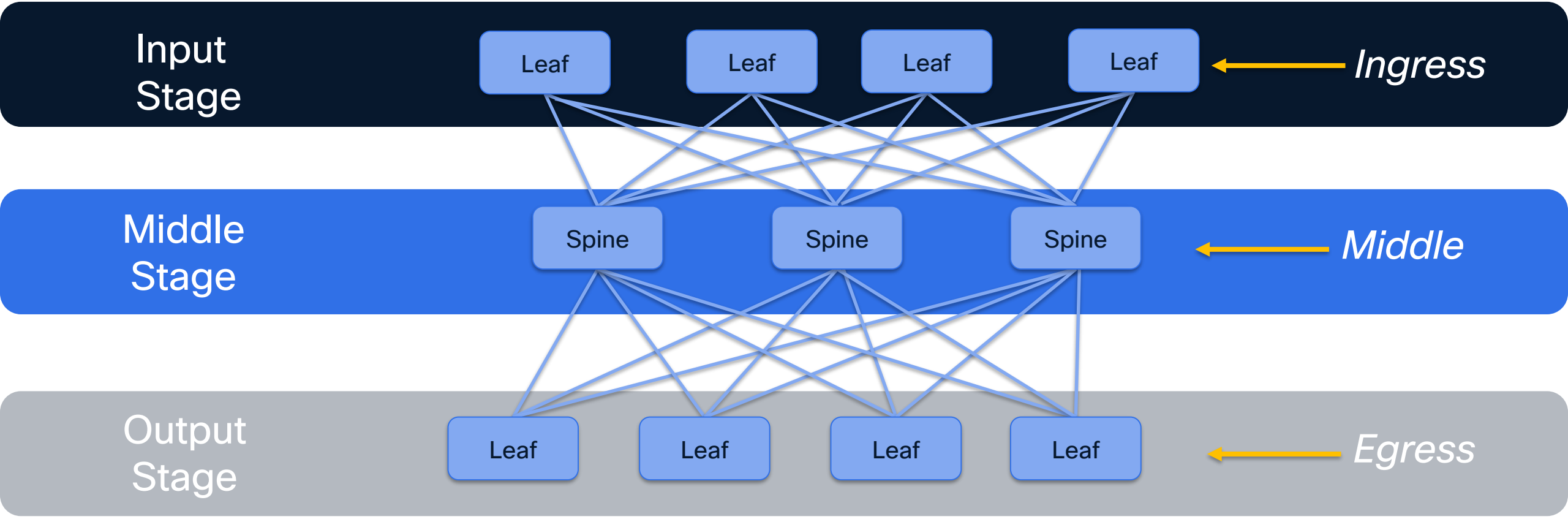
Evolution of the Data Center Network

Newer Traffic Patterns – The Need for Network Overlays



Evolution of the Data Center Network

3-Stage Clos Fabric

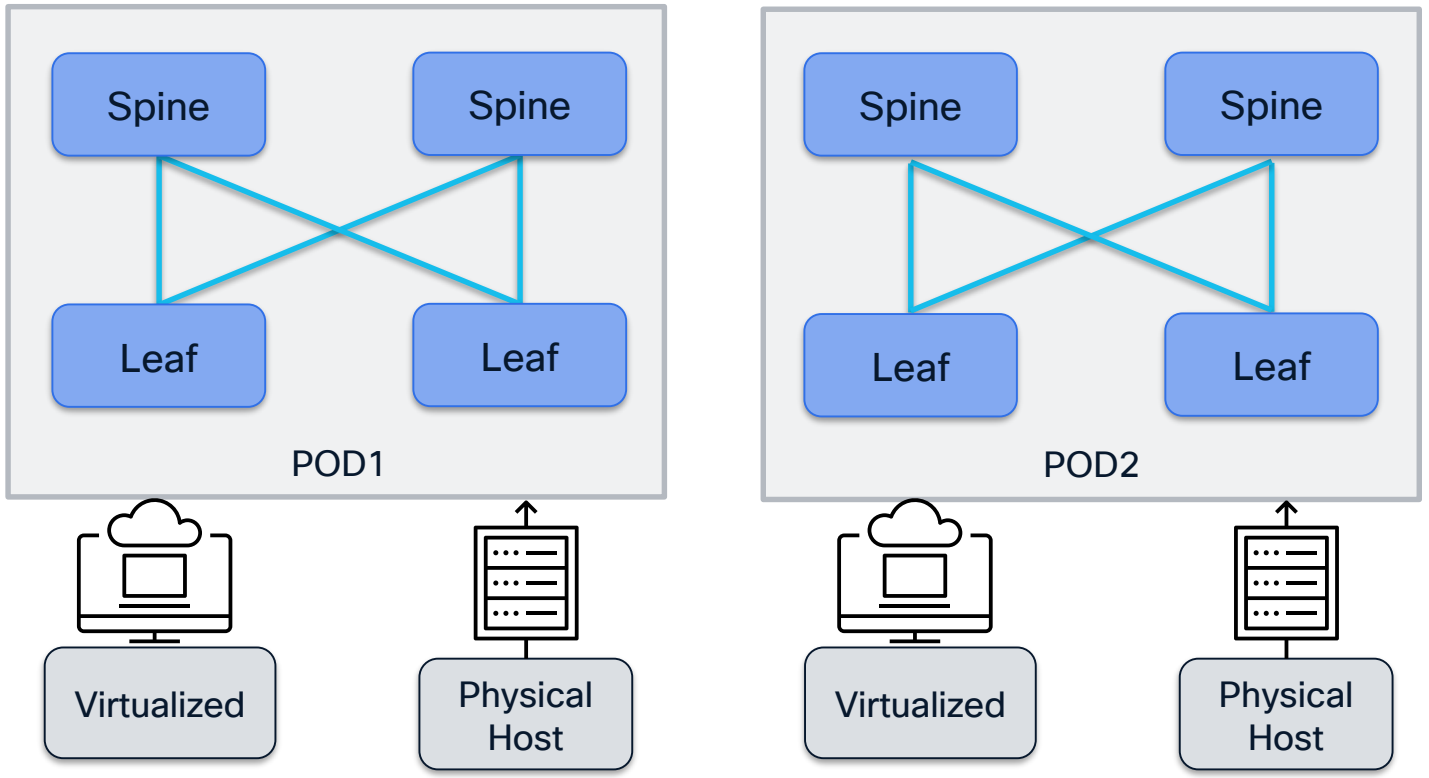


Invented by Edson Erwin in 1938
March 1953: A Study of Non-Blocking Switching Networks. (Clos, Charles)

3-Stage Clos Fabric

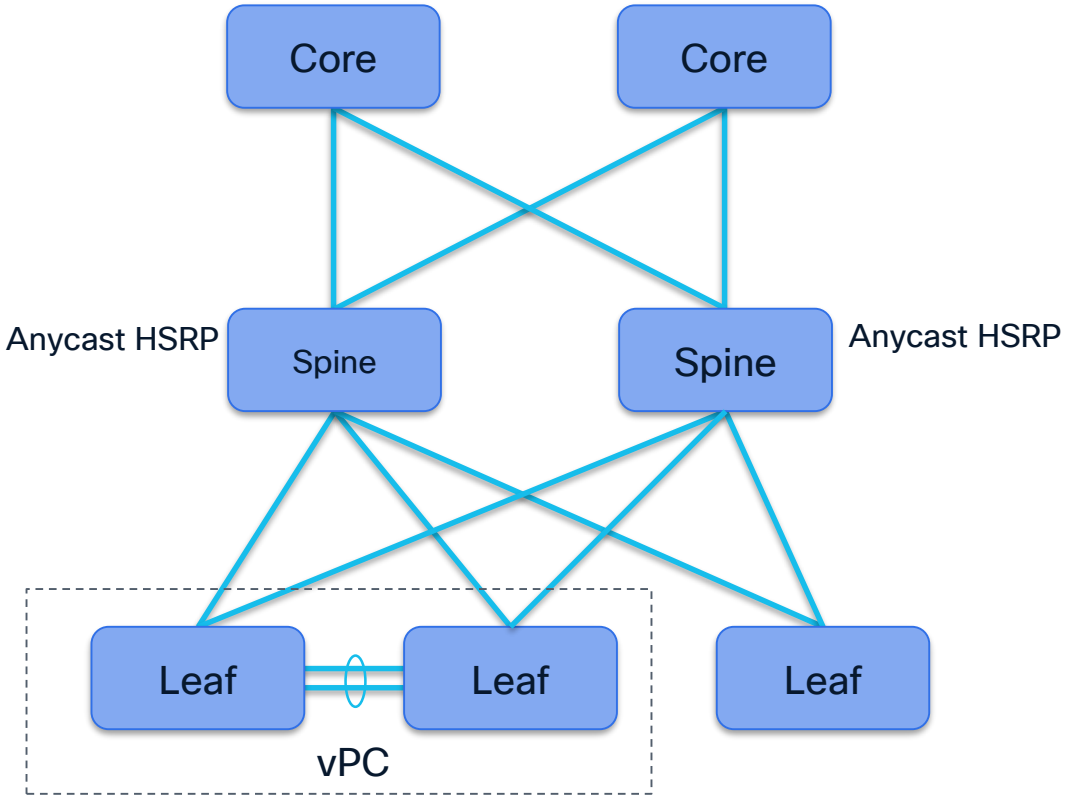
Evolution of the Data Center Network

Folded 3-Stage Fabric

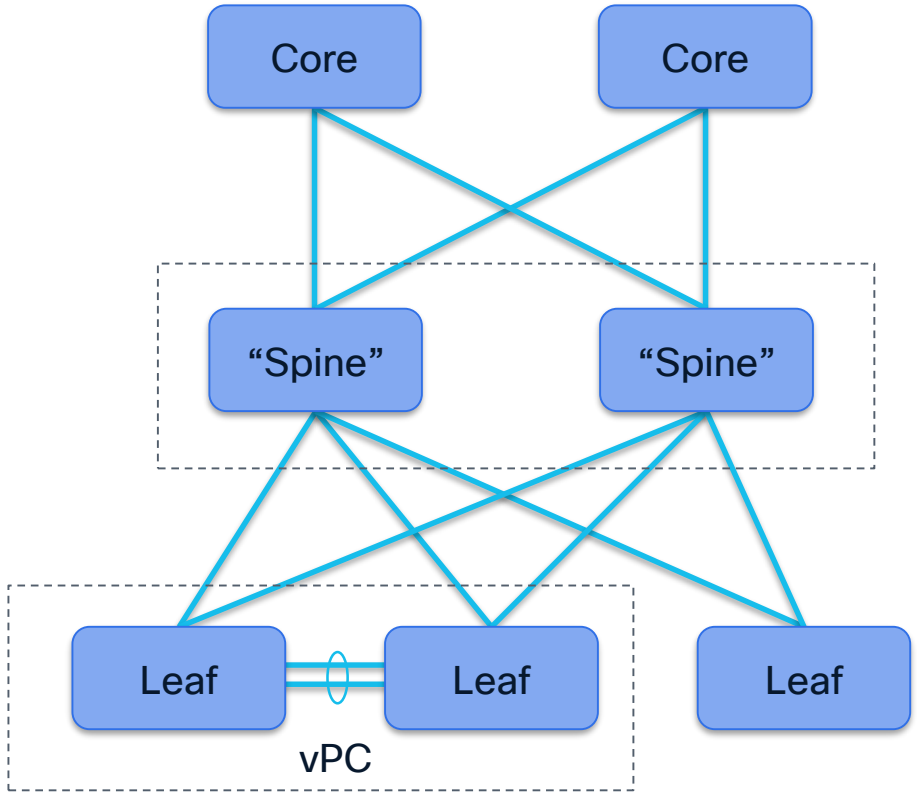


Evolution of the Data Center Network

FabricPath (MAC-in-MAC)



VXLAN Flood & Learn (MAC-in-IP)



Data Center Network Challenges

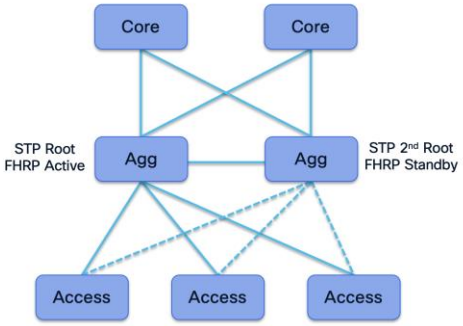
Legacy Architectures

Convergence Time:
STP high convergence time
TCN MAC Flush

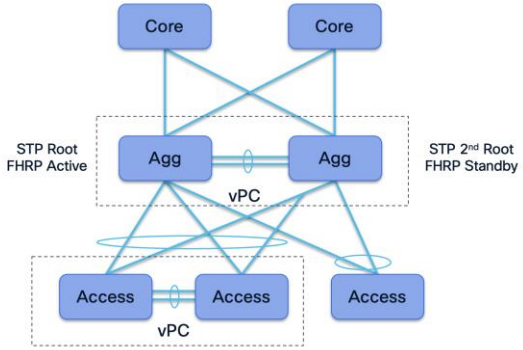
Unused Links:
STP Redundant links in blocked state

Rigid Network Services:
L4-L7 Services placed at Distribution Layer

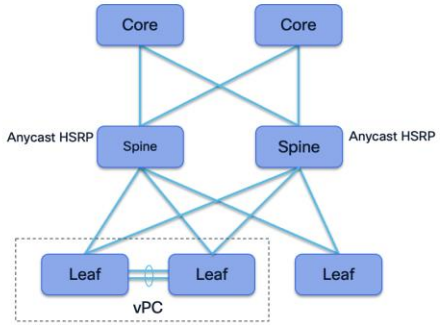
Classic Spanning-Tree



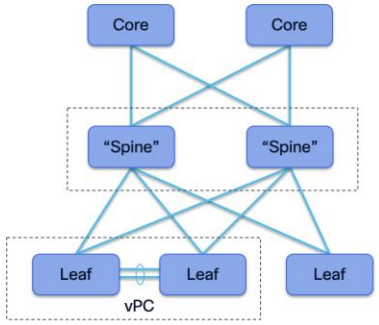
vPC and Spanning-Tree



FabricPath (MAC-in-MAC)



VXLAN Flood & Learn (MAC-in-IP)



Suboptimal Forwarding:
Path defined by
Root Switch or FHRP Active

OpEx/CapEx:
Expensive to upgrade
Scale up with large Chassis

Limited Scale:
No Control Plane
Limited Workload Mobility

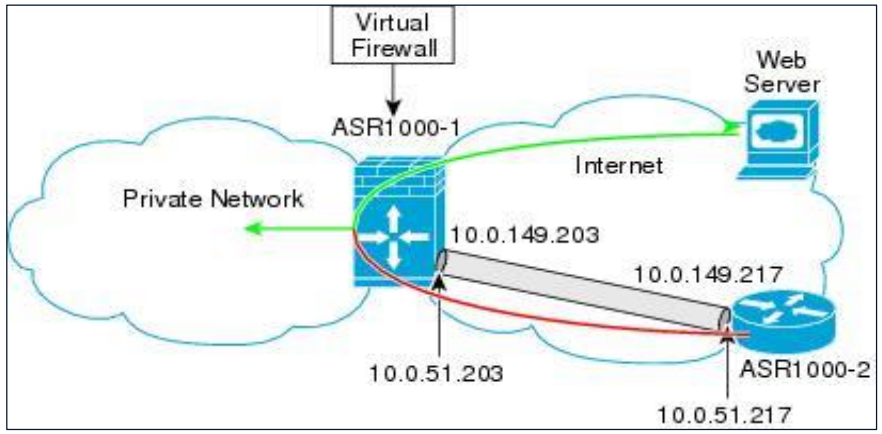
"Spine" = Not Really a Spine

The Need for Network Overlays

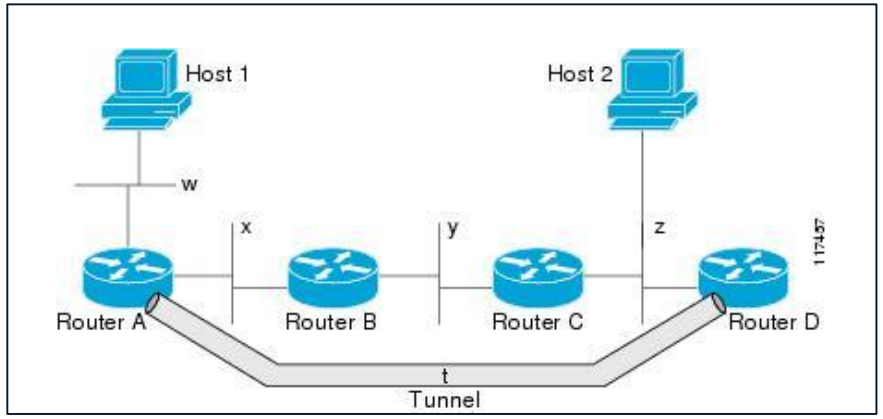
Examples of Overlays – Rewind

DID YOU KNOW?

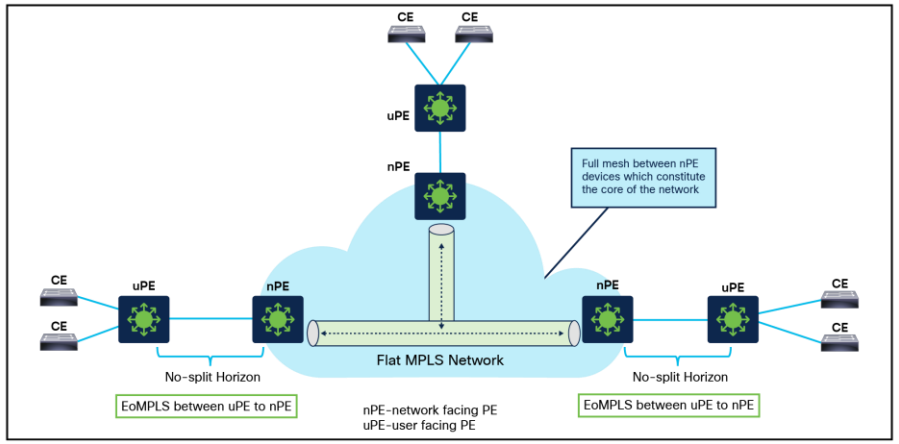
IPSec



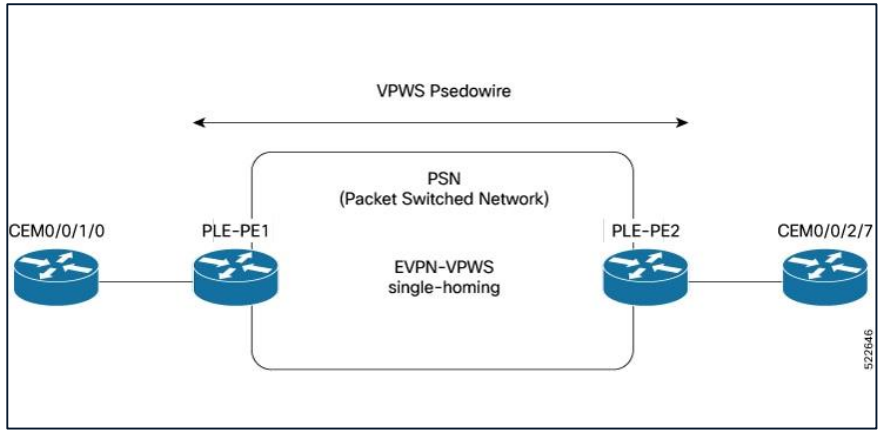
GRE (Generic Route Encapsulation)



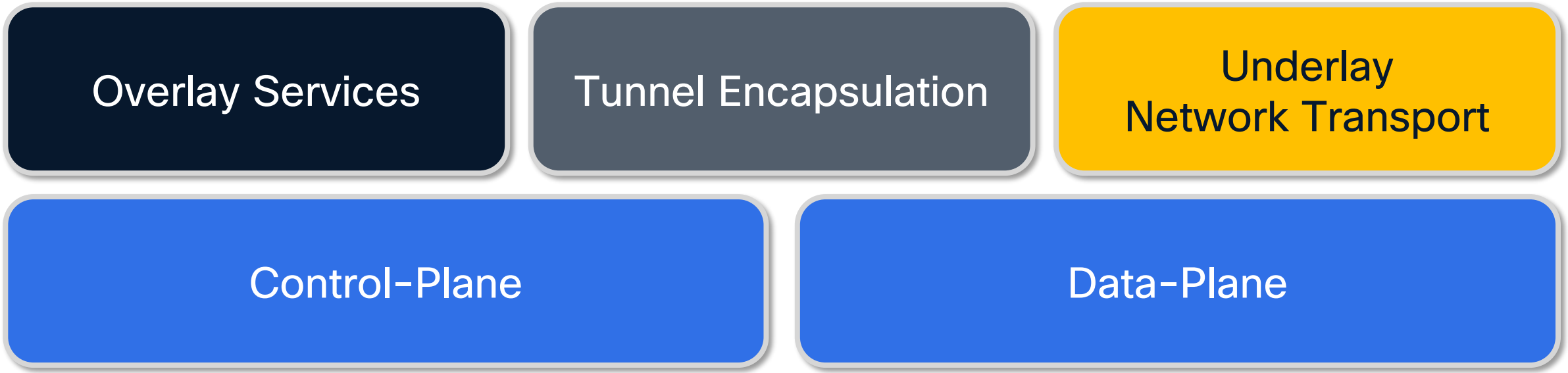
MPLS



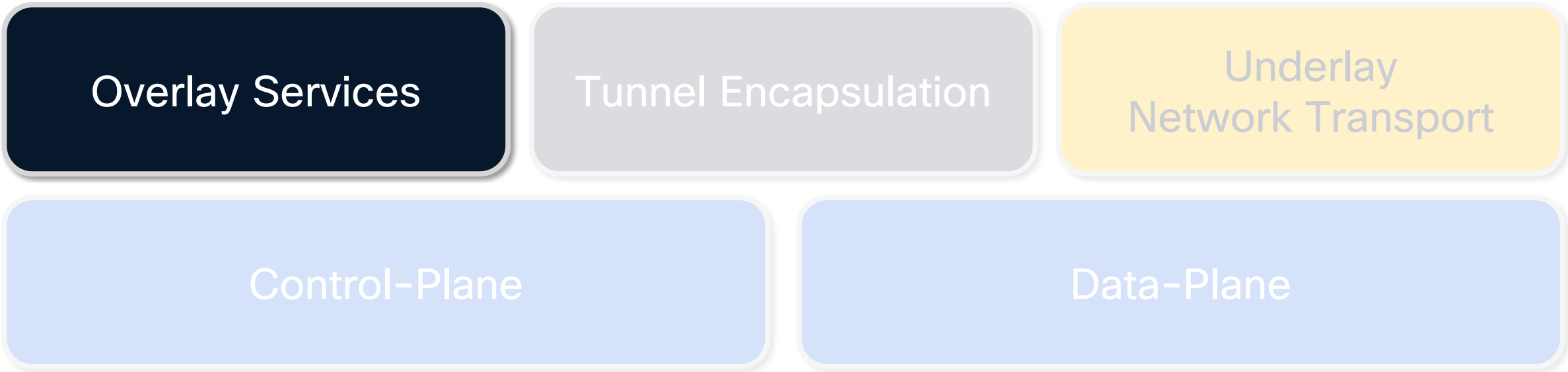
VPWS – Virtual Private Wire Service



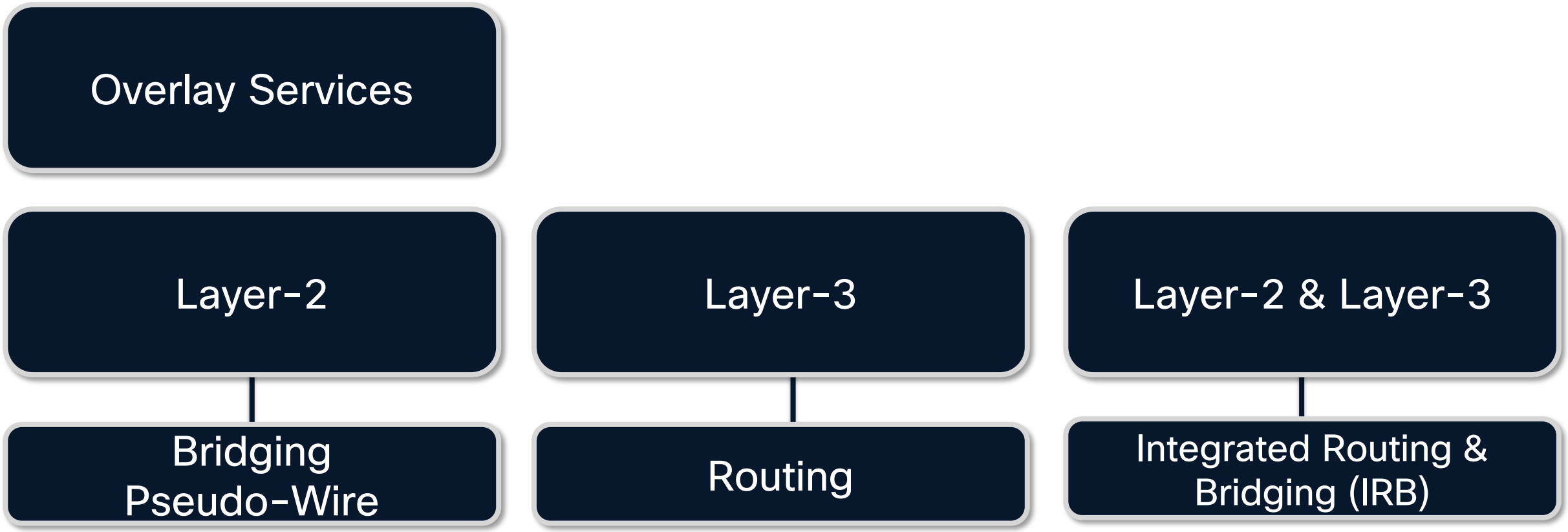
Overlay Taxonomy



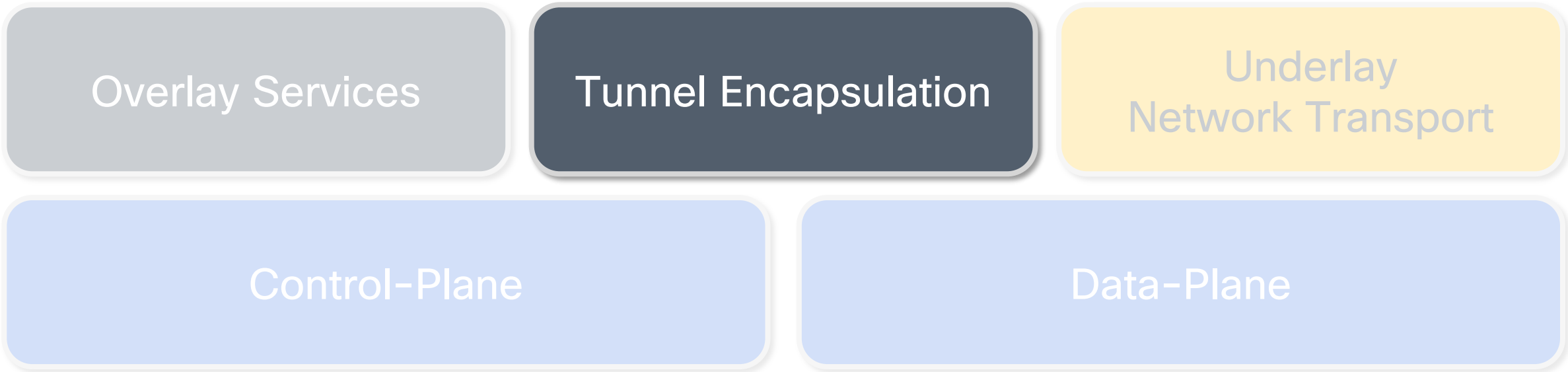
Overlay Services



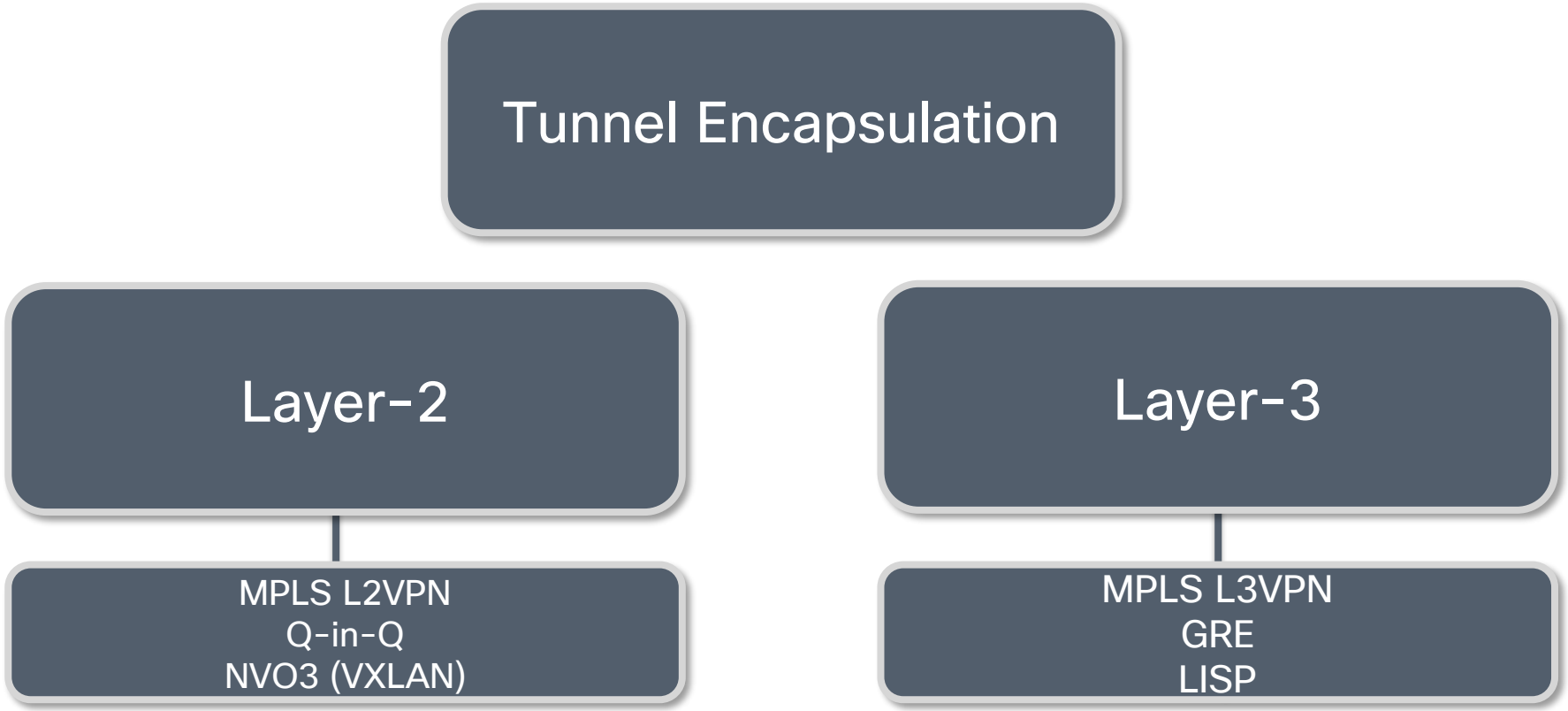
Overlay Services



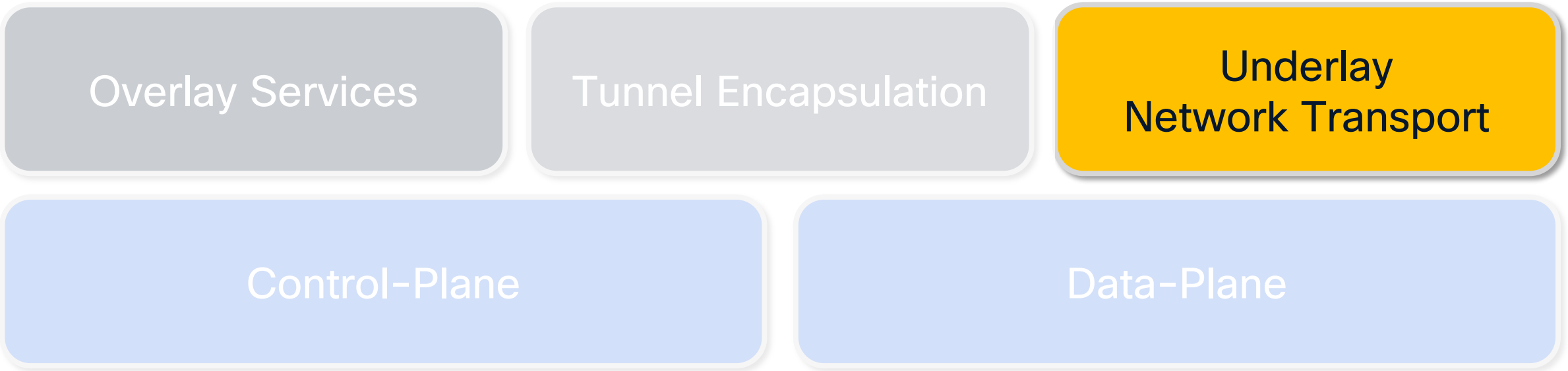
Tunnel Encapsulation



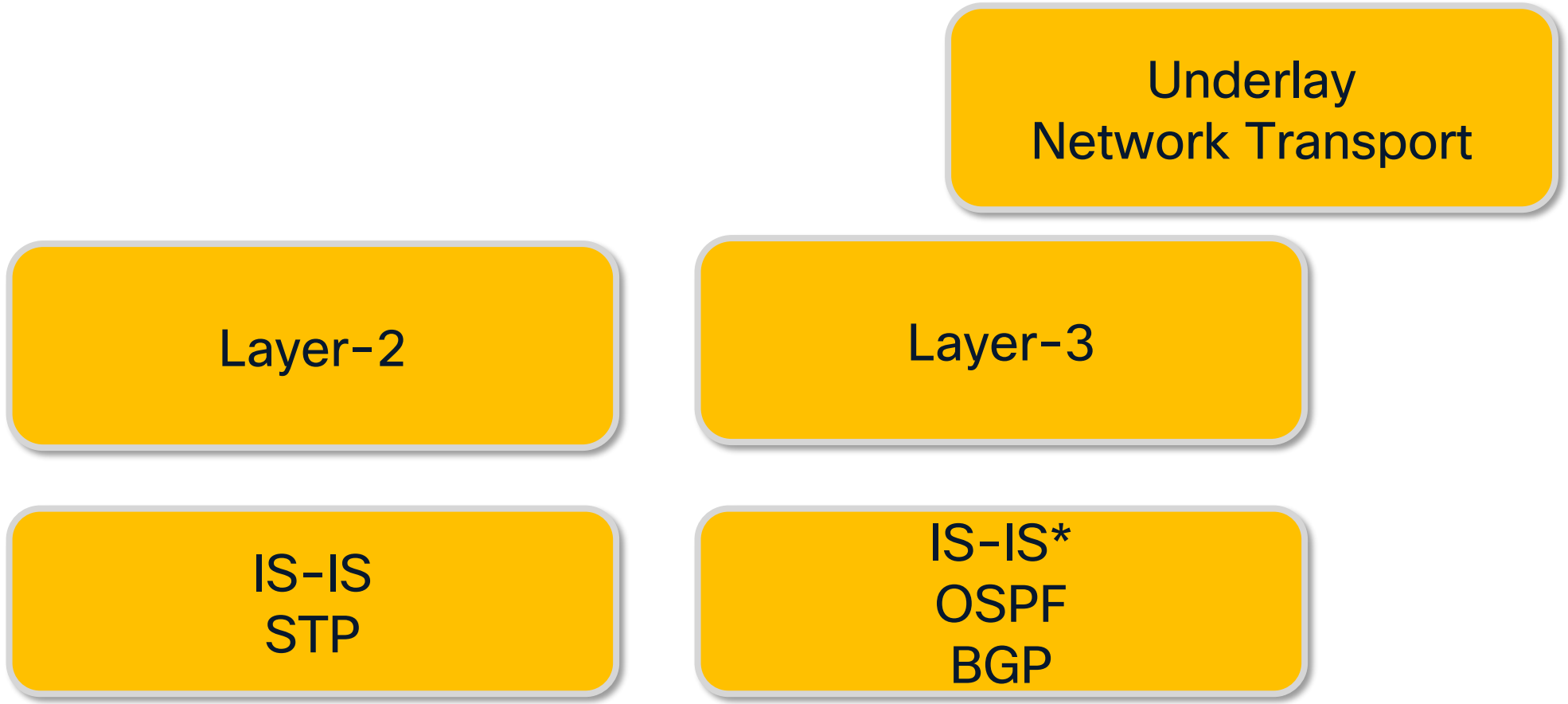
Tunnel Encapsulation



Underlay Network Transport

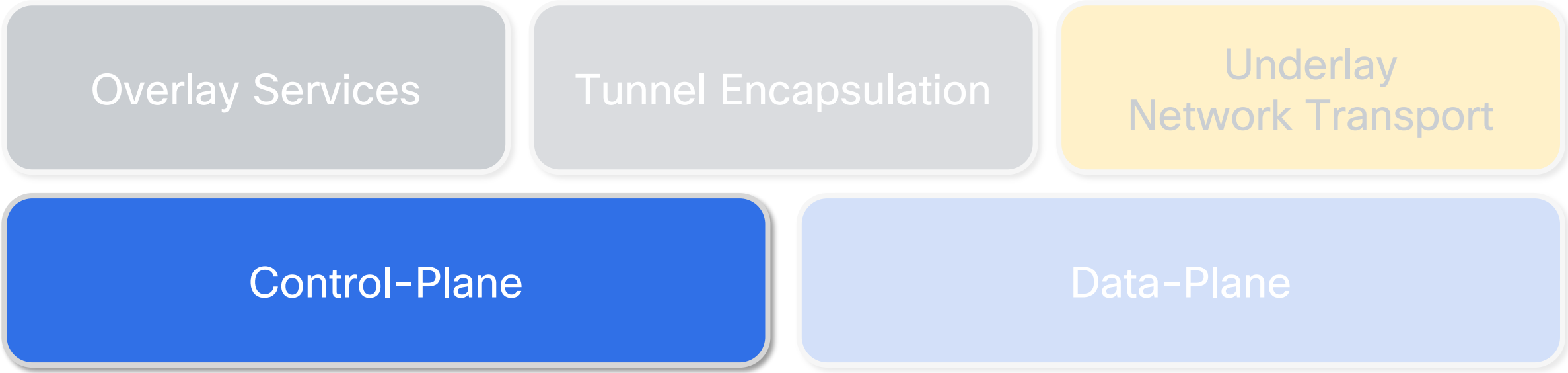


Underlay Network Transport



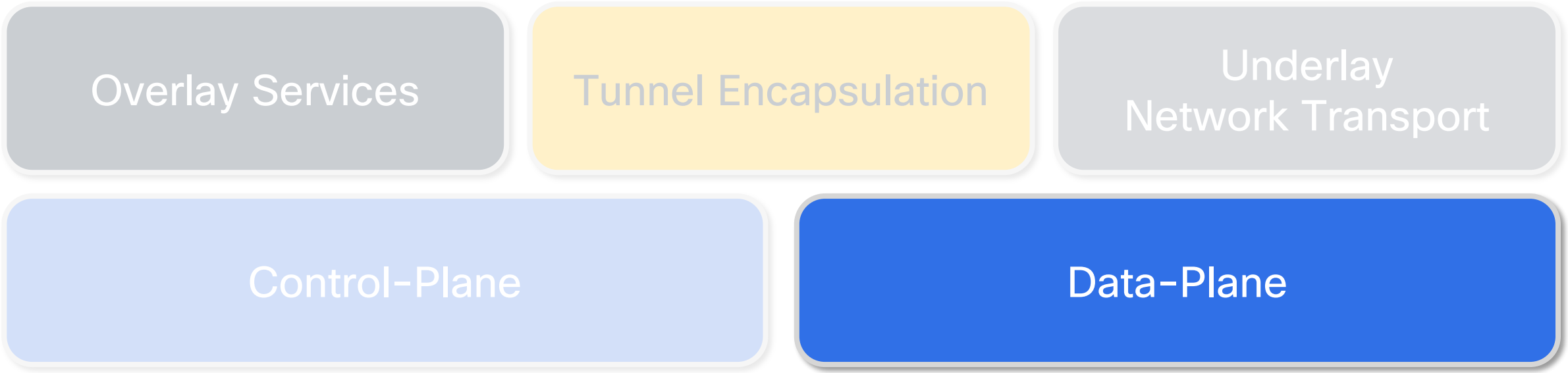
*IS-IS operates at L2 but transports L2 & L3

Control-Plane

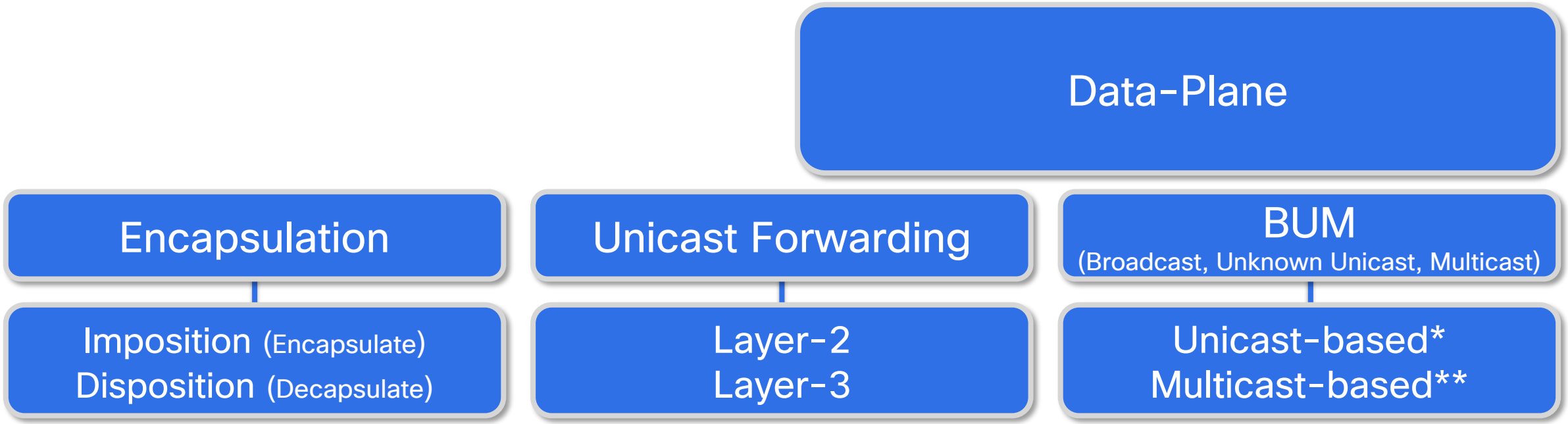


Control-Plane

Data-Plane



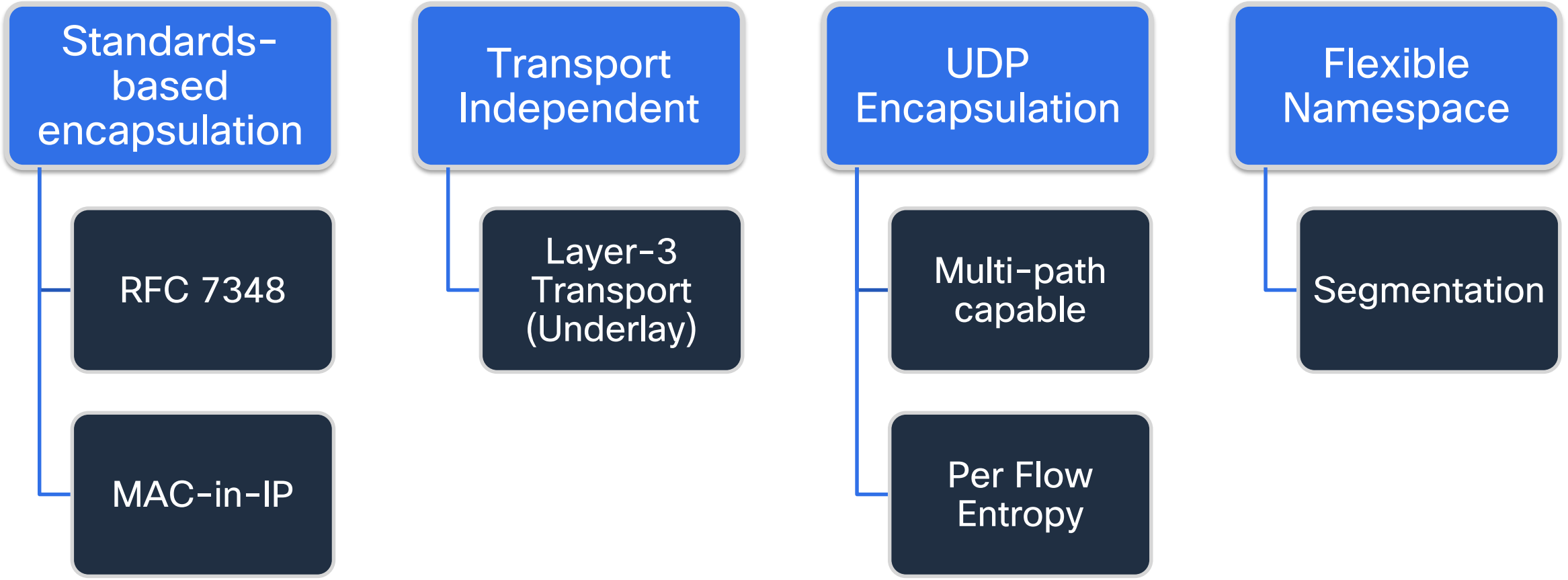
Data-Plane



*Ingress/Head-End Replication
**PIM

VXLAN:
Virtual Extensible LAN

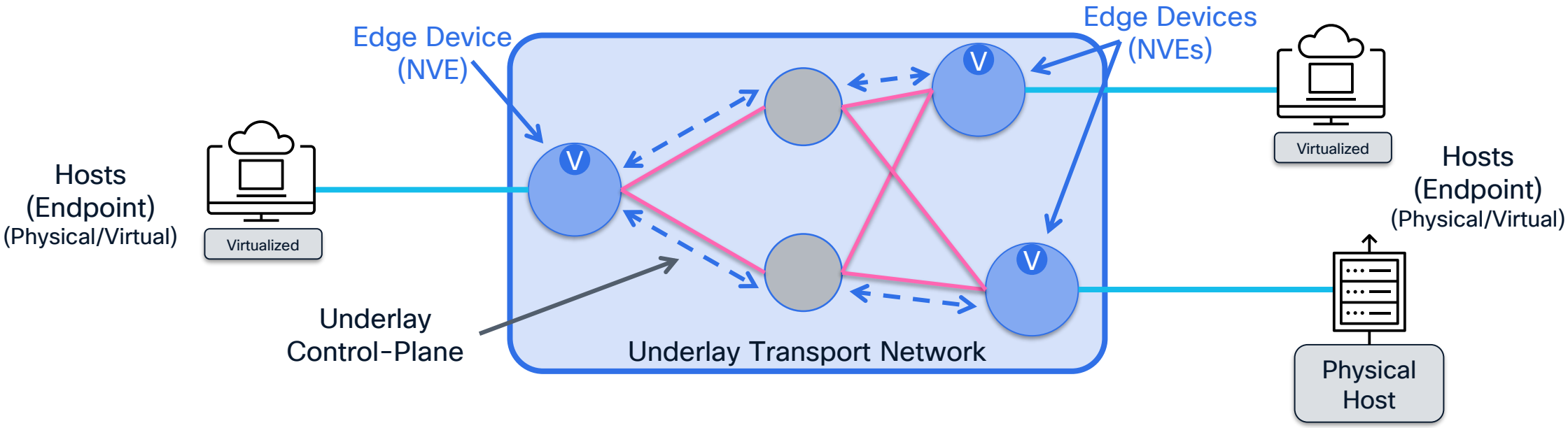
What is VXLAN?



Underlay Taxonomy

- Edge Devices host the VTEP
- Responsible for the encapsulation and decapsulation of the VXLAN Header

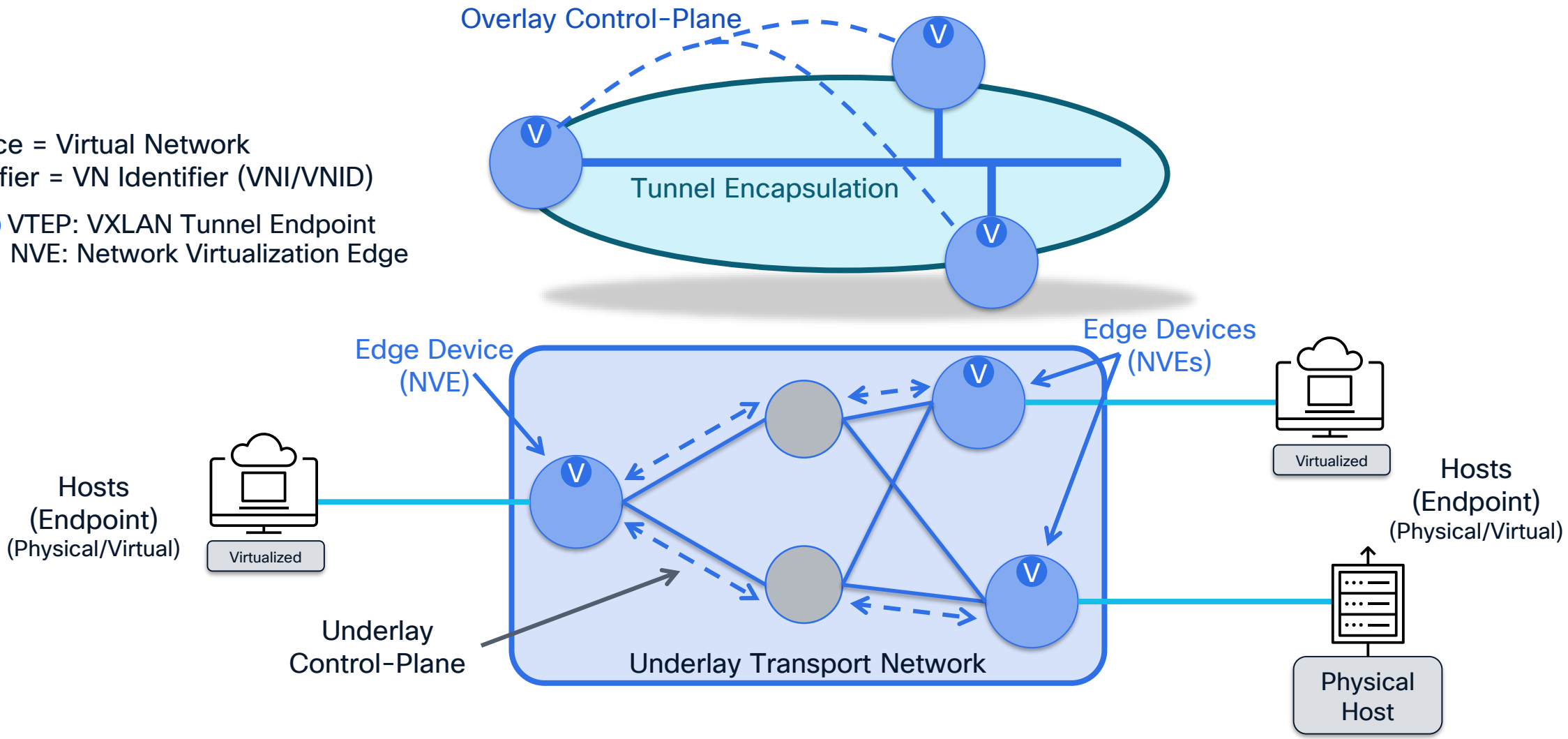
V VTEP: VXLAN Tunnel Endpoint
NVE: Network Virtualization Edge



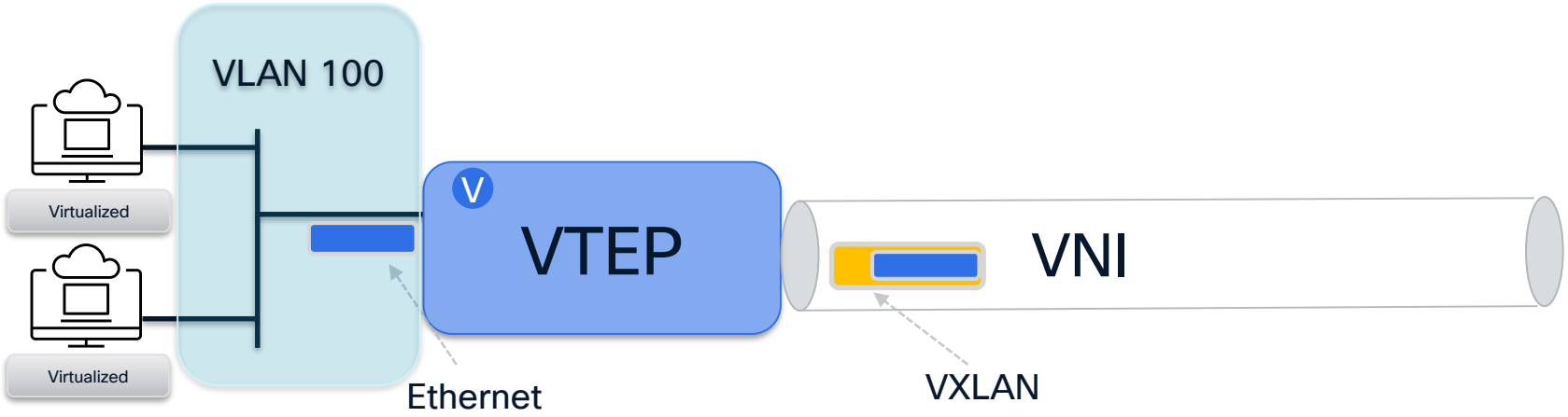
Overlay Taxonomy

Service = Virtual Network
Identifier = VN Identifier (VNI/VNID)

V VTEP: VXLAN Tunnel Endpoint
NVE: Network Virtualization Edge

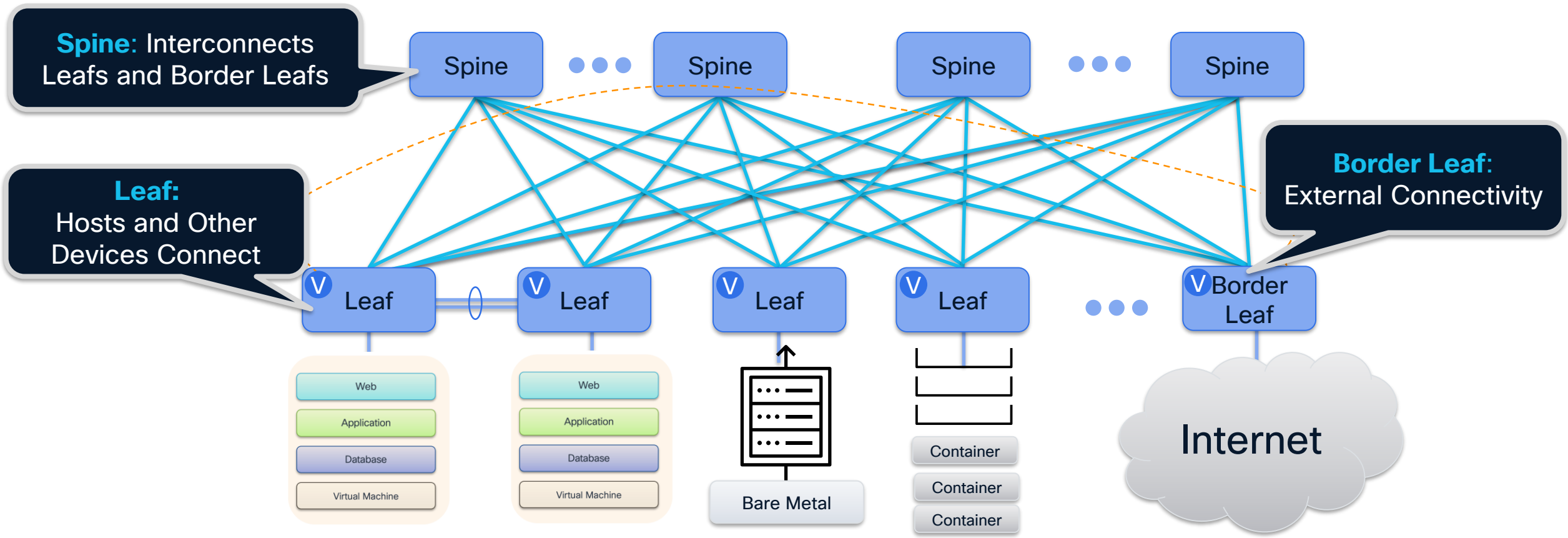


VXLAN Tunnel Endpoint (VTEP)

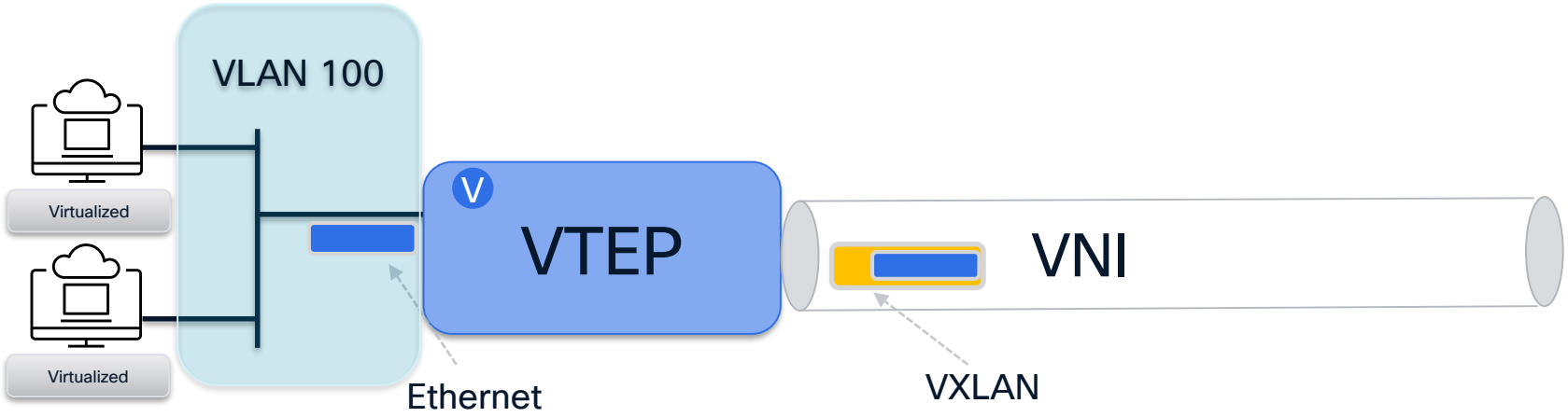


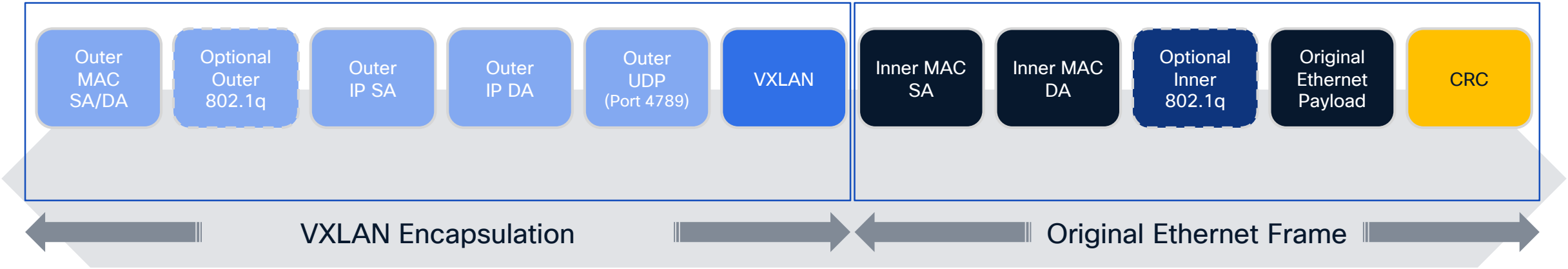
- VXLAN Tunnel Endpoint – Network Virtualization Edge
- Each VTEP is uniquely identified by an IP Address
- VTEP Discovers or learns remote VTEPs, and end hosts attached to them
- VTEP bridges when forwarding packets within the same VNI and Routes for Inter-VNI traffic

VTEP Device Roles



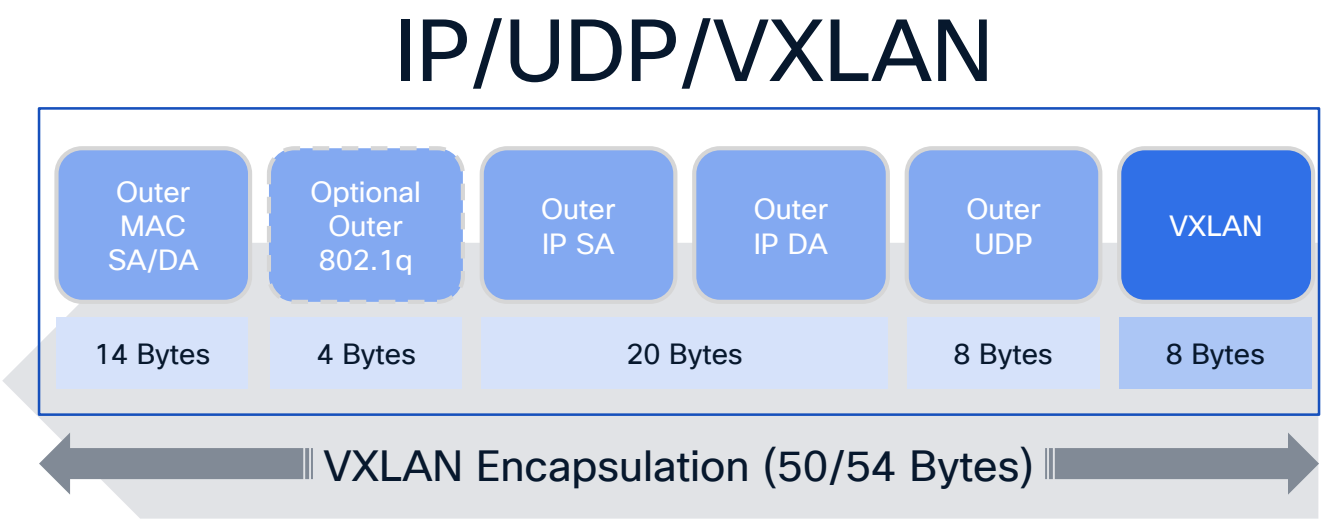
VXLAN Tunnel Endpoint (VTEP)





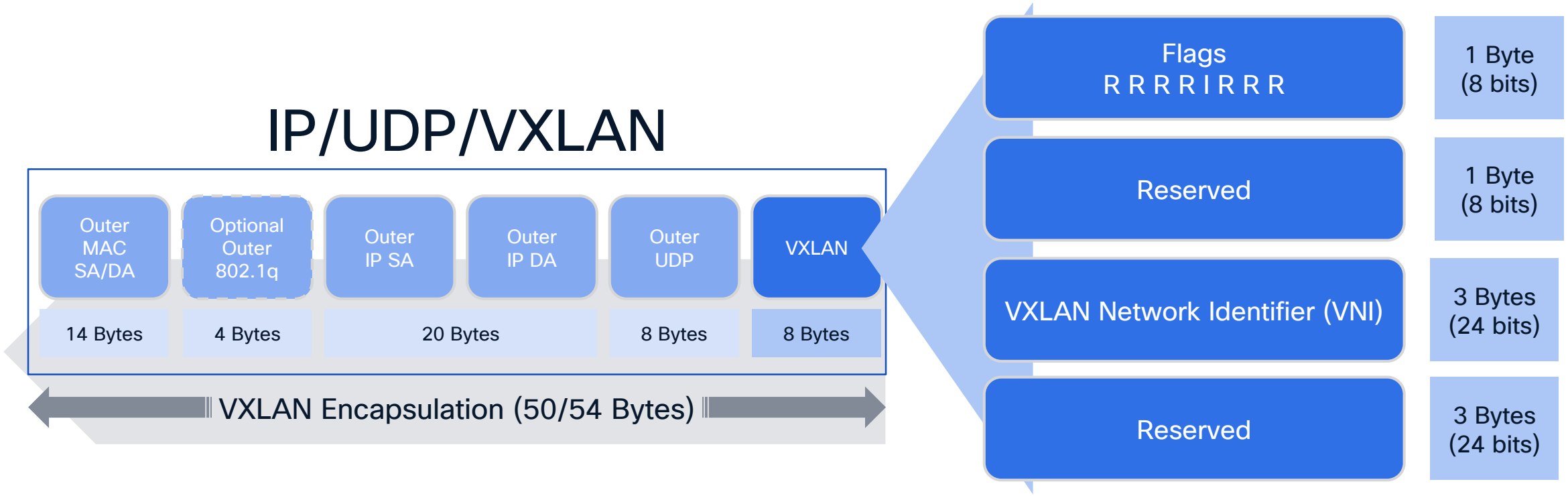
VXLAN Packet Format

- VXLAN uses MAC in UDP encapsulation

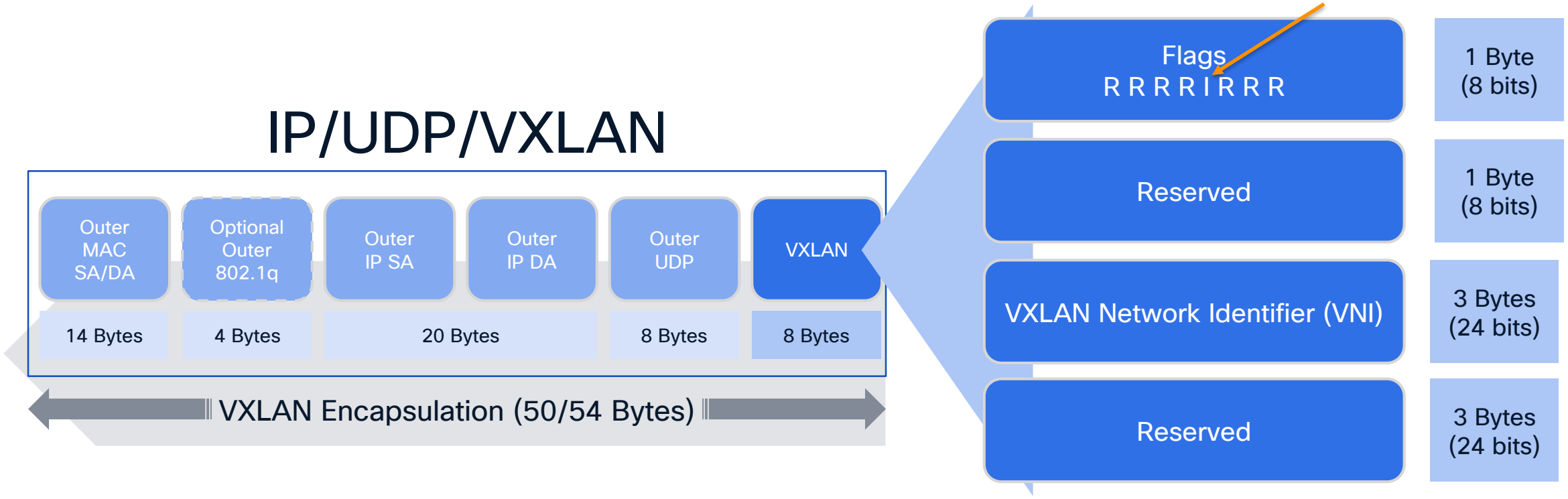


- Adds UDP and VXLAN Header before original Ethernet Frame

VXLAN Header Details

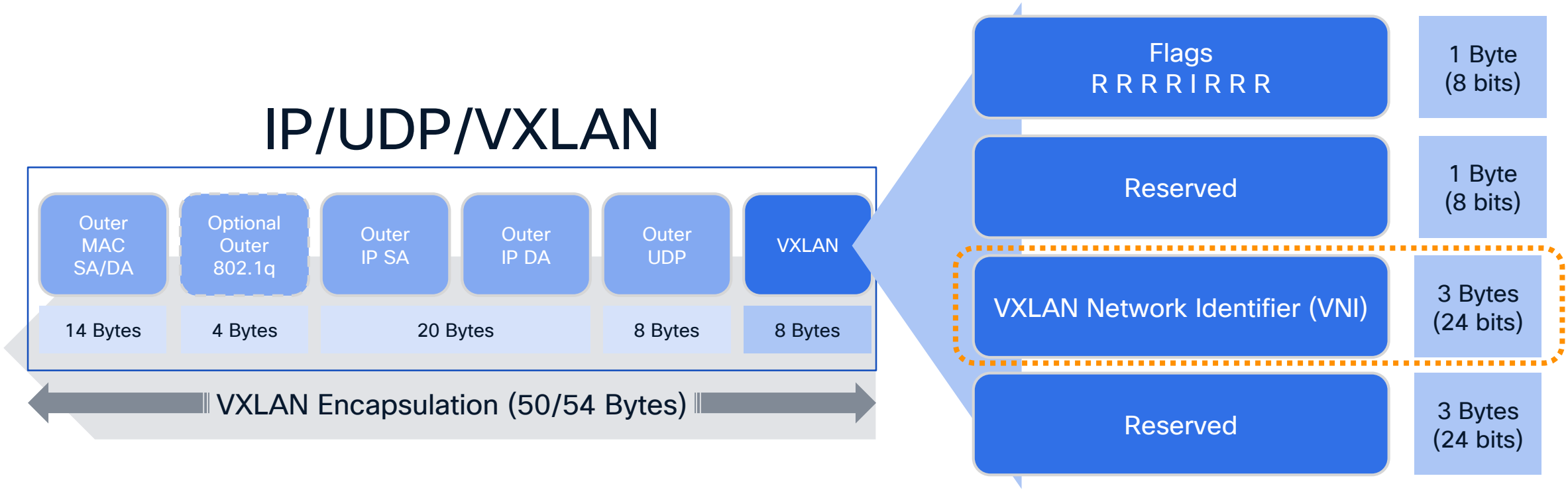


VXLAN Header Details



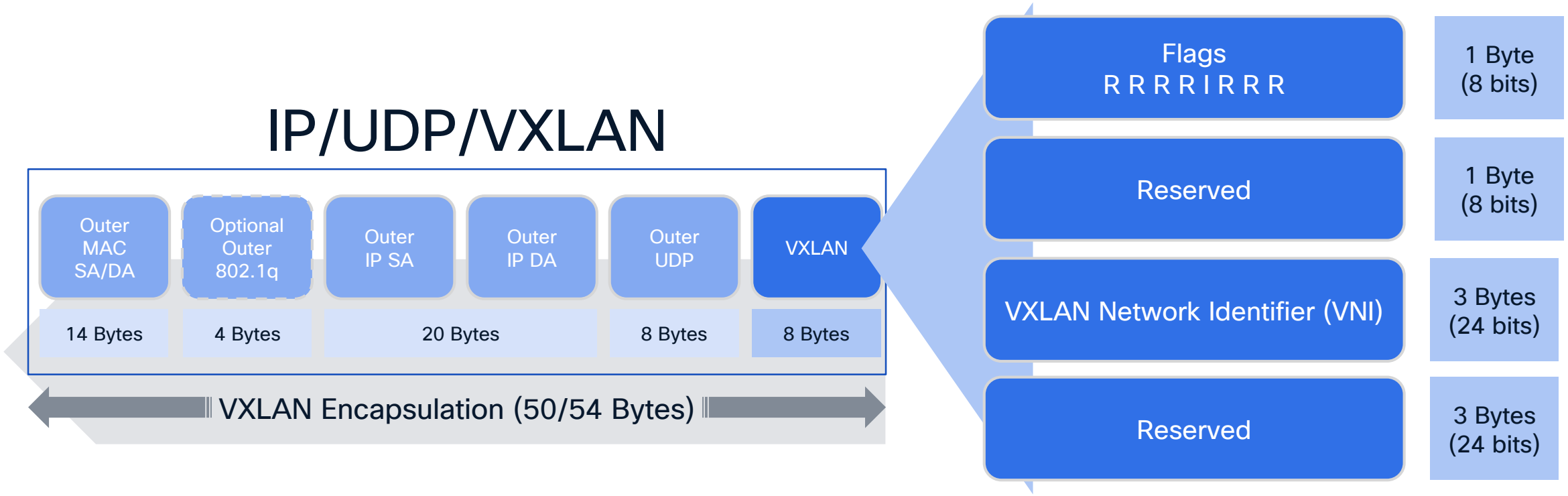
- Flags Field: I-flag (set to 1) for valid VNI. Other flags remain as R (set to 0)

VXLAN Header Details



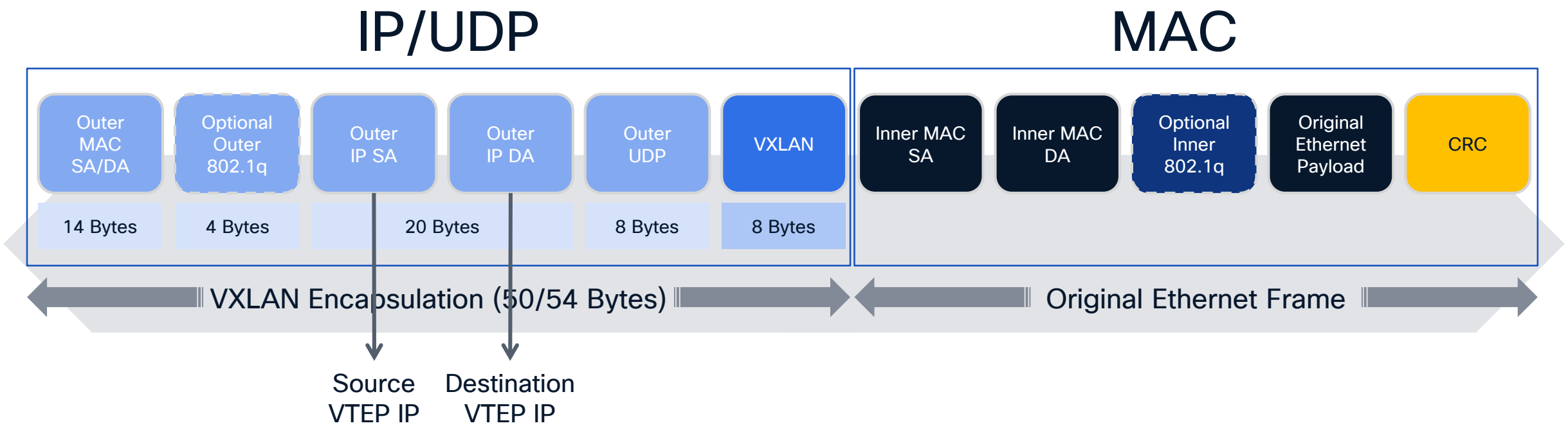
- Flags Field: I-flag (set to 1) for valid VNI. Other flags remain as R (set to 0)
- VNI Field: Allows VNI 1-16,777,215 (some implementation only 4096-16,777,215)

VXLAN Header Details



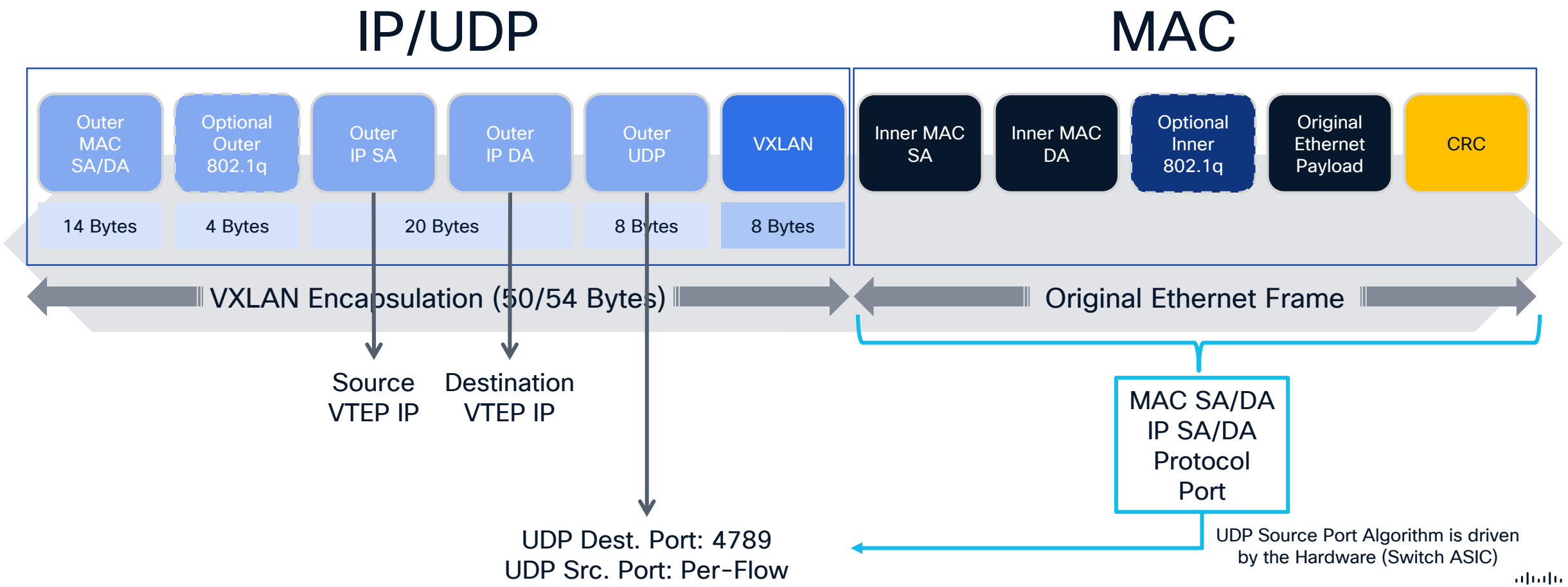

```
vlan 100
  vn-segment 50000
```

Transport Independence

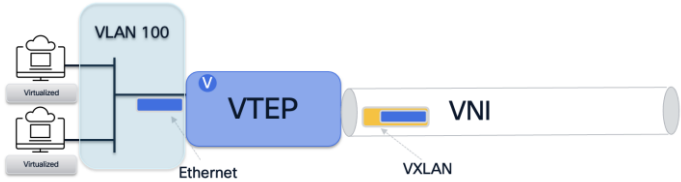


Multipath Capable

Traffic Entropy



Packet Capture - Host to VTEP




```
Frame 2: 102 bytes on wire (816 bits), 102 bytes captured (816 bits)
Ethernet II, Src: 52:0d:cf:cf:1b:08 (52:0d:cf:cf:1b:08), Dst: 20:20:00:00:00:aa (20:20:00:00:00:aa)
   Destination: 20:20:00:00:00:aa (20:20:00:00:00:aa)
   Source: 52:0d:cf:cf:1b:08 (52:0d:cf:cf:1b:08)
   Type: 802.1Q Virtual LAN (0x8100)
   [Stream index: 1]
802.1Q Virtual LAN, PRI: 0, DEI: 0, ID: 100
   000. .... .... .... = Priority: Best Effort (default) (0)
   ...0 .... .... .... = DEI: Ineligible
   .... 0000 0110 0100 = ID: 100
   Type: IPv4 (0x0800)
Internet Protocol Version 4, Src: 100.100.100.2, Dst: 200.200.200.2
   0100 .... = Version: 4
   .... 0101 = Header Length: 20 bytes (5)
   Differentiated Services Field: 0x00 (DSCP: CS0, ECN: Not-ECT)
   Total Length: 84
   Identification: 0xbd56 (48470)
   000. .... = Flags: 0x0
   ...0 0000 0000 0000 = Fragment Offset: 0
   Time to Live: 255
   Protocol: ICMP (1)
   Header Checksum: 0xa520 [validation disabled]
   [Header checksum status: Unverified]
   Source Address: 100.100.100.2
   Destination Address: 200.200.200.2
   [Stream index: 0]
Internet Control Message Protocol
   Type: 8 (Echo (ping) request)
   Code: 0
   Checksum: 0x15f4 [correct]
   [Checksum Status: Good]
   Identifier (BE): 13605 (0x3525)
   Identifier (LE): 9525 (0x2535)
   Sequence Number (BE): 52225 (0xcc01)
   Sequence Number (LE): 460 (0x01cc)
   [Response frame: 3]
   Timestamp from icmp data: Nov  7, 2024 08:14:13.468214000 IST
   [Timestamp from icmp data (relative): 58158.750333000 seconds]
   Data (40 bytes)
```

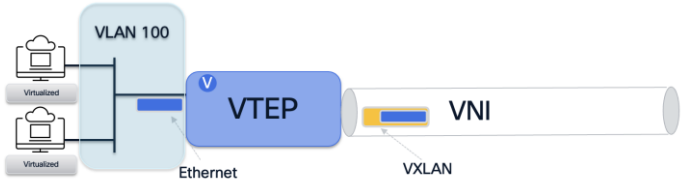
VLAN 100

Ethernet Header

IP Header

Payload

Packet Capture – VTEP to Spine




```
> Frame 2: 148 bytes on wire (1184 bits), 148 bytes captured (1184 bits)
> Ethernet II, Src: 52:0d:18:e0:1b:08 (52:0d:18:e0:1b:08), Dst: 52:03:7e:03:1b:08 (52:03:7e:03:1b:08)
v Internet Protocol Version 4, Src: 10.3.0.1, Dst: 10.3.0.2
    0100 .... = Version: 4
    .... 0101 = Header Length: 20 bytes (5)
  > Differentiated Services Field: 0x00 (DSCP: CS0, ECN: Not-ECT)
    Total Length: 134
    Identification: 0x0000 (0)
  > 000. .... = Flags: 0x0
    ...0 0000 0000 0000 = Fragment Offset: 0
    Time to Live: 255
    Protocol: UDP (17)
    Header Checksum: 0xa75e [validation disabled]
    [Header checksum status: Unverified]
    Source Address: 10.3.0.1
    Destination Address: 10.3.0.2
    [Stream index: 0]
> User Datagram Protocol, Src Port: 53936, Dst Port: 4789
v Virtual eXtensible Local Area Network
  > Flags: 0x0800, VXLAN Network ID (VNI)
    Group Policy ID: 0
    VXLAN Network Identifier (VNI): 50000
    Reserved: 0
v Ethernet II, Src: 52:0d:18:e0:1b:08 (52:0d:18:e0:1b:08), Dst: 52:11:5f:78:1b:08 (52:11:5f:78:1b:08)
  > Destination: 52:11:5f:78:1b:08 (52:11:5f:78:1b:08)
  > Source: 52:0d:18:e0:1b:08 (52:0d:18:e0:1b:08)
    Type: IPv4 (0x0800)
    [Stream index: 1]
v Internet Protocol Version 4, Src: 100.100.100.2, Dst: 200.200.200.2
    0100 .... = Version: 4
    .... 0101 = Header Length: 20 bytes (5)
  > Differentiated Services Field: 0x00 (DSCP: CS0, ECN: Not-ECT)
    Total Length: 84
    Identification: 0xc086 (49286)
  > 000. .... = Flags: 0x0
    ...0 0000 0000 0000 = Fragment Offset: 0
    Time to Live: 254
    Protocol: ICMP (1)
    Header Checksum: 0xa2f0 [validation disabled]
    [Header checksum status: Unverified]
    Source Address: 100.100.100.2
    Destination Address: 200.200.200.2
    [Stream index: 1]
> Internet Control Message Protocol
```

IP Header

VTEP IP Addresses

UDP Header

VNI 50000

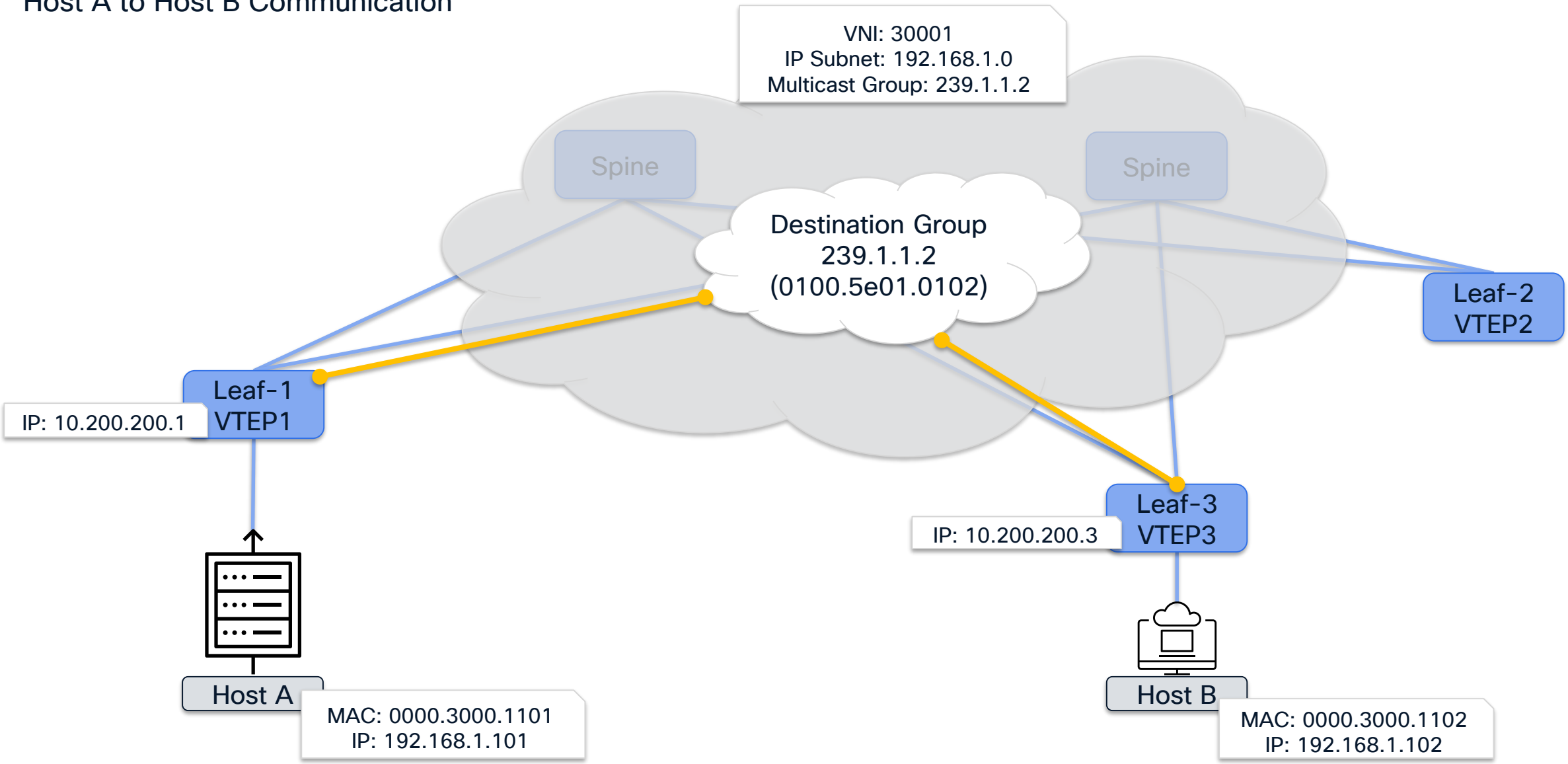
VXLAN Encapsulation

Payload

Data Plane
Flood and Learn

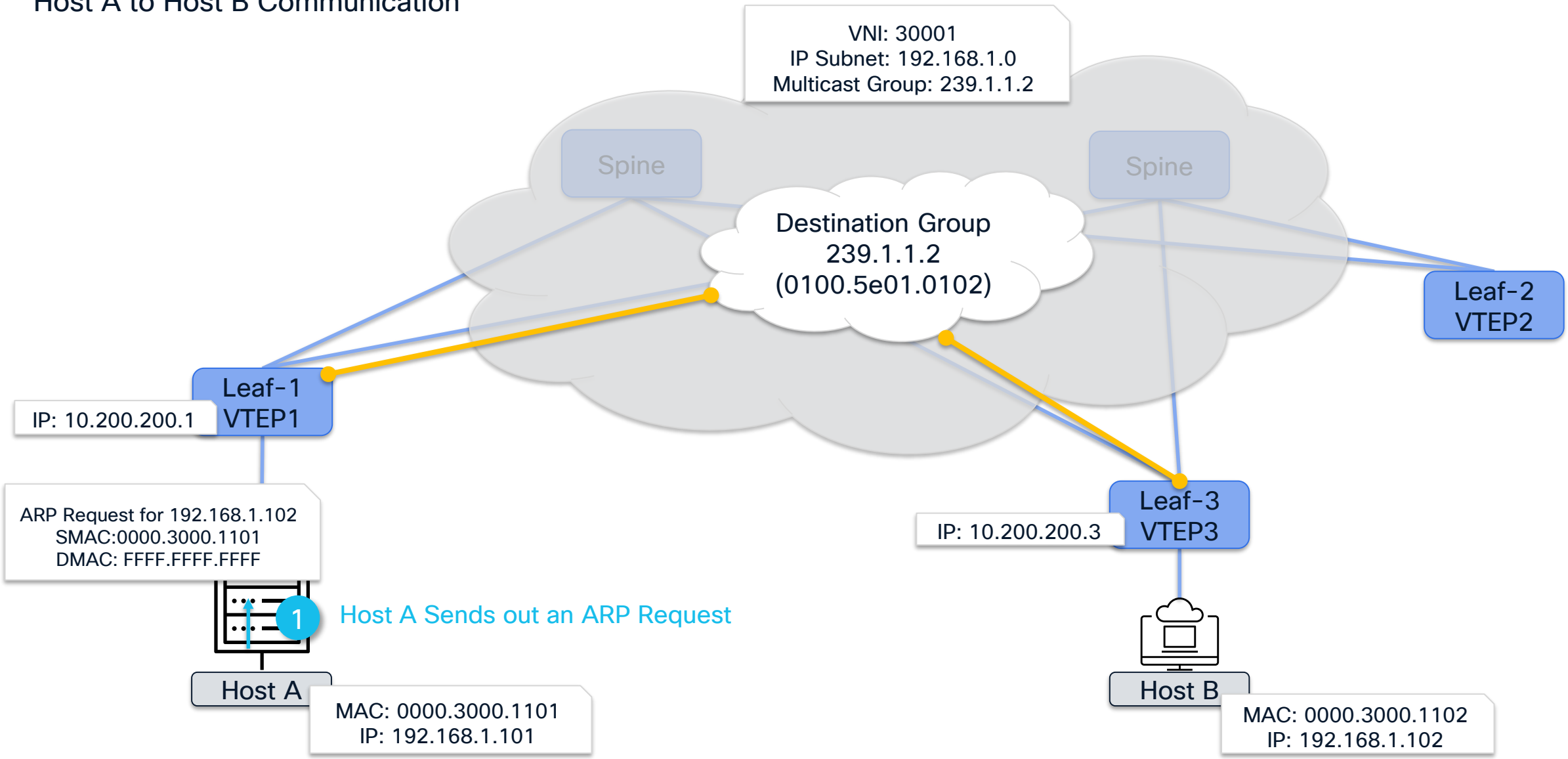
VXLAN Flood and Learn

Host A to Host B Communication



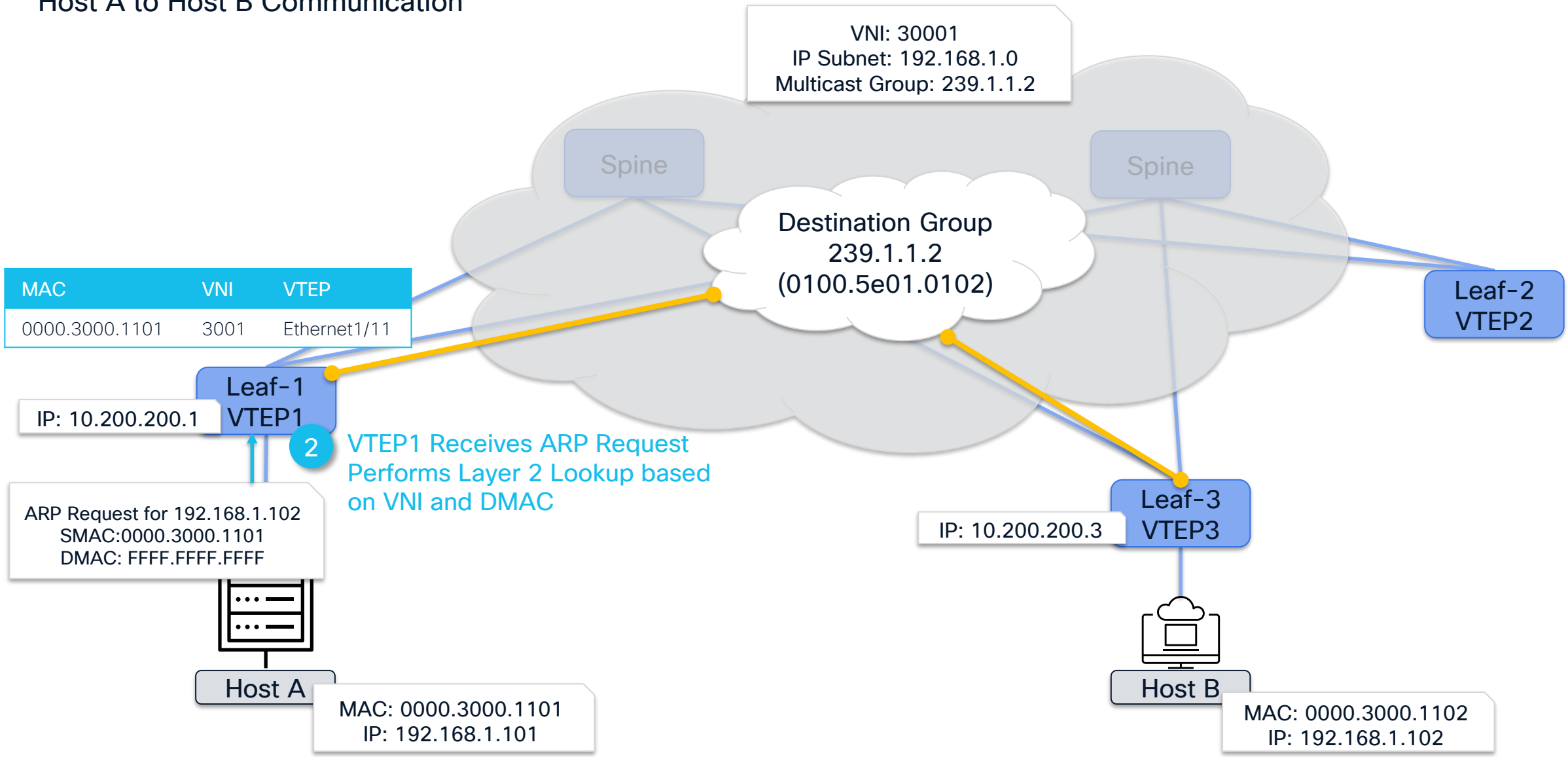
VXLAN Flood and Learn

Host A to Host B Communication



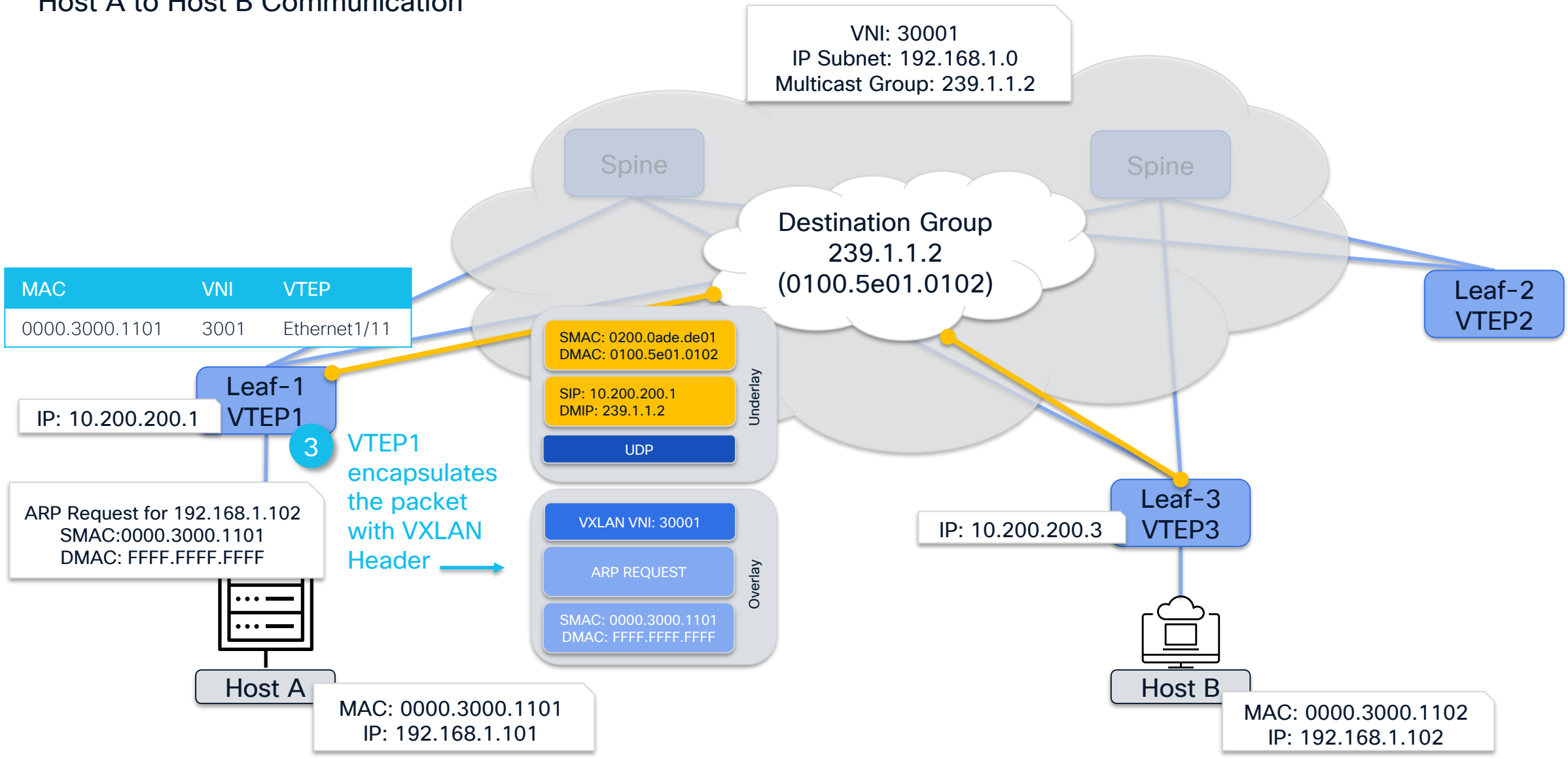
VXLAN Flood and Learn

Host A to Host B Communication



VXLAN Flood and Learn

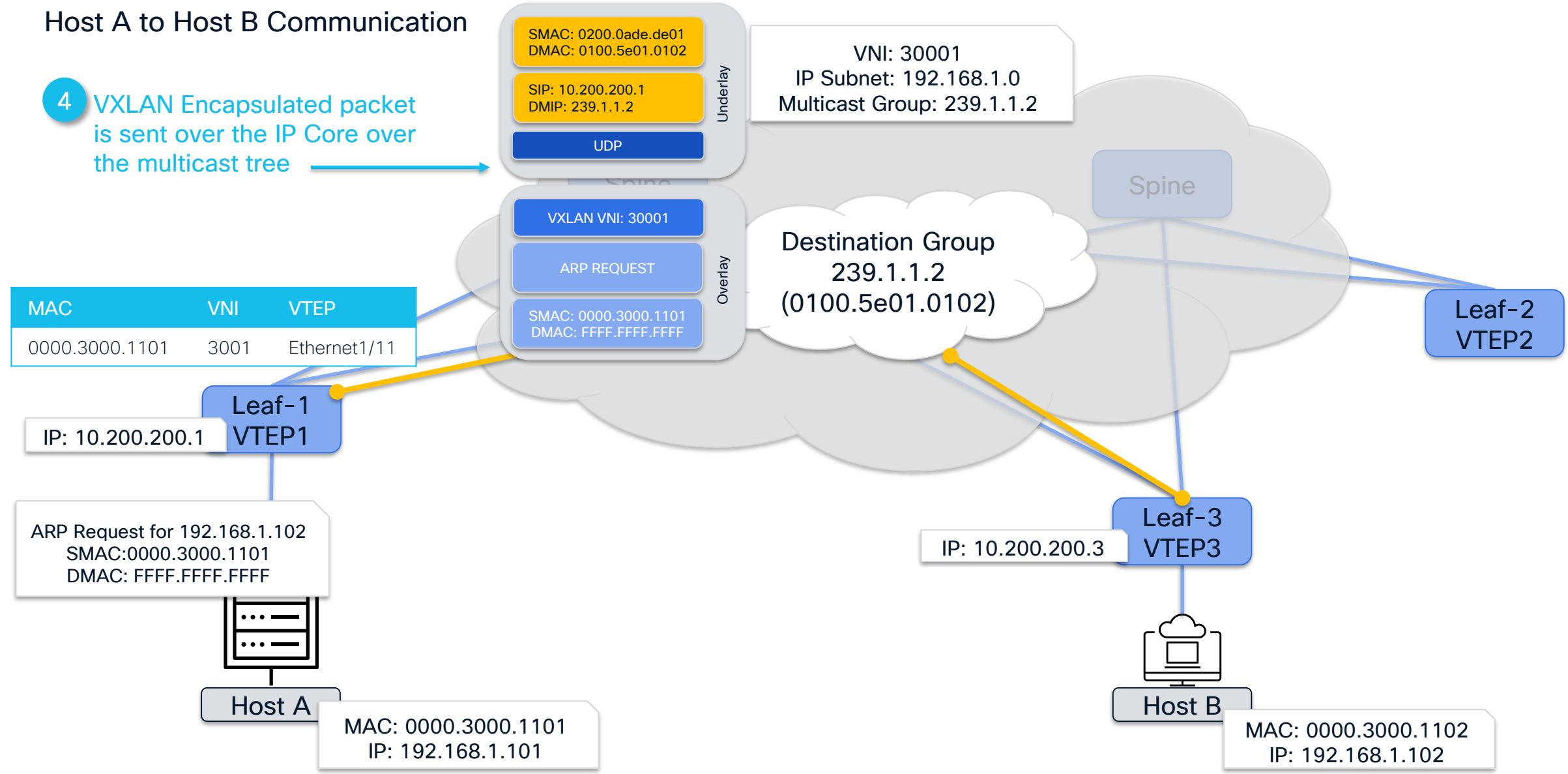
Host A to Host B Communication



VXLAN Flood and Learn

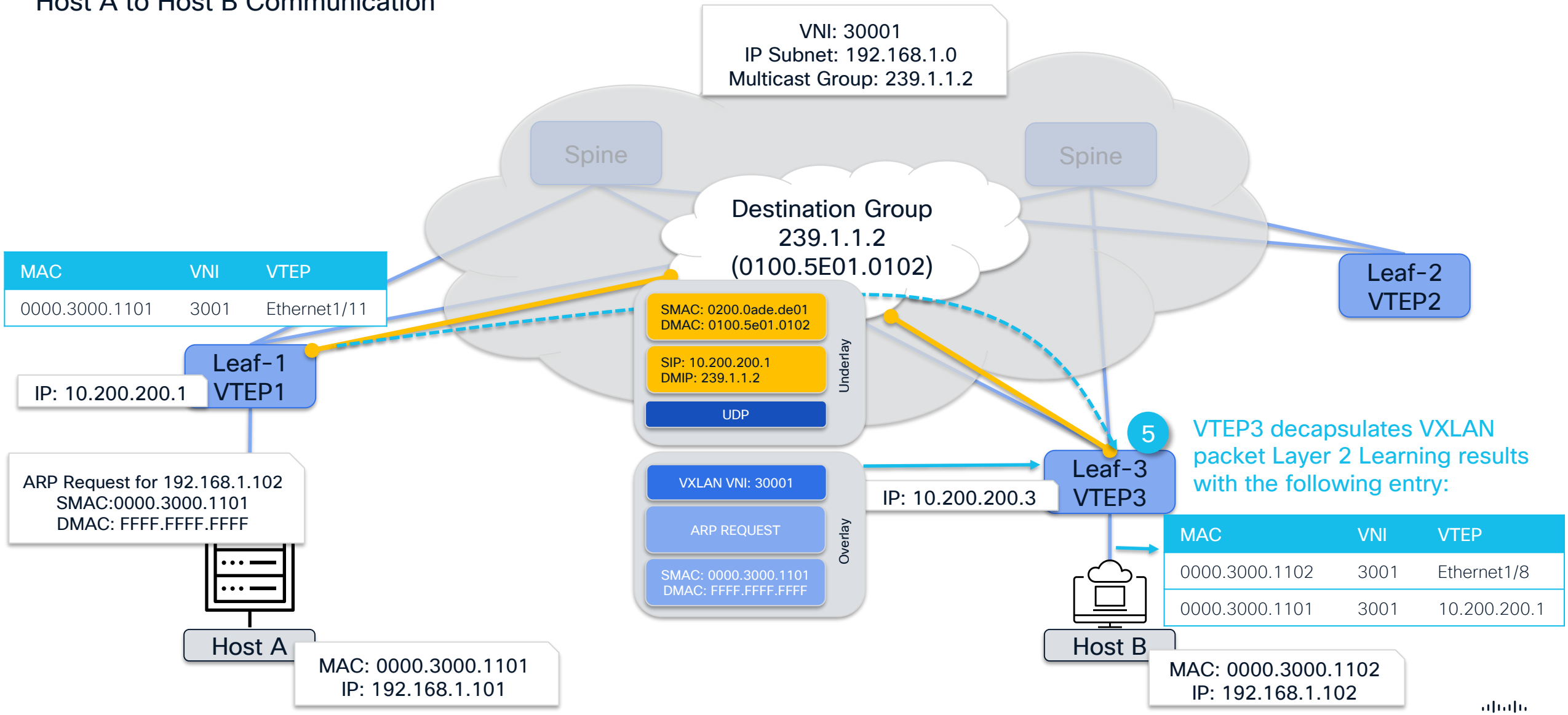
Host A to Host B Communication

4 VXLAN Encapsulated packet is sent over the IP Core over the multicast tree



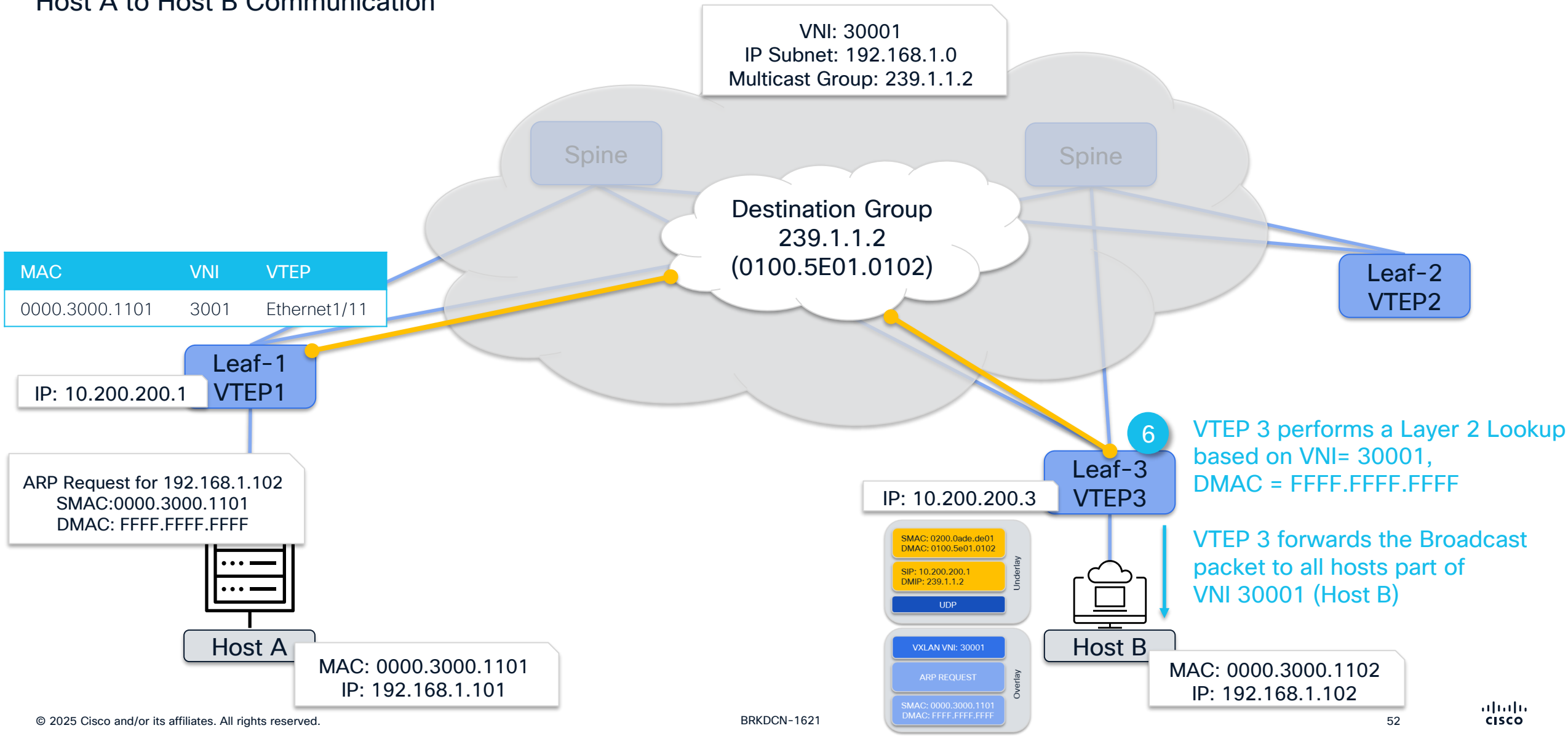
VXLAN Flood and Learn

Host A to Host B Communication



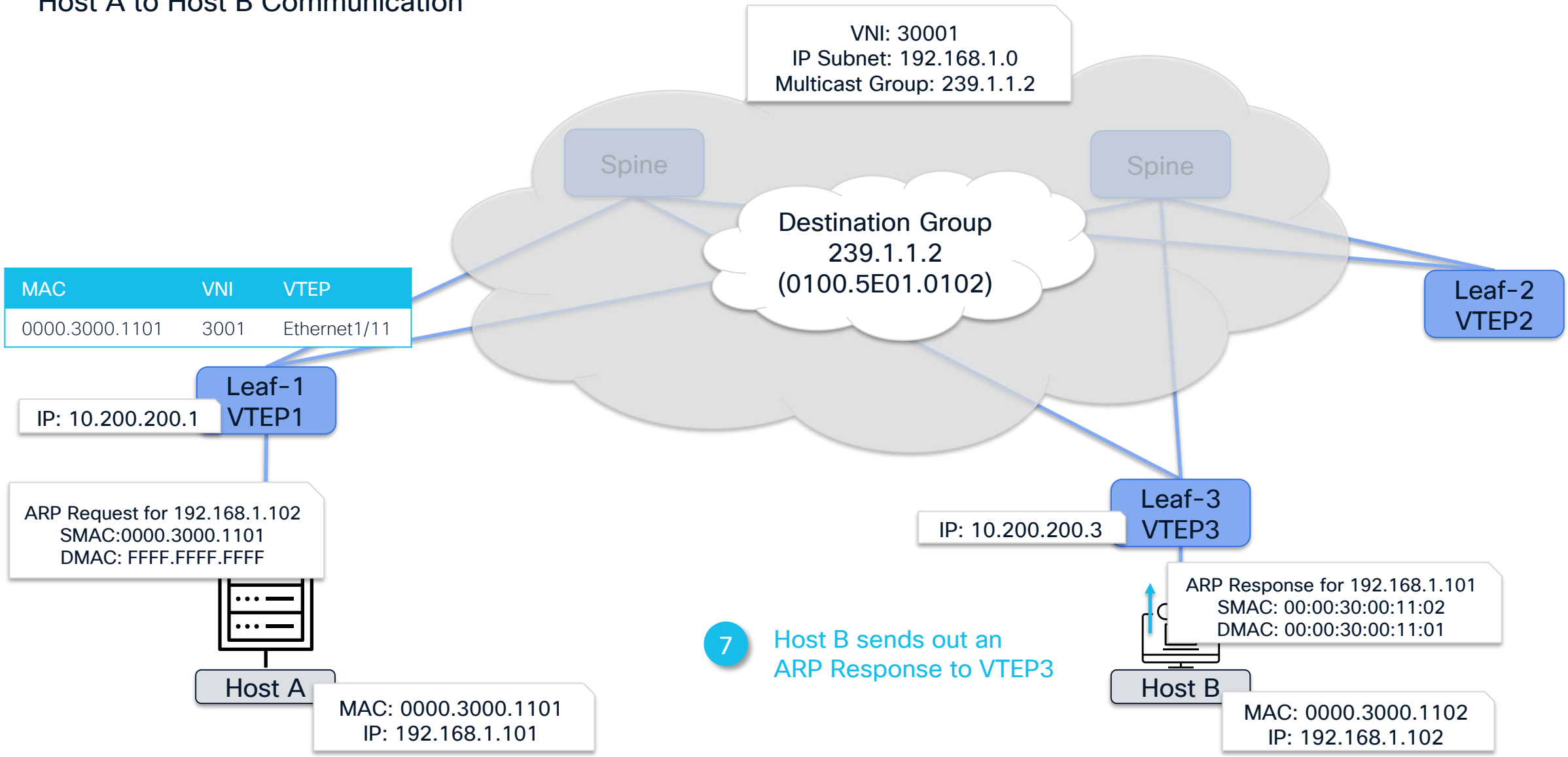
VXLAN Flood and Learn

Host A to Host B Communication



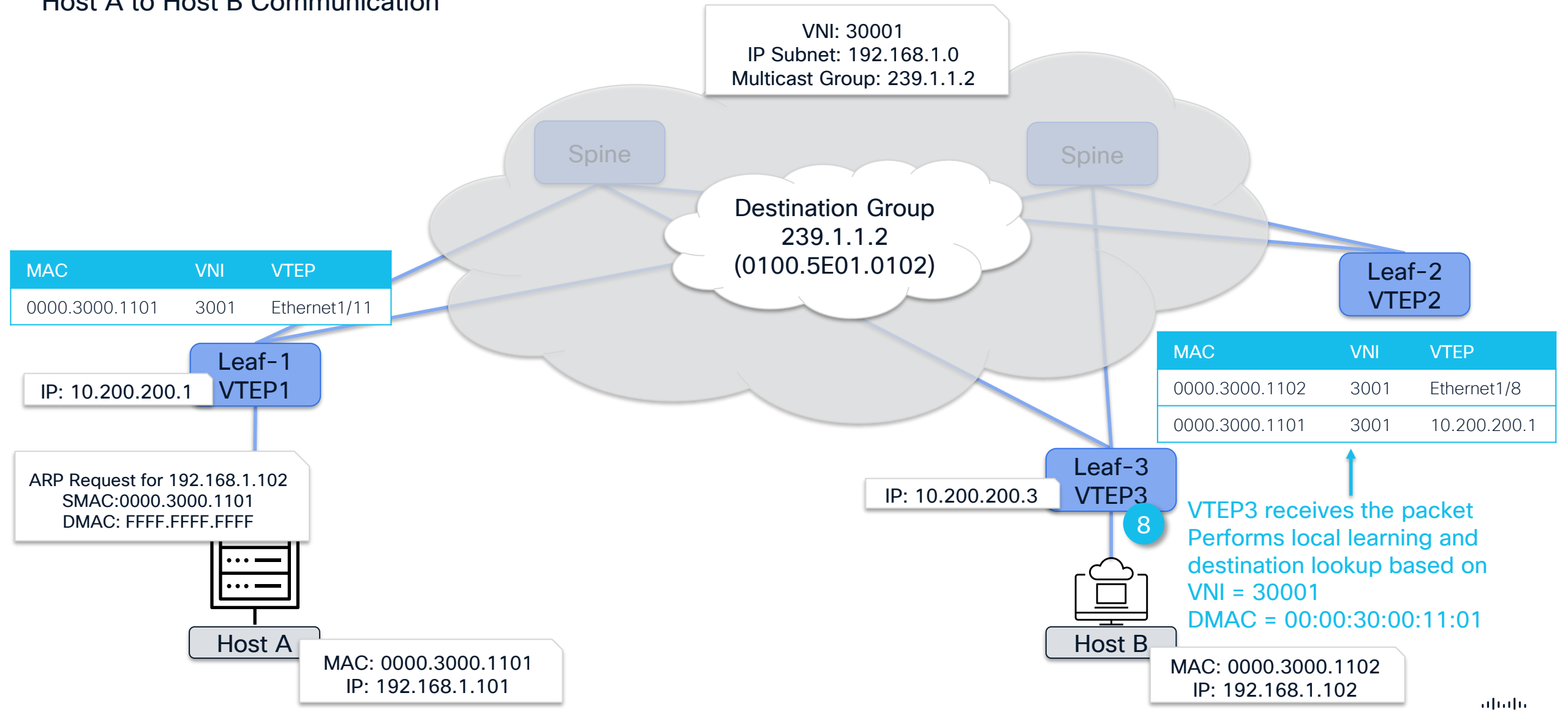
VXLAN Flood and Learn

Host A to Host B Communication



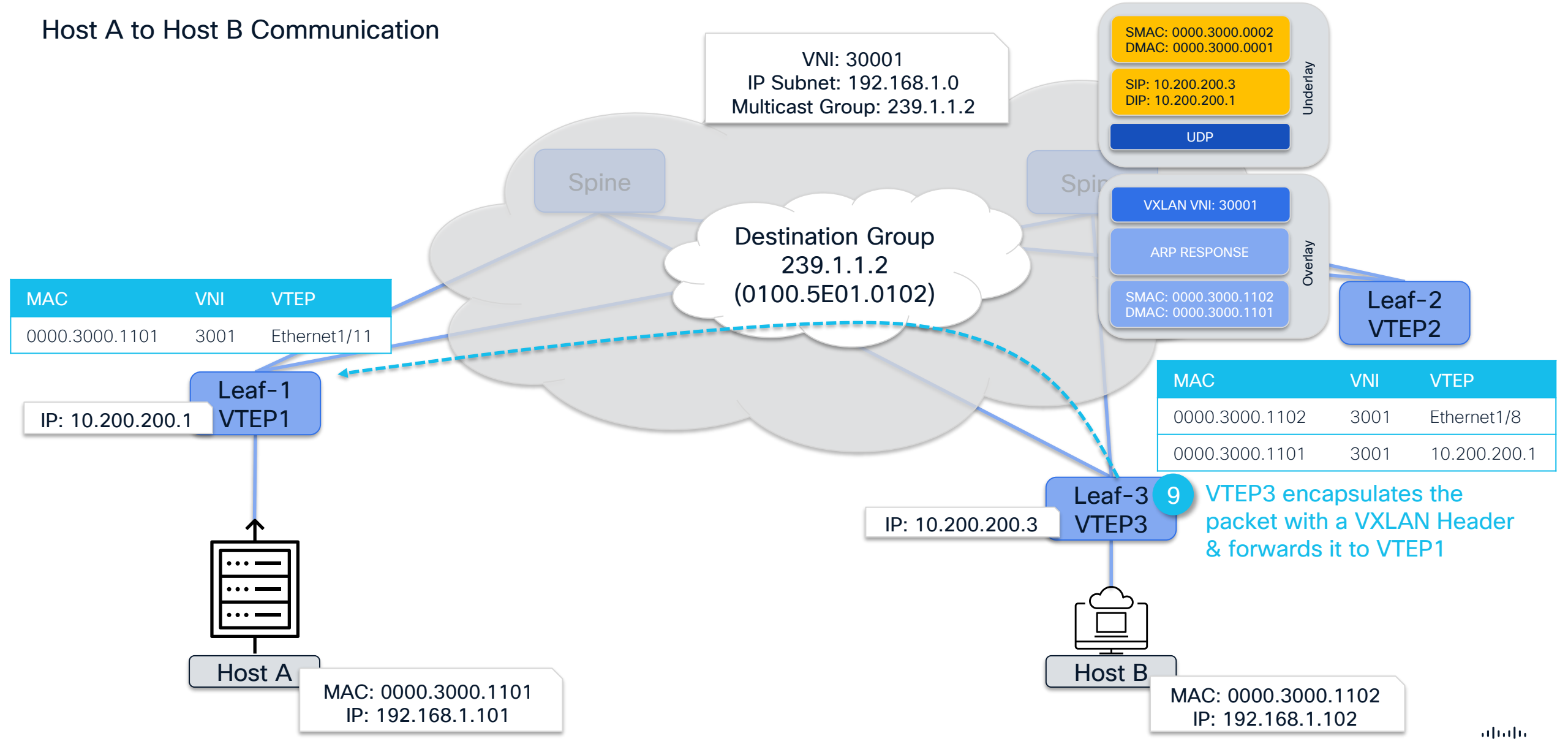
VXLAN Flood and Learn

Host A to Host B Communication



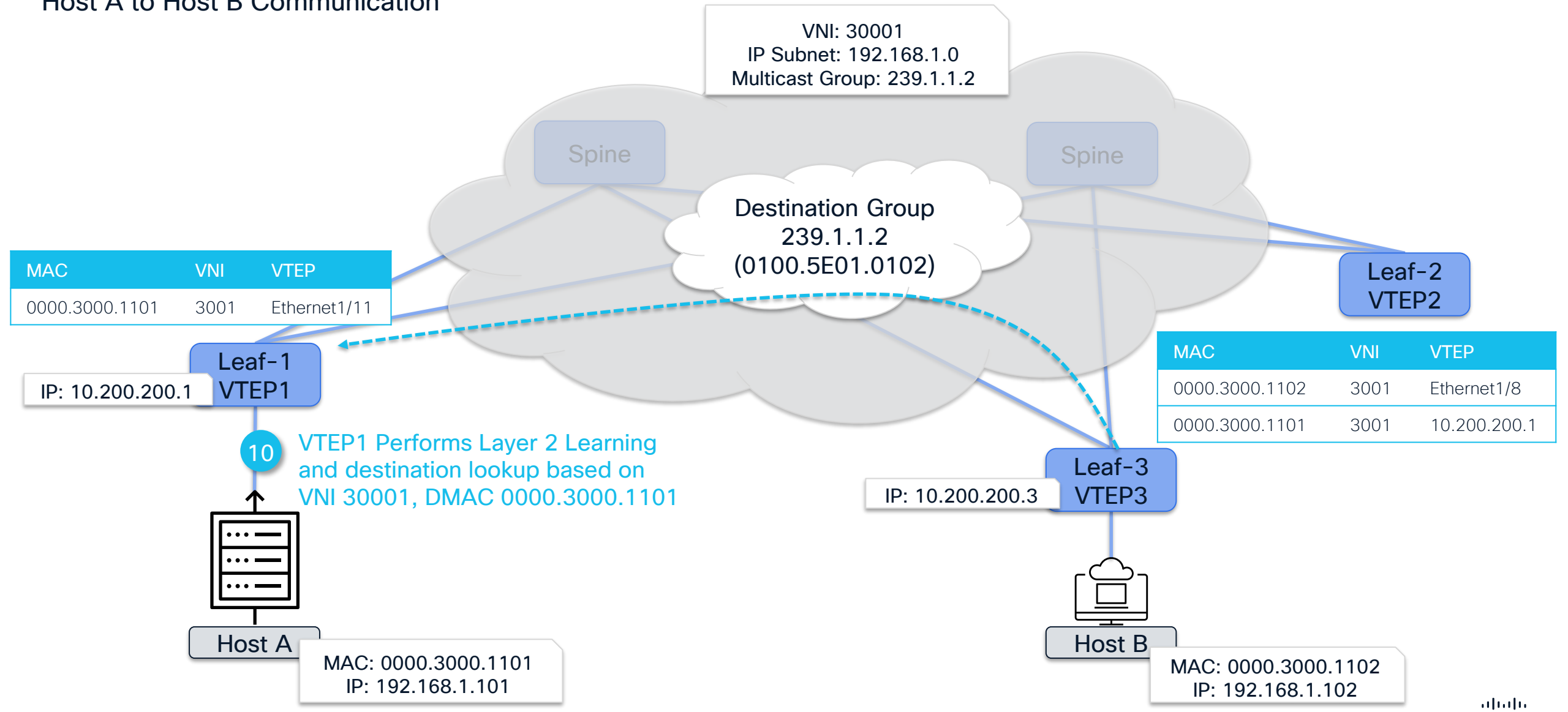
VXLAN Flood and Learn

Host A to Host B Communication



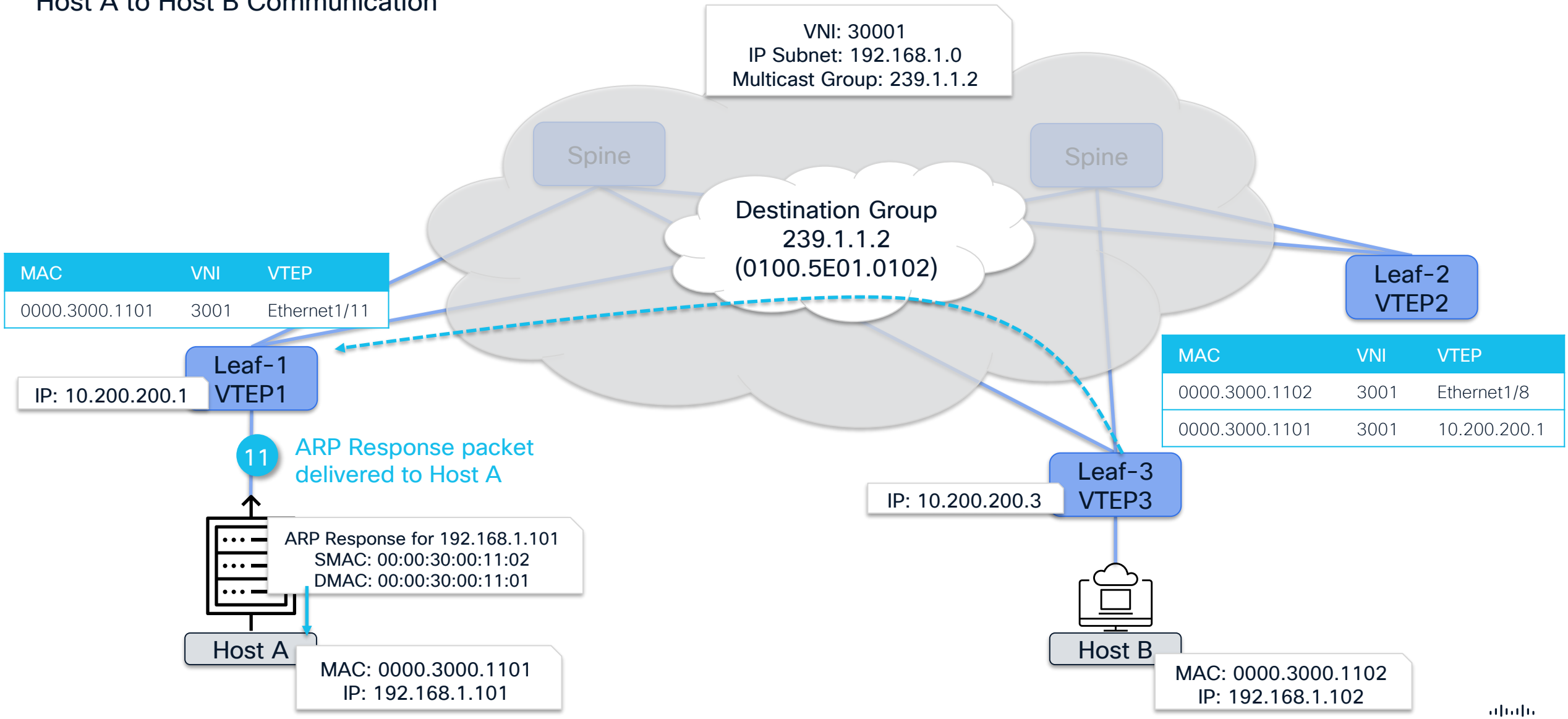
VXLAN Flood and Learn

Host A to Host B Communication



VXLAN Flood and Learn

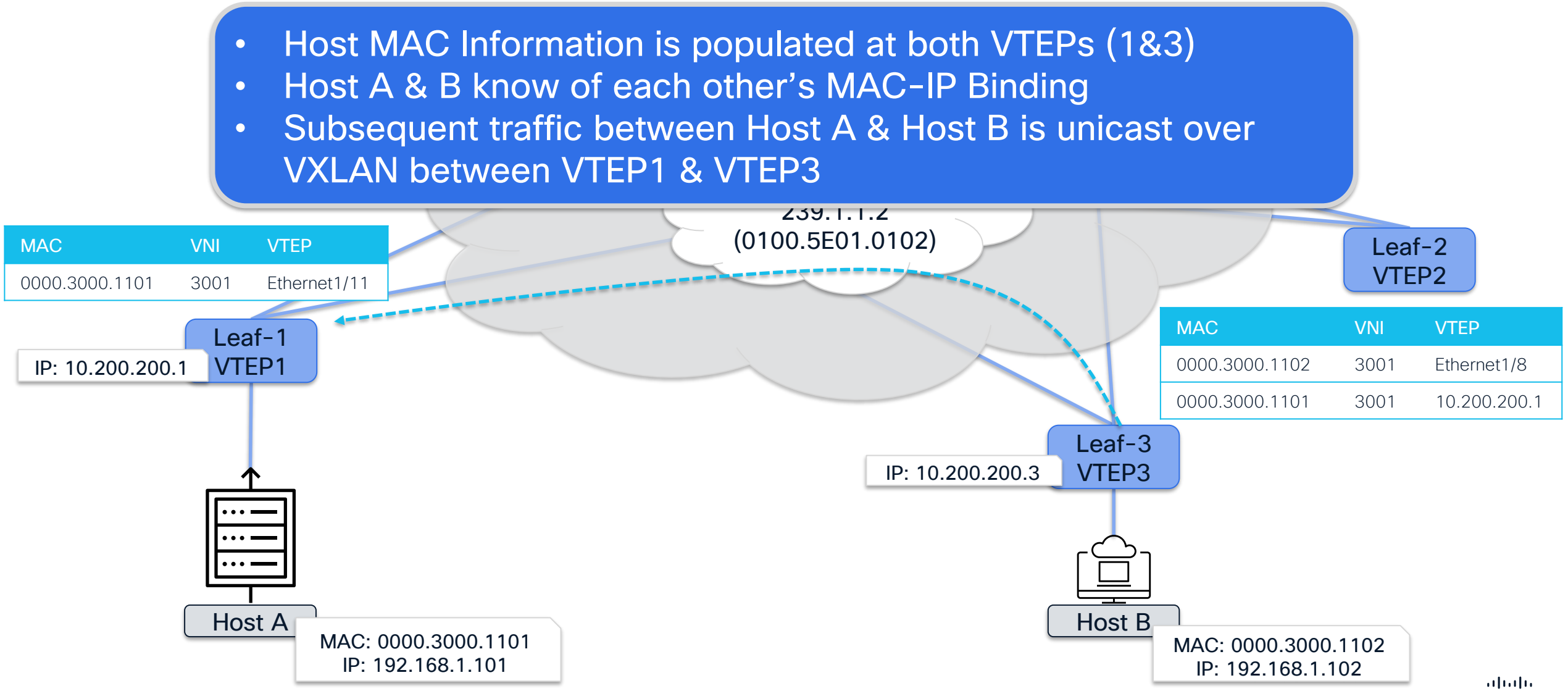
Host A to Host B Communication



VXLAN Flood and Learn

Host A to Host B Communication

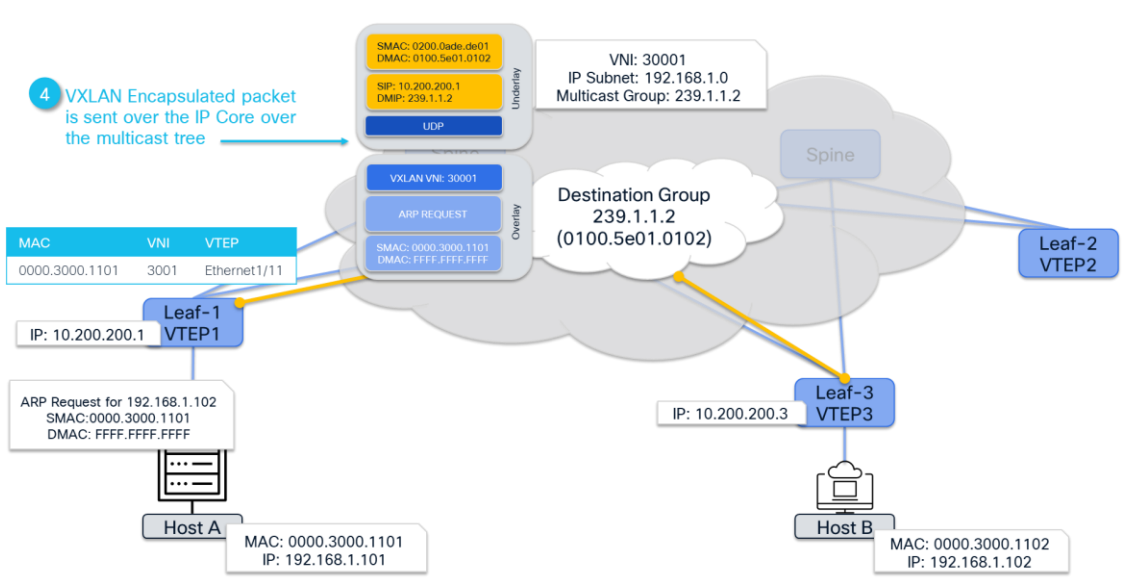
- Host MAC Information is populated at both VTEPs (1&3)
- Host A & B know of each other's MAC-IP Binding
- Subsequent traffic between Host A & Host B is unicast over VXLAN between VTEP1 & VTEP3

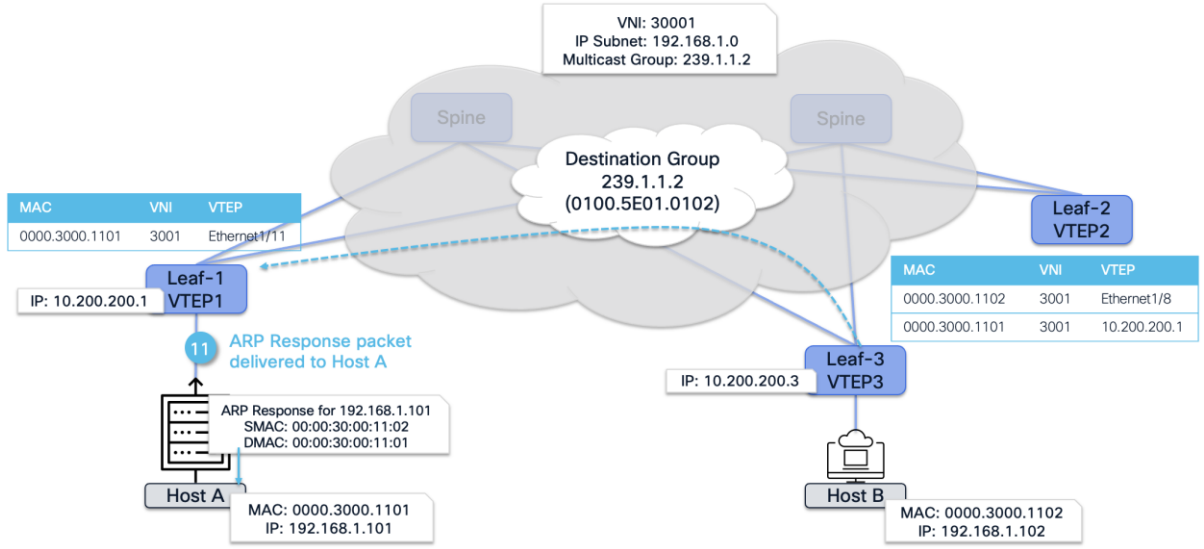


Flood & Learn Mechanism

Broadcast, Unknown Unicast, Multicast (BUM)Traffic

Multidestination Traffic is “Flooded” over the VXLAN Tunnel between VTEPs to “Learn” about the Host MAC addresses located behind the VTEPs so subsequent communication is delivered via Unicast

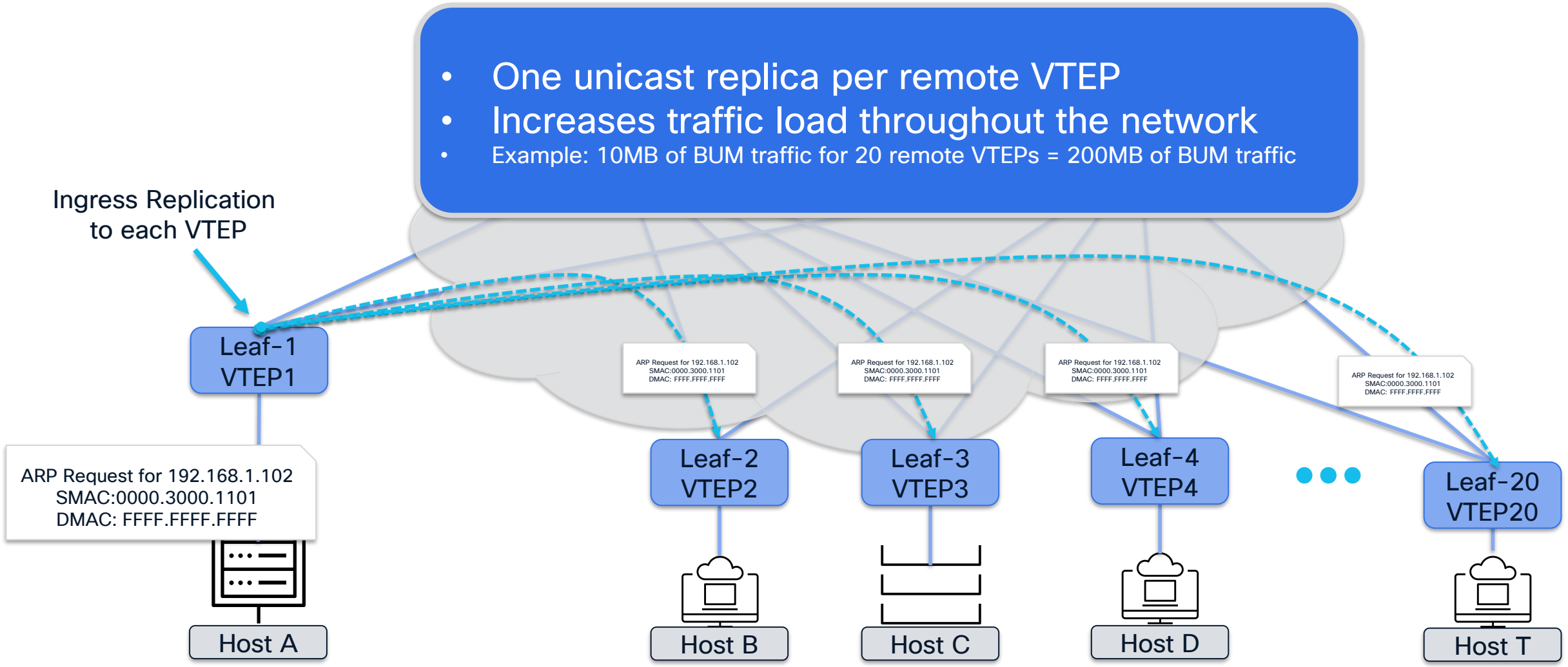




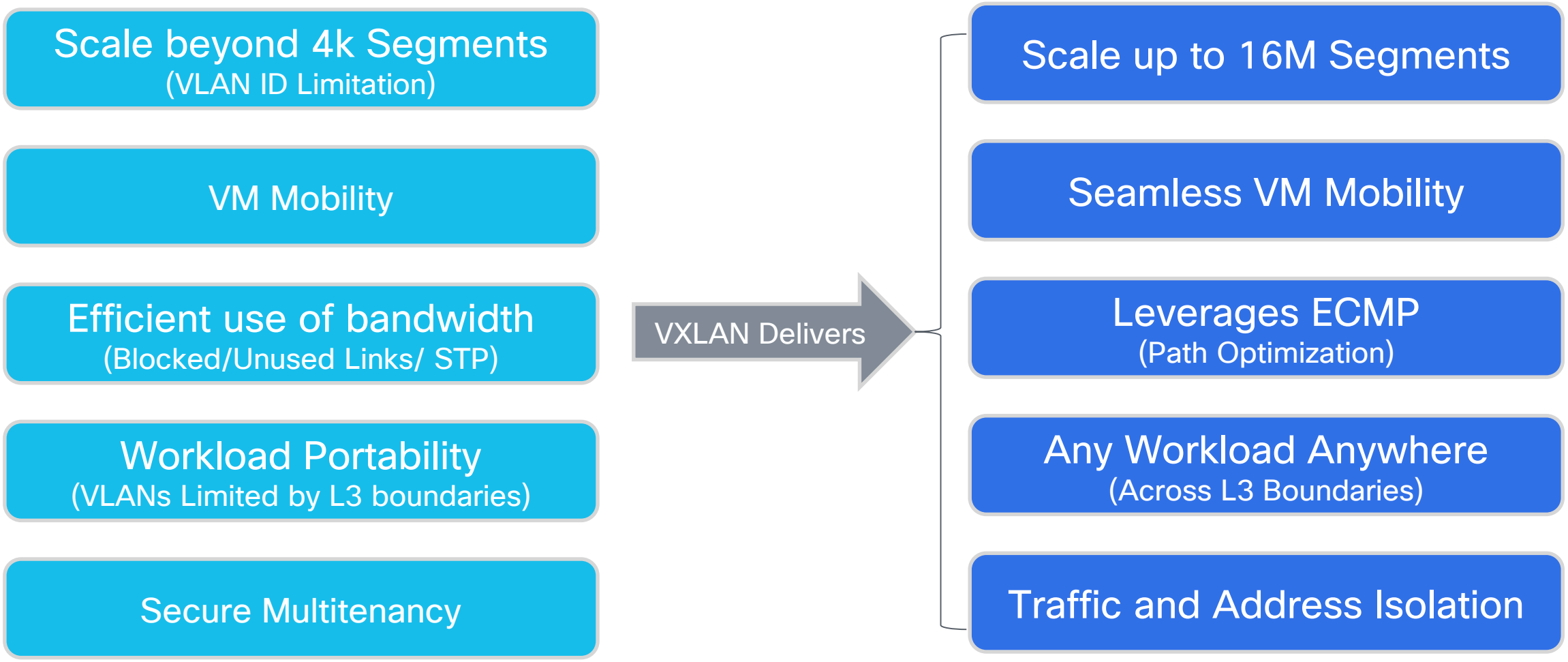
See hidden slides for Packet Walk

Ingress Replication or Head-End Replication

Flood & Learn Mechanism



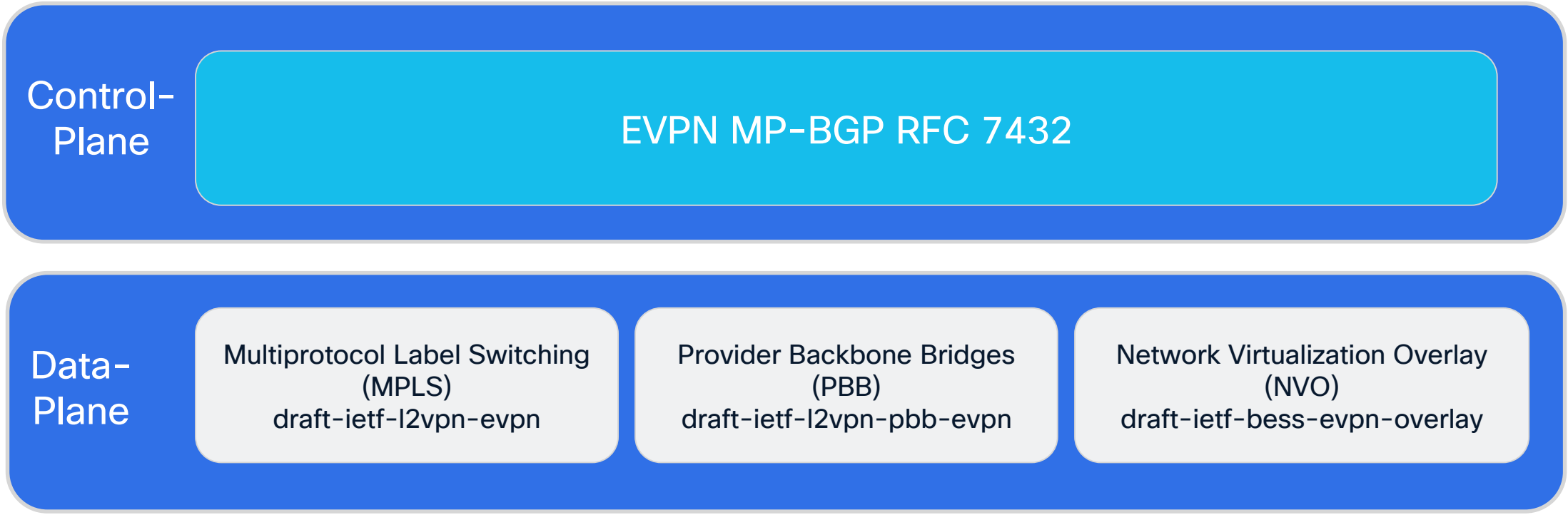
Why VXLAN? How did we get here?



Control Plane
EVPN

EVPN in the Data Center

IETF RFC/Draft for Control and Data Plane



- EVPN over NVO Tunnels (i.e., VXLAN) for Data Center Fabric Encapsulations
- Provides Layer-2 and Layer-3 Overlays over IP Networks

BGP EVPN Route Types - RFC 7432

For your reference 

RFC/Draft	Route Type	Description
RFC 7432	1	Ethernet Auto-Discovery EAD Route
	2	MAC/IP Advertisement Route
	3	Inclusive Multicast Ethernet Tag Route
	4	Ethernet Segment Route
draft-ietf-bess-evpn-prefix-advertisement	5	IP Prefix Route

BGP EVPN *Route Types - RFC 7432

For your reference

RFC/Draft	Route Type	Description	
RFC 7432	1	Ethernet Auto-Discovery EAD Route	BGP Based Multi-Homing Mass Withdrawal/Aliasing
	2	MAC/IP Advertisement Route	L2 VNI MAC or MAC-IP from L2 MAC Learning or ARP
	3	Inclusive Multicast Ethernet Tag Route	Dynamic Peer Discovery for EVPN Ingress Replication
	4	Ethernet Segment Route	BGP Based Multi-Homing BUM DF Election/Split-Horizon
draft-ietf-bess-evpn-prefix-advertisement	5	IP Prefix Route	IETF Draft, Advertise IP Prefixes

*Type 6 - Selective Multicast Ethernet Tag Route
*Type 7/8 - IGMP Sync Routes

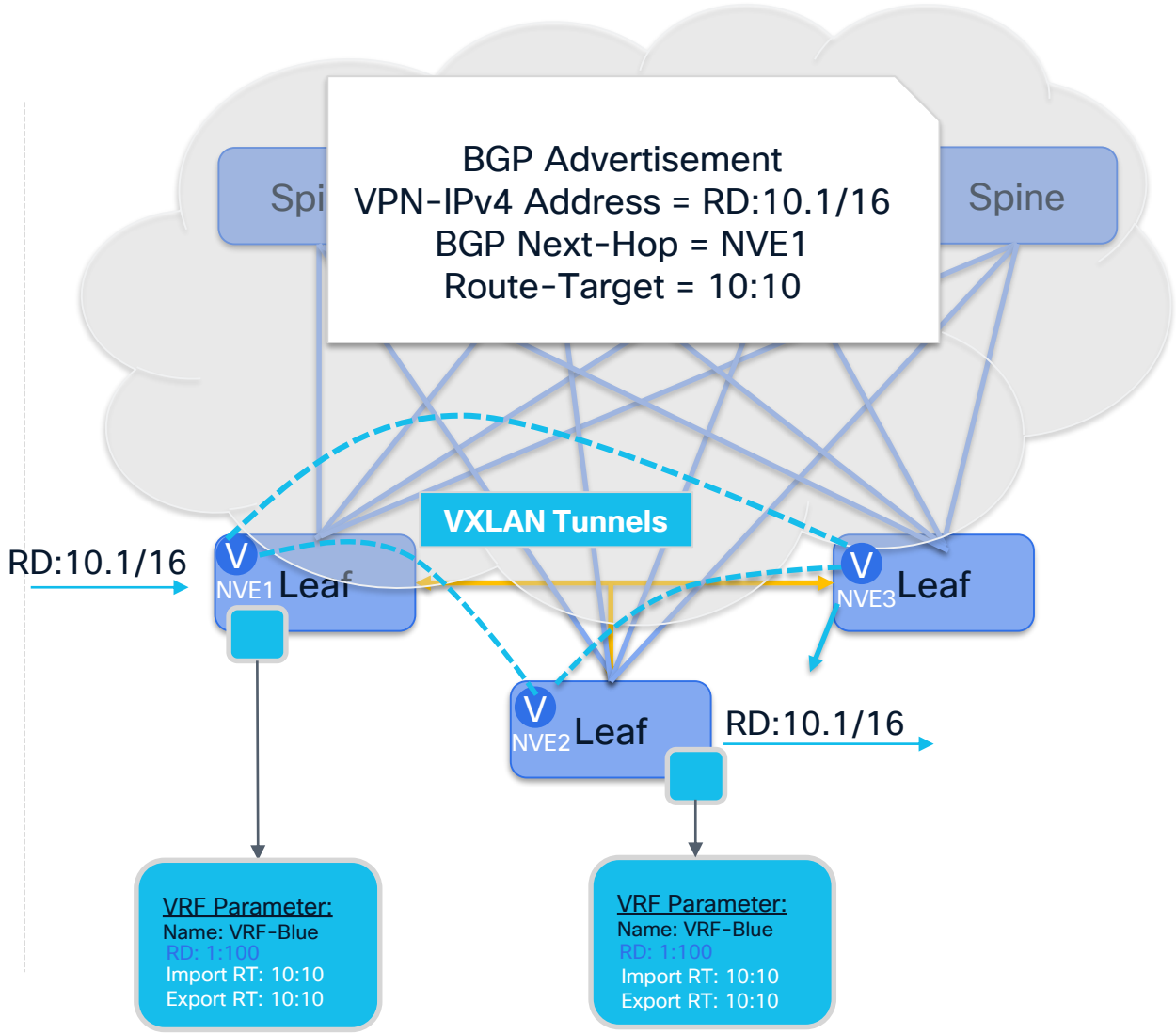
MP-BGP VPN Terminology

VPN Address Family
A Multi-Protocol BGP Extension to Distribute VPN Routes

Virtual Routing & Forwarding (VRF)
Overlay MAC or IP Routes are stored in dedicated VRF tables (MAC or IP VRF)

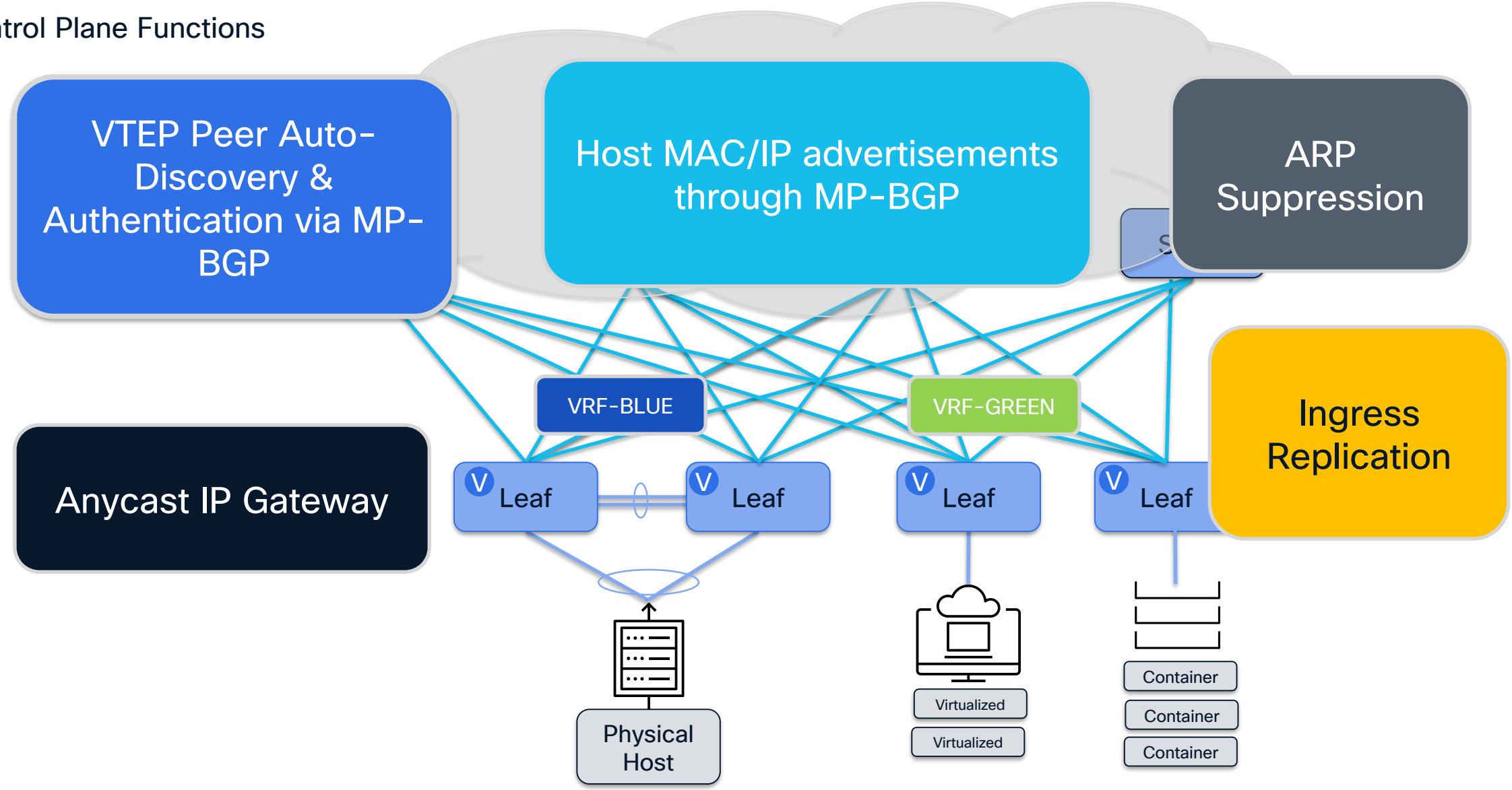
Route Distinguisher (RD)
8-byte Field, VRF Parameter
Value to make the MAC or IP Routes unique (RD + VPN prefix)

Route Target (RT)
8-byte Field, VRF Parameter
Value for the Import/Export Rules of VPN Routes (MAC or IP) into the VRF



EVPN in the Data Center

Control Plane Functions



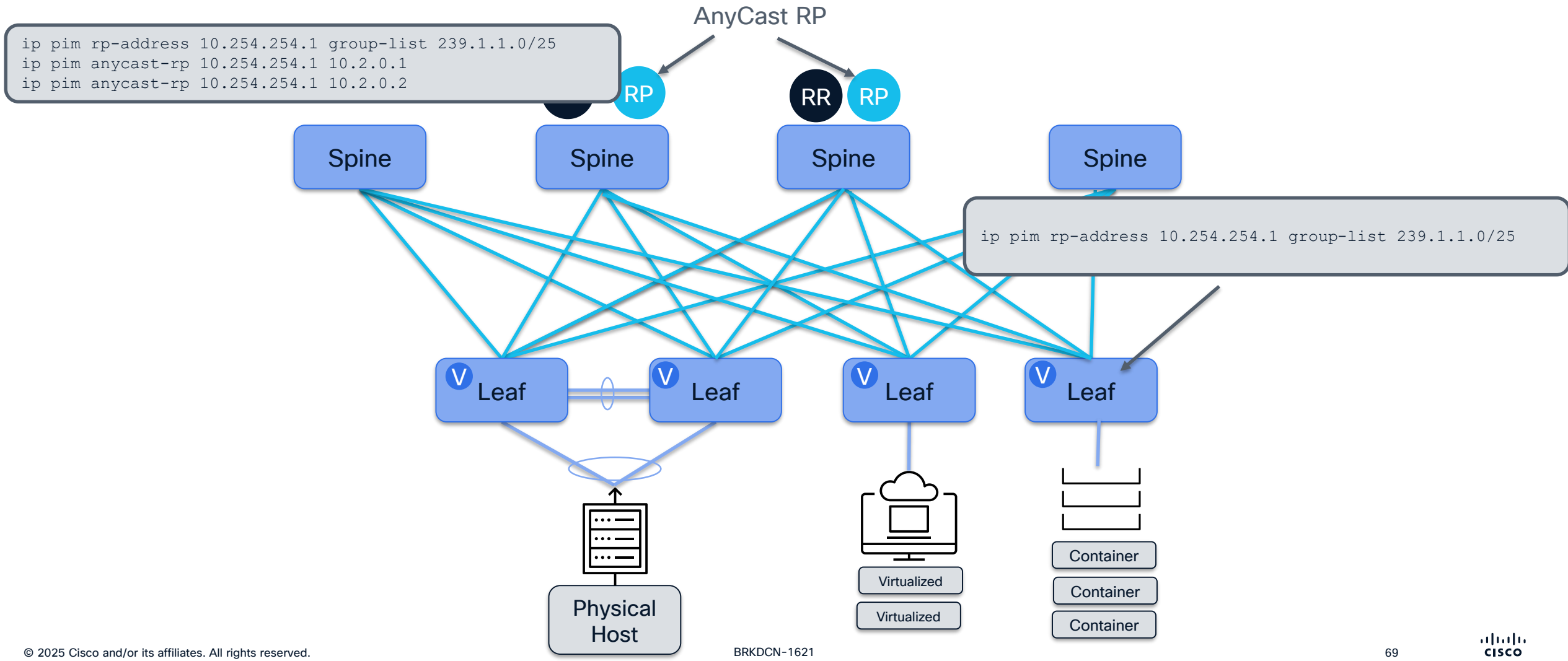
Configuring the EVPN VXLAN Fabric

VXLAN BGP EVPN

Network Underlay

RR Route Reflector

RP Rendezvous Point



VXLAN BGP EVPN

Network Underlay

RR Route Reflector

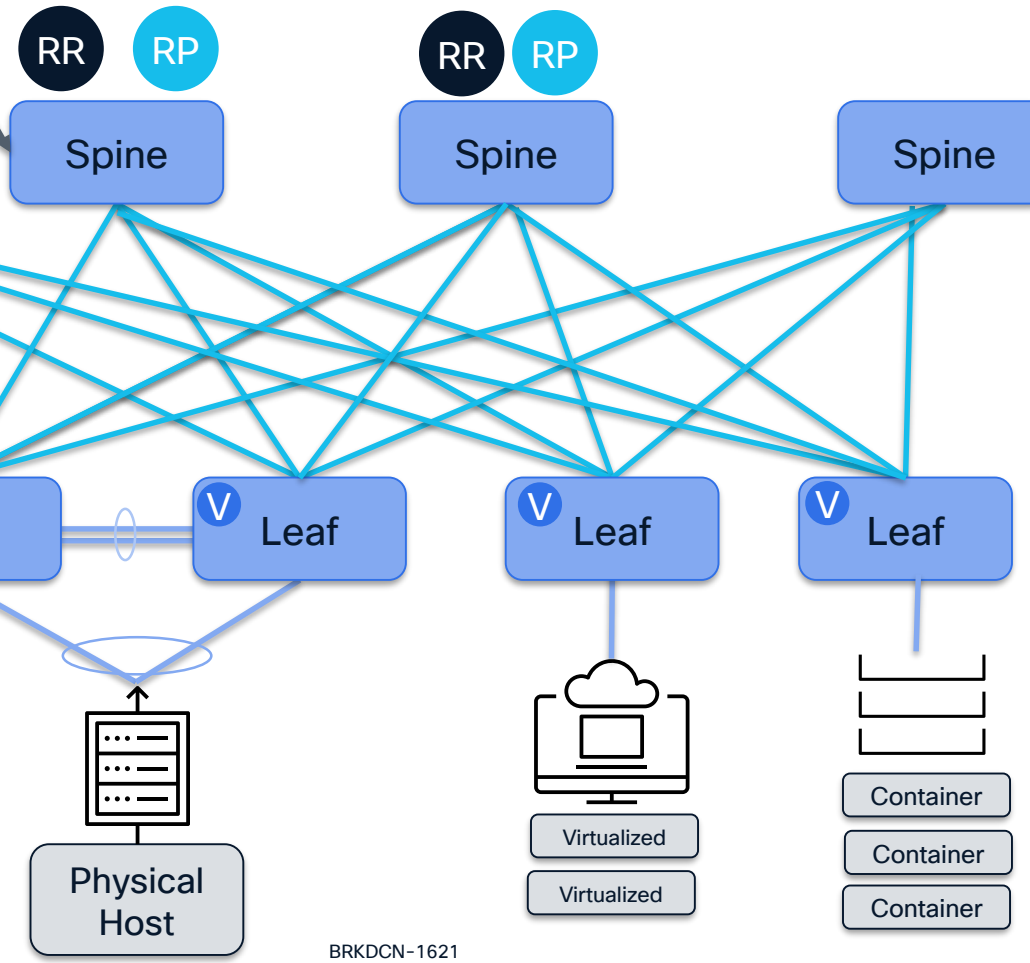
RP Rendezvous Point

```
interface Ethernet1/2
  description connected-to-leaf-1-Ethernet1/1
  no switchport
  mtu 9216
  ip address 10.4.0.5/30
  ip ospf network point-to-point
  ip router ospf UNDERLAY area 0.0.0.0
  ip pim sparse-mode
  no shutdown
```

```
interface loopback0
  description Routing loopback
interface
  ip address 10.2.0.7/32
  ip router ospf UNDERLAY area 0.0.0.0
  ip pim sparse-mode
```

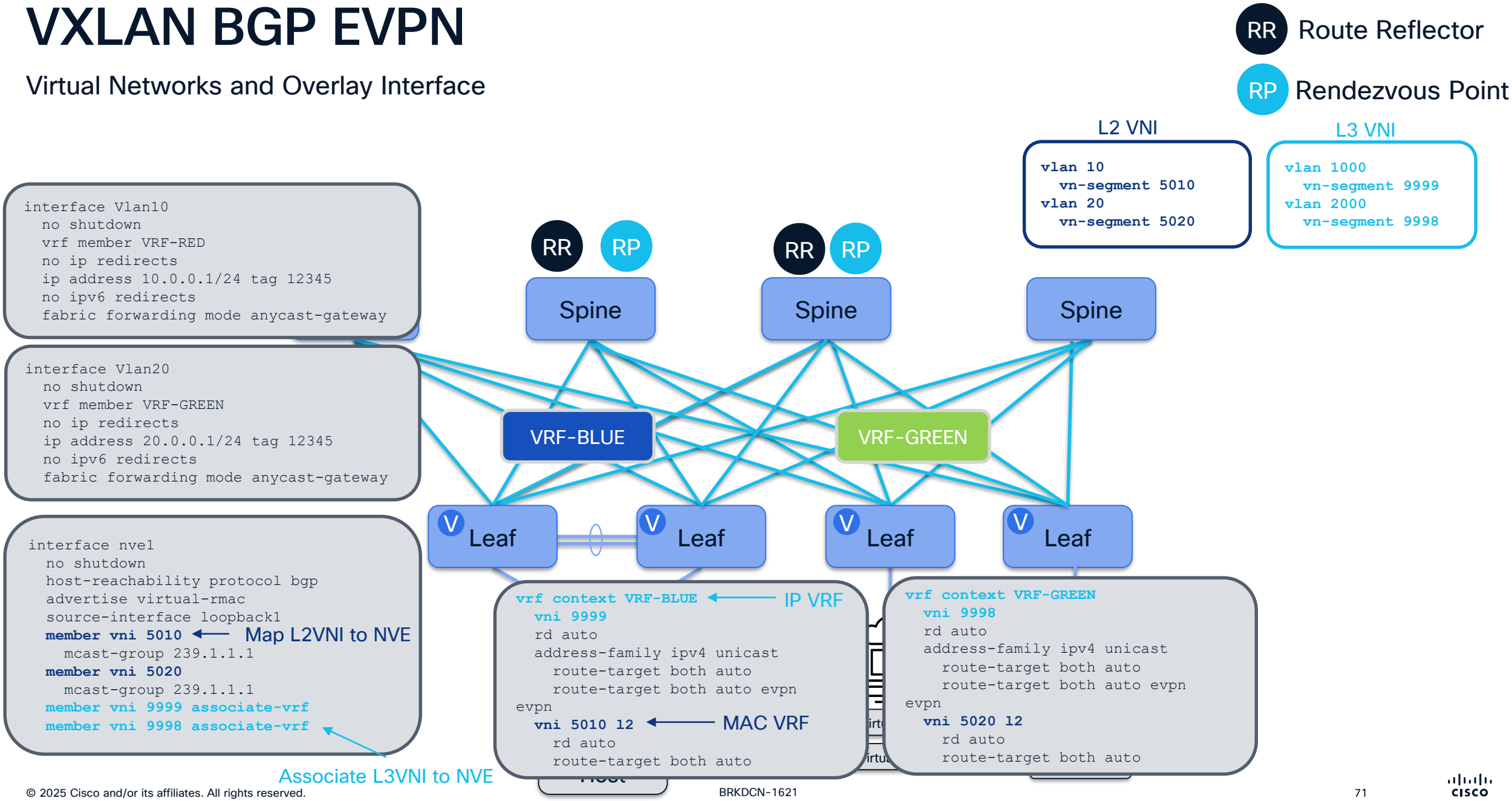
```
router ospf UNDERLAY
router-id 10.2.0.7
```

```
interface Ethernet1/2
  description connected-to-spine-1-Ethernet1/1
  no switchport
  mtu 9216
  ip address 10.4.0.6/30
  ip ospf network point-to-point
  ip router ospf UNDERLAY area 0.0.0.0
  ip pim sparse-mode
  no shutdown
```



VXLAN BGP EVPN

Virtual Networks and Overlay Interface



VXLAN BGP EVPN

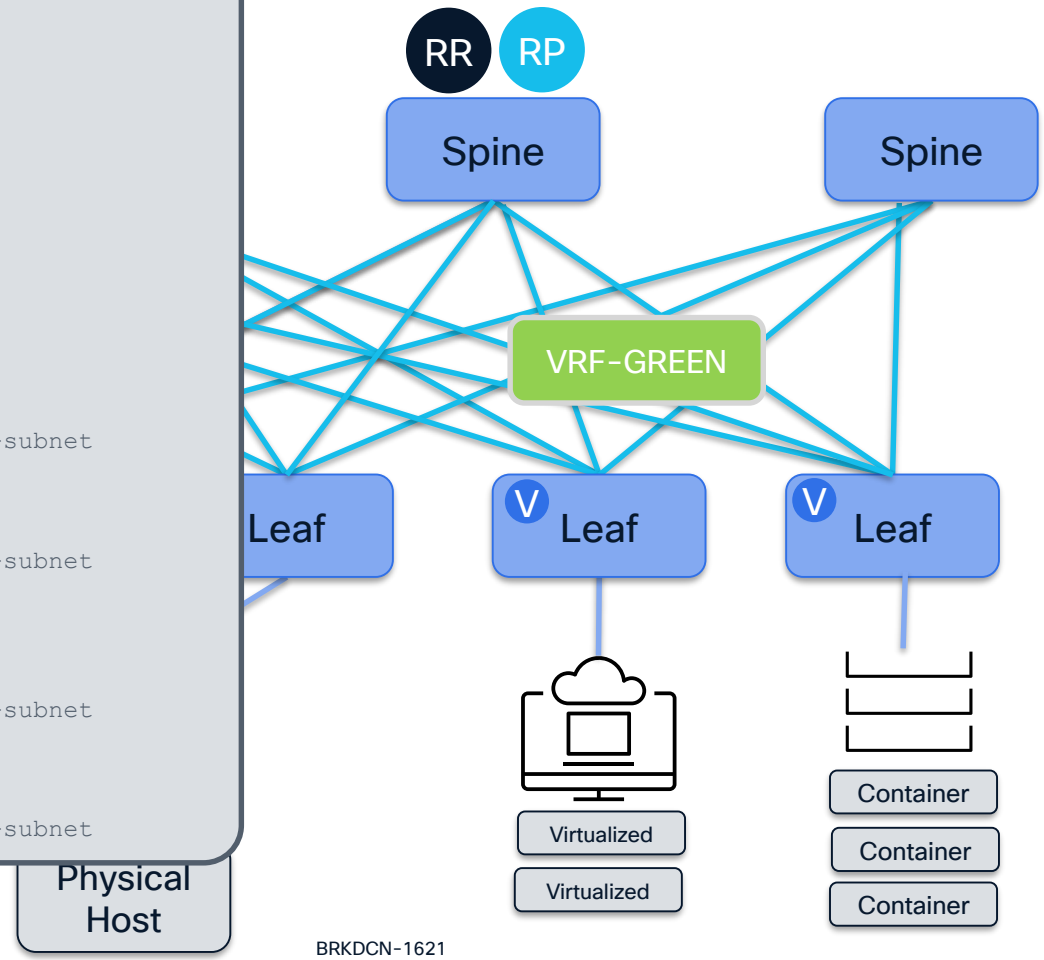
BGP on Leaf and Spine

RR Route Reflector

RP Rendezvous Point

Leaf BGP Config

```
router bgp 65010
  router-id 10.2.0.7
  address-family l2vpn evpn
    advertise-pip
  neighbor 10.2.0.1
    remote-as 65010
    update-source loopback0
    address-family l2vpn evpn
      send-community
      send-community extended
  neighbor 10.2.0.2
    remote-as 65010
    update-source loopback0
    address-family l2vpn evpn
      send-community
      send-community extended
vrf VRF-RED
    address-family ipv4 unicast
      advertise l2vpn evpn
      redistribute direct route-map fabric-rmap-redist-subnet
      maximum-paths ibgp 2
    address-family ipv6 unicast
      advertise l2vpn evpn
      redistribute direct route-map fabric-rmap-redist-subnet
      maximum-paths ibgp 2
vrf VRF-GREEN
    address-family ipv4 unicast
      advertise l2vpn evpn
      redistribute direct route-map fabric-rmap-redist-subnet
      maximum-paths ibgp 2
    address-family ipv6 unicast
      advertise l2vpn evpn
      redistribute direct route-map fabric-rmap-redist-subnet
```



Spine BGP Config

```
router bgp 65010
  router-id 10.2.0.2
  neighbor 10.2.0.3
    remote-as 65010
    update-source loopback0
    address-family l2vpn evpn
      send-community
      send-community extended
      route-reflector-client
  neighbor 10.2.0.4
    remote-as 65010
    update-source loopback0
    address-family l2vpn evpn
      send-community
      send-community extended
      route-reflector-client
  neighbor 10.2.0.5
    remote-as 65010
    update-source loopback0
    address-family l2vpn evpn
      send-community
      send-community extended
      route-reflector-client
```

VXLAN BGP EVPN

Distributed Anycast Gateway

RR Route Reflector

RP Rendezvous Point

The same anycast gateway Virtual IP Address and MAC address are configured on all VTEPs in the VNI

VLAN to VNI Mapping

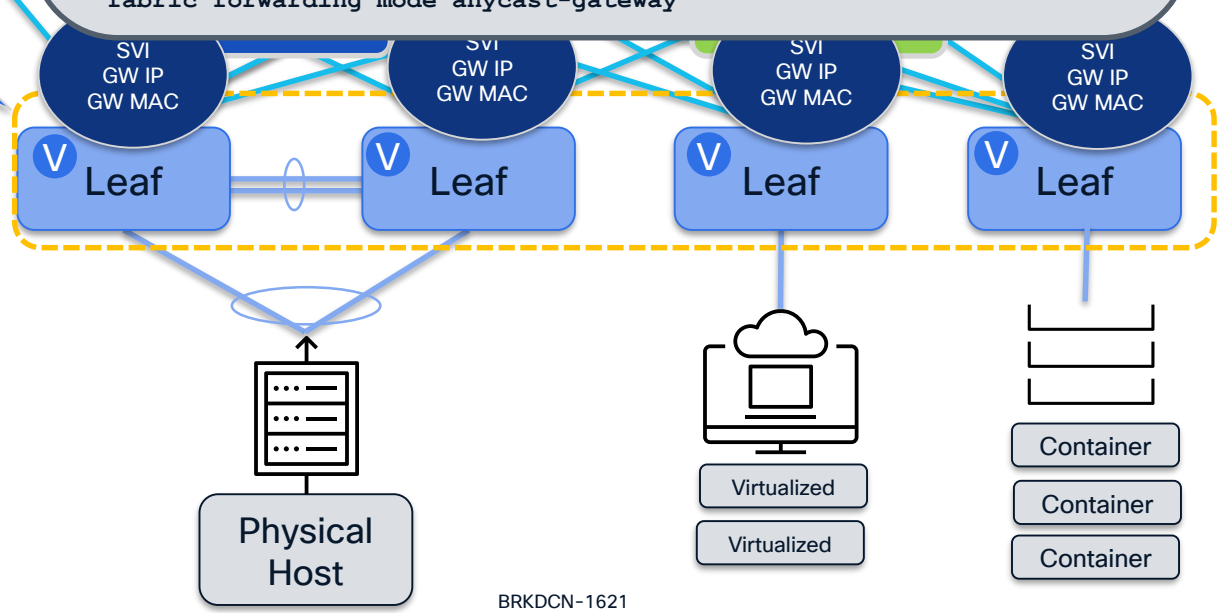
```
vlan 20
  vn-segment 5020
```

Anycast Gateway MAC Address- Identically configured on all VTEPs

```
fabric forwarding anycast-gateway-mac 0002.0002.0002
```

Distributed IP Anycast Gateway (SVI) - Identically configured on all VTEPs

```
interface Vlan20
  no shutdown
  vrf member VRF-GREEN
  no ip redirects
  ip address 20.0.0.1/24 tag 12345
  no ipv6 redirects
  fabric forwarding mode anycast-gateway
```

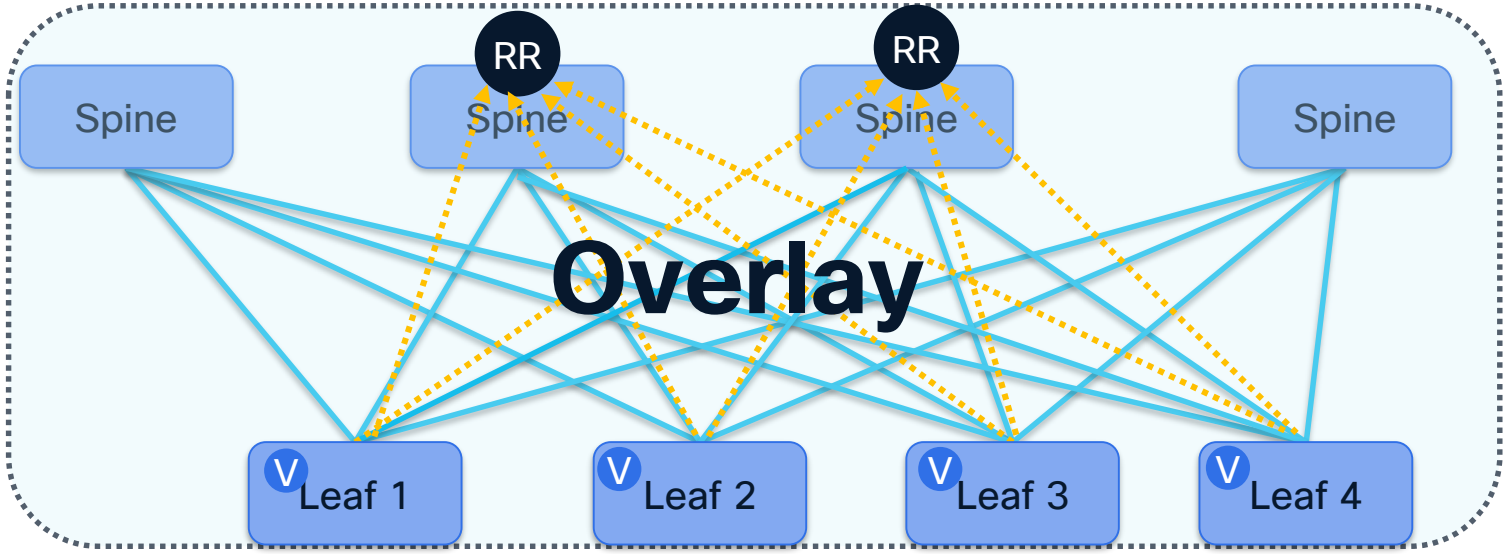


EVPN Peer and Host Discovery
L2VNI

VXLAN BGP EVPN

Host and Subnet Distribution

RR Route Reflector

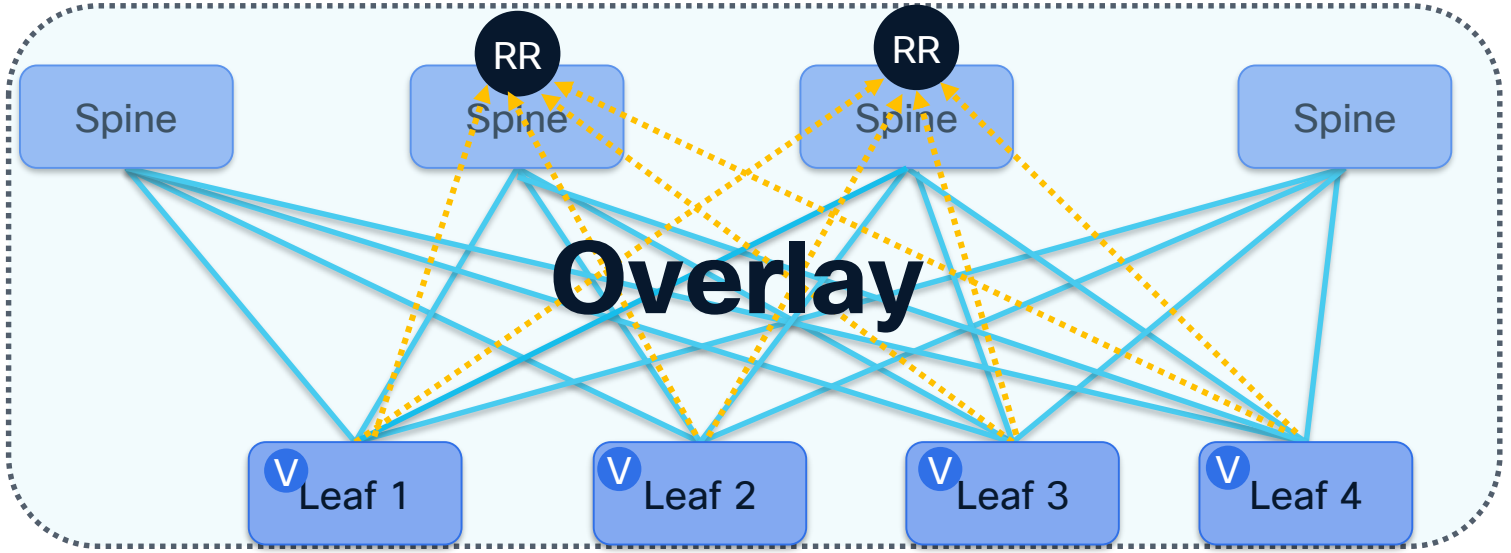


- Host Route Distribution decoupled from the Underlay protocol
- Use Multi-protocol-BGP (MP-BGP) on the leaf nodes to distribute internal Host/Subnet Routes and external reachability information
- Route-Reflectors (RR) deployed for scaling purposes

VXLAN BGP EVPN

Host and Subnet Routes

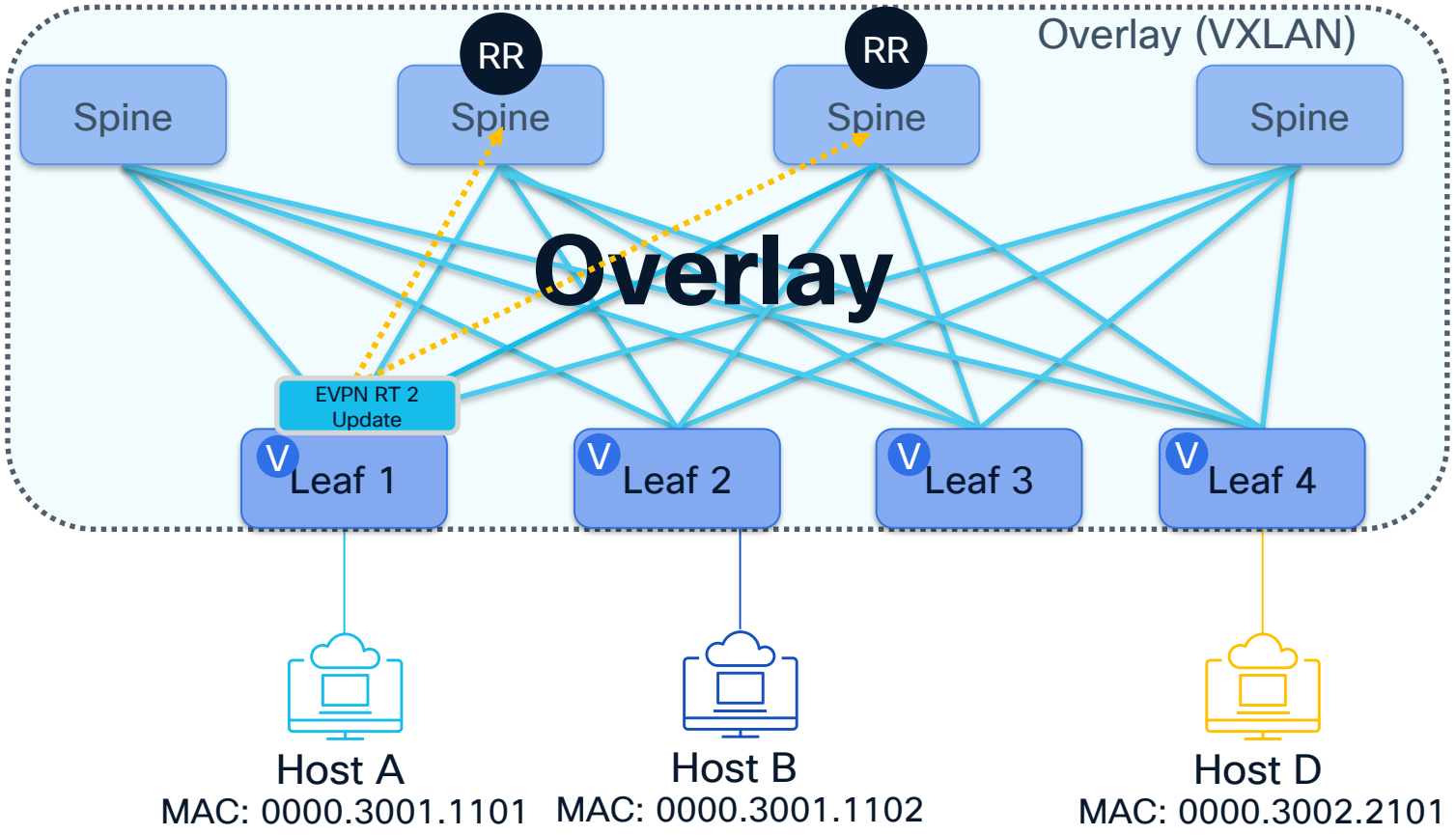
RR Route Reflector



- BGP EVPN NLRI
  (Network Layer Reachability Information)
  BGP Update Format
- Host MAC (Route Type 2)
  - MAC only, Single VNI, Single Route Target
- Host MAC + IP (Route Type 2)
  - MAC and IP, Two VNI, Two Route Targets, Router MAC
- Internal and External Subnet Prefixes (Route Type 5)
  - IP Subnet Prefix, Single VNI, Single Route Target

EVPN Control Plane

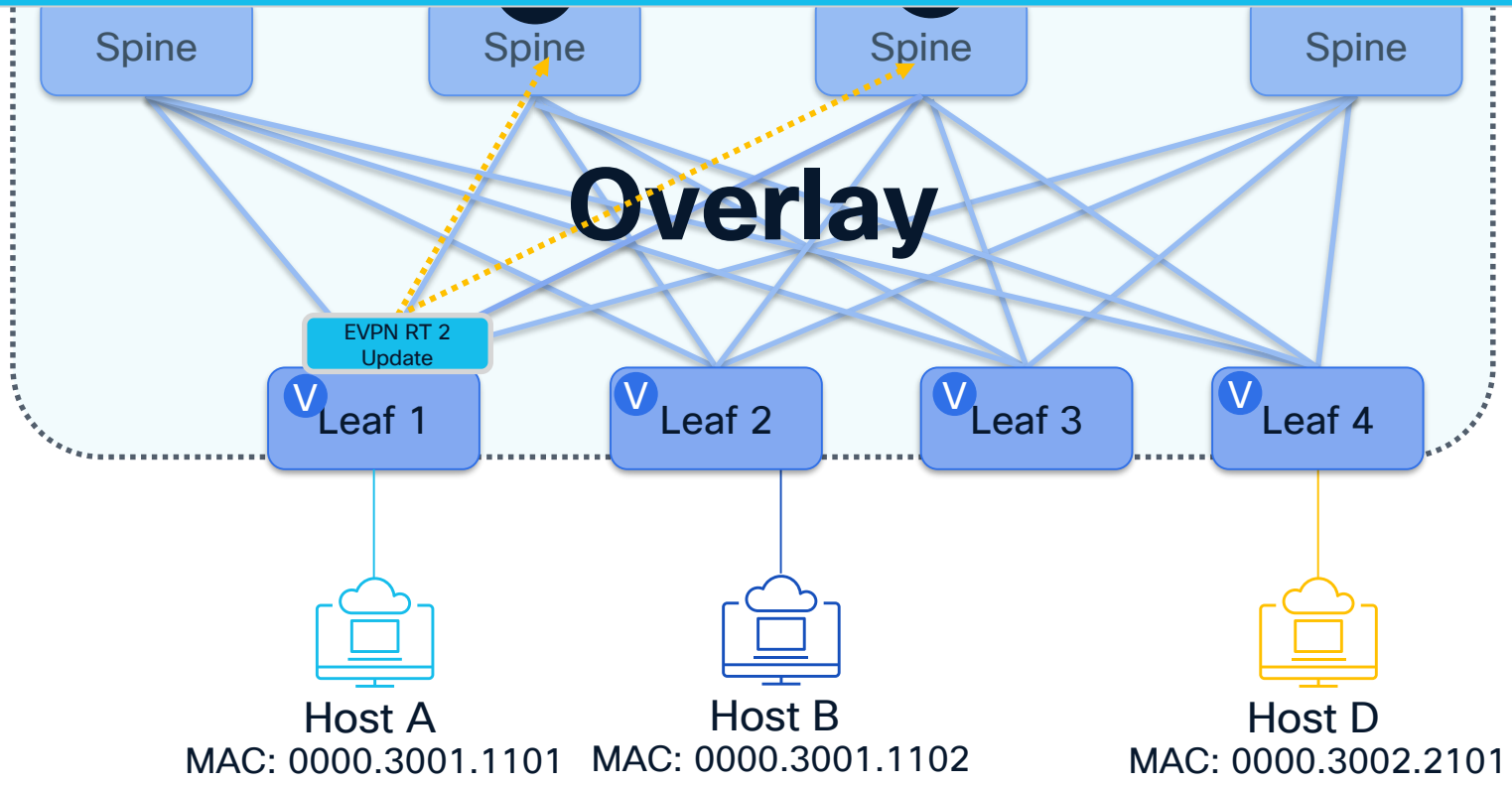
Host Advertisements



EVPN Control Plane

Host Advertisements

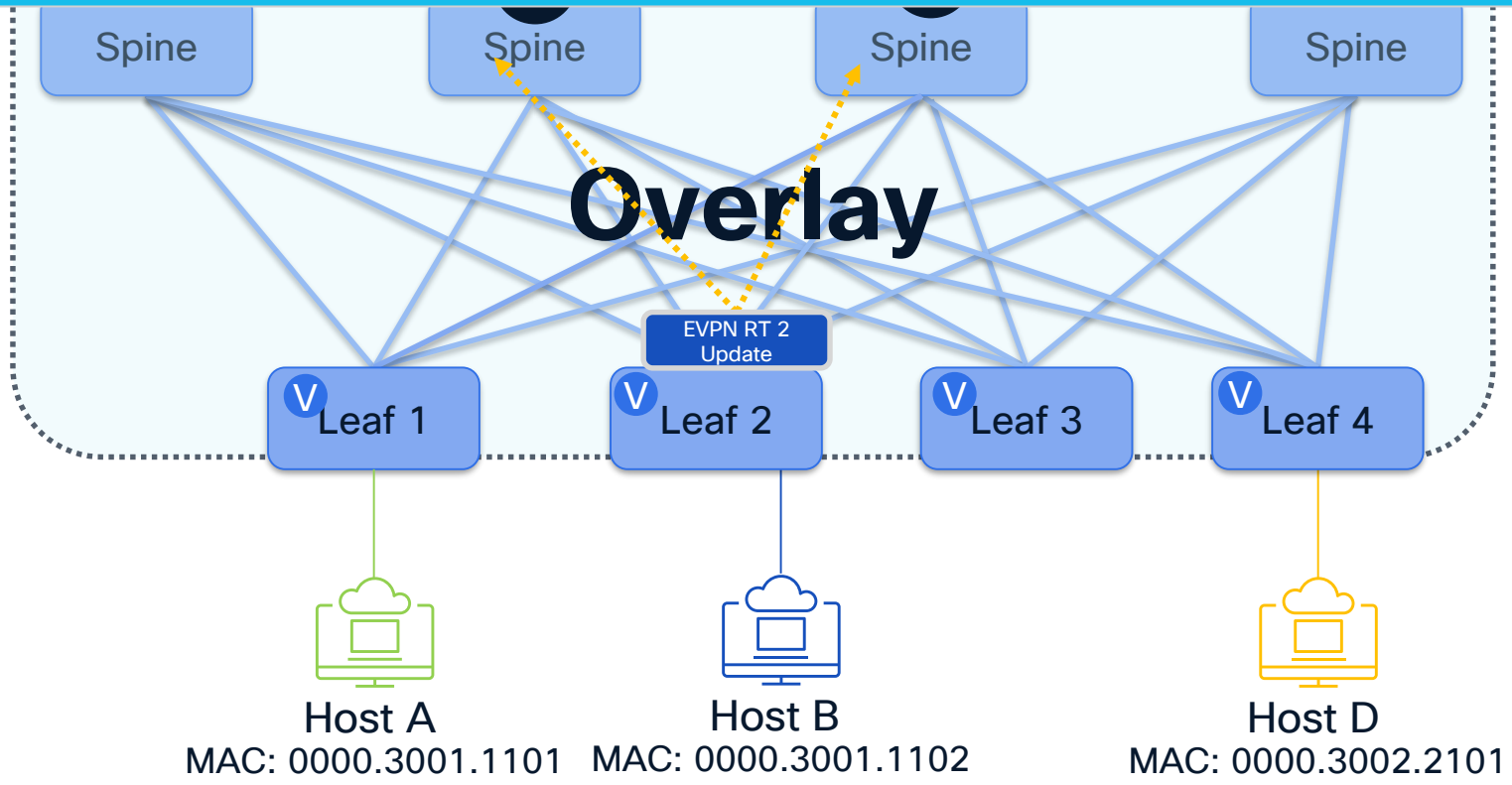
Type	MAC / Length	L2VNI / RT	IP / Length	L3VNI / RT	Next-Hop	Seq
2	0000.3001.1101 / 48	3001, 65500:3001	-	-	10.200.200.101	0



EVPN Control Plane

Host Advertisements

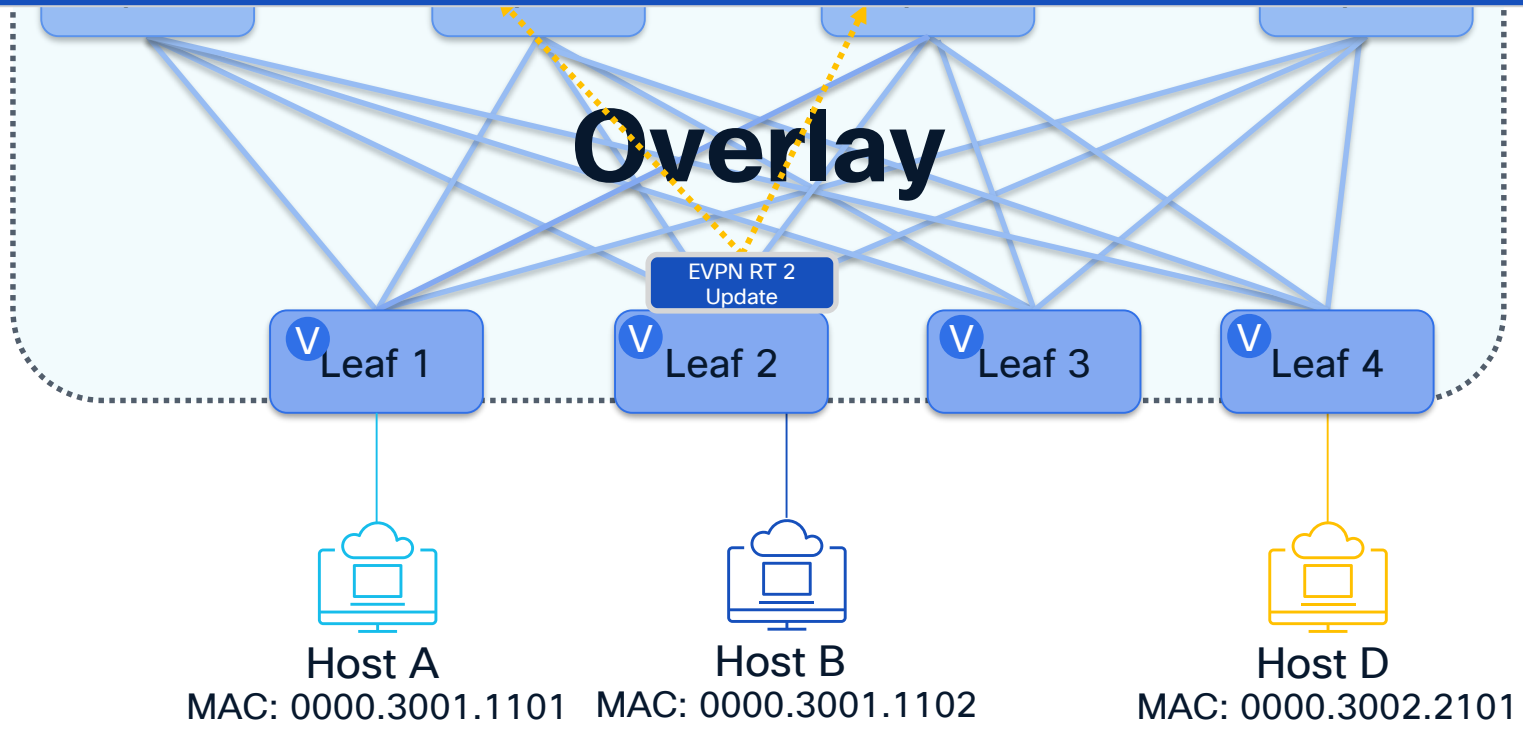
Type	MAC / Length	L2VNI / RT	IP / Length	L3VNI / RT	Next-Hop	Seq
2	0000.3001.1101 / 48	3001, 65500:3001	-	-	10.200.200.101	0



EVPN Control Plane

Host Advertisements

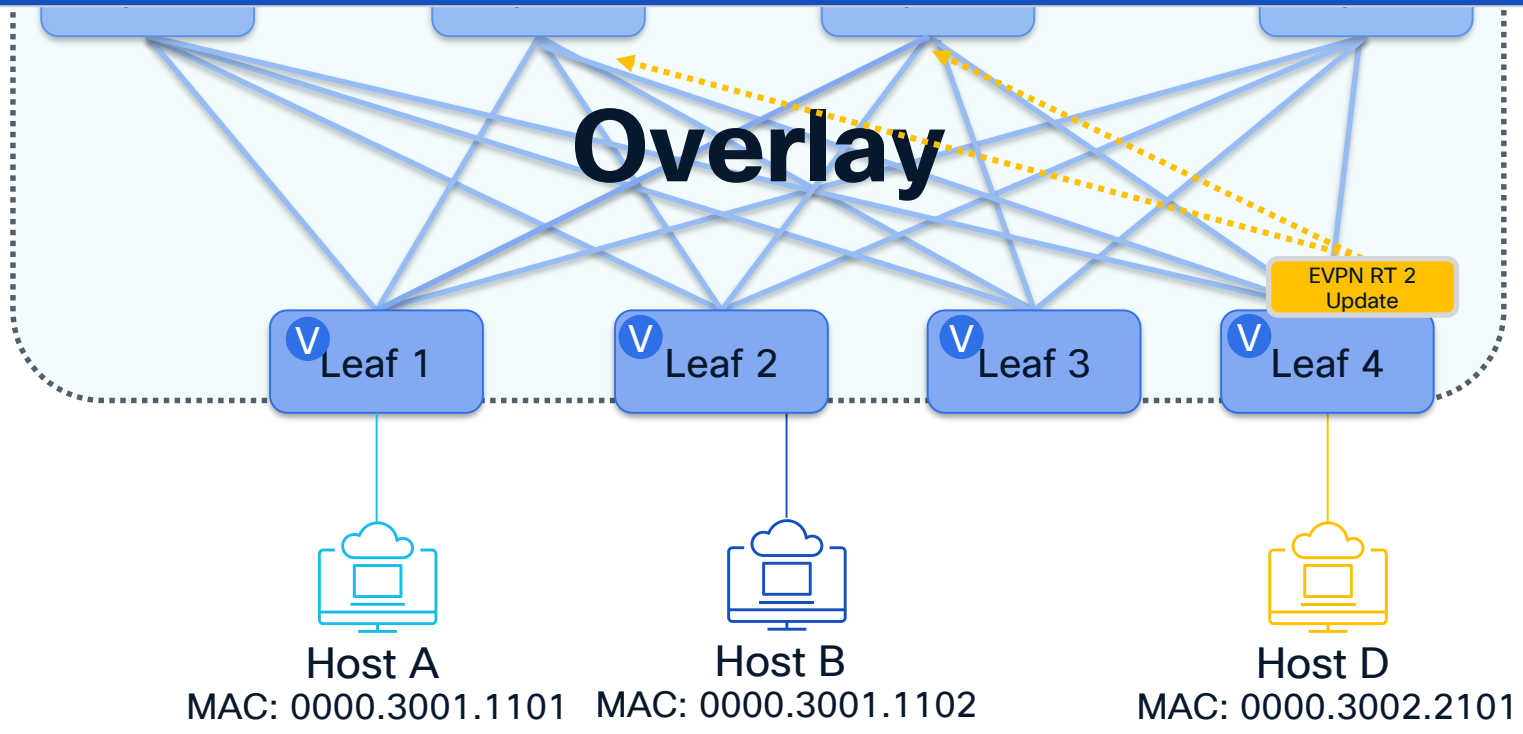
Type	MAC / Length	L2VNI / RT	IP / Length	L3VNI / RT	Next-Hop	Seq
2	0000.3001.1101 / 48	3001, 65500:3001	-	-	10.200.200.101	0
2	0000.3001.1102 / 48	3001, 65500:3001	-	-	10.200.200.102	0



EVPN Control Plane

Host Advertisements

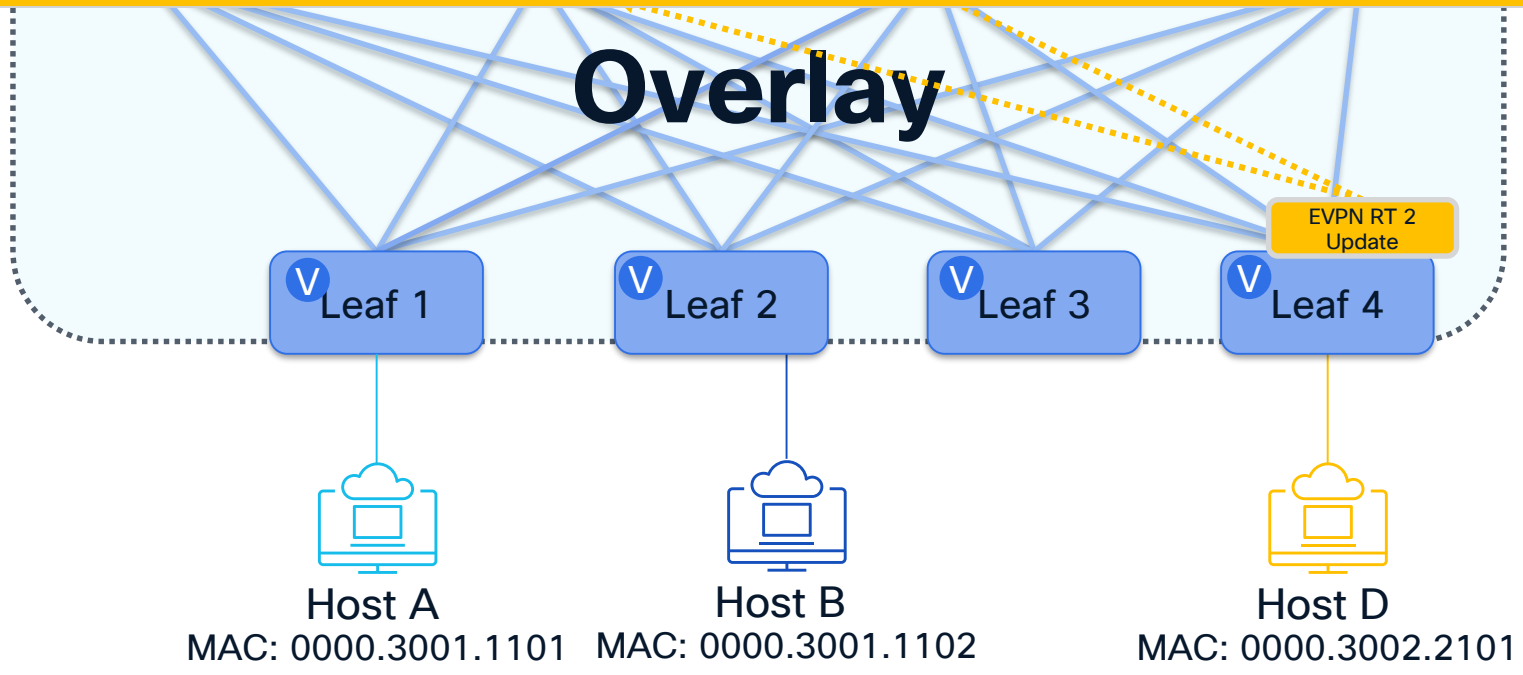
Type	MAC / Length	L2VNI / RT	IP / Length	L3VNI / RT	Next-Hop	Seq
2	0000.3001.1101 / 48	3001, 65500:3001	-	-	10.200.200.101	0
2	0000.3001.1102 / 48	3001, 65500:3001	-	-	10.200.200.102	0



EVPN Control Plane

Host Advertisements

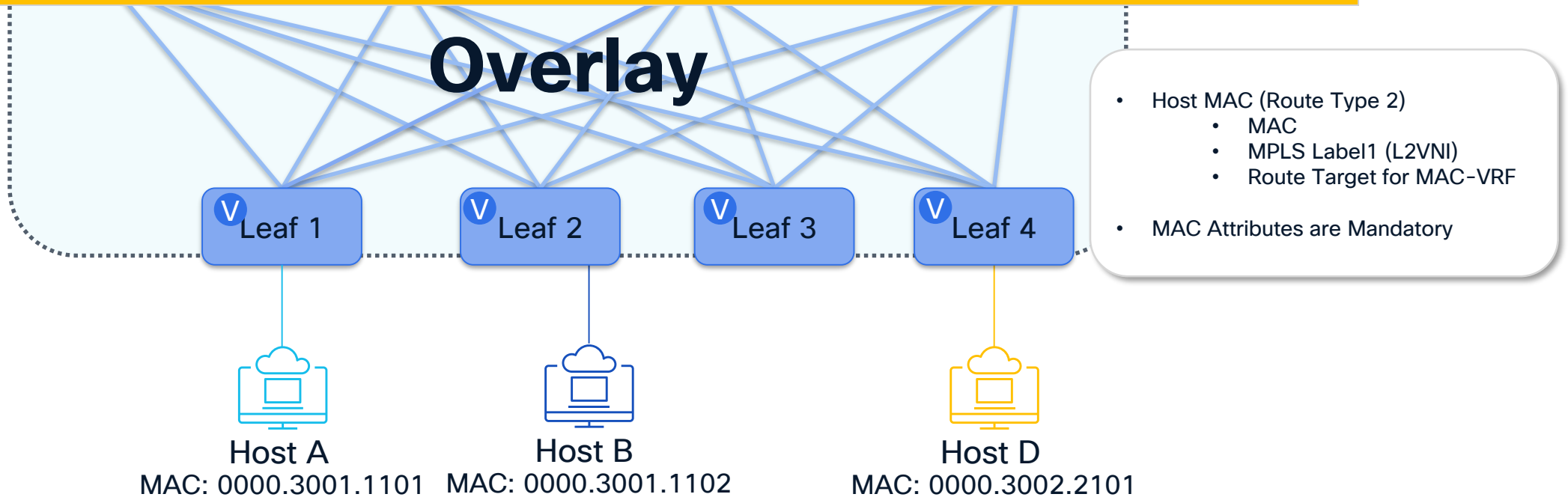
Type	MAC / Length	L2VNI / RT	IP / Length	L3VNI / RT	Next-Hop	Seq
2	0000.3001.1101 / 48	3001, 65500:3001	-	-	10.200.200.101	0
2	0000.3001.1102 / 48	3001, 65500:3001	-	-	10.200.200.102	0
2	0000.3002.2101 / 48	3002, 65500:3001	-	-	10.200.200.104	0



EVPN Control Plane

Host Advertisements

Type	MAC / Length	L2VNI / RT	IP / Length	L3VNI / RT	Next-Hop	Seq
2	0000.3001.1101 / 48	3001, 65500:3001	-	-	10.200.200.101	0
2	0000.3001.1102 / 48	3001, 65500:3001	-	-	10.200.200.102	0
2	0000.3002.2101 / 48	3002, 65500:3001	-	-	10.200.200.104	0

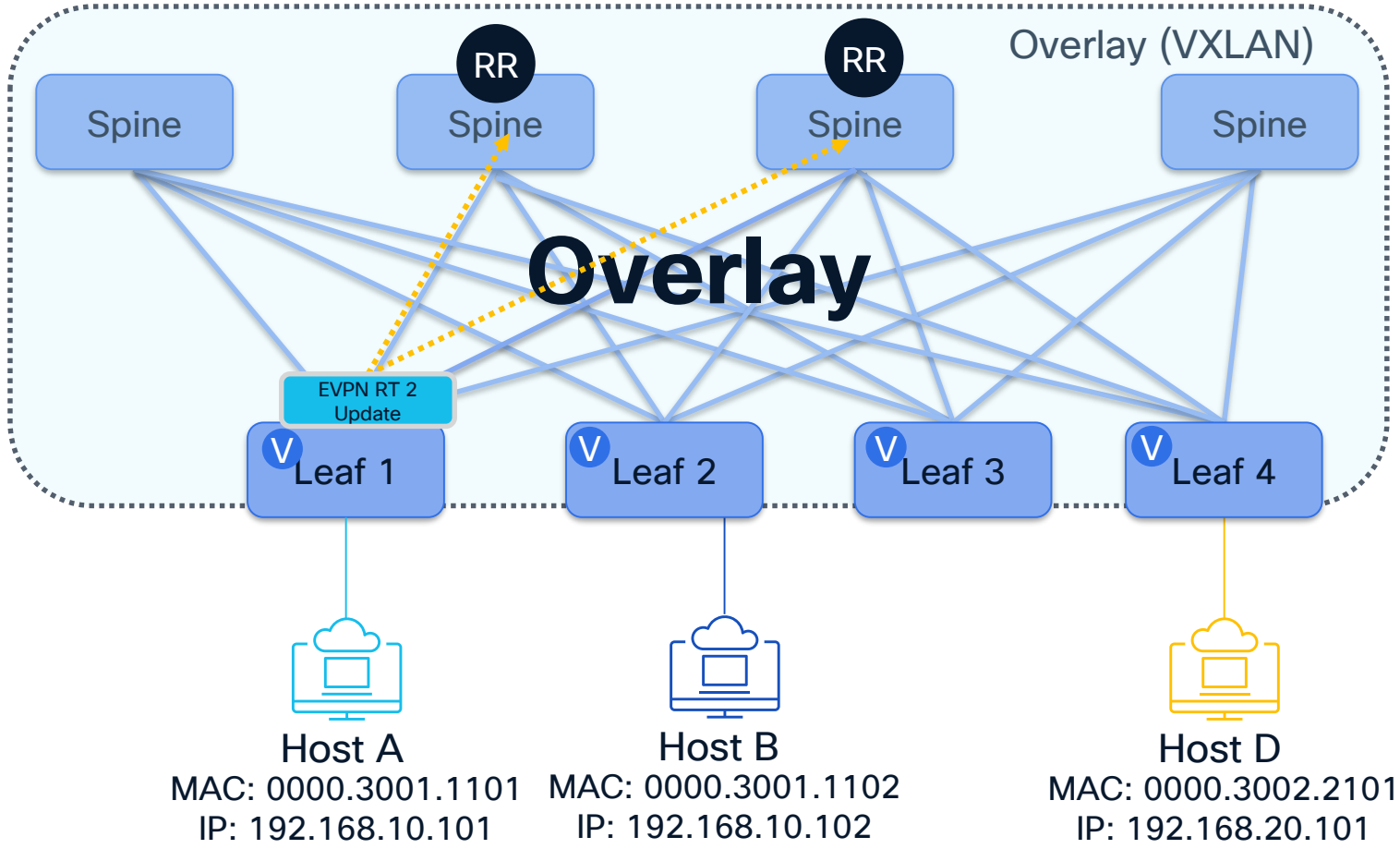


- Host MAC (Route Type 2)
  - MAC
  - MPLS Label1 (L2VNI)
  - Route Target for MAC-VRF
- MAC Attributes are Mandatory

EVPN Peer and Host Discovery
L3VNI

EVPN Control Plane

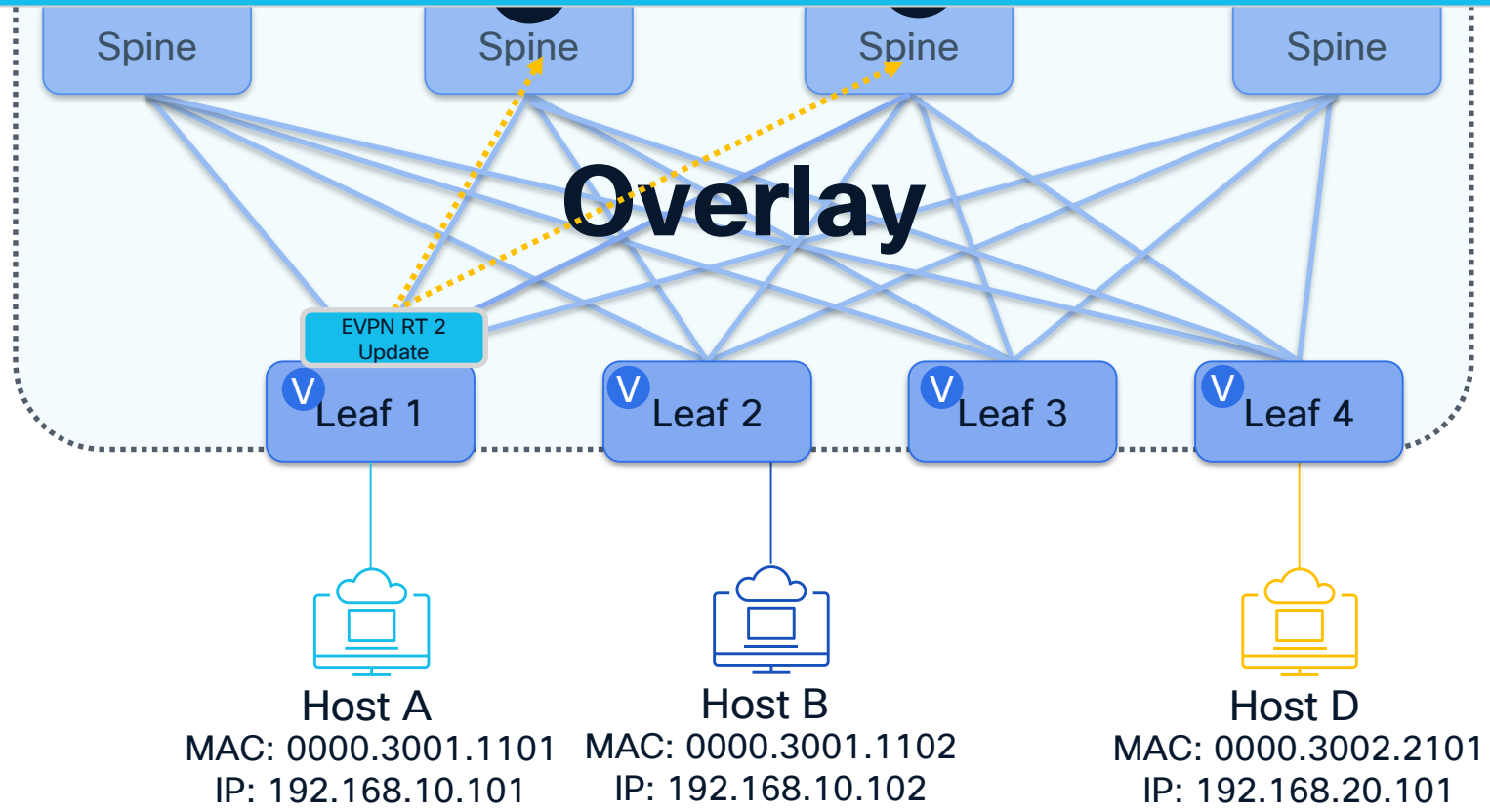
Host Advertisements



EVPN Control Plane

Host Advertisements

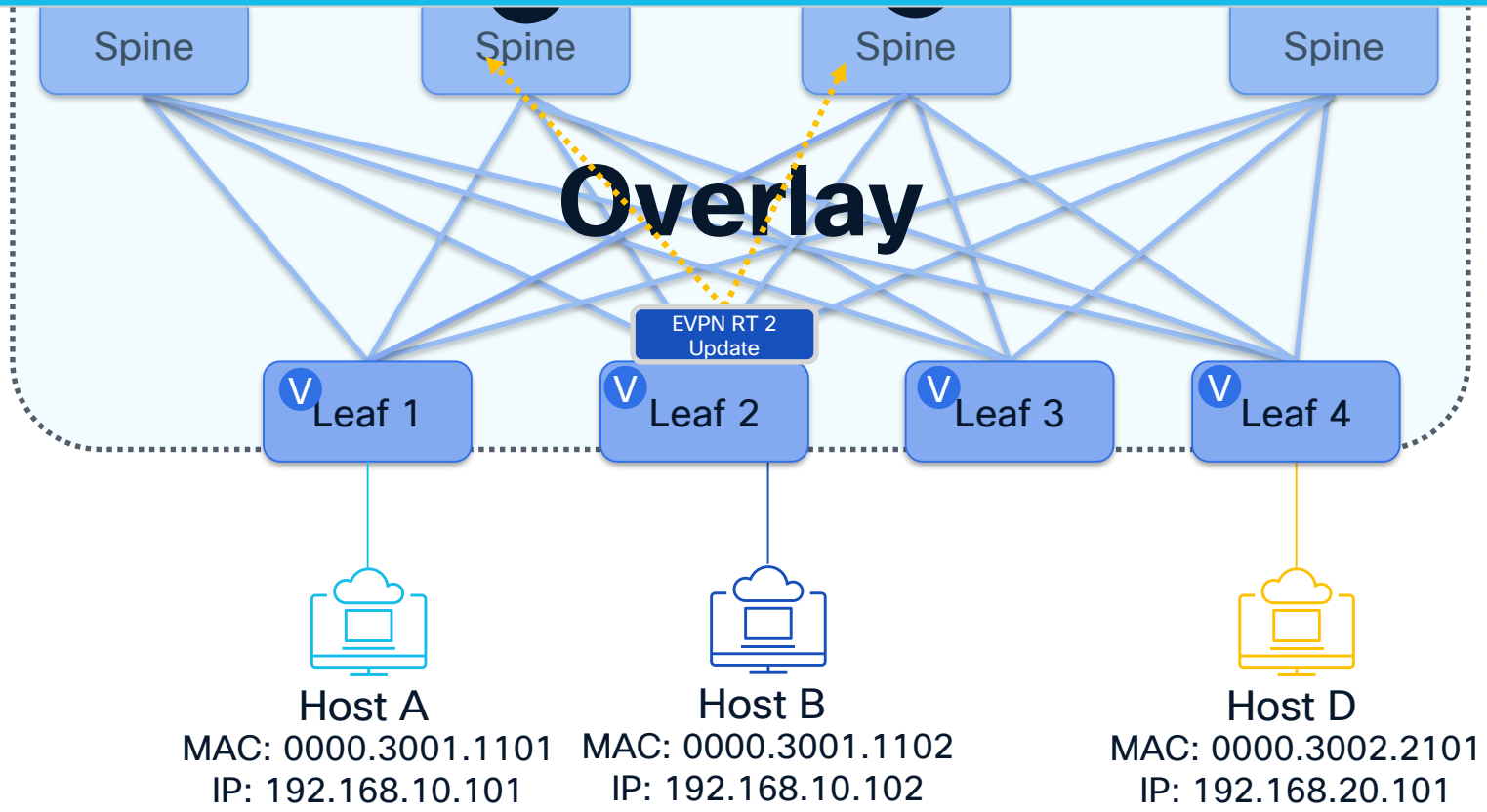
Type	MAC / Length	L2VNI / RT	IP / Length	L3VNI / RT	Next-Hop	Seq
2	0000.3001.1101 / 48	3001, 65500:3001	192.168.10.101/32	5000, 65500:5000	10.200.200.101	0



EVPN Control Plane

Host Advertisements

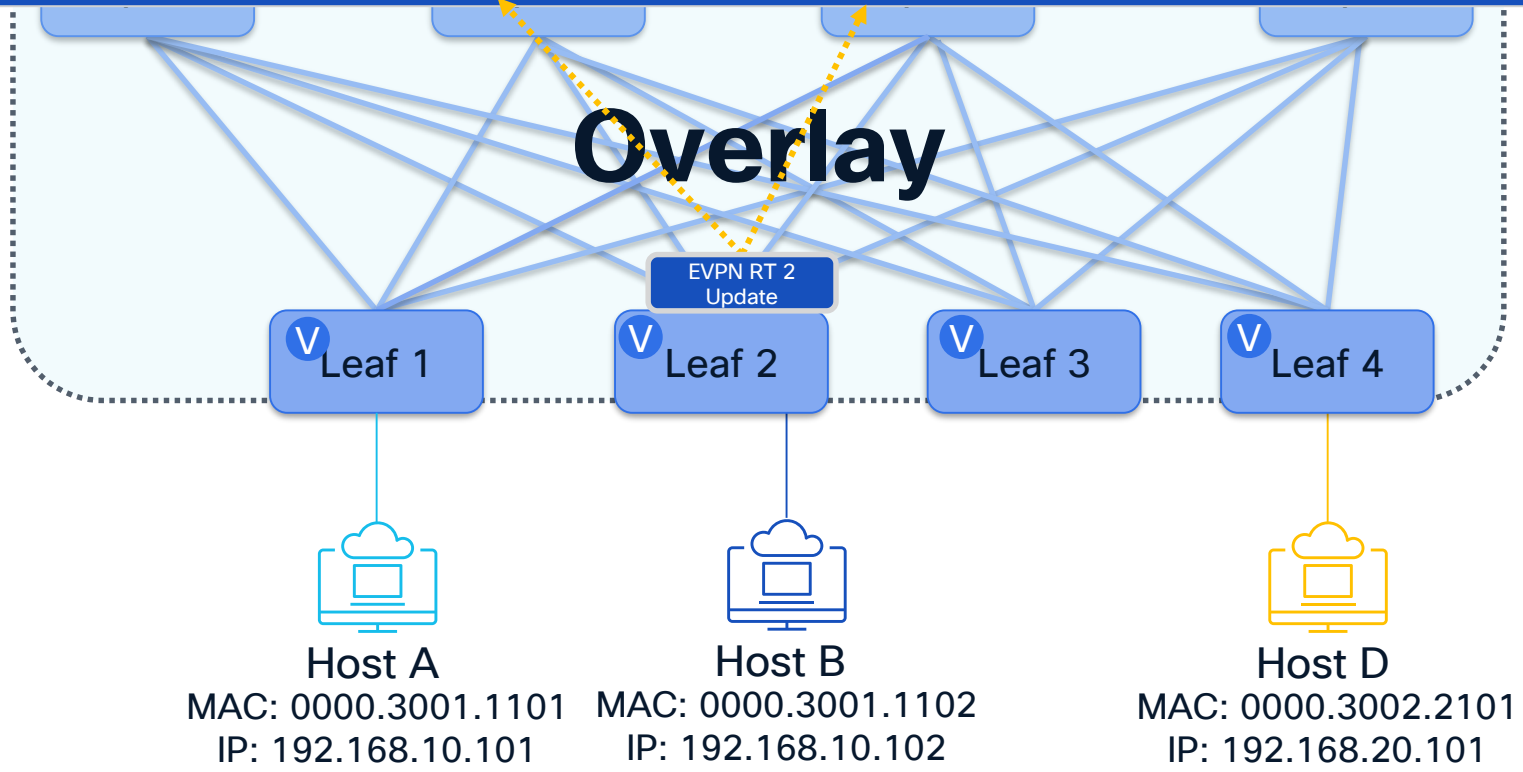
Type	MAC / Length	L2VNI / RT	IP / Length	L3VNI / RT	Next-Hop	Seq
2	0000.3001.1101 / 48	3001, 65500:3001	192.168.10.101/32	5000, 65500:5000	10.200.200.101	0



EVPN Control Plane

Host Advertisements

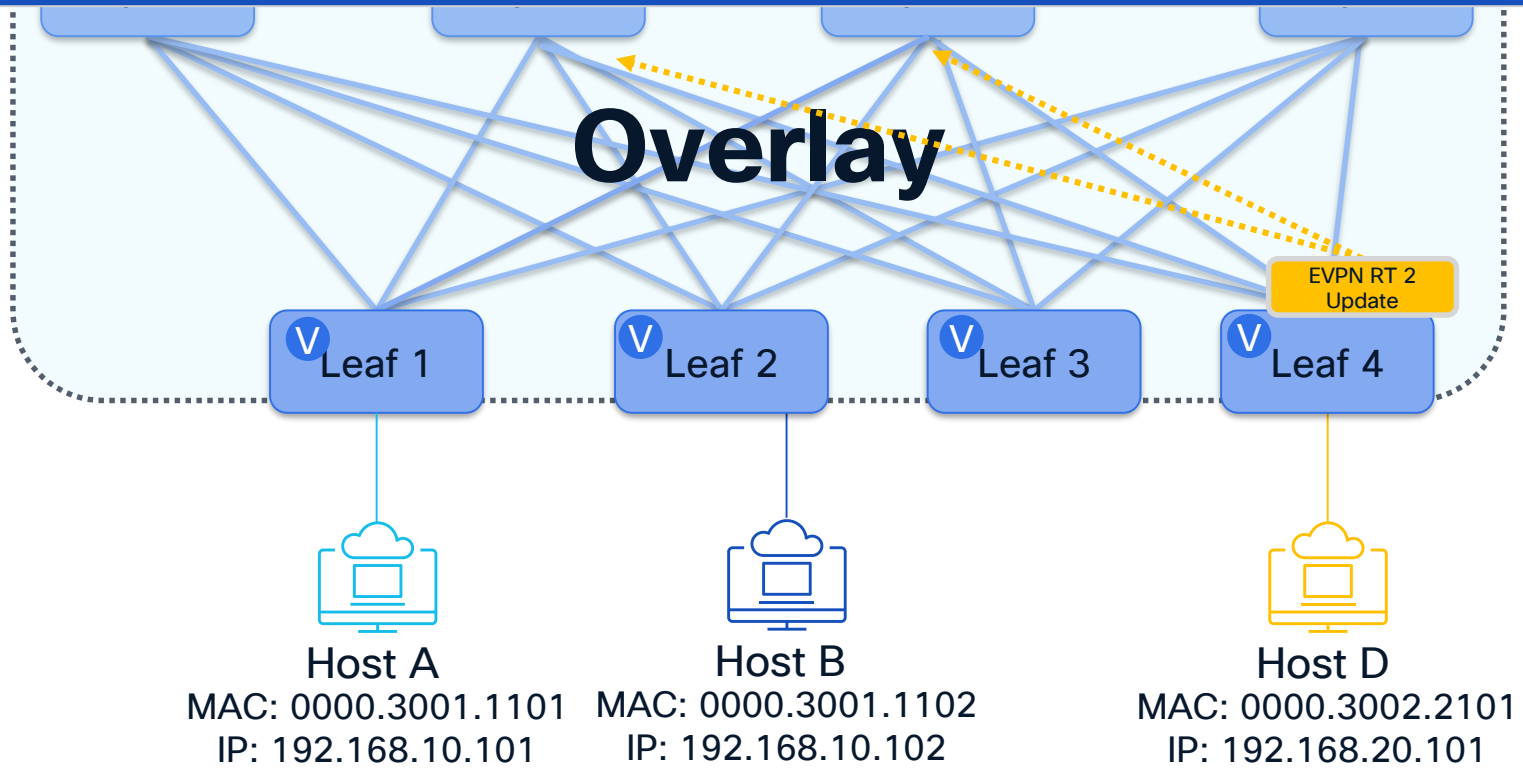
Type	MAC / Length	L2VNI / RT	IP / Length	L3VNI / RT	Next-Hop	Seq
2	0000.3001.1101 / 48	3001, 65500:3001	192.168.10.101/32	5000, 65500:5000	10.200.200.101	0
2	0000.3001.1102 / 48	3001, 65500:3001	192.168.10.102/32	5000, 65500:5000	10.200.200.102	0



EVPN Control Plane

Host Advertisements

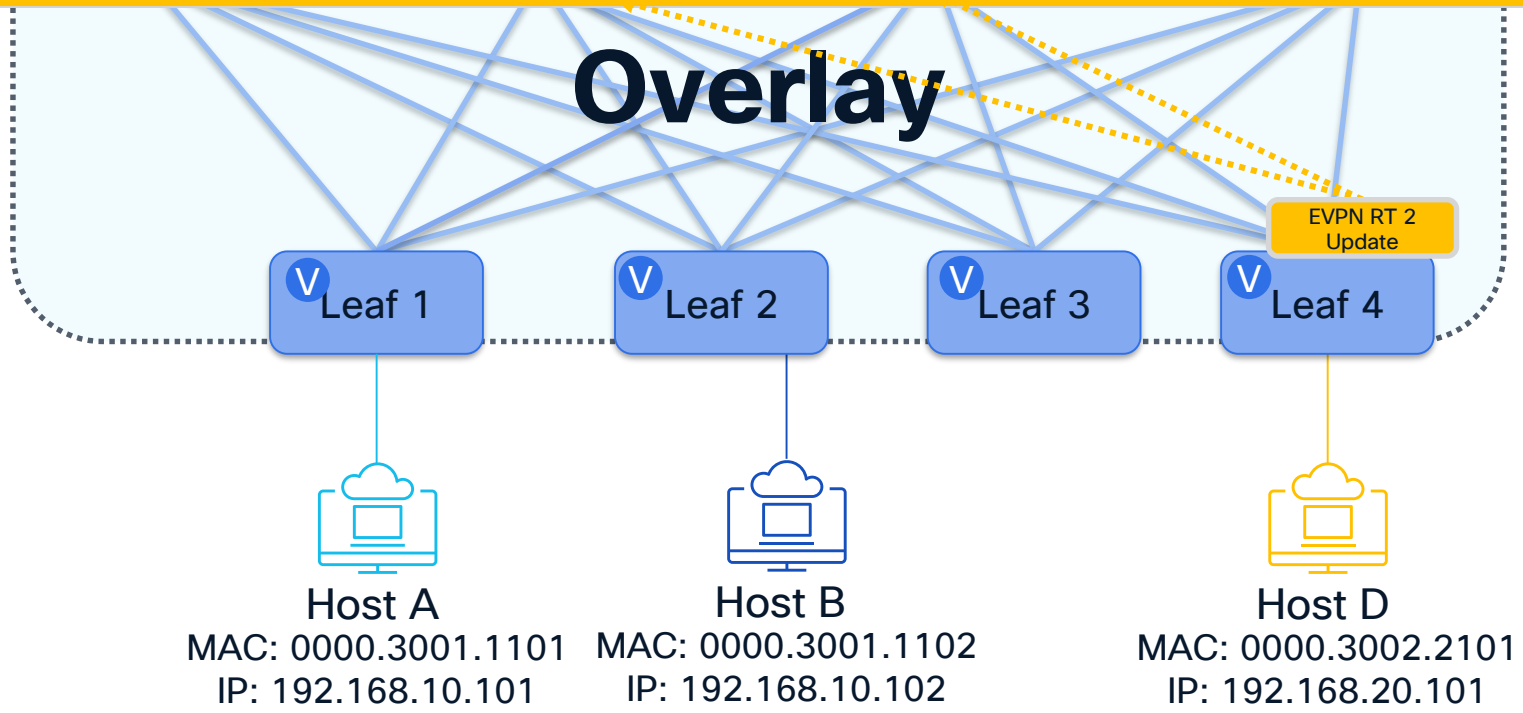
Type	MAC / Length	L2VNI / RT	IP / Length	L3VNI / RT	Next-Hop	Seq
2	0000.3001.1101 / 48	3001, 65500:3001	192.168.10.101/32	5000, 65500:5000	10.200.200.101	0
2	0000.3001.1102 / 48	3001, 65500:3001	192.168.10.102/32	5000, 65500:5000	10.200.200.102	0



EVPN Control Plane

Host Advertisements

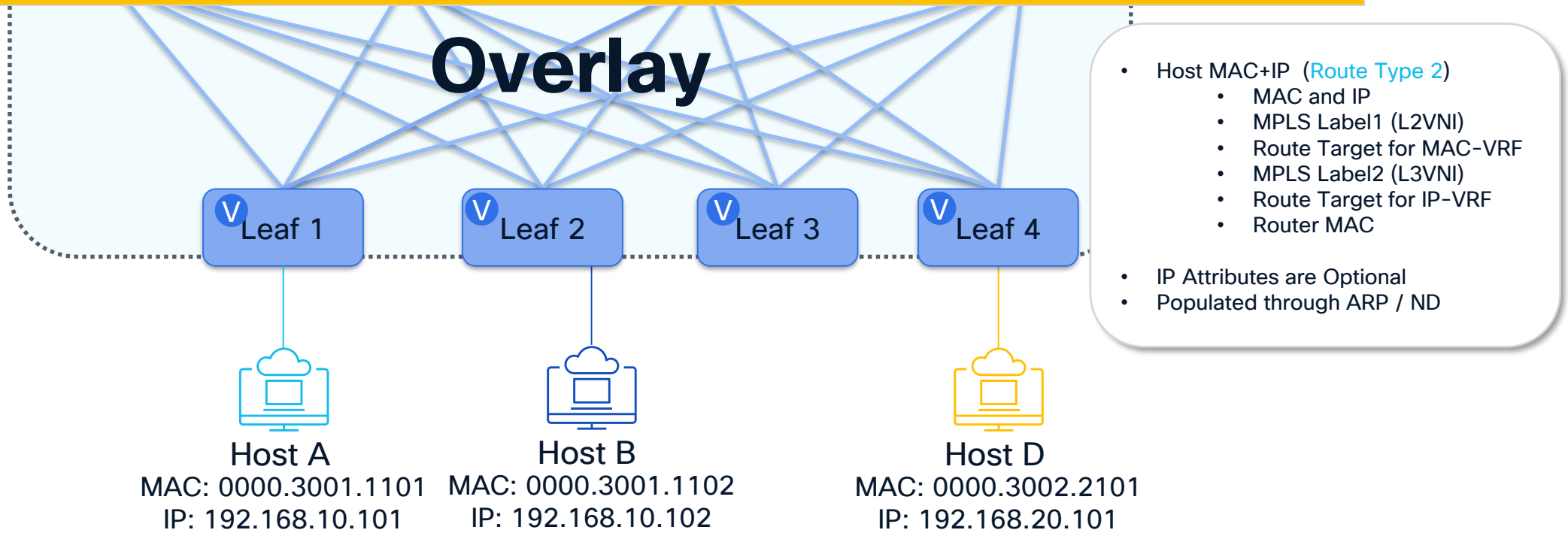
Type	MAC / Length	L2VNI / RT	IP / Length	L3VNI / RT	Next-Hop	Seq
2	0000.3001.1101 / 48	3001, 65500:3001	192.168.10.101/32	5000, 65500:5000	10.200.200.101	0
2	0000.3001.1102 / 48	3001, 65500:3001	192.168.10.102/32	5000, 65500:5000	10.200.200.102	0
2	0000.3002.2101 / 48	3002, 65500:3001	192.168.20.101/32	5000, 65500:5000	10.200.200.104	0



EVPN Control Plane

Host Advertisements

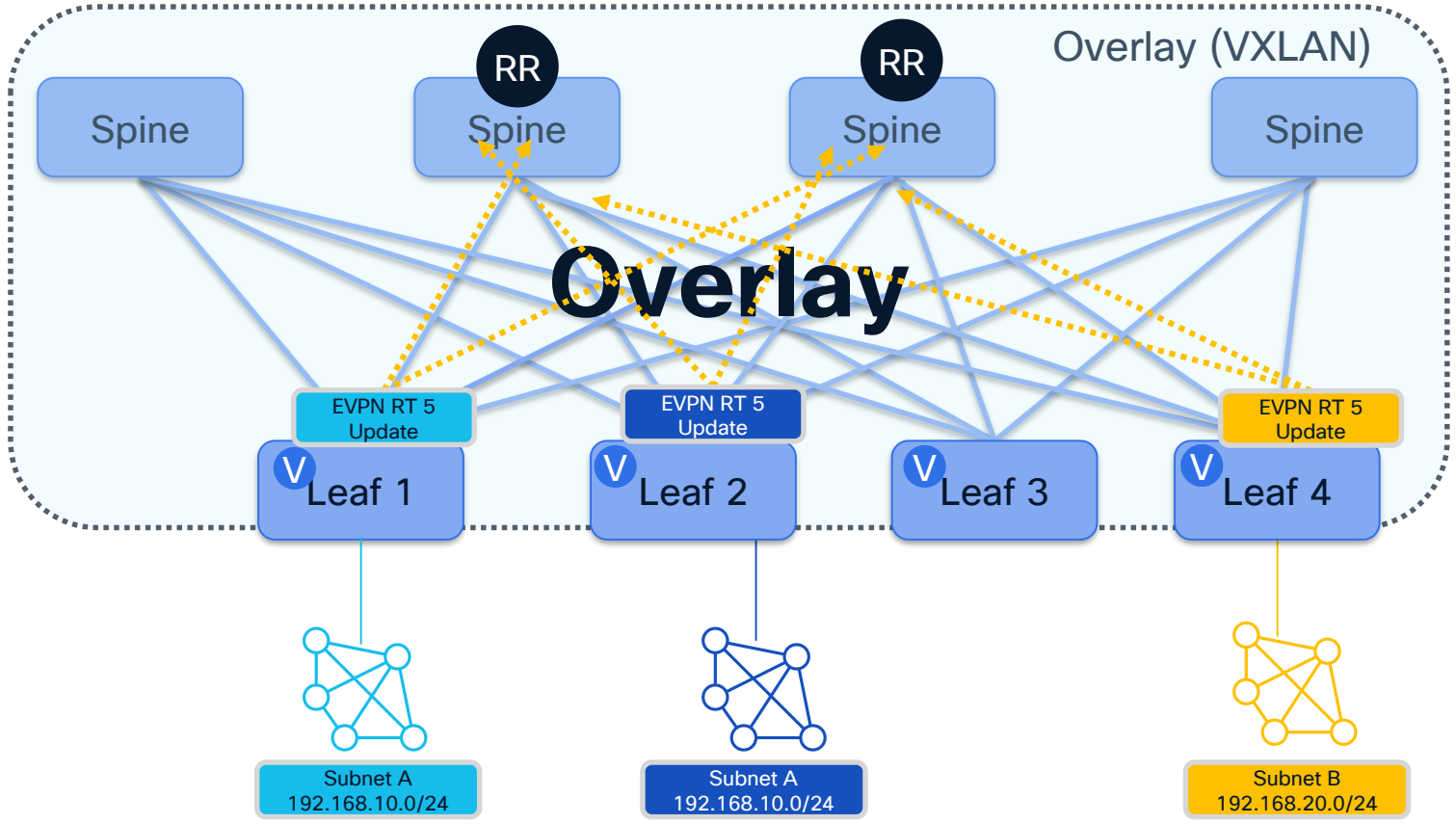
Type	MAC / Length	L2VNI / RT	IP / Length	L3VNI / RT	Next-Hop	Seq
2	0000.3001.1101 / 48	3001, 65500:3001	192.168.10.101/32	5000, 65500:5000	10.200.200.101	0
2	0000.3001.1102 / 48	3001, 65500:3001	192.168.10.102/32	5000, 65500:5000	10.200.200.102	0
2	0000.3002.2101 / 48	3002, 65500:3001	192.168.20.101/32	5000, 65500:5000	10.200.200.104	0



- Host MAC+IP (Route Type 2)
  - MAC and IP
  - MPLS Label1 (L2VNI)
  - Route Target for MAC-VRF
  - MPLS Label2 (L3VNI)
  - Route Target for IP-VRF
  - Router MAC
- IP Attributes are Optional
- Populated through ARP / ND

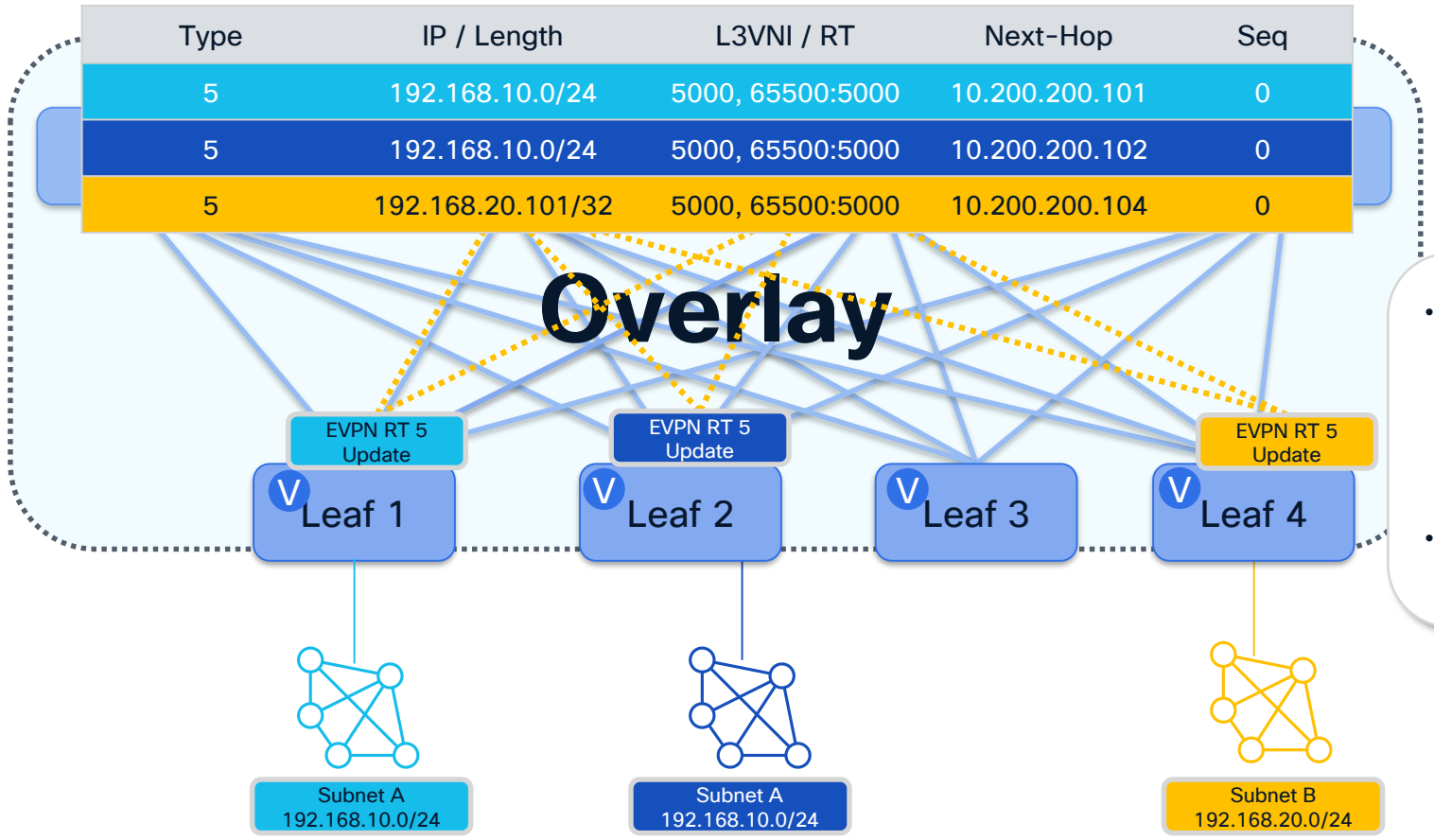
EVPN Control Plane

Subnet Advertisements



EVPN Control Plane

Subnet Advertisements

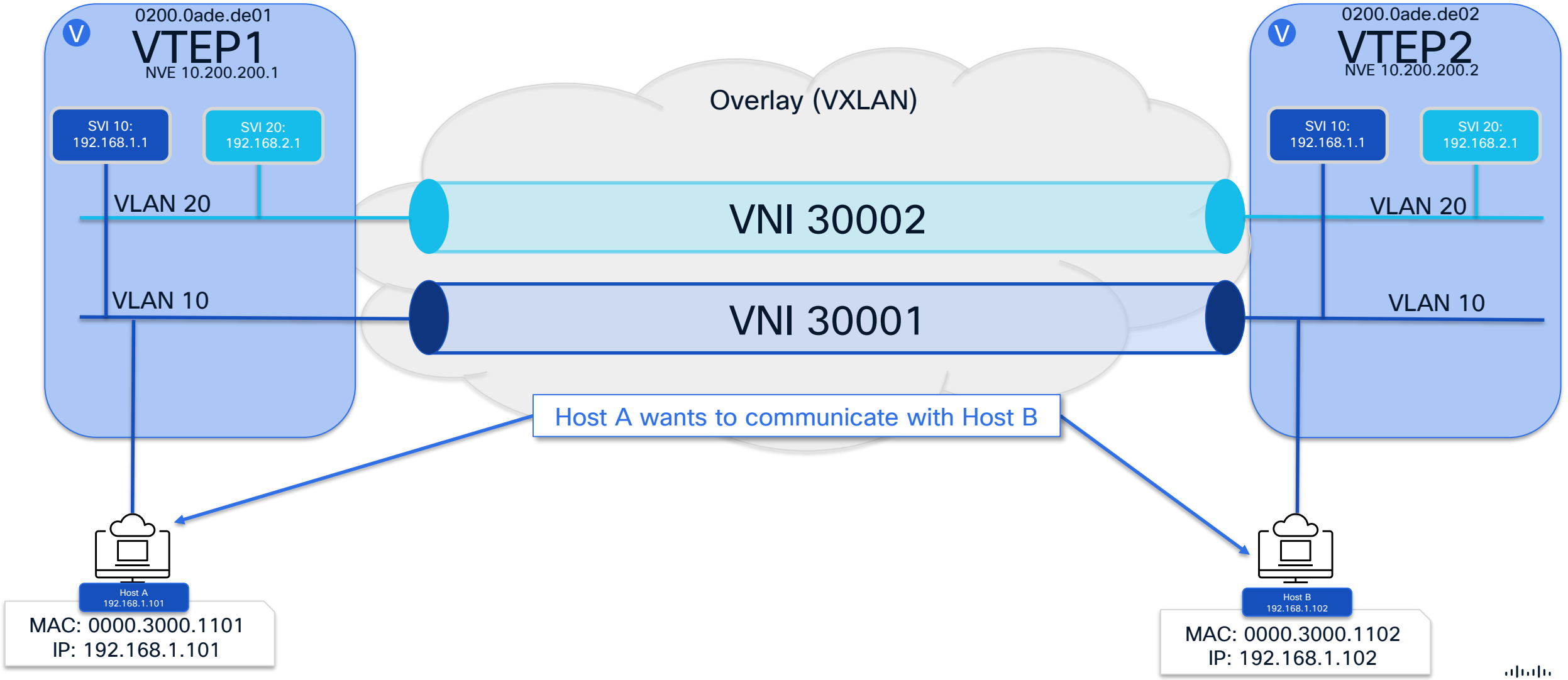


- Internal and External Subnet Prefixes (Route Type 5)
  - IP Prefix
  - MPLS Label (L3VNI)
  - Route Target for IP-VRF
  - Router MAC
- Populated through External Routing Protocol

Packet Walk
Host Communication in the Same VLAN/VNI

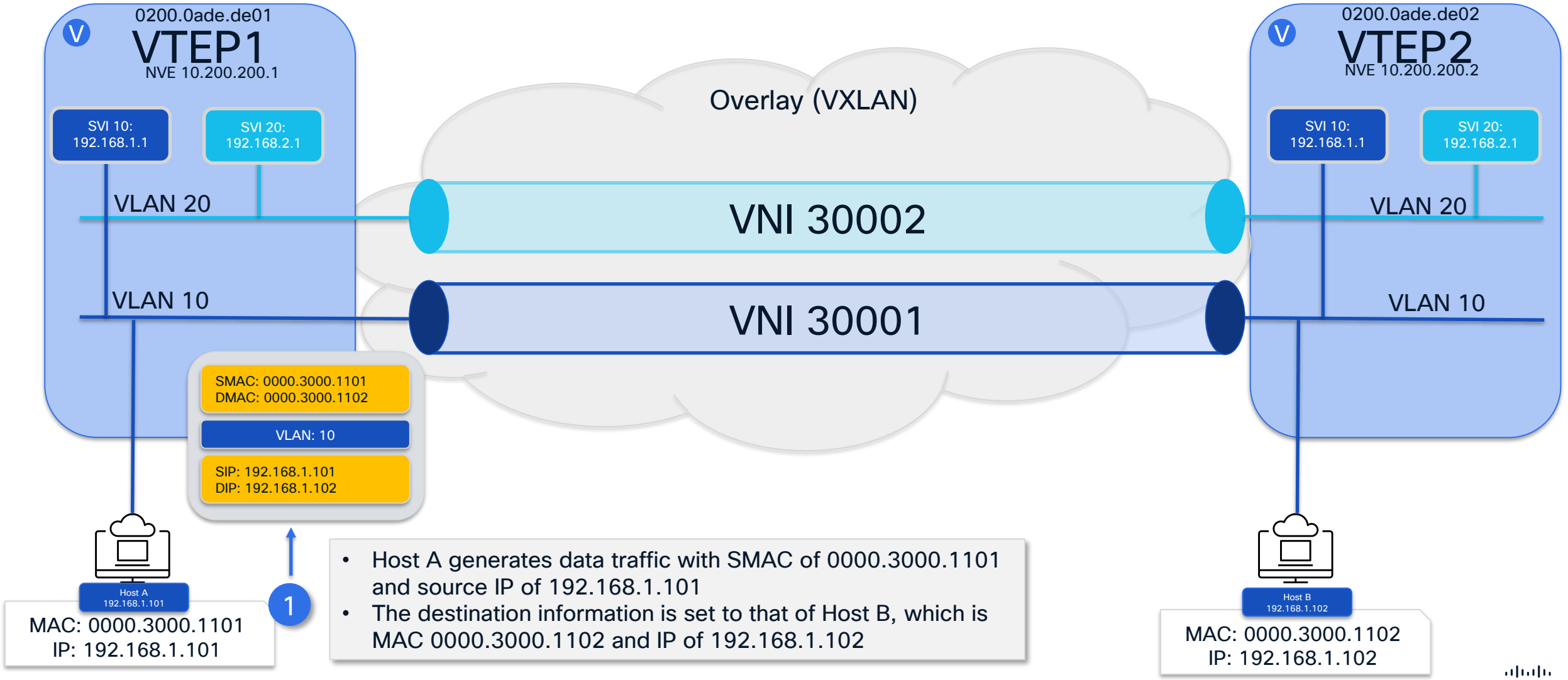
Communication between hosts in the same VLAN/VNI

Packet Walk – ARP Cache Populated



Communication between hosts in the same VLAN/VNI

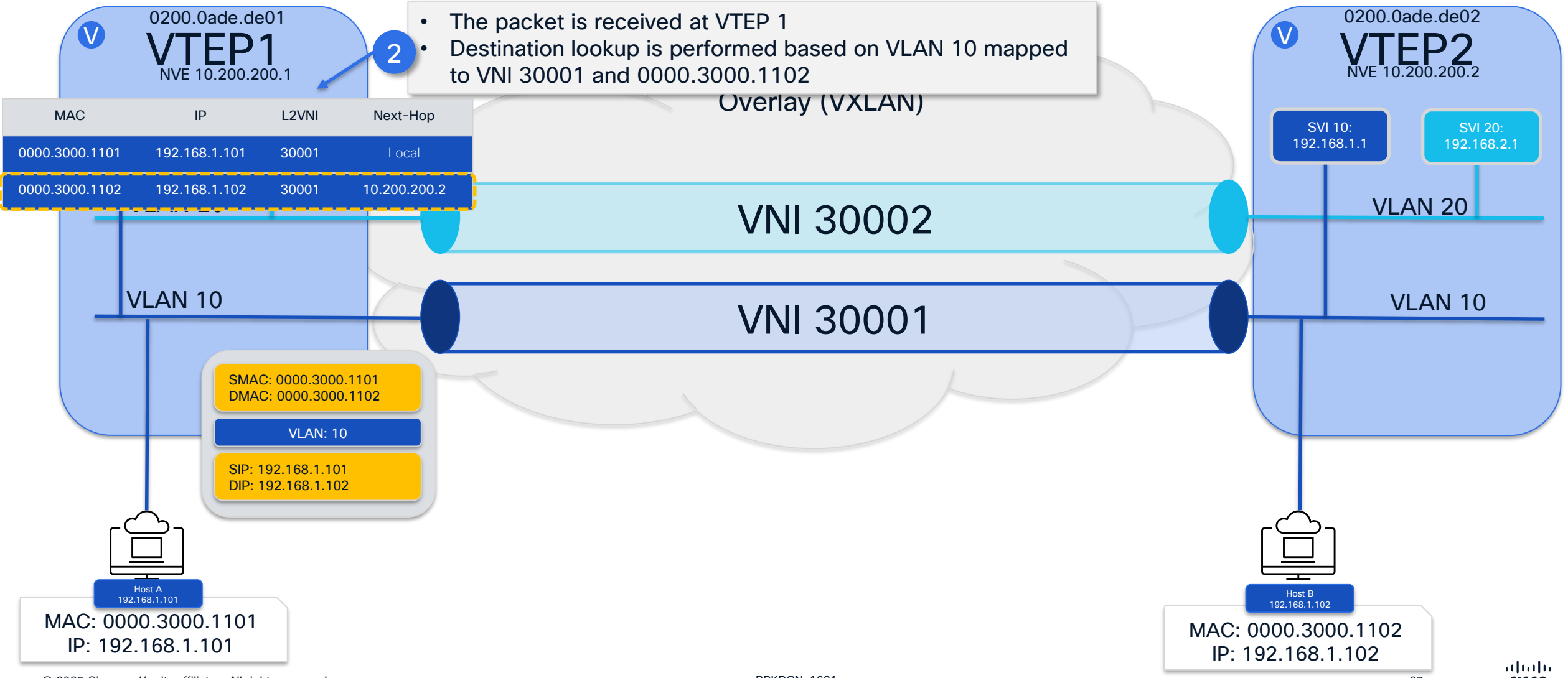
Packet Walk – ARP Cache Populated



- Host A generates data traffic with SMAC of 0000.3000.1101 and source IP of 192.168.1.101
- The destination information is set to that of Host B, which is MAC 0000.3000.1102 and IP of 192.168.1.102

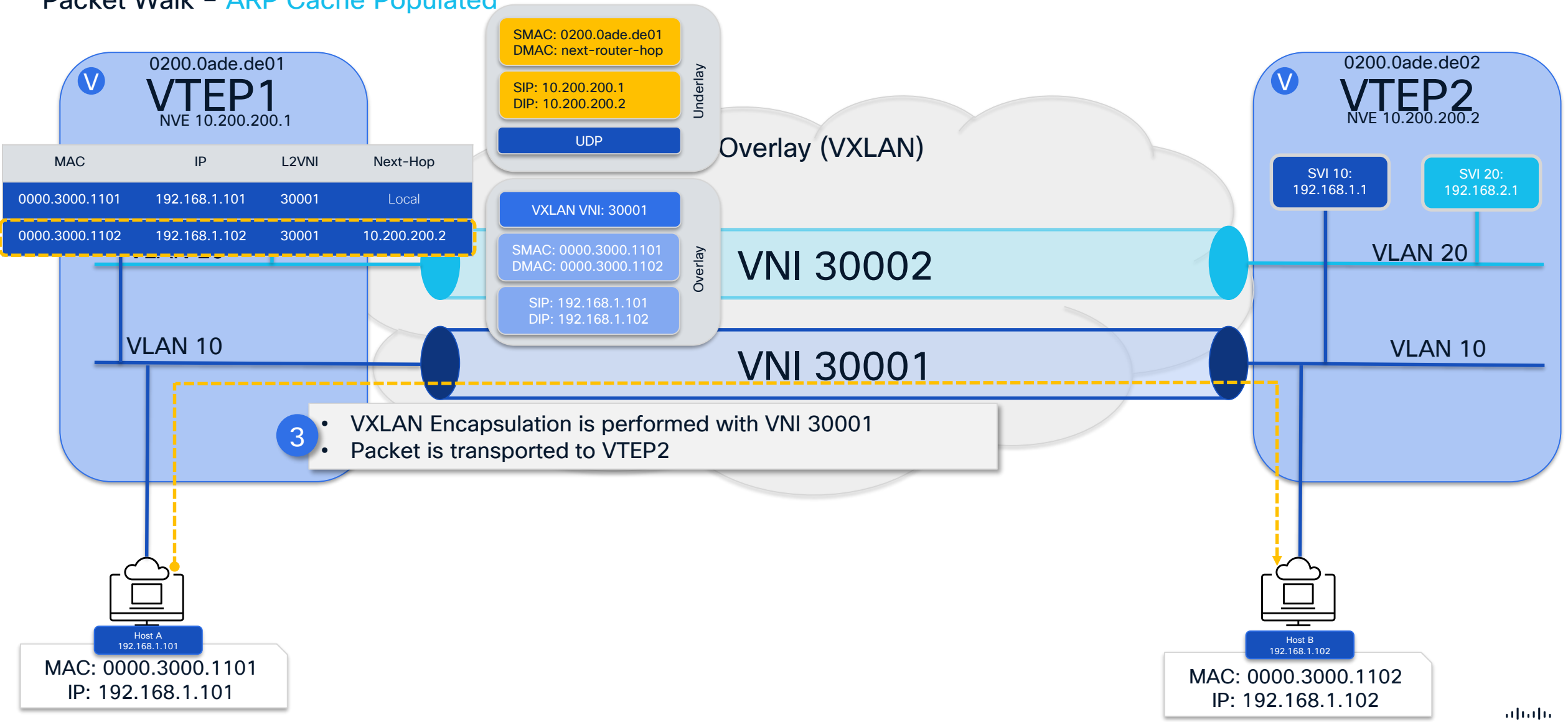
Communication between hosts in the same VLAN/VNI

Packet Walk – ARP Cache Populated



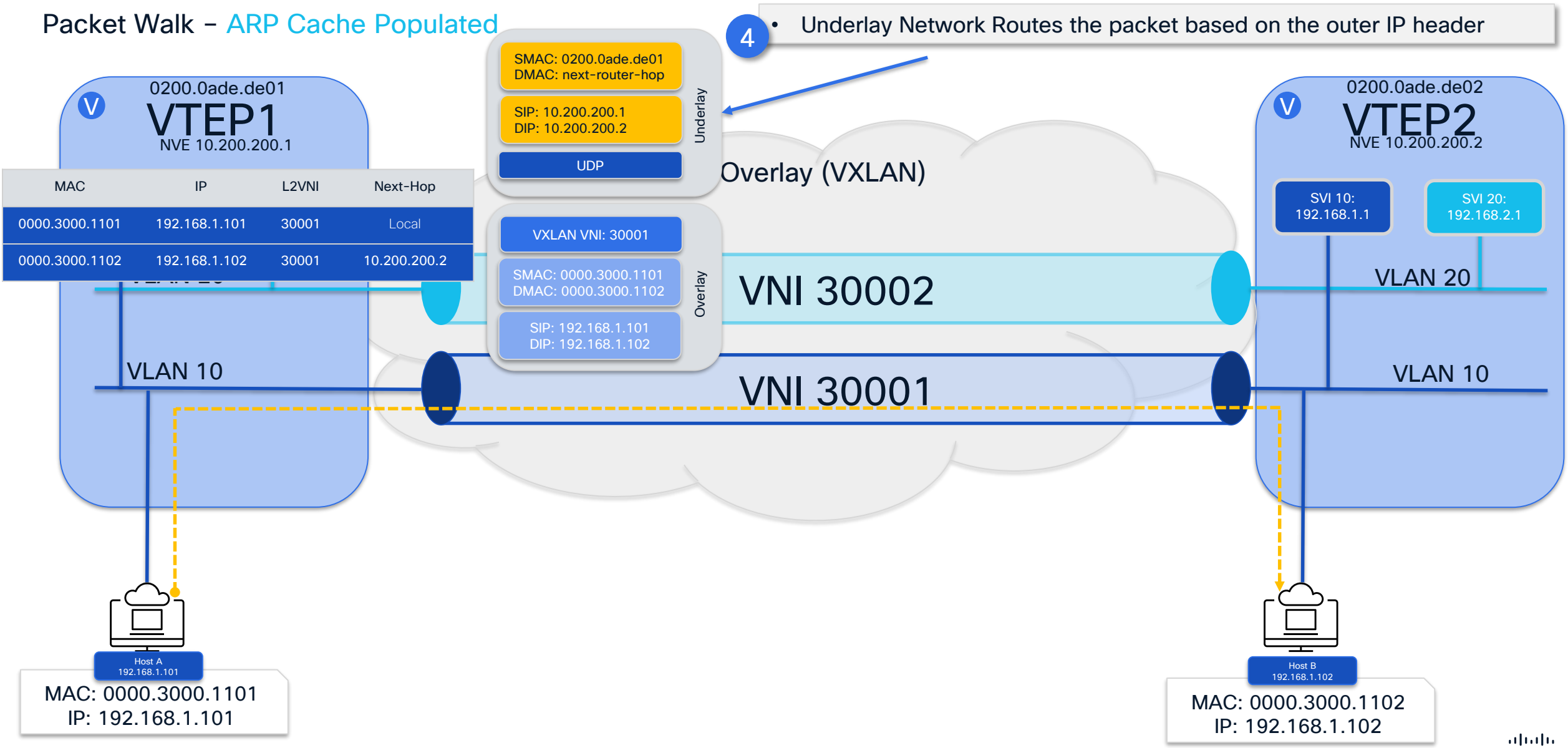
Communication between hosts in the same VLAN/VNI

Packet Walk – ARP Cache Populated



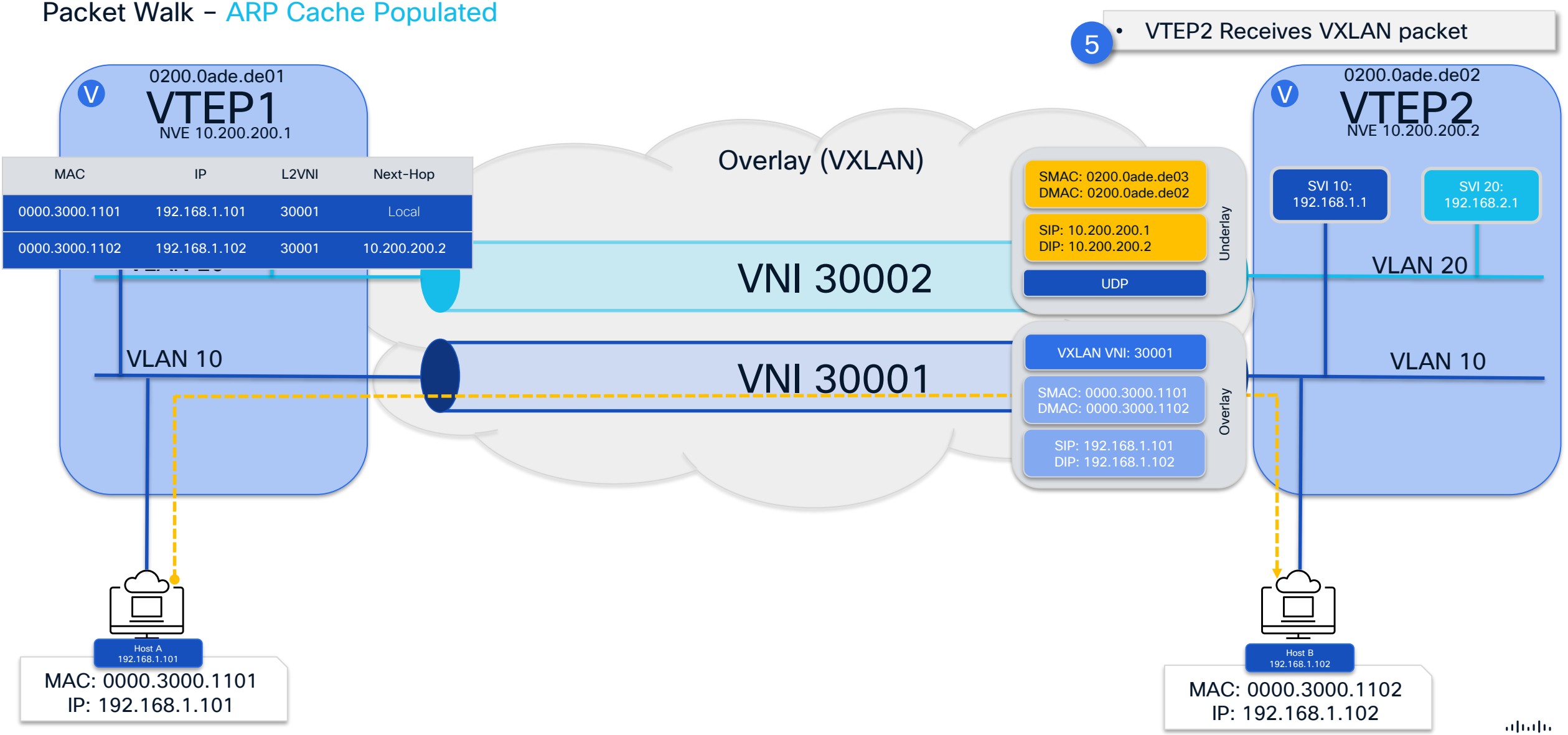
Communication between hosts in the same VLAN/VNI

Packet Walk – ARP Cache Populated



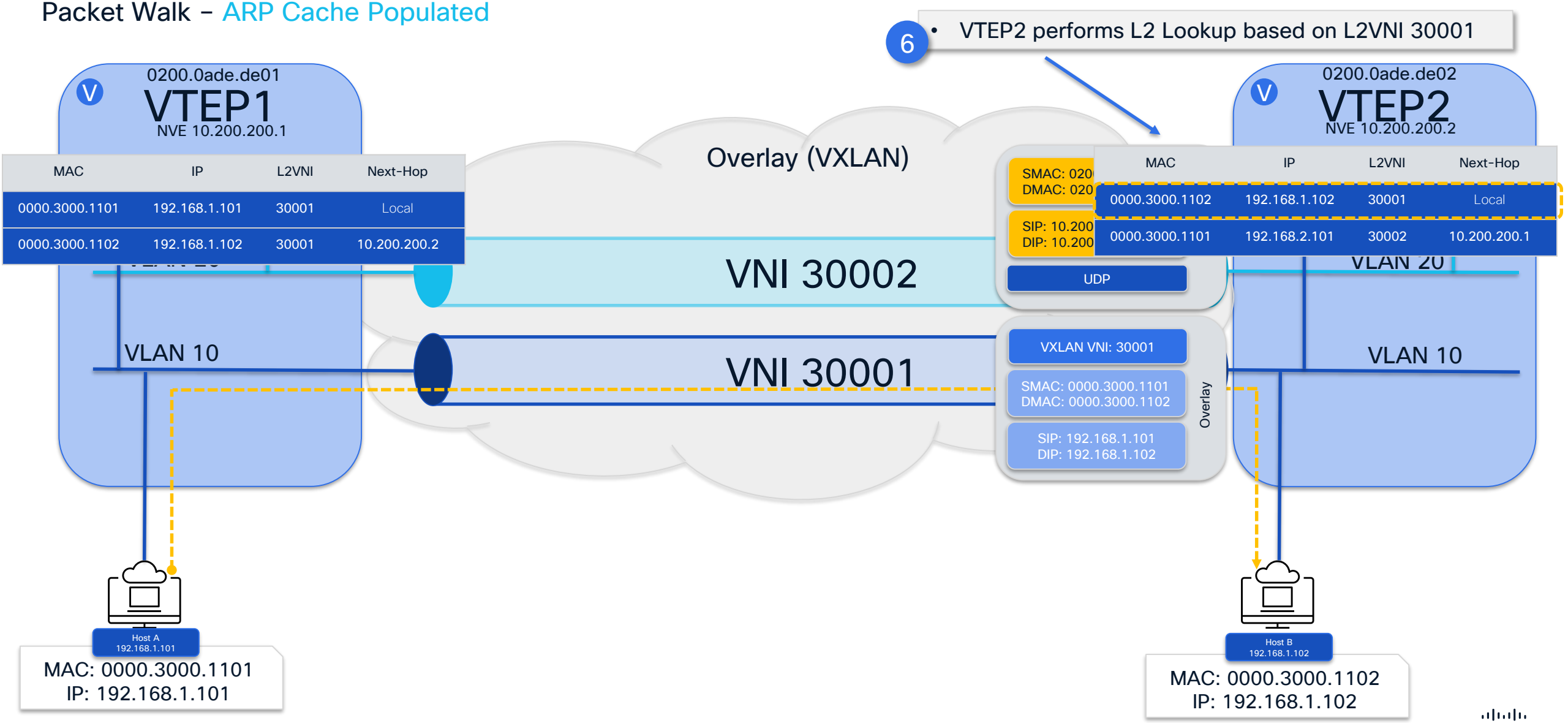
Communication between hosts in the same VLAN/VNI

Packet Walk – ARP Cache Populated



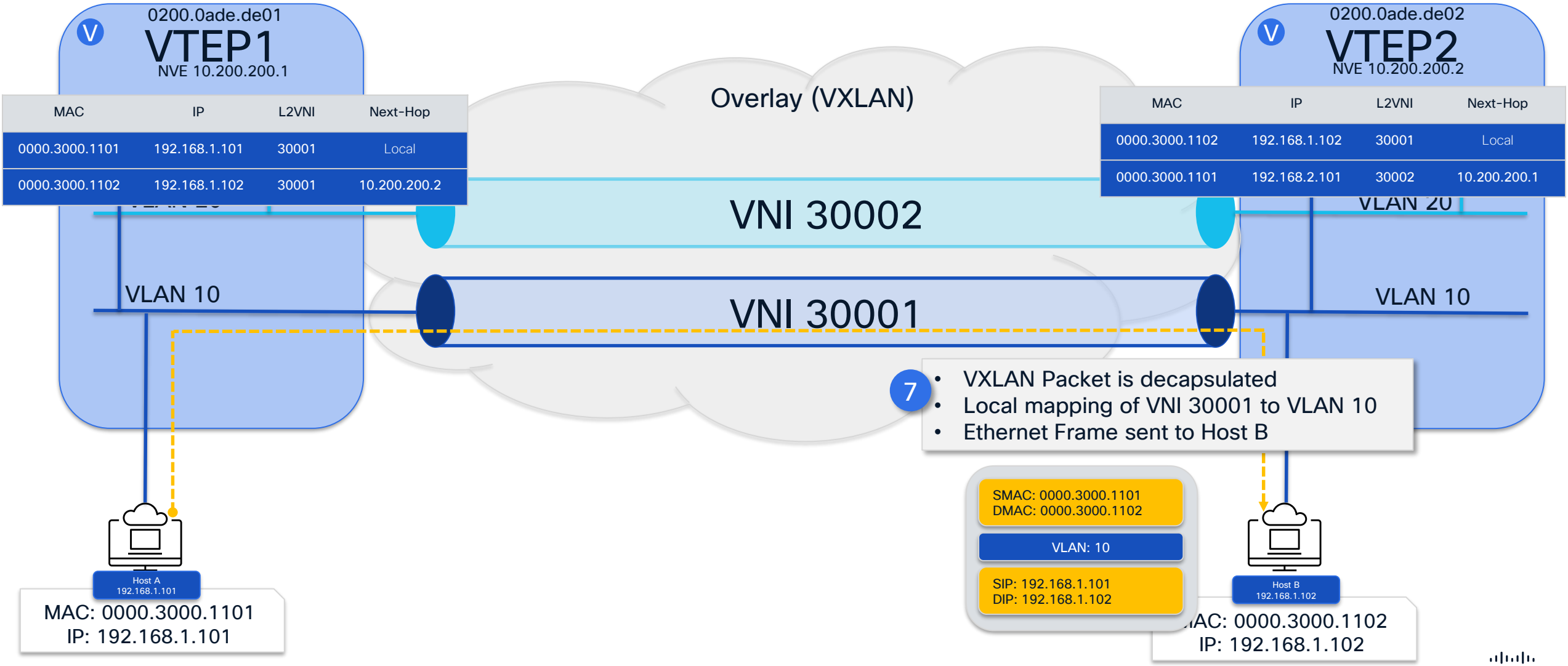
Communication between hosts in the same VLAN/VNI

Packet Walk – ARP Cache Populated



Communication between hosts in the same VLAN/VNI

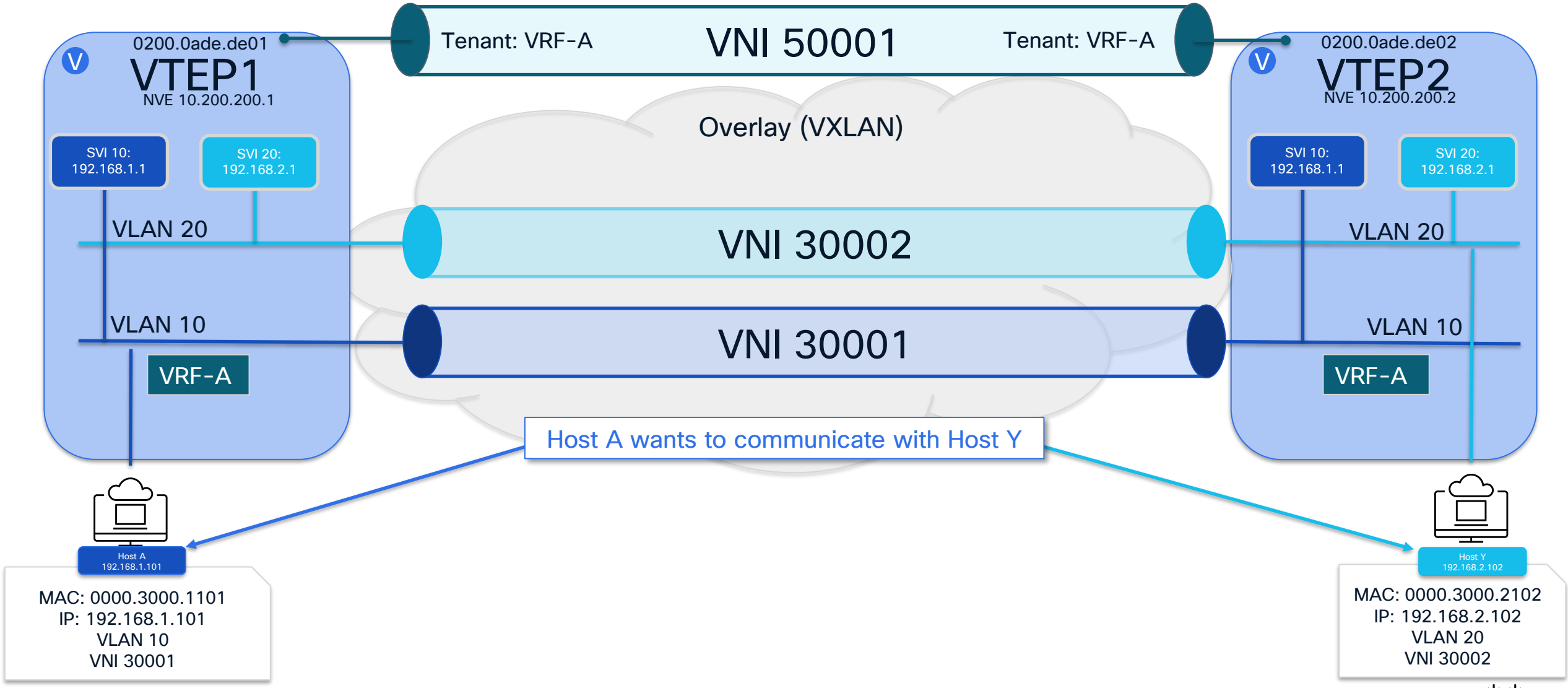
Packet Walk – ARP Cache Populated



Packet Walk
Host Communication in
Different VLAN/VNI

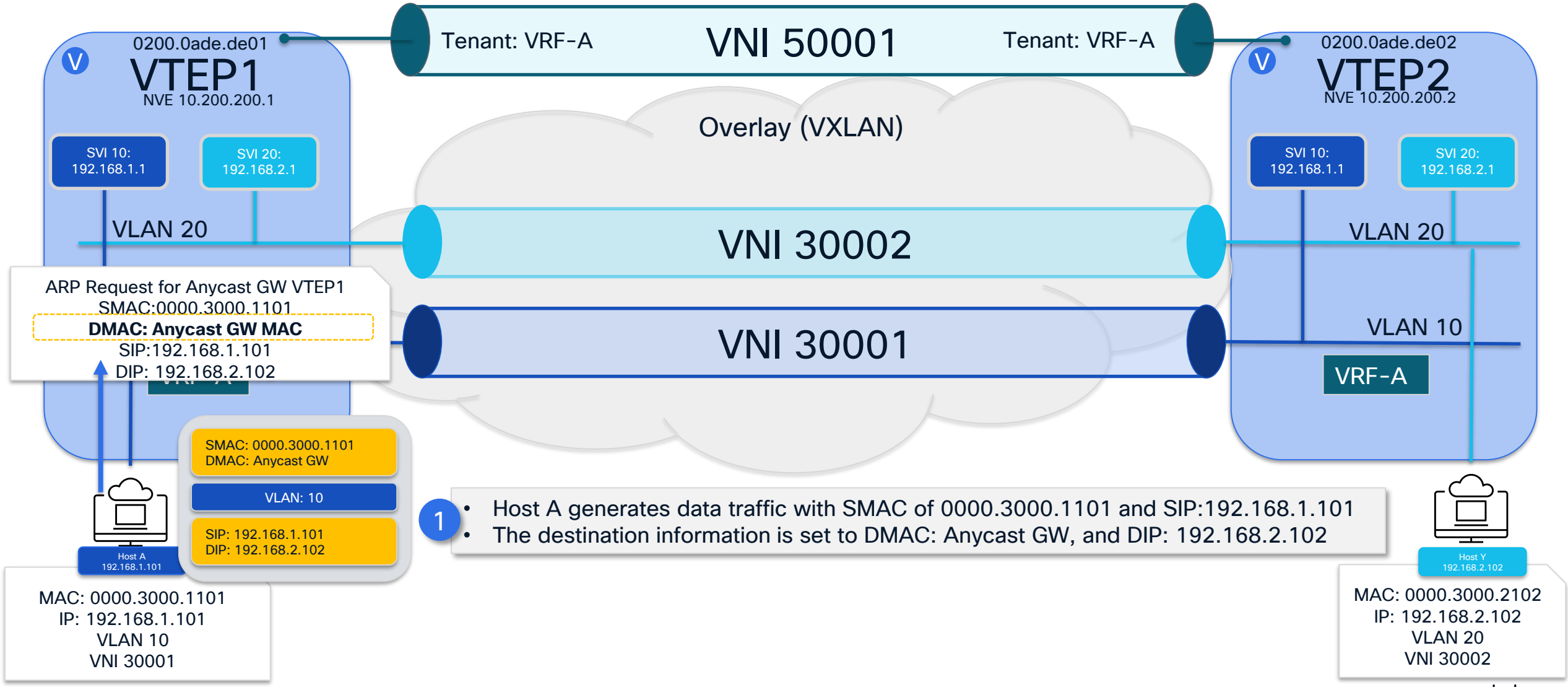
Communication between hosts in different VLAN/VNI (L3)

Packet Walk



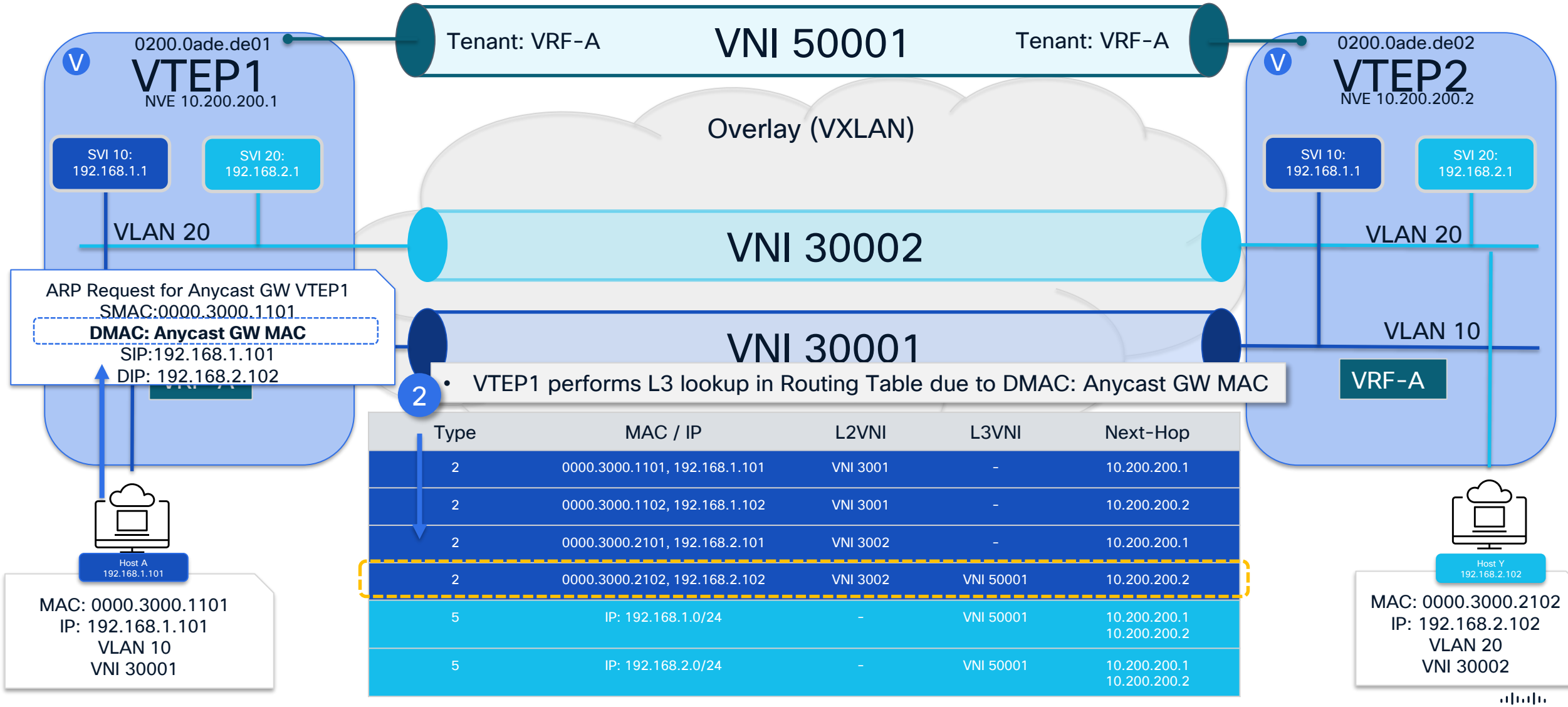
Communication between hosts in different VLAN/VNI (L3)

Packet Walk



Communication between hosts in different VLAN/VNI (L3)

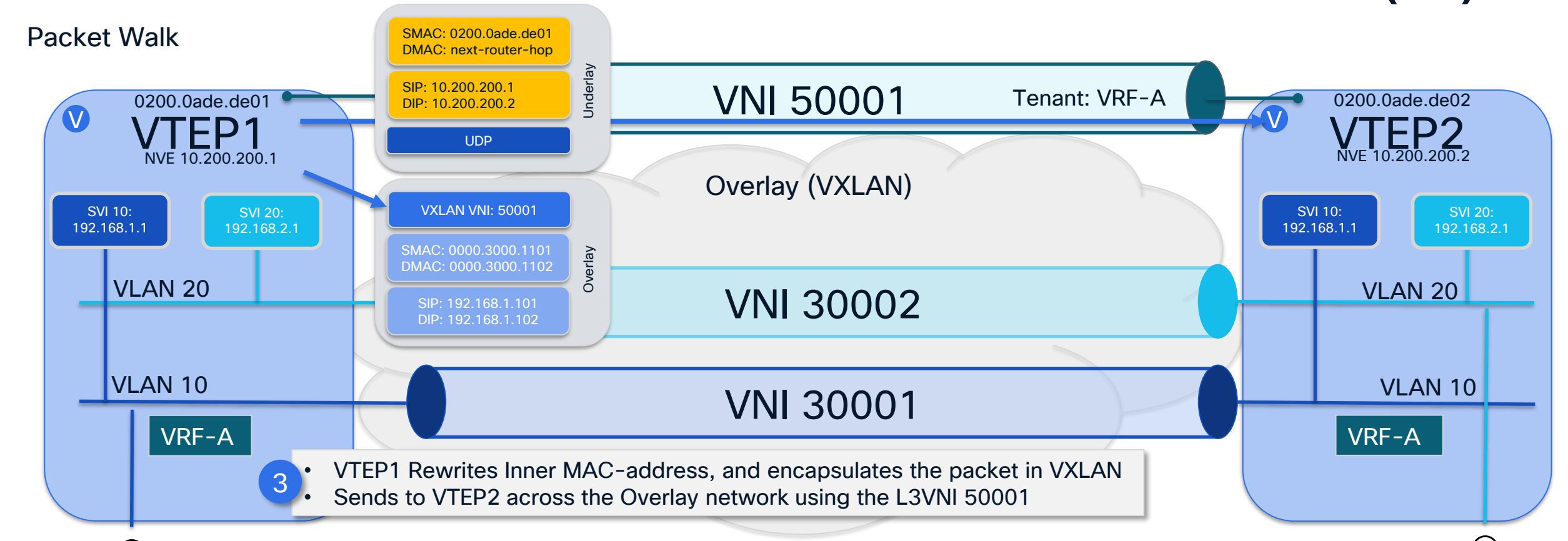
Packet Walk



Type	MAC / IP	L2VNI	L3VNI	Next-Hop
2	0000.3000.1101, 192.168.1.101	VNI 3001	-	10.200.200.1
2	0000.3000.1102, 192.168.1.102	VNI 3001	-	10.200.200.2
2	0000.3000.2101, 192.168.2.101	VNI 3002	-	10.200.200.1
2	0000.3000.2102, 192.168.2.102	VNI 3002	VNI 50001	10.200.200.2
5	IP: 192.168.1.0/24	-	VNI 50001	10.200.200.1 10.200.200.2
5	IP: 192.168.2.0/24	-	VNI 50001	10.200.200.1 10.200.200.2

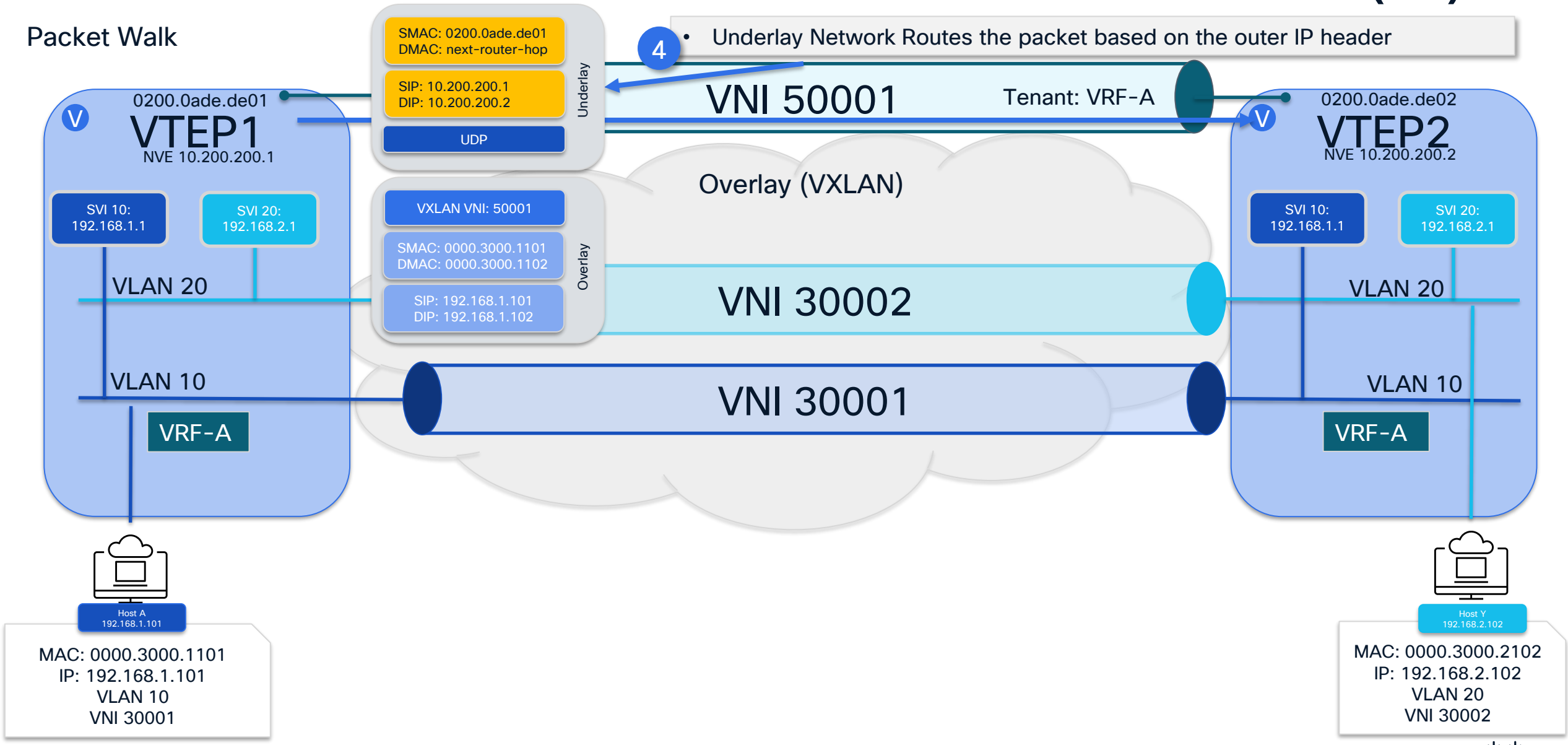
Communication between hosts in different VLAN/VNI (L3)

Packet Walk



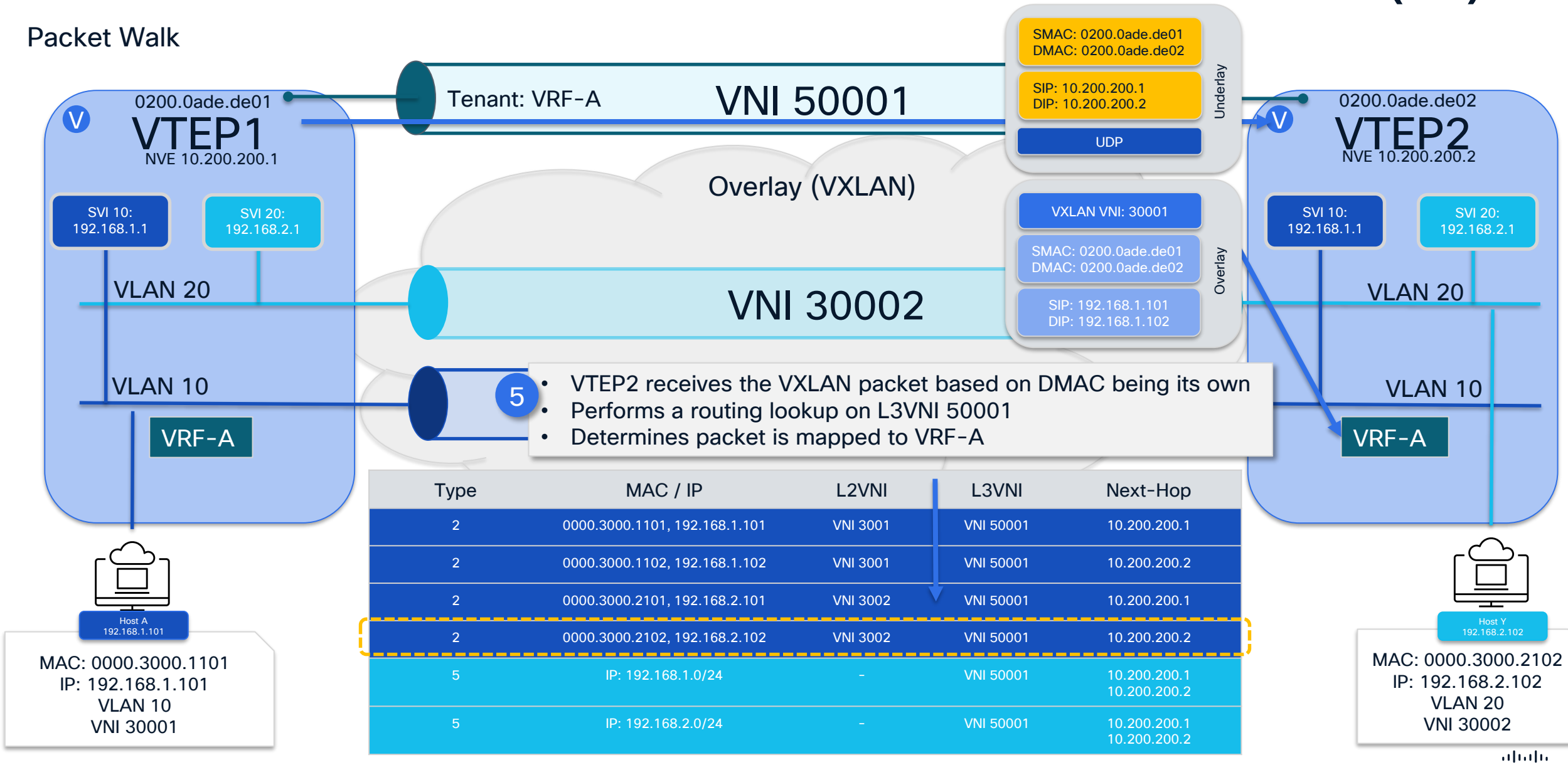
Communication between hosts in different VLAN/VNI (L3)

Packet Walk



Communication between hosts in different VLAN/VNI (L3)

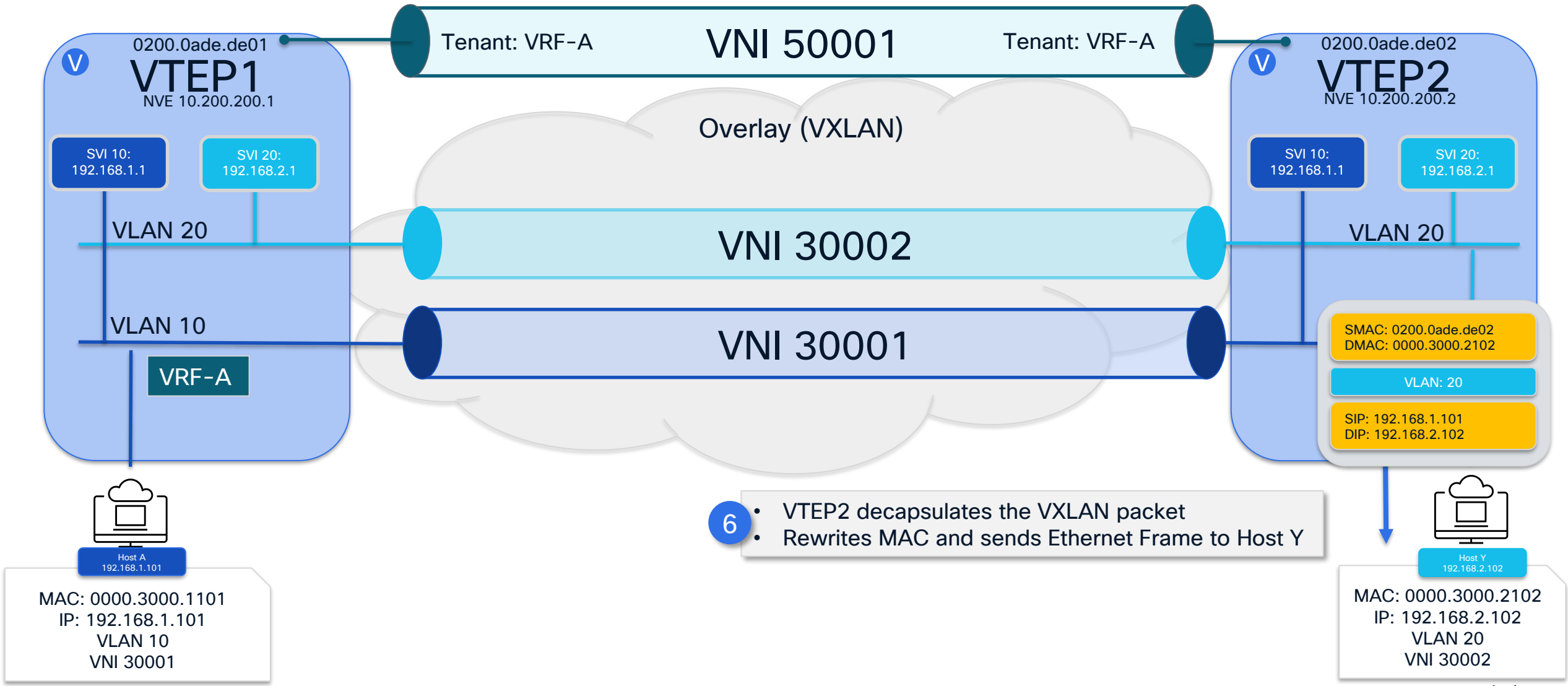
Packet Walk



Type	MAC / IP	L2VNI	L3VNI	Next-Hop
2	0000.3000.1101, 192.168.1.101	VNI 3001	VNI 50001	10.200.200.1
2	0000.3000.1102, 192.168.1.102	VNI 3001	VNI 50001	10.200.200.2
2	0000.3000.2101, 192.168.2.101	VNI 3002	VNI 50001	10.200.200.1
2	0000.3000.2102, 192.168.2.102	VNI 3002	VNI 50001	10.200.200.2
5	IP: 192.168.1.0/24	-	VNI 50001	10.200.200.1 10.200.200.2
5	IP: 192.168.2.0/24	-	VNI 50001	10.200.200.1 10.200.200.2

Communication between hosts in different VLAN/VNI (L3)

Packet Walk



Conclusion

Control-Plane Comparison

EVPN - Ethernet VPN (RFC8365)

- A Better Ethernet over IP
  - No Spanning-Tree (terminates at NVE)
  - Endpoint Learning is based on BGP exchange (EVPN uses BGP)
  - Provides Integrated Routing & Bridging (IRB)
    - Distributed Anycast Gateway for Default Gateway
    - Uses a Layer-3 VPN approach like MPLS L3VPN
  - And there is much more in EVPN!

https://datatracker.ietf.org/doc/html/rfc8365

Flood&Learn (RFC7348)

- Ethernet over IP
  - No Spanning-Tree (terminates at NVE)
  - Endpoint Learning is based on Flood and Learn (it's in the name)
  - Requires Extra Work for Routing
    - FHRP for Default Gateway
    - Over-the-Top VRF-lite for Prefix Routing (or use the Underlay?!)

https://datatracker.ietf.org/doc/html/rfc7348

Conclusion

- Data Center Network Challenges and Evolution
- Virtual Private Networks and Overlay Taxonomy
- VXLAN enables scalable Data Center Networks
- BGP EVPN as a better Control-Plane
  - Provides a robust control plane enabling multi-tenancy, VM mobility, and optimizes forwarding

Recommended Cisco Live Sessions

- Introduction to NDFC: Simplifying Management of Your Data Center – **BRKDCN-1619**
- Migrate from Your Existing Network to VXLAN EVPN – **BRKDCN-2951**
- Multi-Tier Fabrics: Network Designs for the Modern Data Center – **BRKDCN-2999**
- Simple Leaf/Spine with a Touch of ToR – Network Designs for the Modern Data Center – **BRKDCN-2967**
- VXLAN EVPN: A Deep Dive into Packet Forwarding – **BRKDCN-3966**
- VXLAN BGP EVPN Multi-Site – **BRKDCN-2913**

Recommended Reading

Building Data Centers with VXLAN BGP EVPN: A Cisco NX-OS Perspective

Cisco Nexus 9000 Series NX-OS VXLAN Configuration Guide, Release 10.4(x)

Migrating Classic Ethernet Environments to VXLAN BGP EVPN

Migrating Cisco FabricPath Environments to VXLAN BGP EVPN White Paper

VXLAN EVPN Multi-Site Design and Deployment White Paper

VXLAN Design with Cisco Nexus 9300 Platform Switches

VXLAN Network with MP-BGP EVPN Control Plane Design Guide

Complete Your Session Evaluations

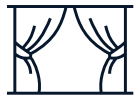
Complete a minimum of 4 session surveys and the Overall Event Survey to be entered in a drawing to win 1 of 5 full conference passes to Cisco Live 2026.

Earn 100 points per survey completed and compete on the Cisco Live Challenge leaderboard.

Level up and earn exclusive prizes!

Complete your surveys in the Cisco Live mobile app.

Continue your education



Visit the Cisco Showcase for related demos

Book your one-on-one Meet the Engineer meeting

Attend the interactive education with DevNet, Capture the Flag, and Walk-in Labs

Visit the On-Demand Library for more sessions at www.CiscoLive.com/on-demand

Contact me at: rlicon@cisco.com

Thank you

cisco Live !


CISCO