
cisco Live!
Let's go
#CiscoLive

The bridge to possible

Introduction to VXLAN

The Future Path of Your Data Center

Richard Licon
Principal Technical Marketing Engineer

BRKDCN-1621

cisco Live!

Cisco Webex App

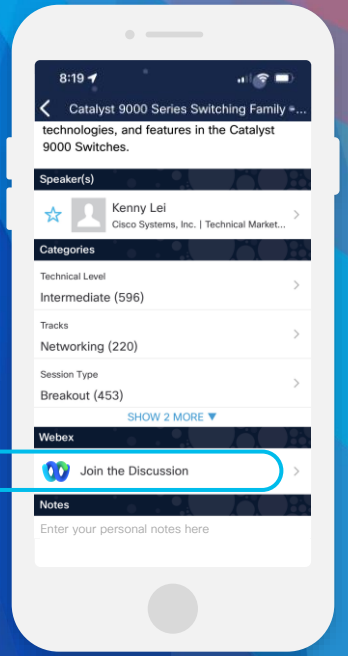
Questions?

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 9, 2023.



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

Session Abstract

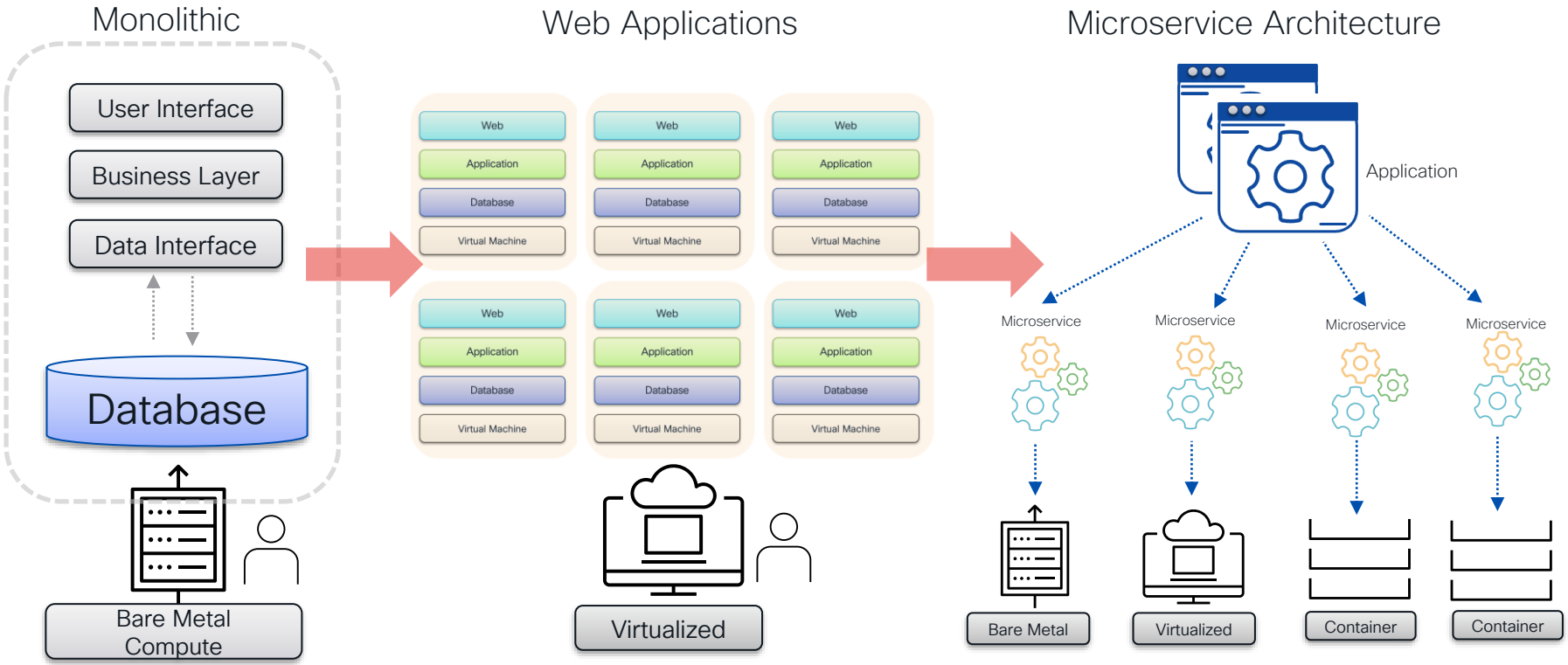
- The need for flexibility in the data center opens the door to network overlays. Several standards have been proposed and implemented to enable virtualized networks to be layered over a physical network. The data center network needs to be flexible enough to support workloads that can move from one 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. Let's learn more about VXLAN and delve deeper into it.
- This session aims at covering what is VXLAN, how it really works, and what problems does it solve. This session will have configuration examples of how we enable a VXLAN Fabric. 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

Agenda

- Evolution of the Data Center
- The Need for network overlays
- What is VXLAN?
- The Need for a Control Plane
- Configuring EVPN VXLAN
- Conclusion

Evolution of the Application Architecture

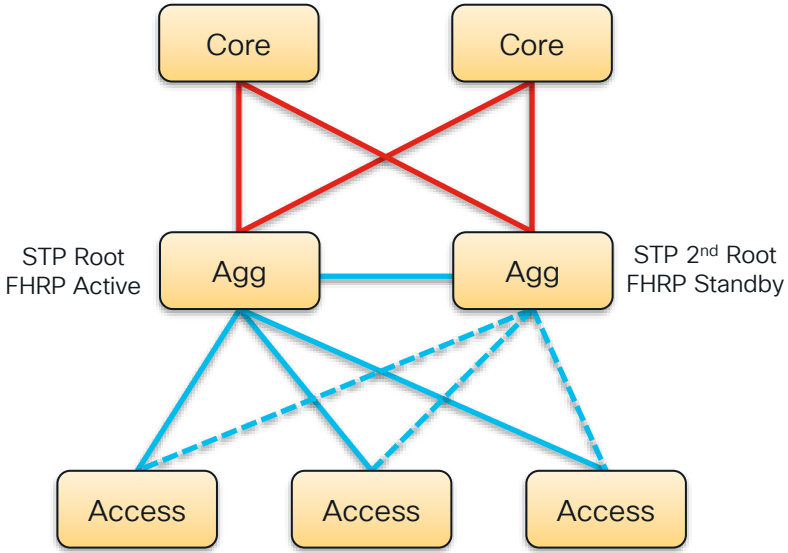
Evolution of the Application Architectures



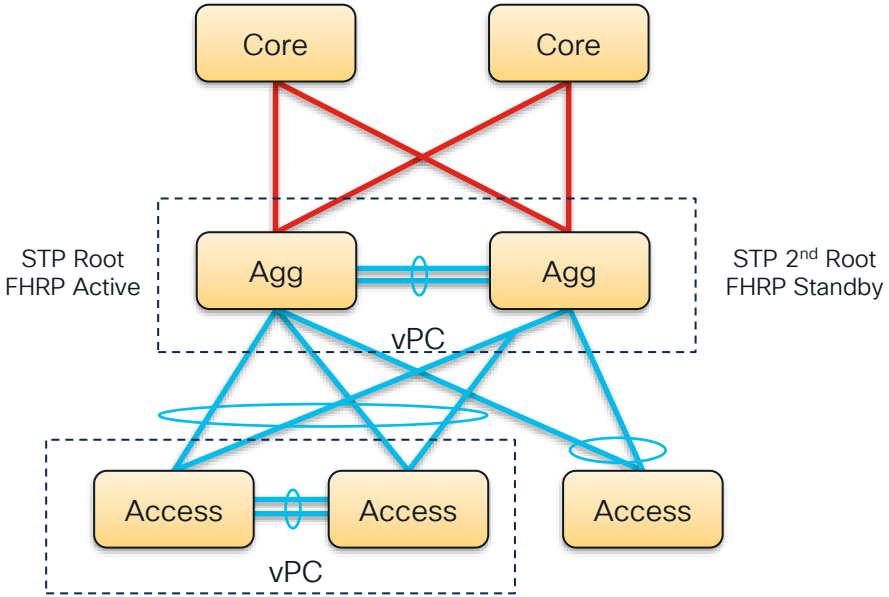
Evolution of the Data Center Network

Evolution of the Data Center Network

Classic Spanning-Tree



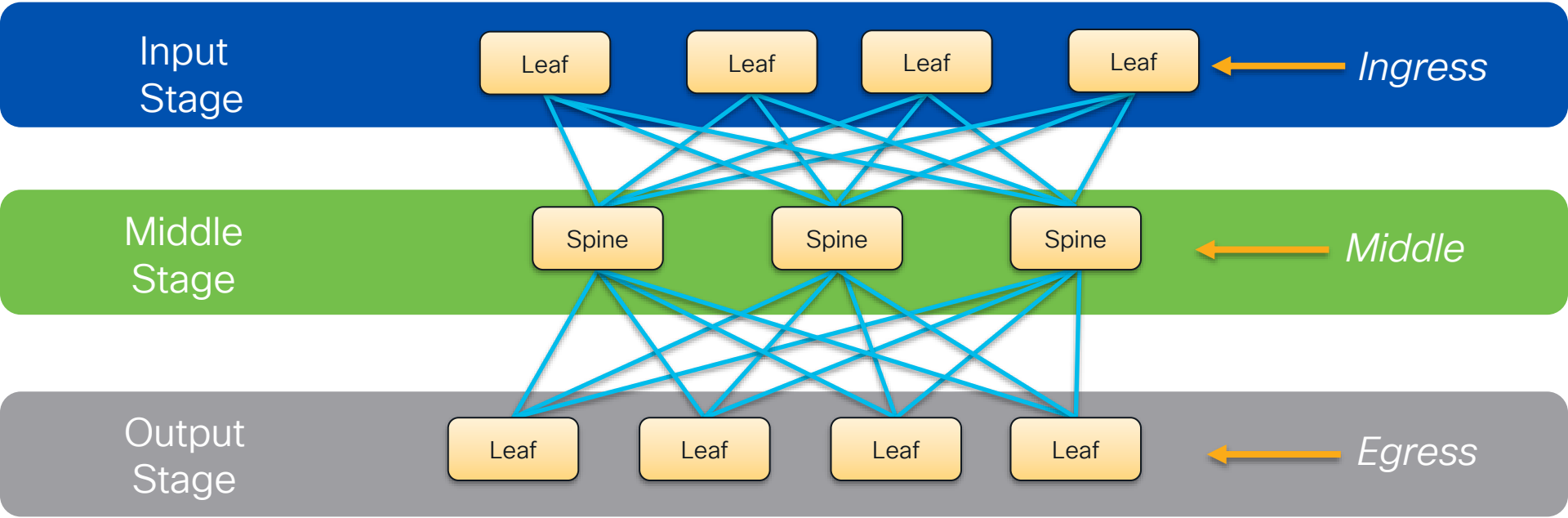
vPC and Spanning-Tree



Evolution of the Data Center Network Architecture: Clos Networks

Evolution of the Data Center Network

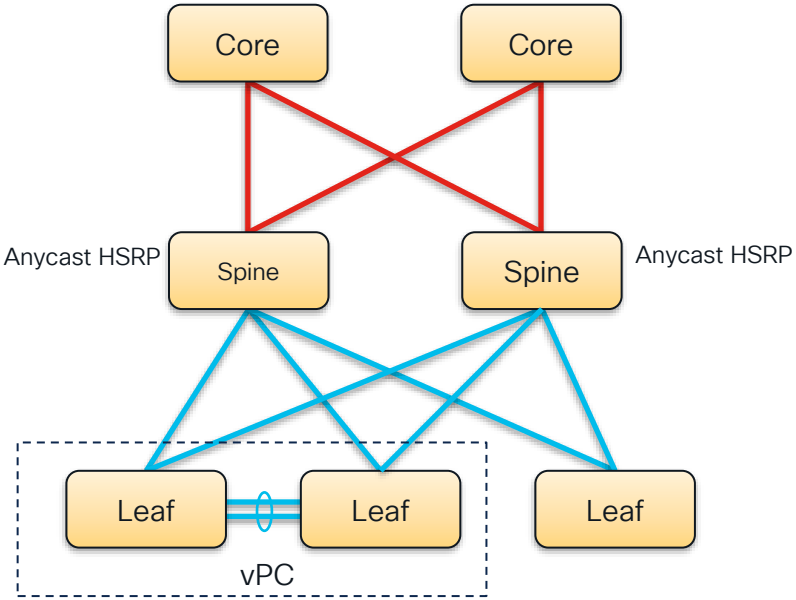
3-Stage Clos Fabric



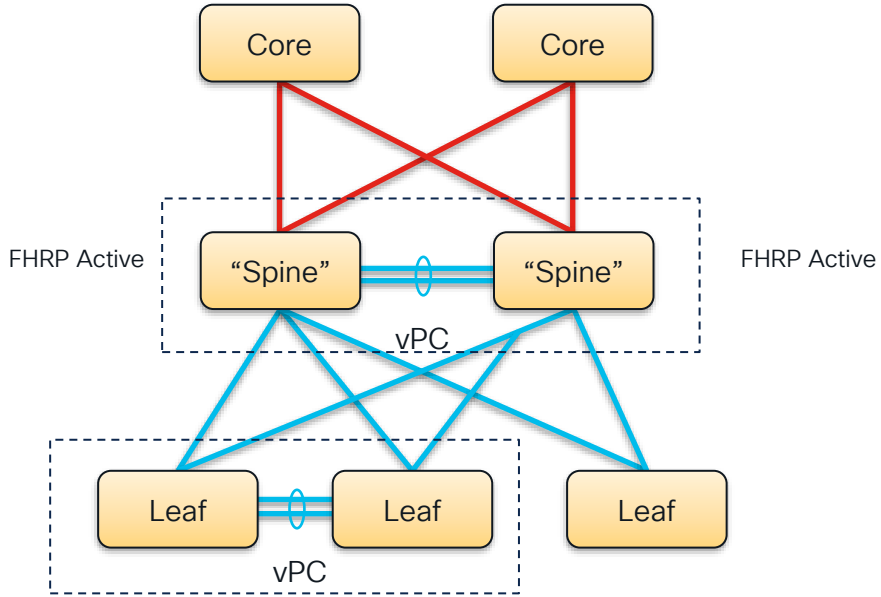
Invented by Edson Erwin in 1938
March 1953: A Study of Non-Blocking Switching Networks. (Clos, Charles)

Evolution of the Data Center Network

FabricPath (MAC-in-MAC)

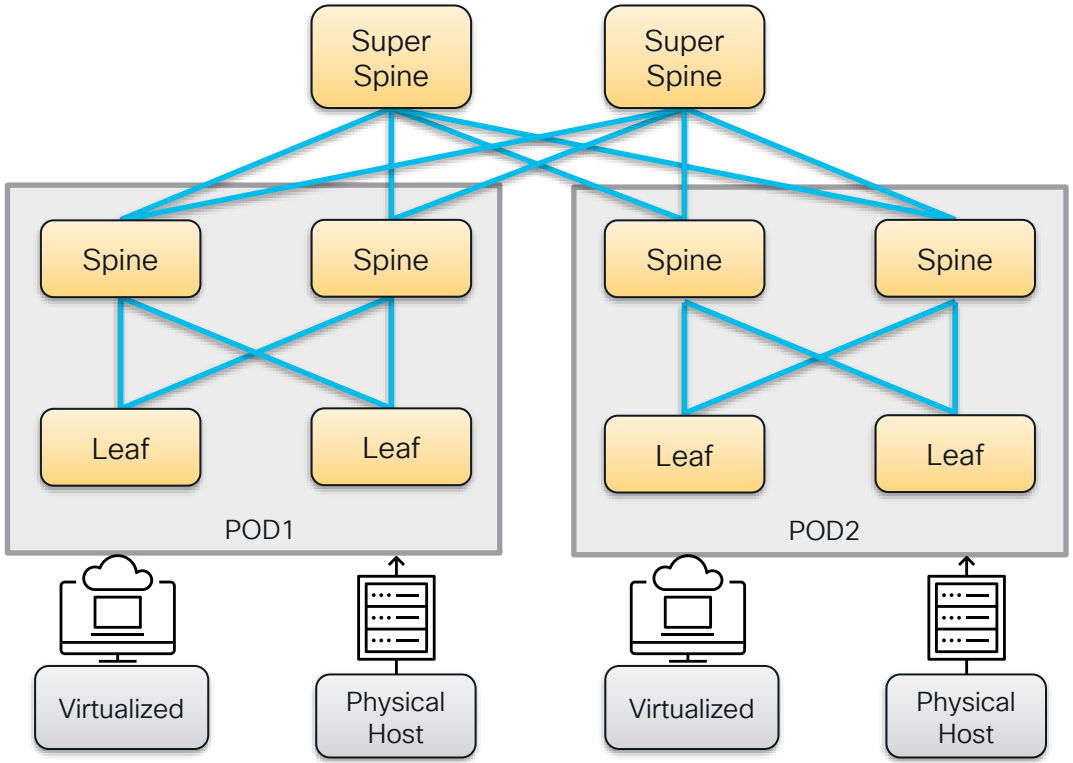


VXLAN Flood & Learn (MAC-in-IP)



Evolution of the Data Center Network

5-Stage Fabric



Data Center Network Challenges

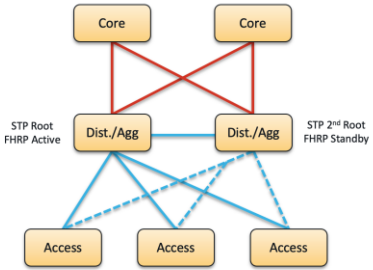
Legacy Architectures

Convergence Time:
STP high convergence time
TCN MAC Flush

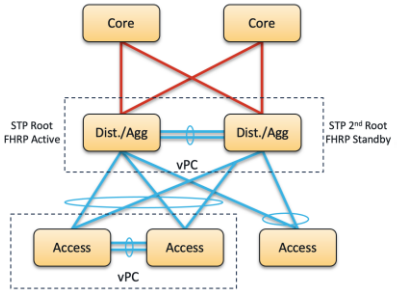
Unused Links:
STP Redundant links in blocked state

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

Classic Spanning-Tree



vPC and Spanning-Tree

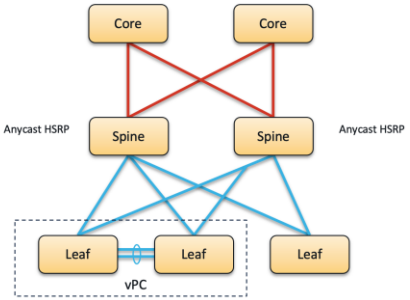


Suboptimal Forwarding:
Path defined by
Root Switch or FHRP Active

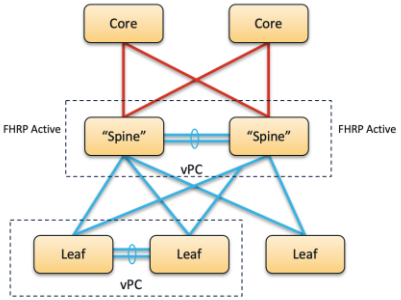
OpEx/CapEx:
Expensive to upgrade
Scale up with large Chassis

Limited Scale:
No Control Plane
Limited Workload Mobility

FabricPath (MAC-in-MAC)



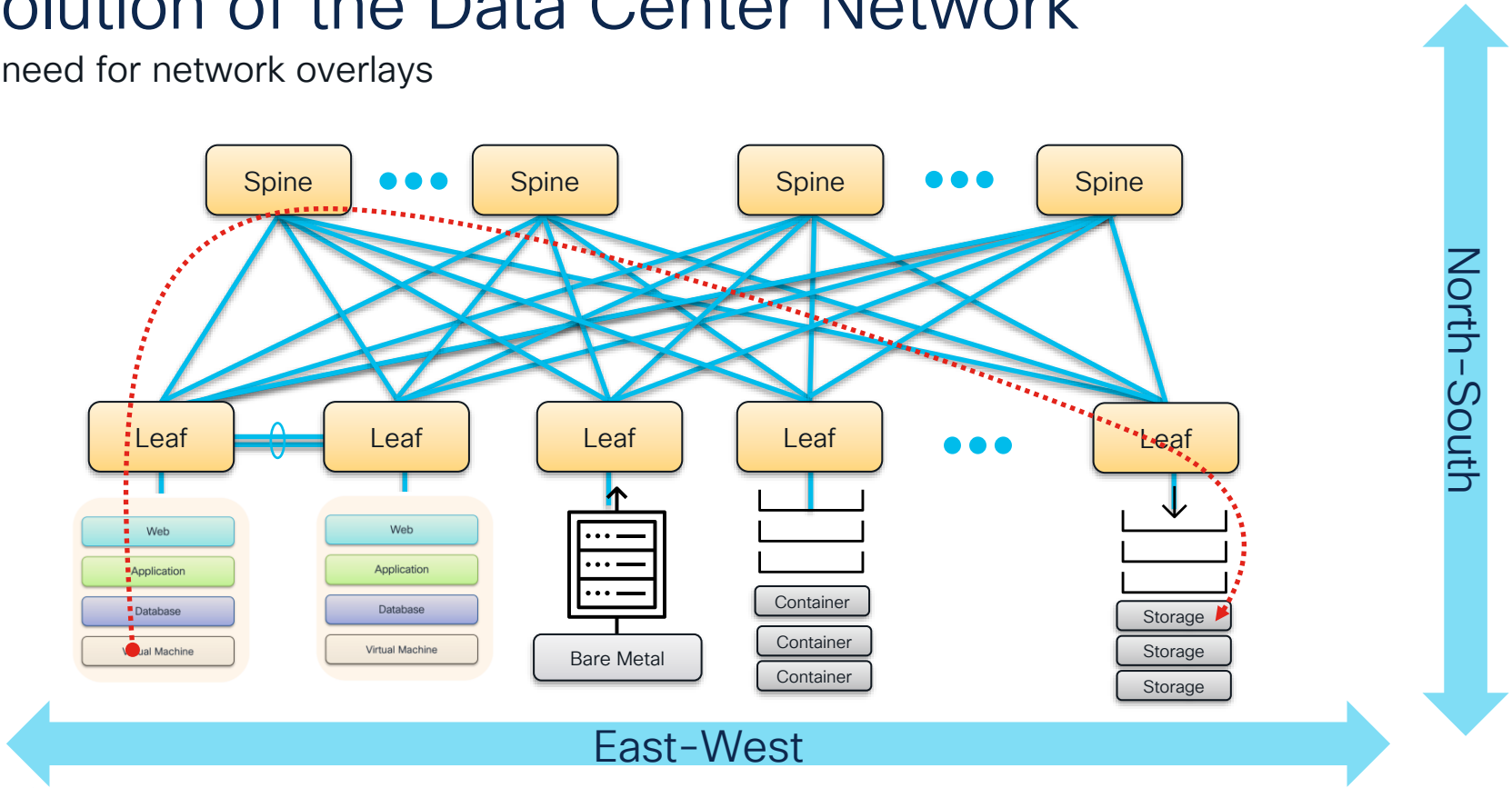
VXLAN Flood&Learn (MAC-in-IP)



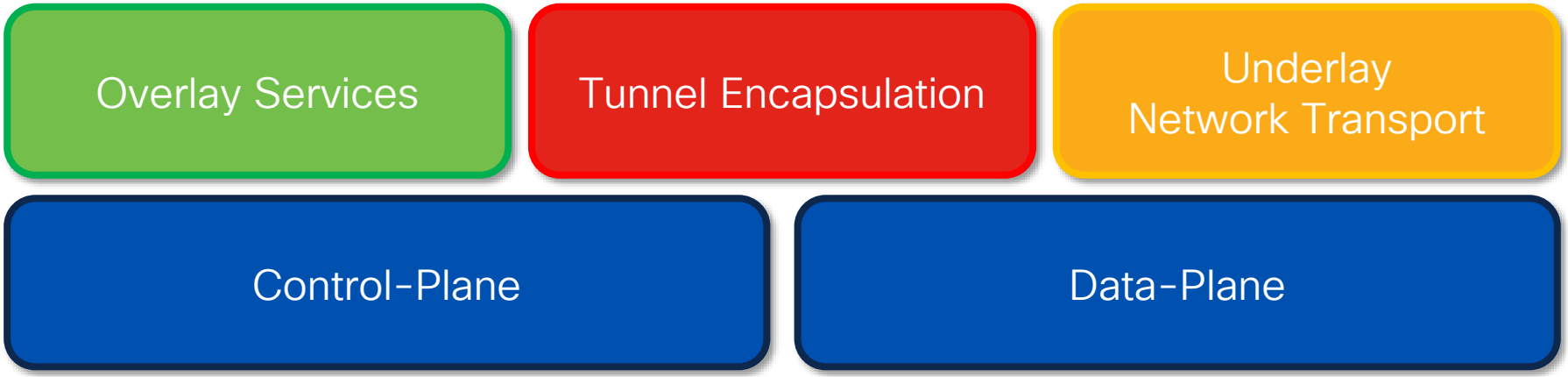
The Need for Network Overlays

Evolution of the Data Center Network

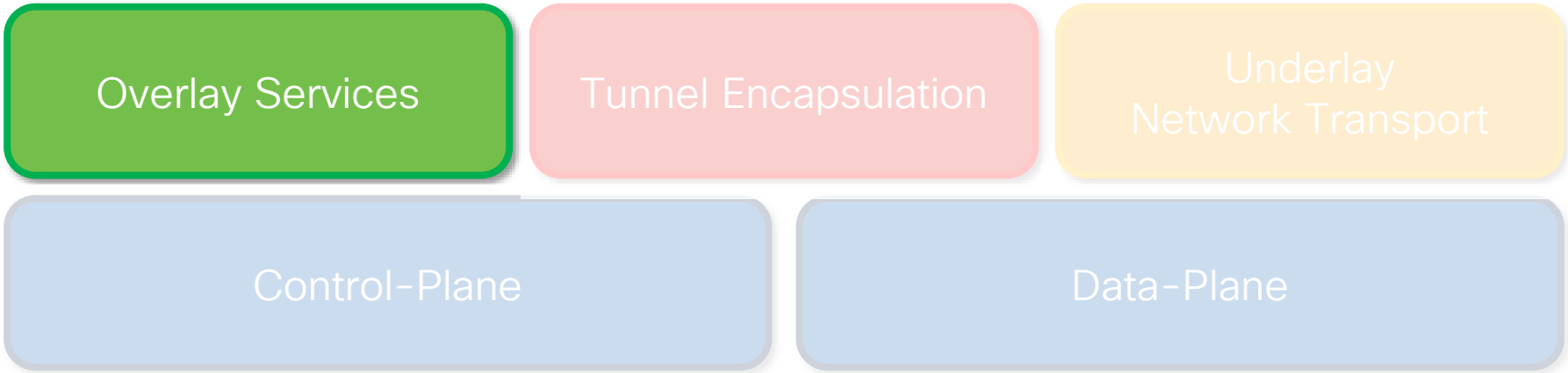
The need for network overlays



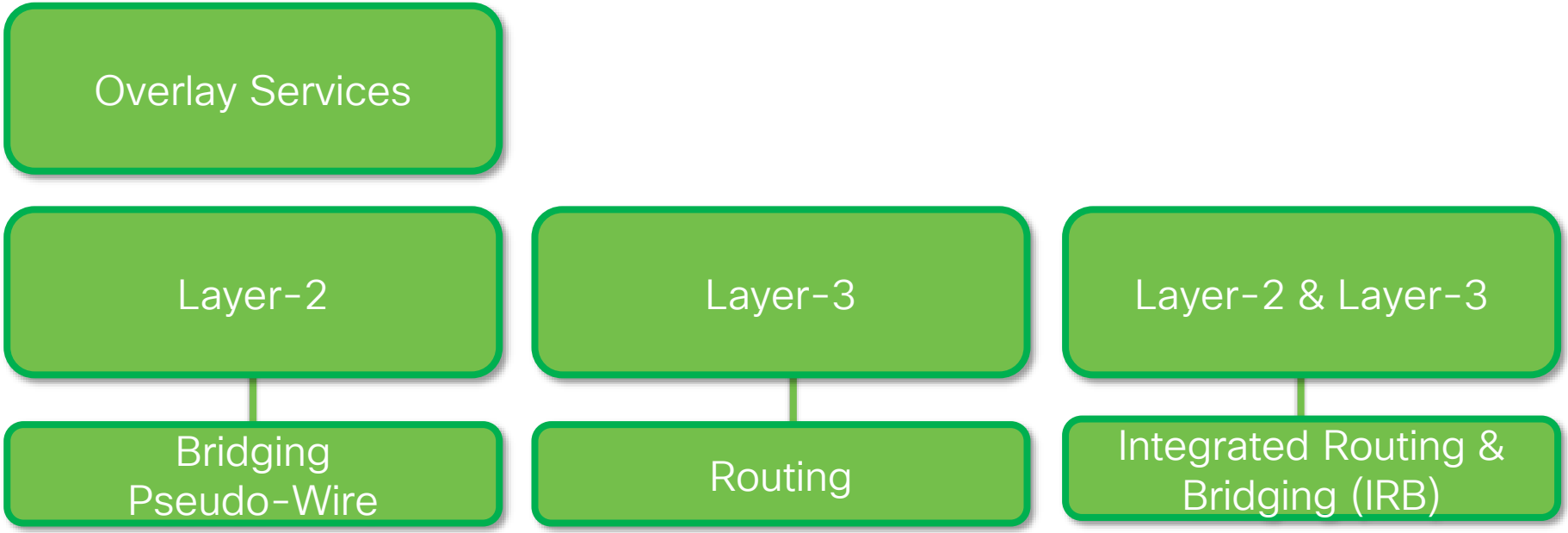
Overlay Taxonomy



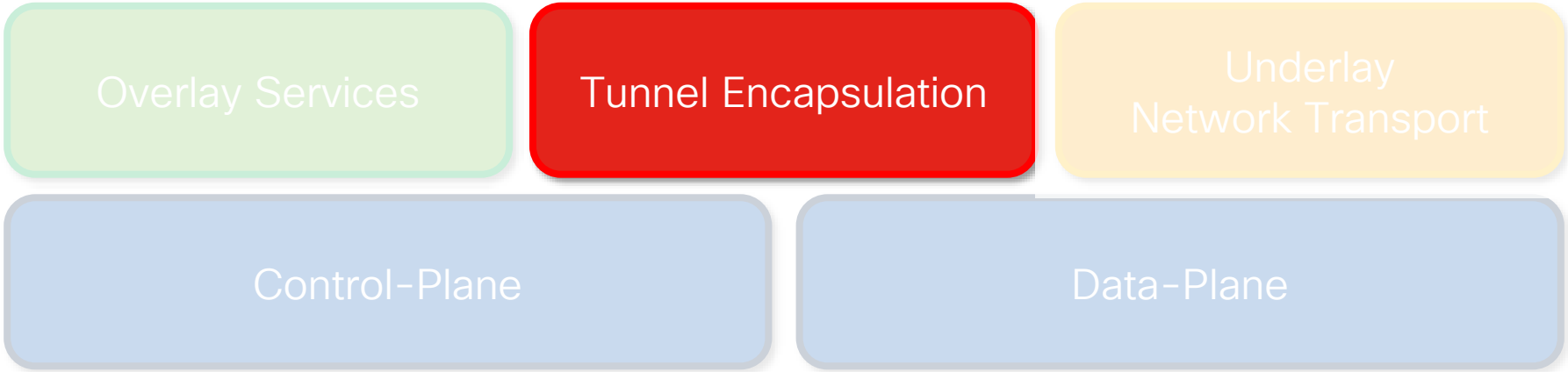
Overlay Services



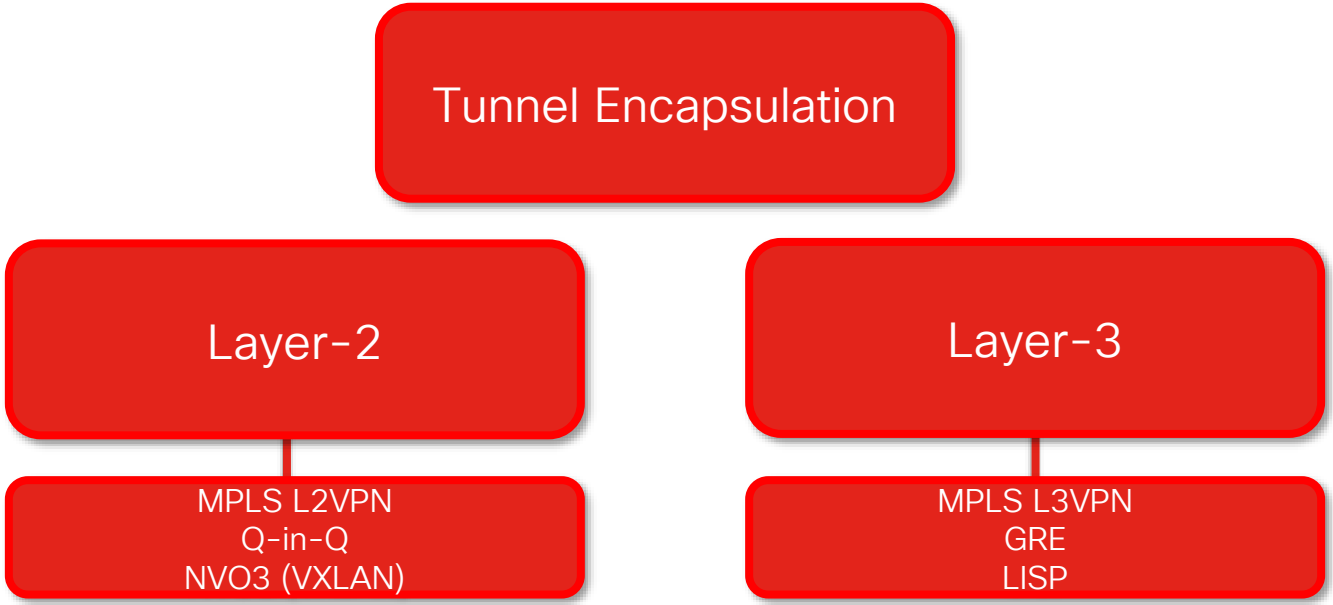
Overlay Services



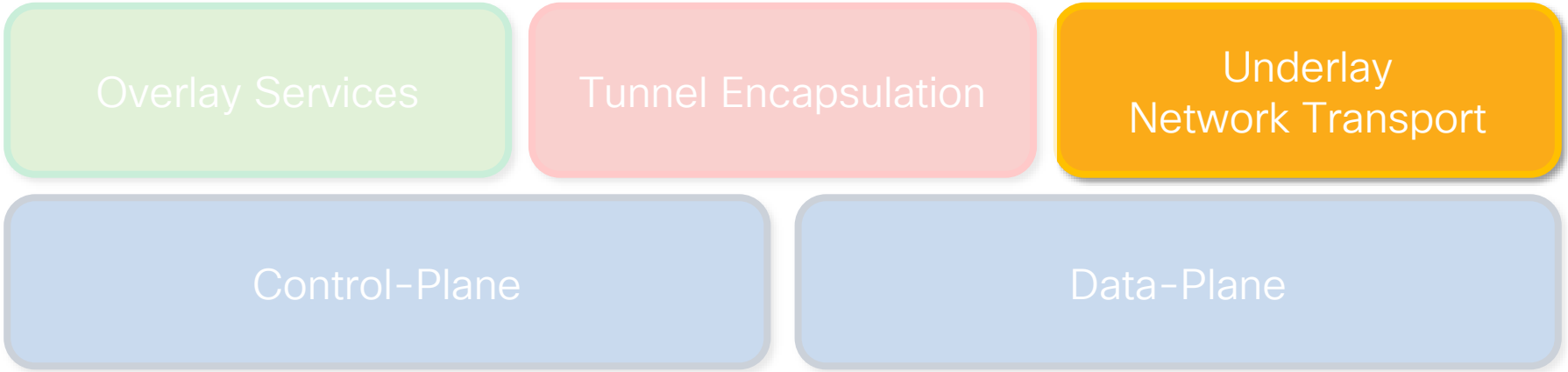
Tunnel Encapsulation



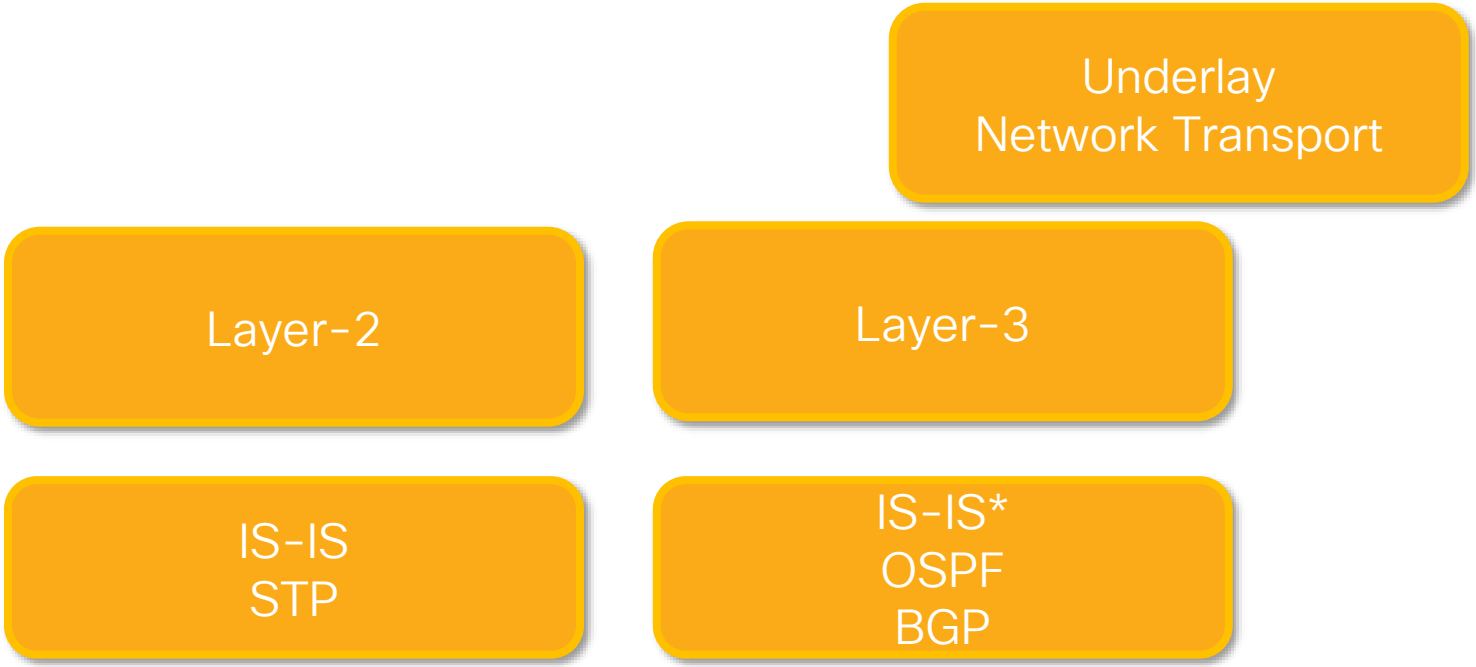
Tunnel Encapsulation



Underlay Network Transport

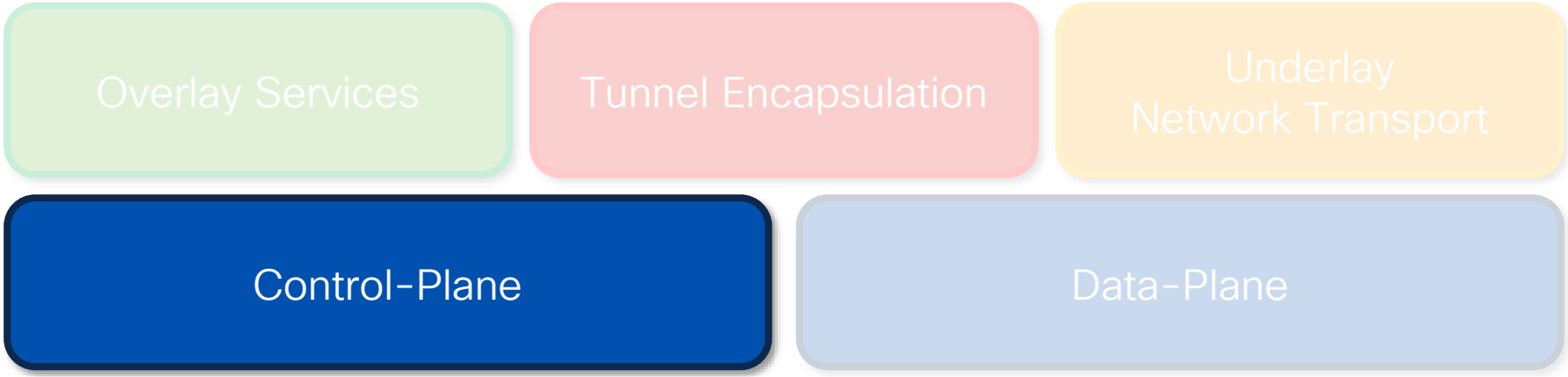


Underlay Network Transport

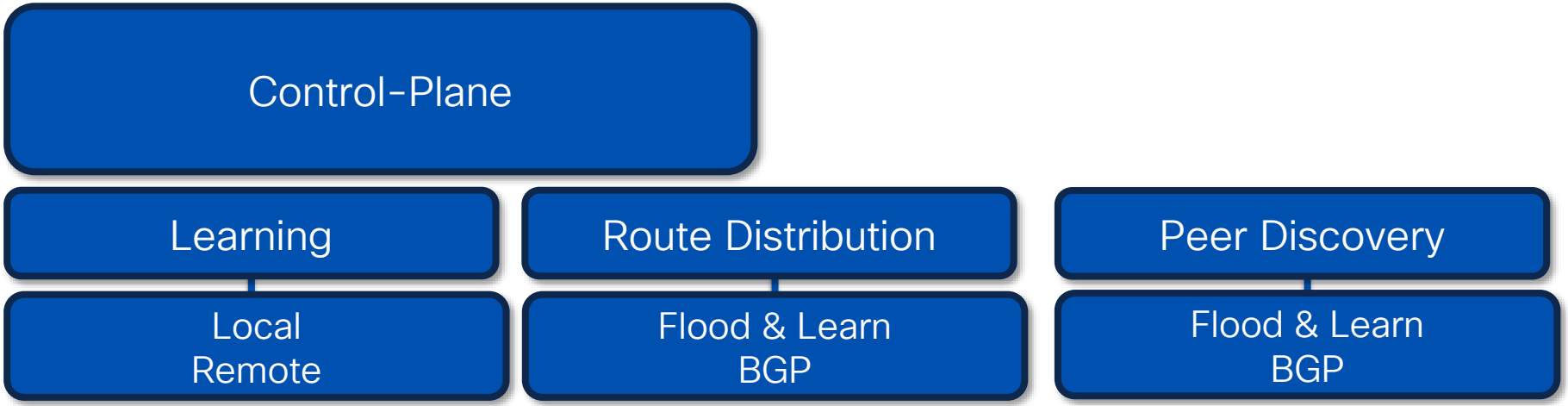


*IS-IS operates at L2 but transports L2 & L3

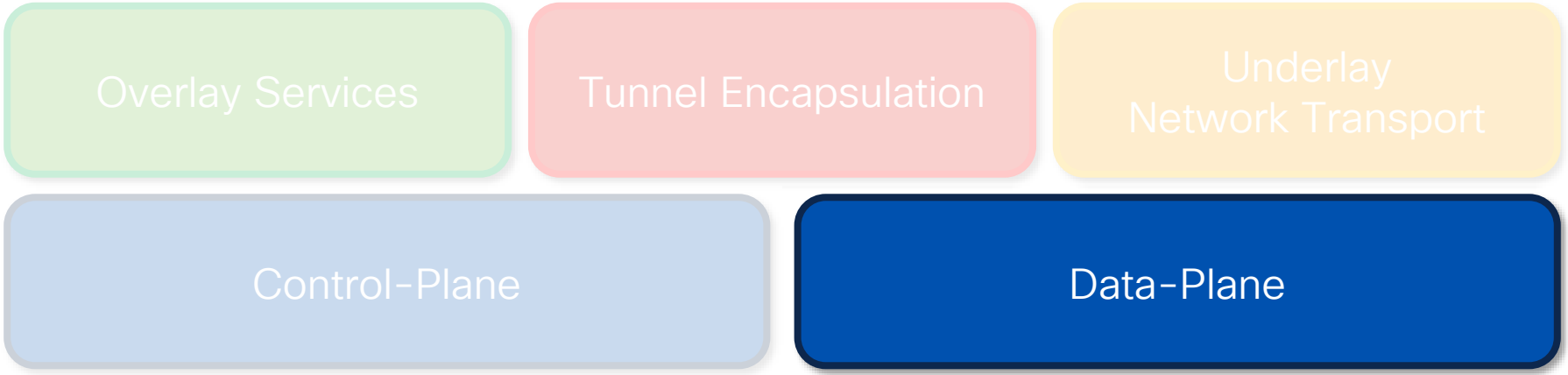
Control-Plane



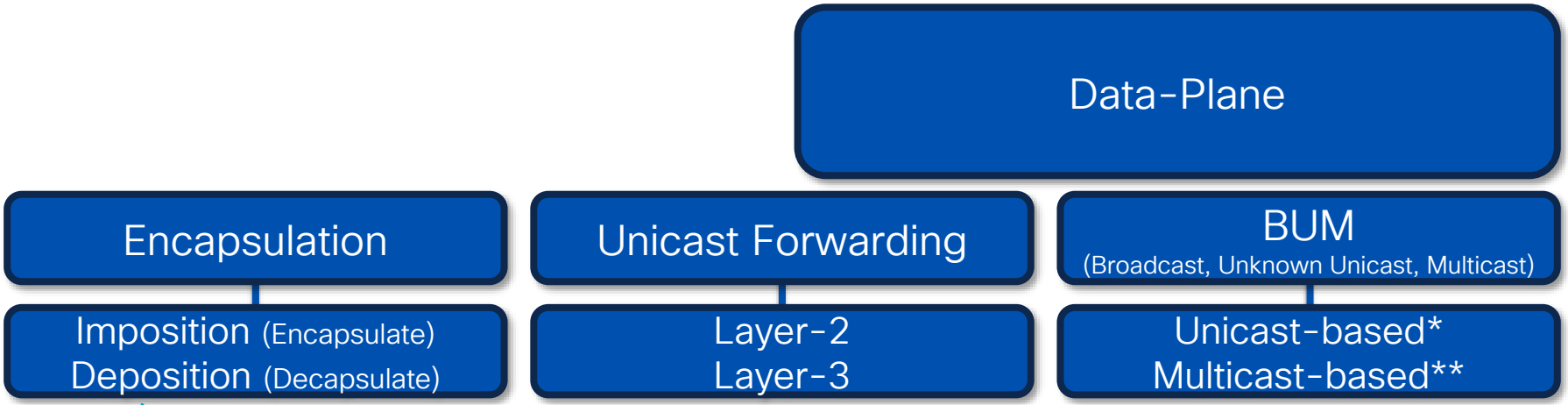
Control-Plane



Data-Plane



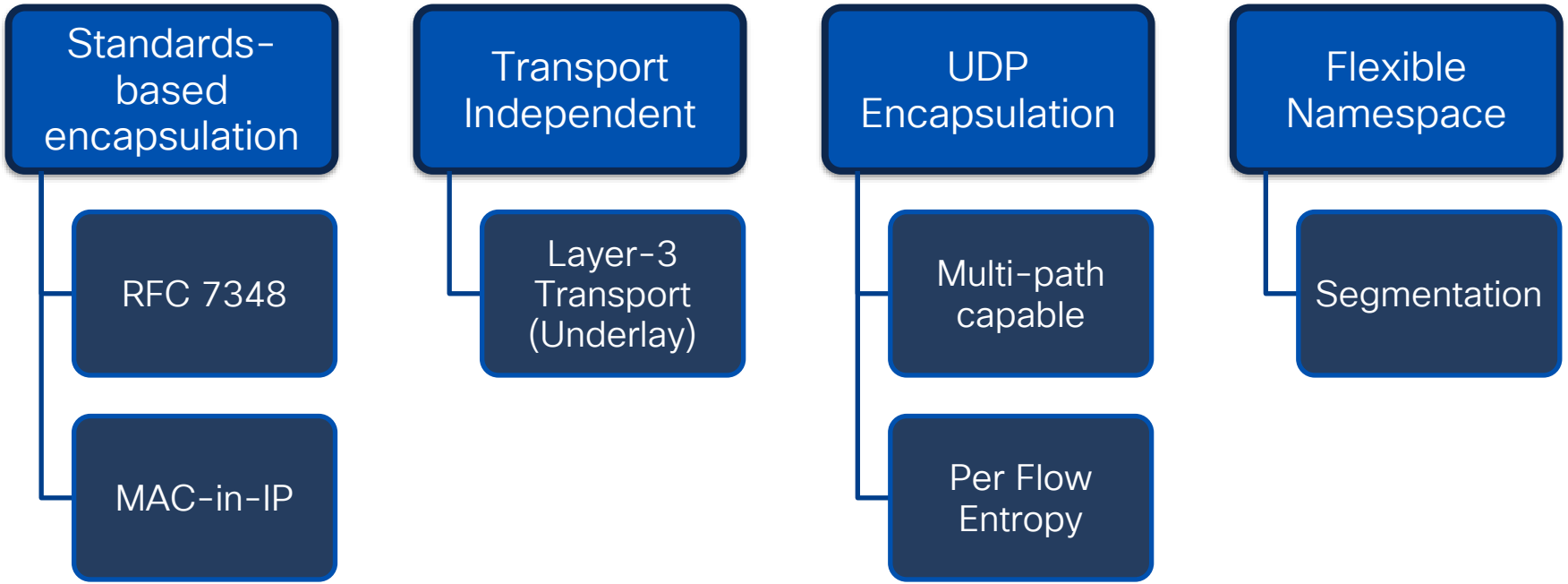
Data-Plane



*Ingress/Head-End Replication
**PIM

VXLAN
Virtual Extensible LAN

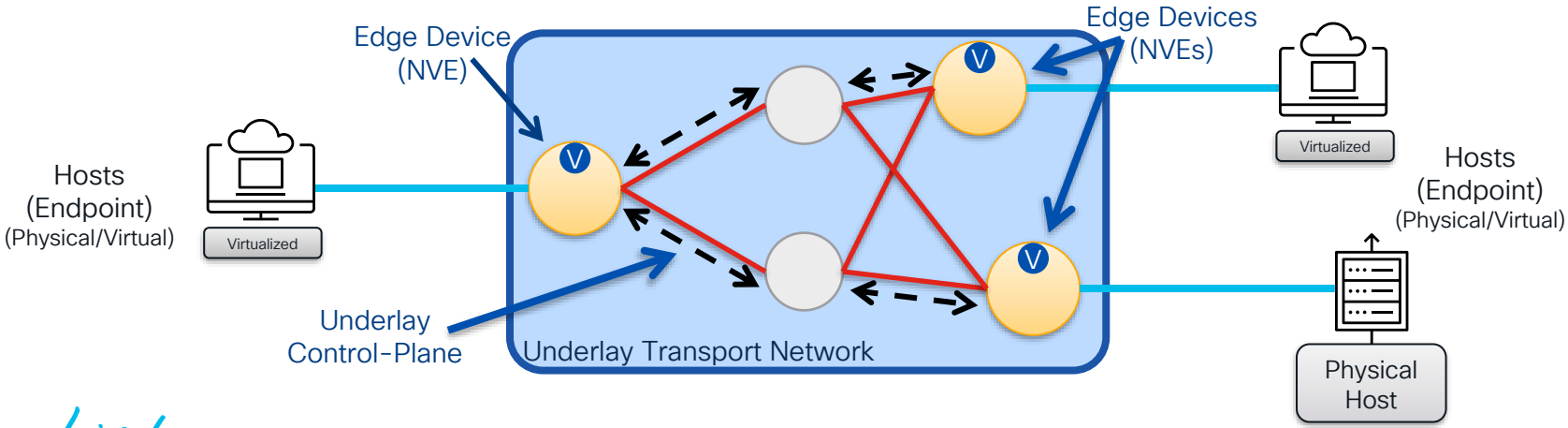
What is VXLAN?



Underlay Taxonomy

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

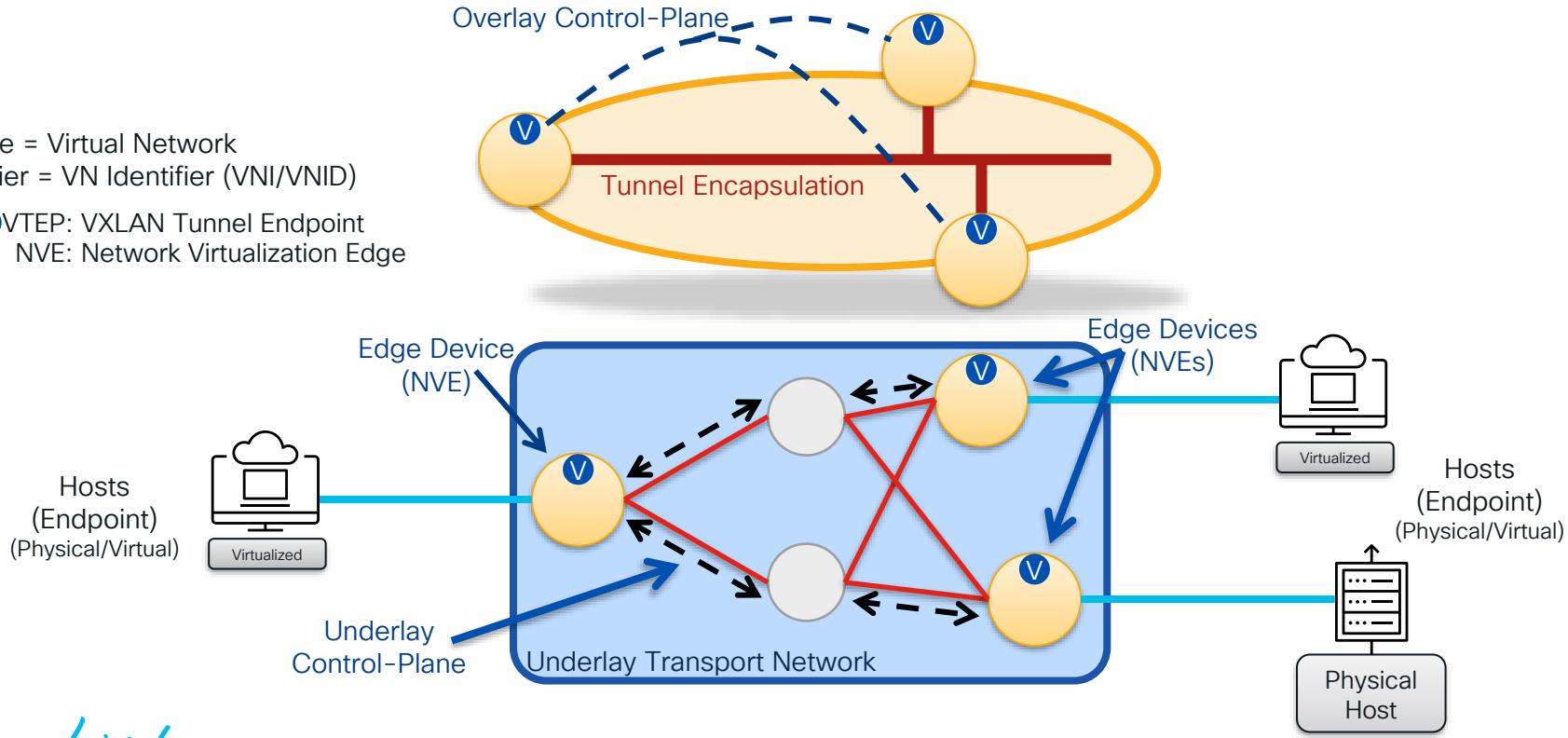
VTEP: VXLAN Tunnel Endpoint
NVE: Network Virtualization Edge



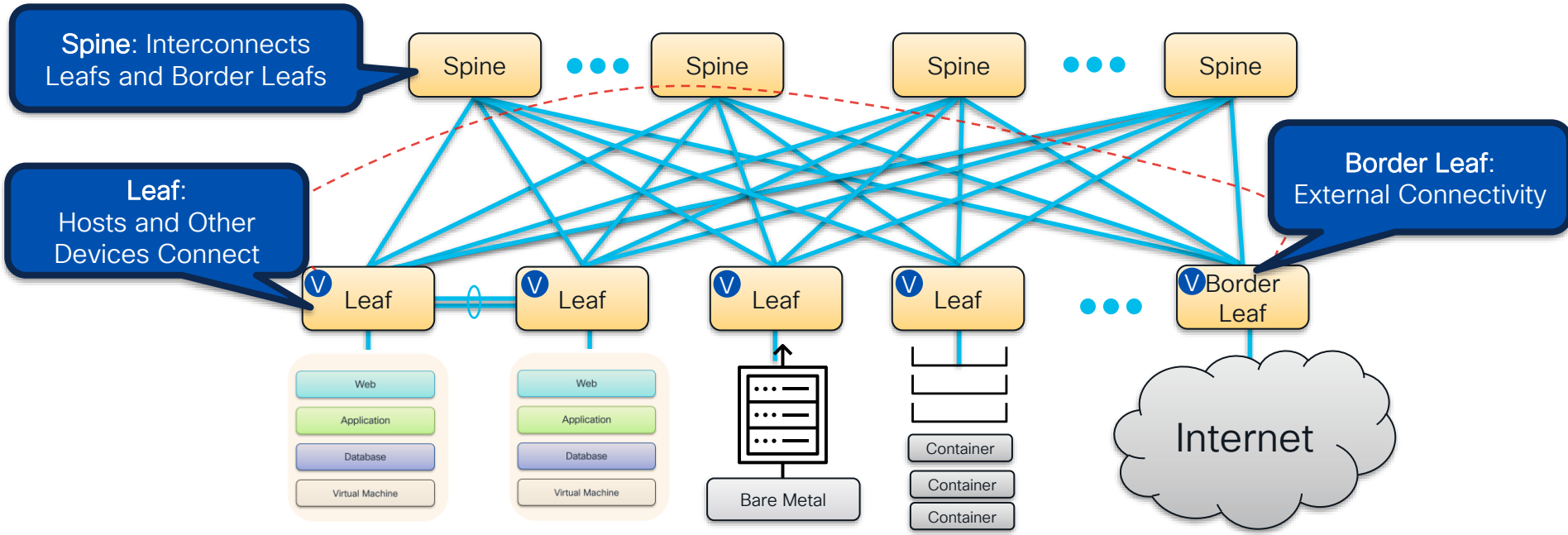
Overlay Taxonomy

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

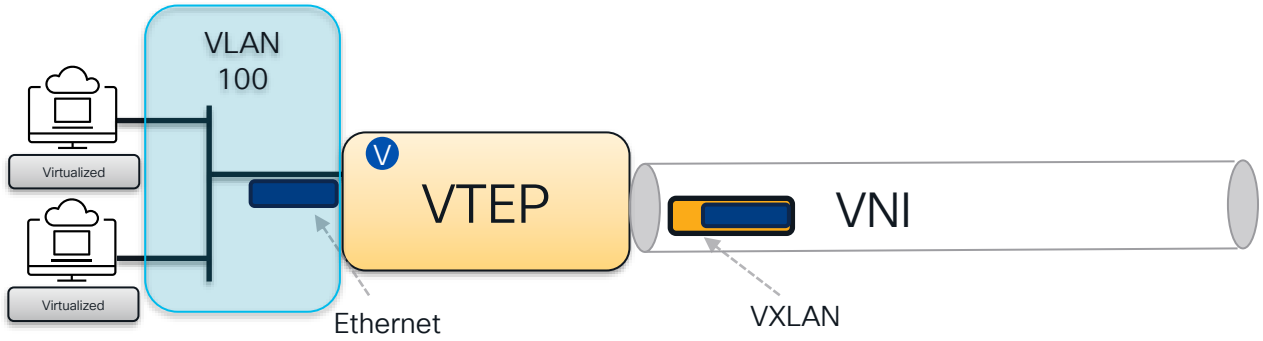
VTEP: VXLAN Tunnel Endpoint
NVE: Network Virtualization Edge



VTEP Device Roles

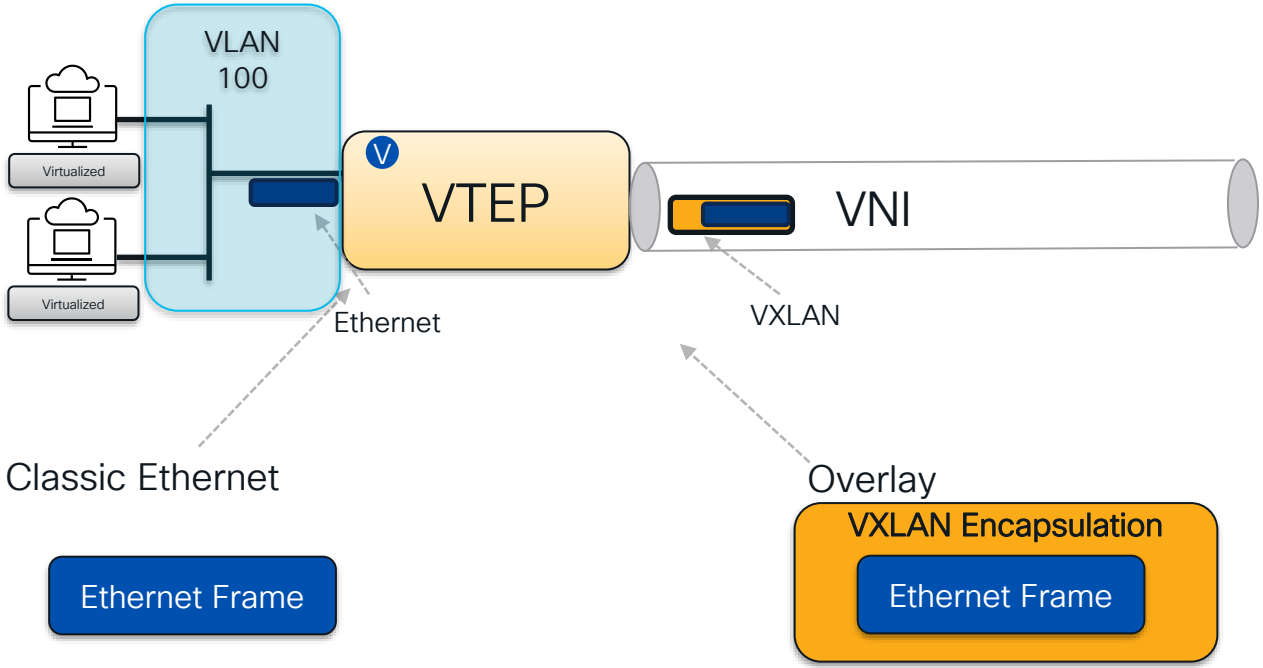


VXLAN (or Virtual) Tunnel Endpoint VTEP

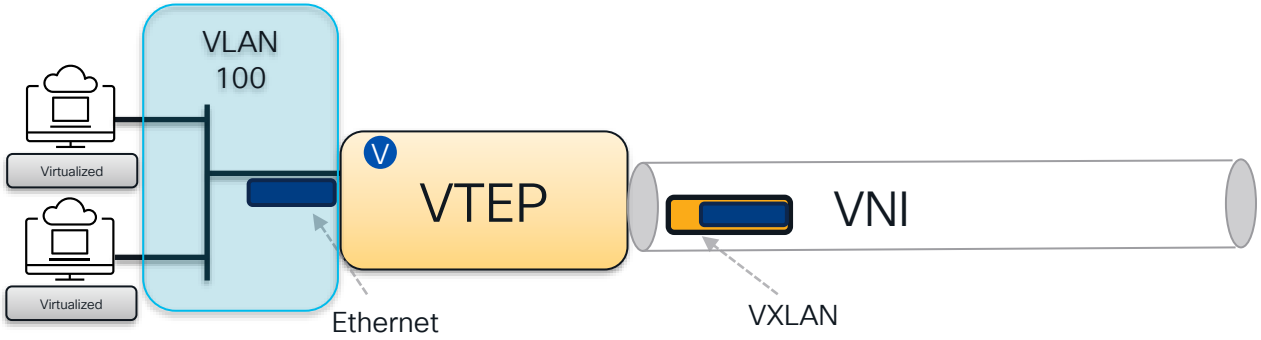


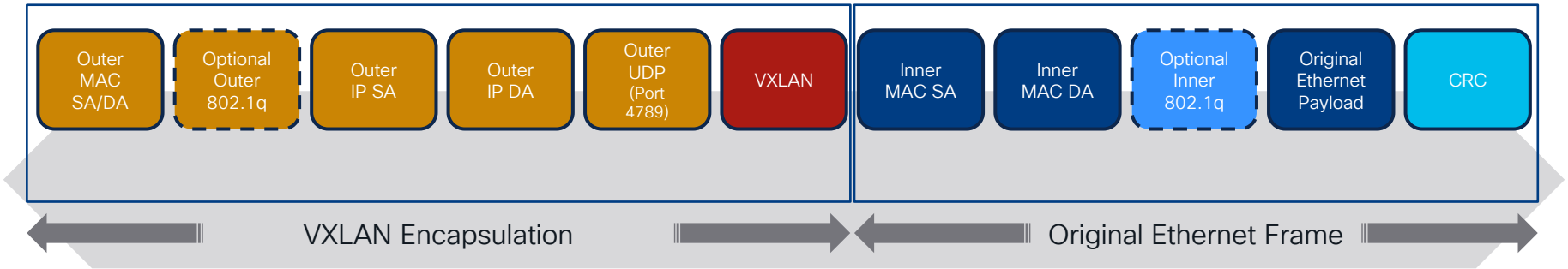
- VXLAN Tunnel Endpoint - Network Virtualization Edge
- Each VTEP is uniquely identified by an IP Address
- VTEP Discovers or learns remote VTEPs, and end hosts attached to them
- VTEP bridges when forwarding packets within the same VNI and Routes for Inter-VNI traffic

VXLAN Encapsulation



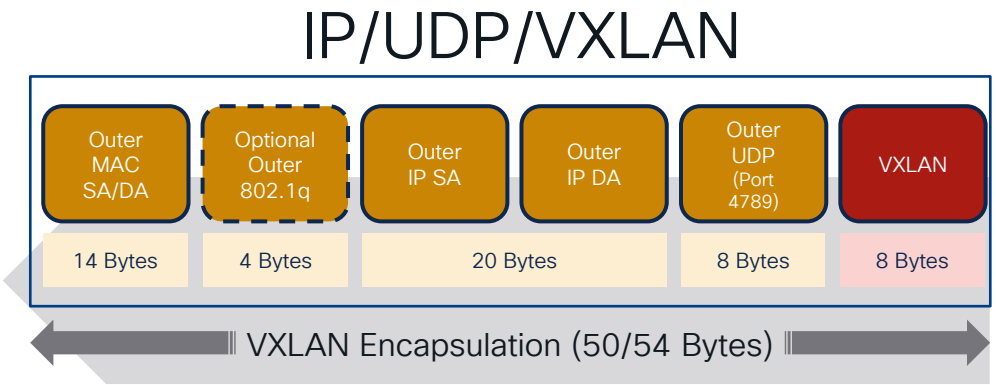
VXLAN Encapsulation / Packet Format





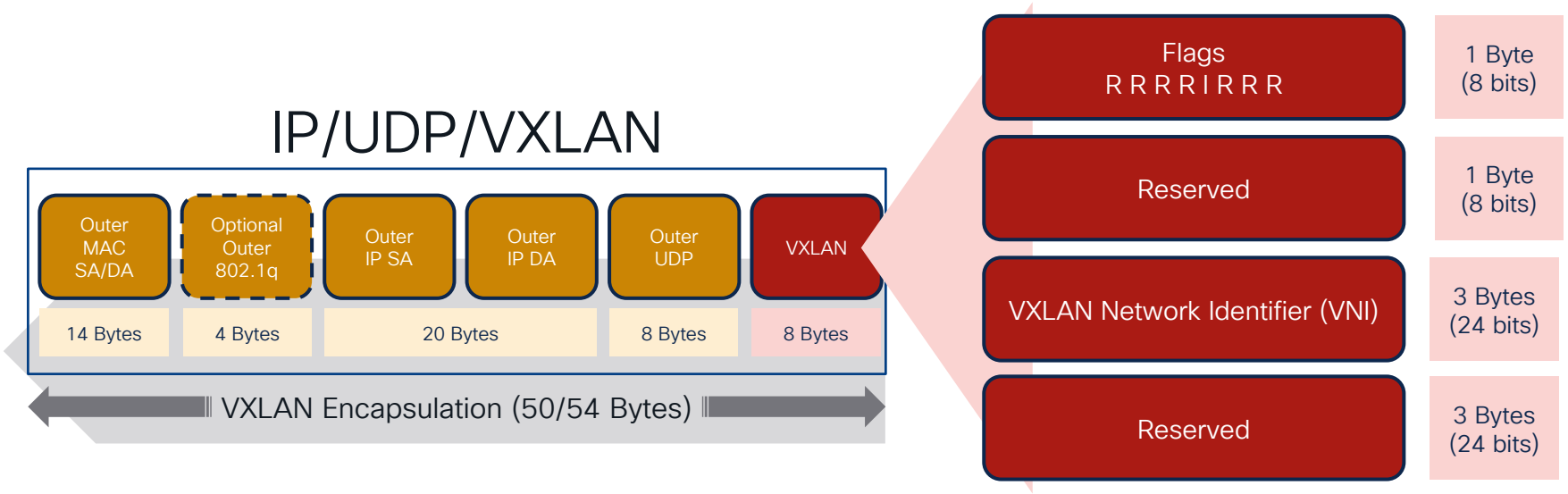
VXLAN Packet Format

- VXLAN uses MAC in UDP encapsulation

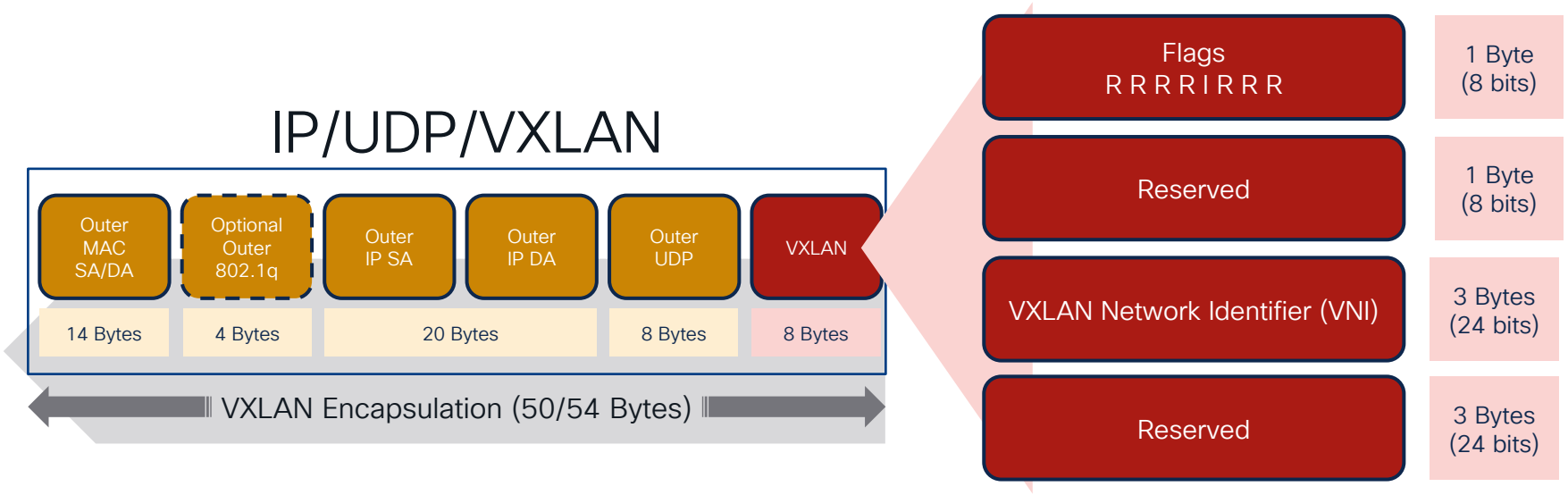


- Adds UDP and VXLAN Header before original Ethernet Frame

VXLAN Header Details

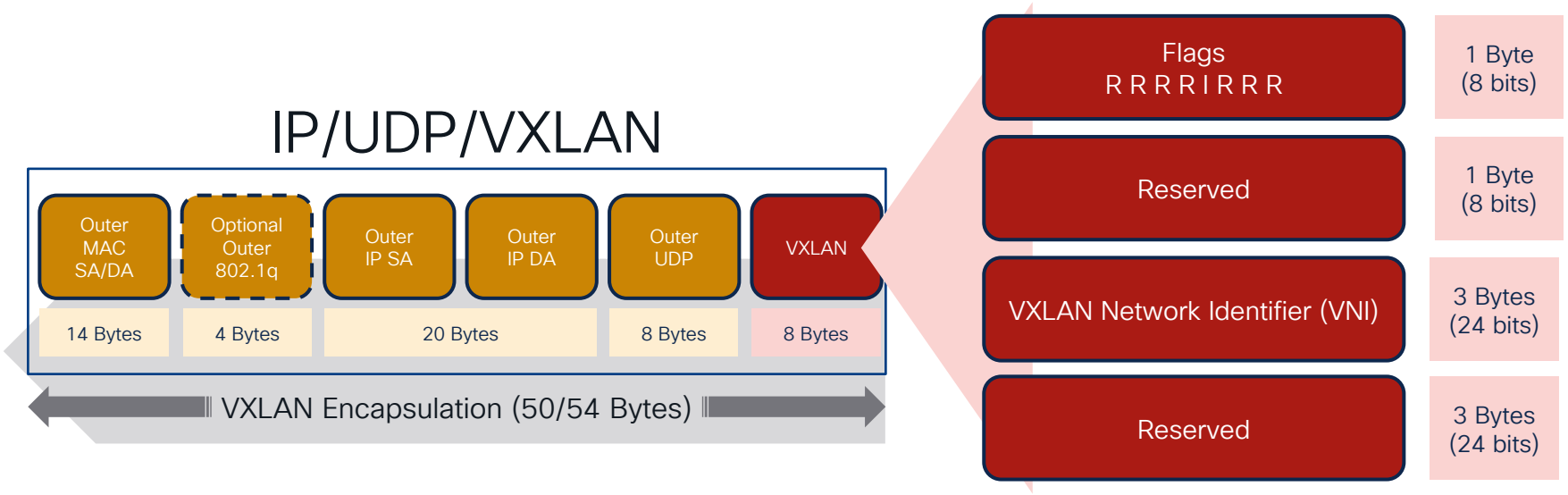


VXLAN Header Details



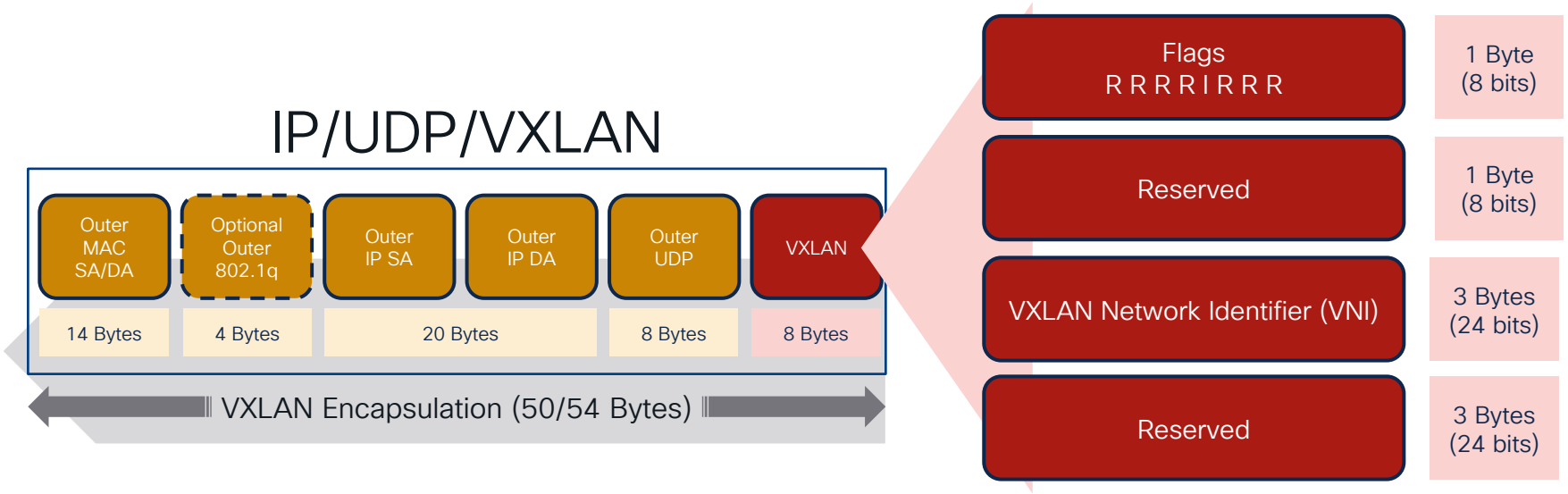
- Flags Field: I-flag (set to 1) for valid VNI. Other flags remain as R (set to 0)

VXLAN Header Details



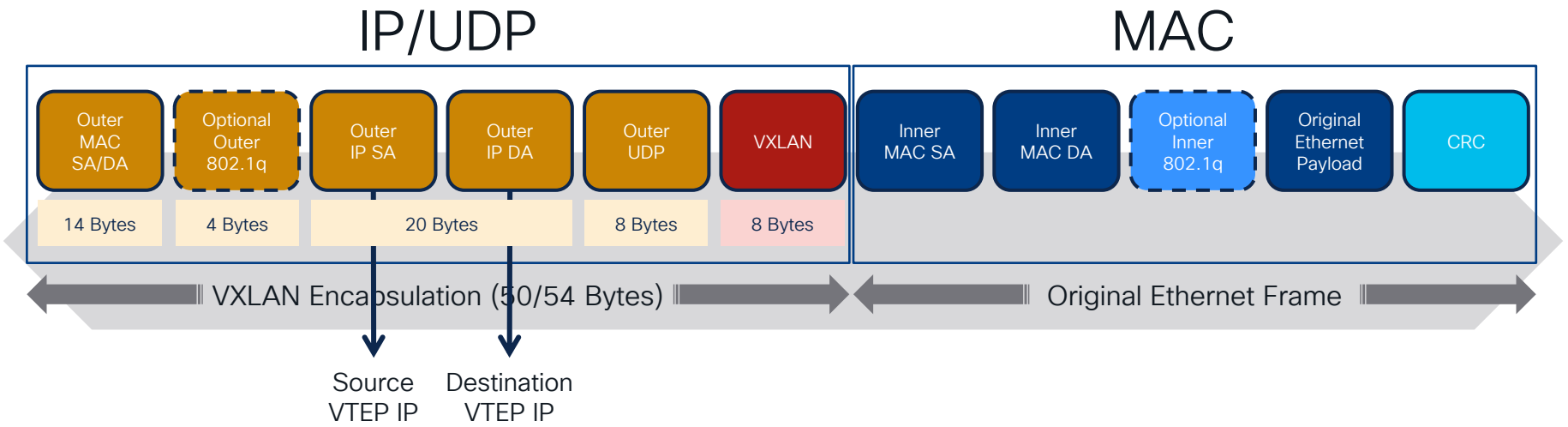
- Flags Field: I-flag (set to 1) for valid VNI. Other flags remain as R (set to 0)
- VNI Field: Allows VNI 1-16,777,215 (some implementation only 4096-16,777,215)

VXLAN Header Details

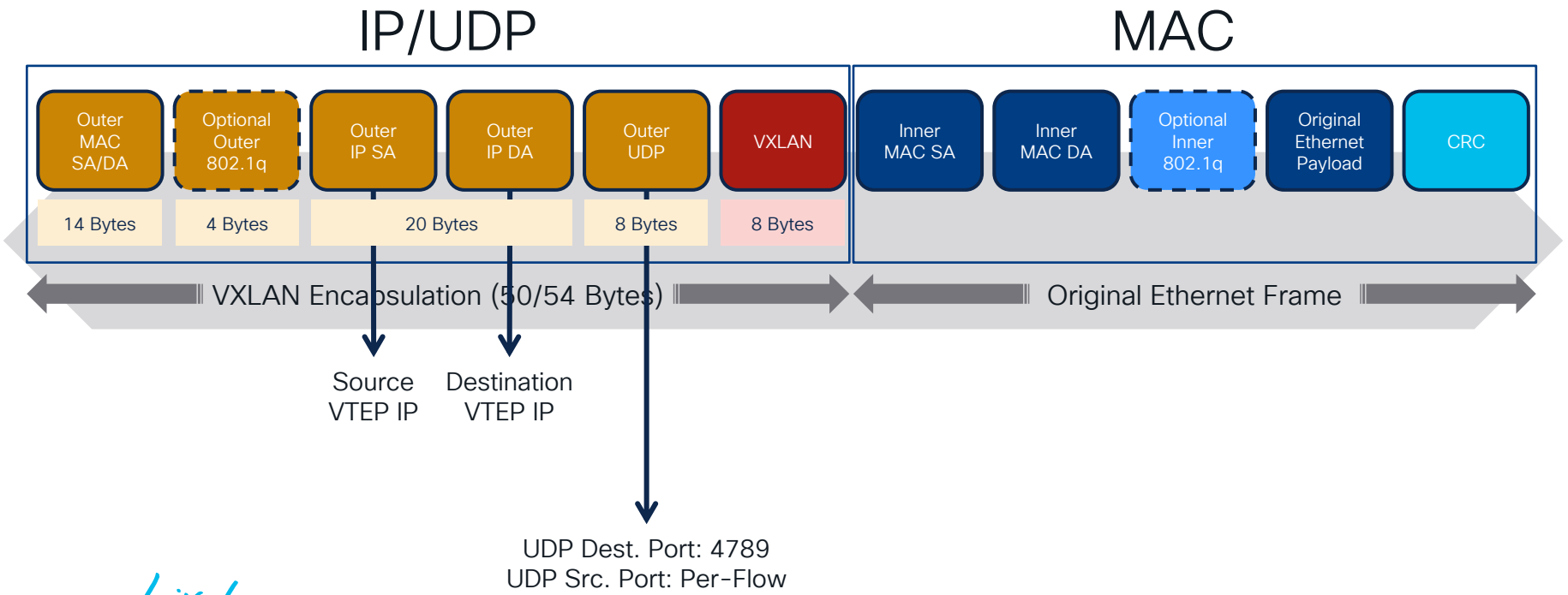



```
vlan 15
  vn-segment 31234
```

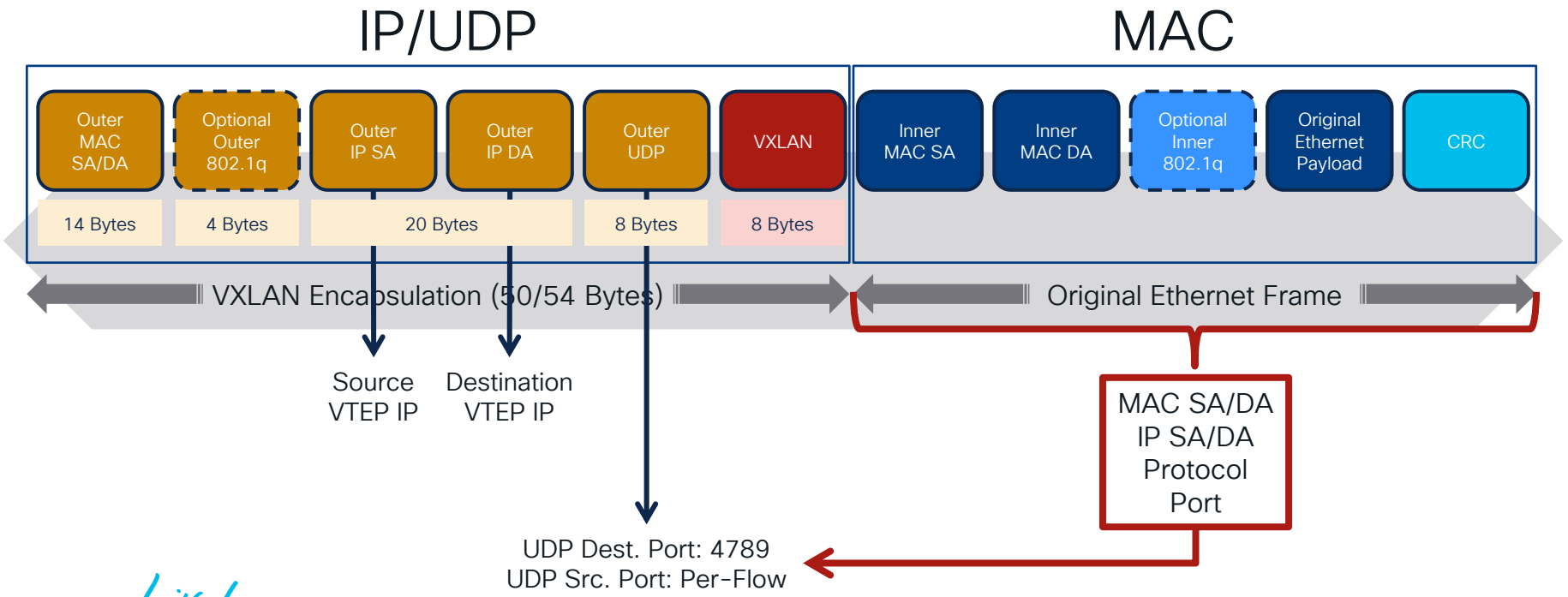
Transport Independence



Multipath Capable



Multipath Capable



Flood and Learn

Two Modes of VXLAN

Flood & Learn

No Control Plane

Limited Scale

Requires Centralized Gateway

VXLAN EVPN

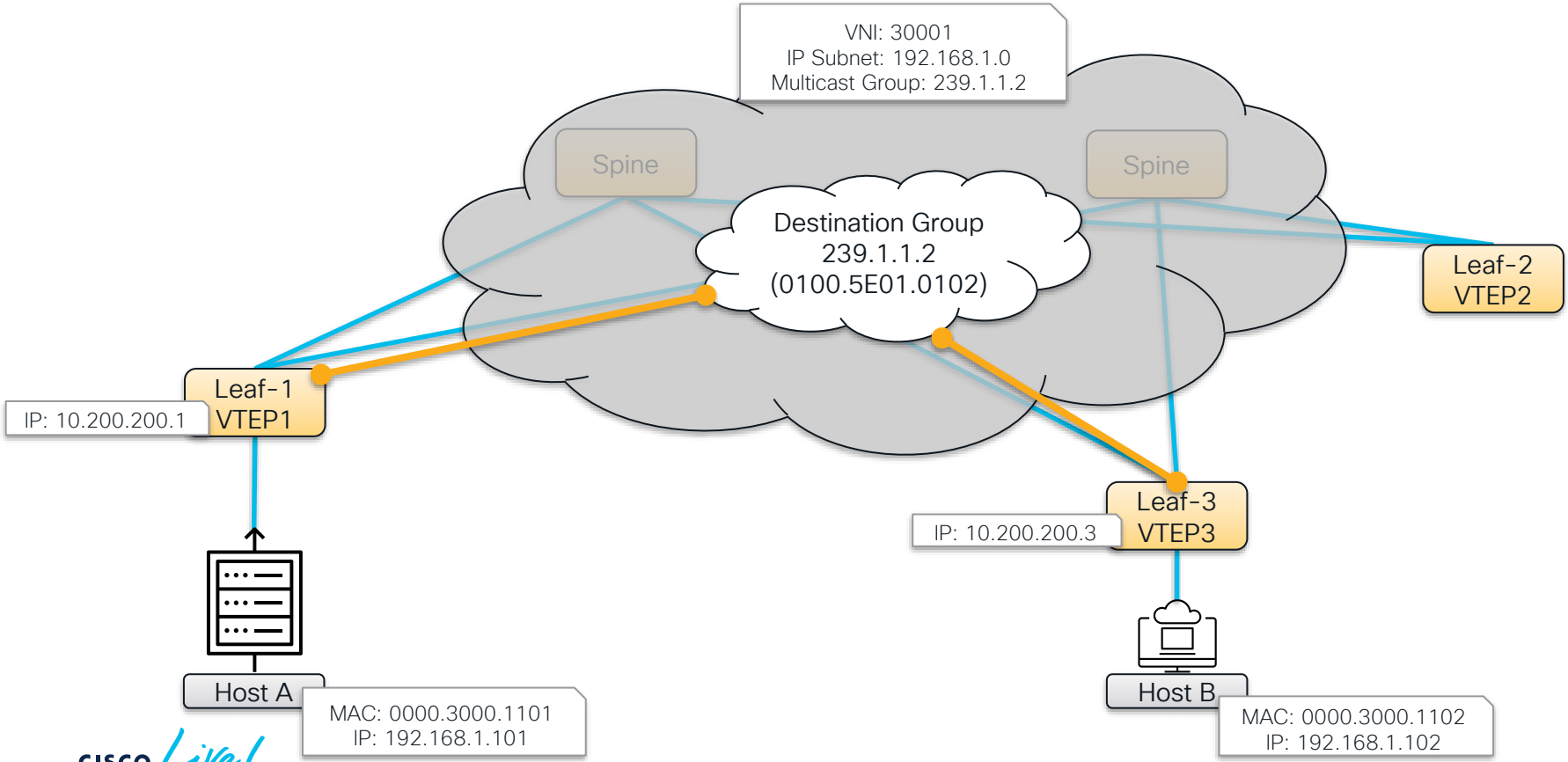
EVPN as the Control Plane

L2/L3 Host & Subnet reachability exchanged

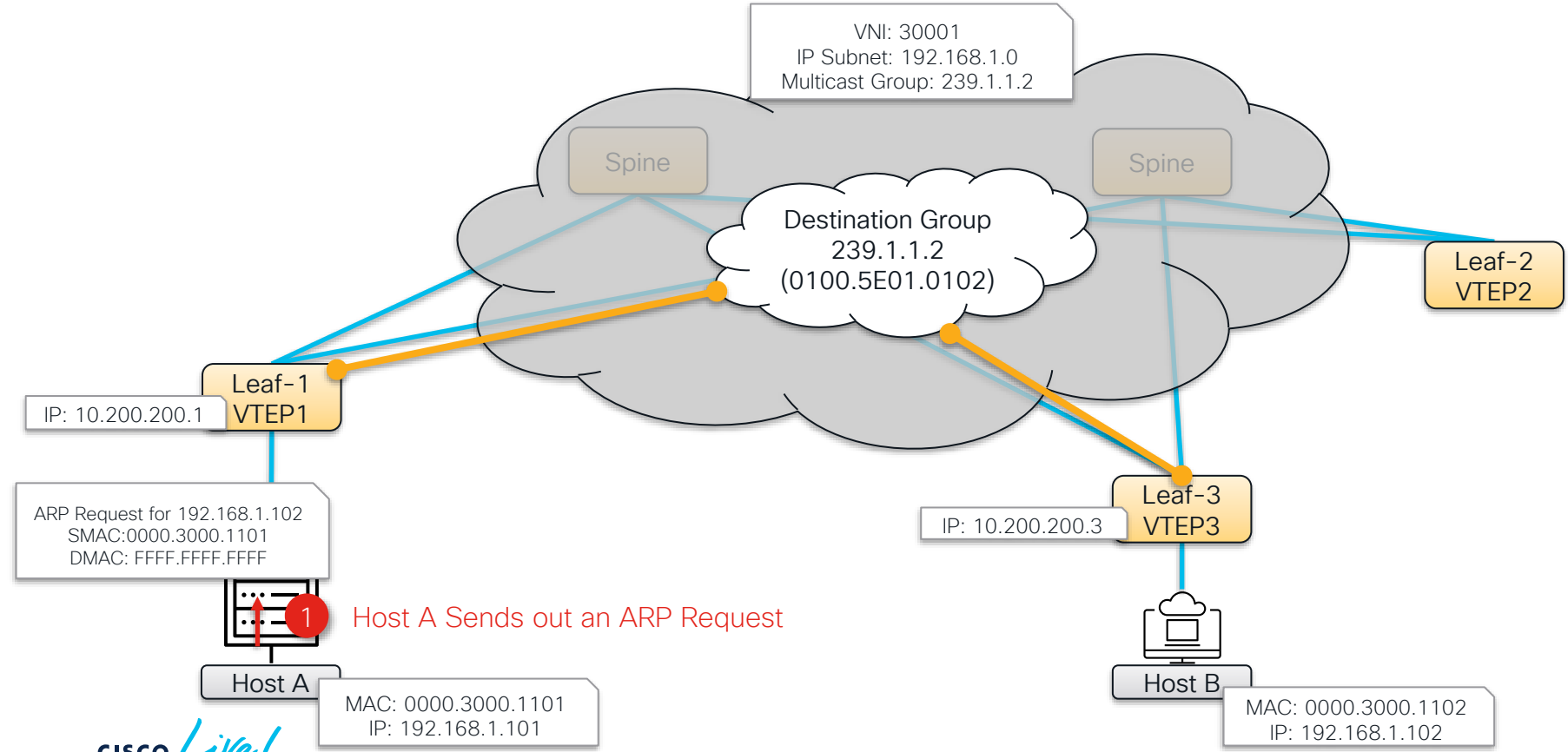
Distributed Anycast Gateway

Optimized Workload Mobility

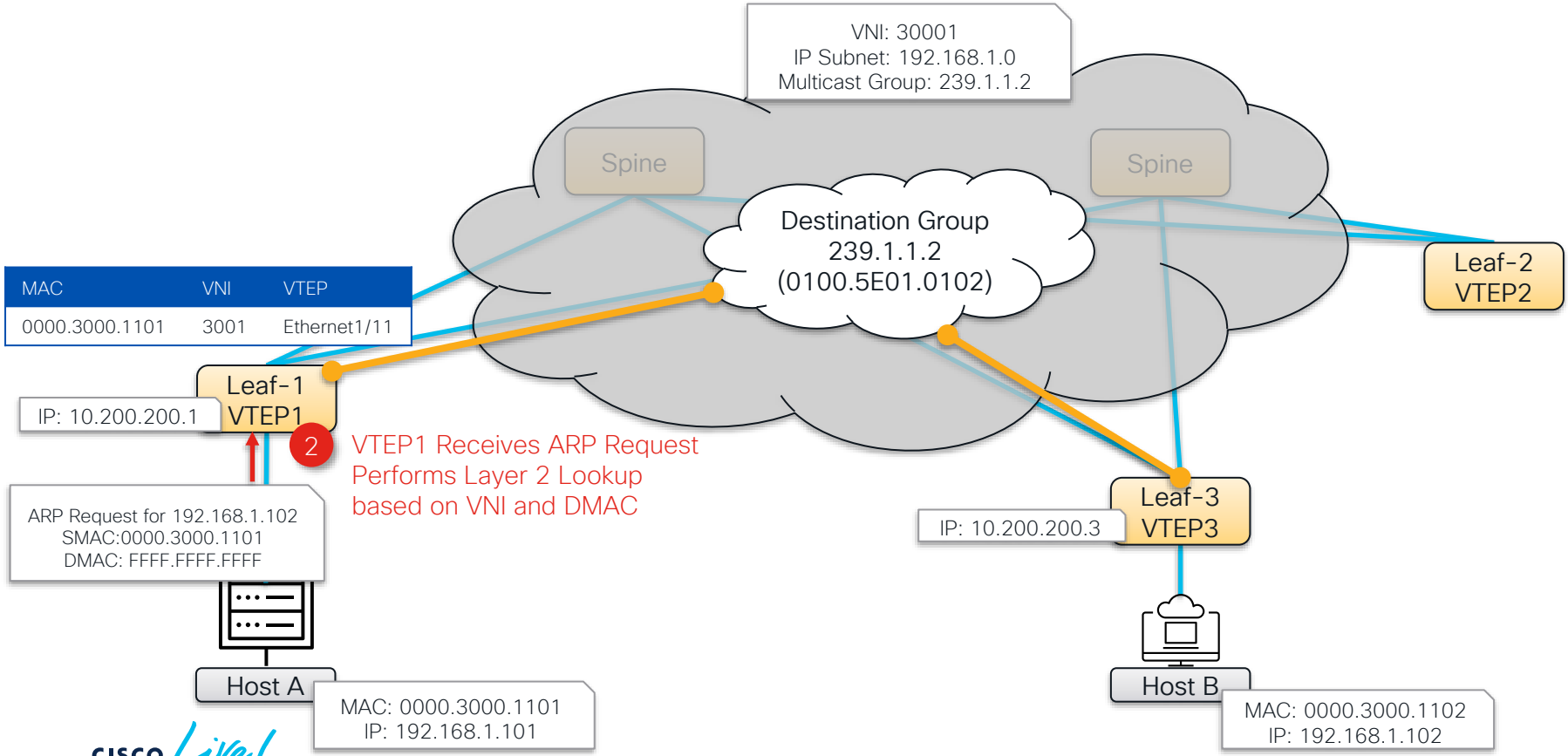
VXLAN Flood and Learn (F&L)
Host A to Host B Communication



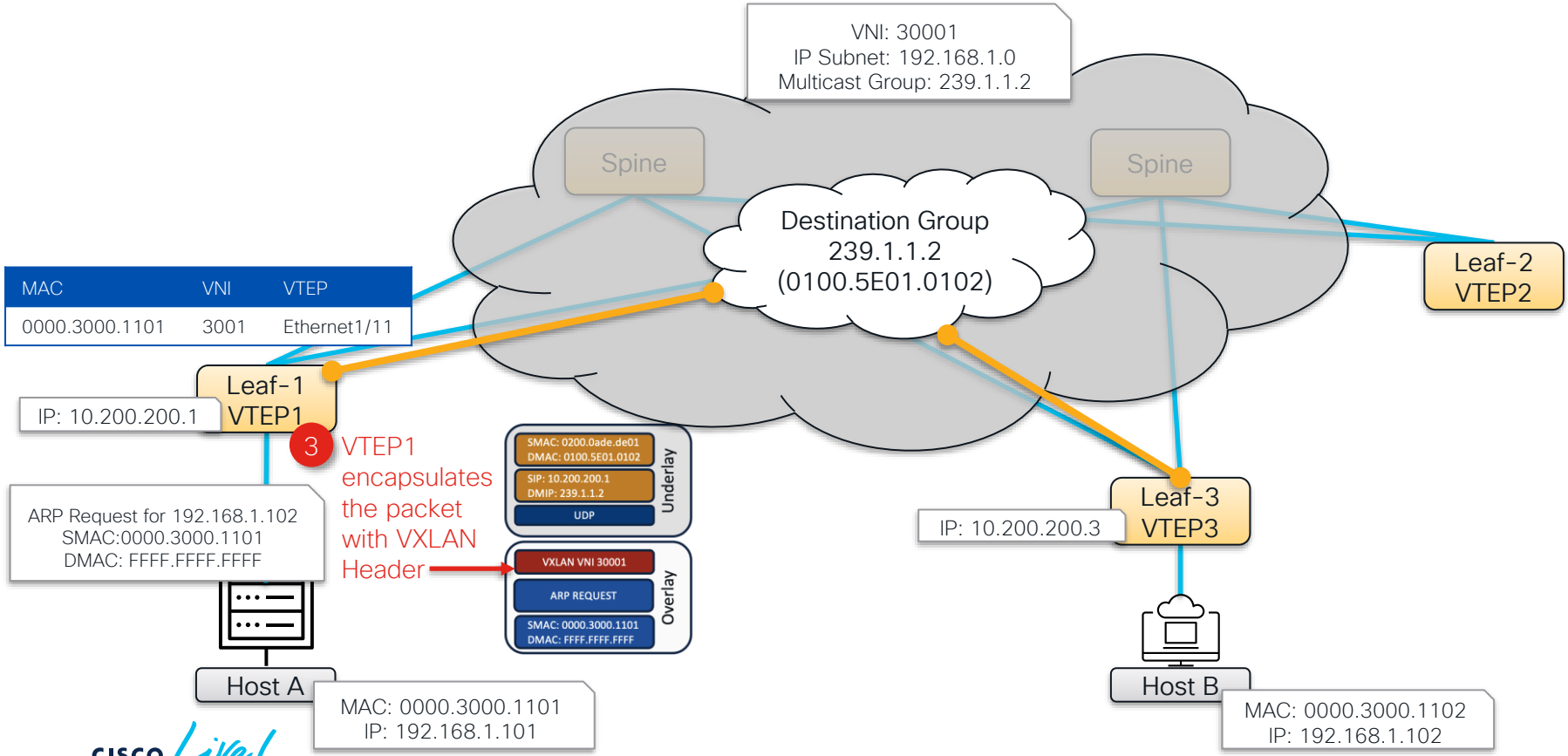
VXLAN Flood and Learn (F&L)
Host A to Host B Communication



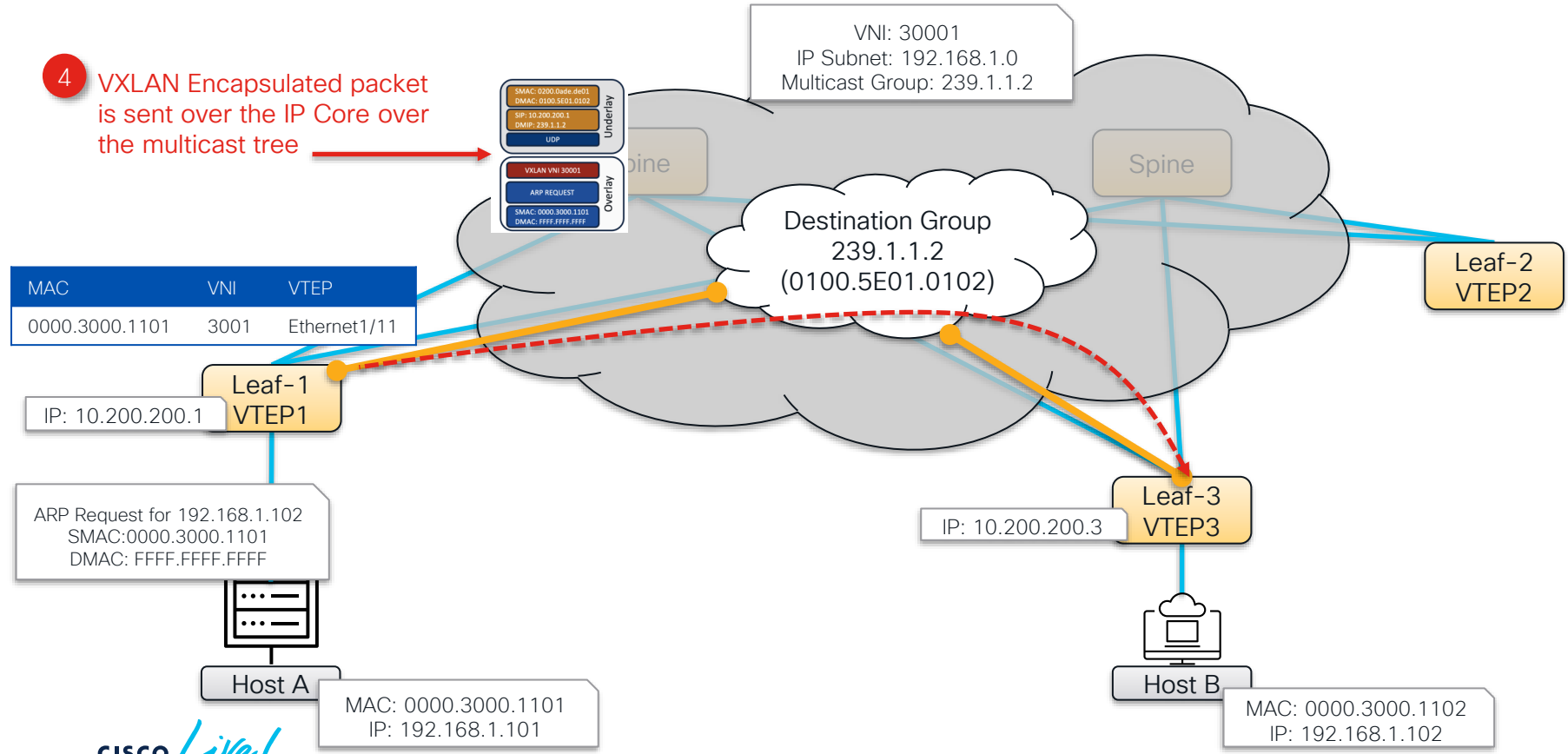
VXLAN Flood and Learn (F&L)
Host A to Host B Communication



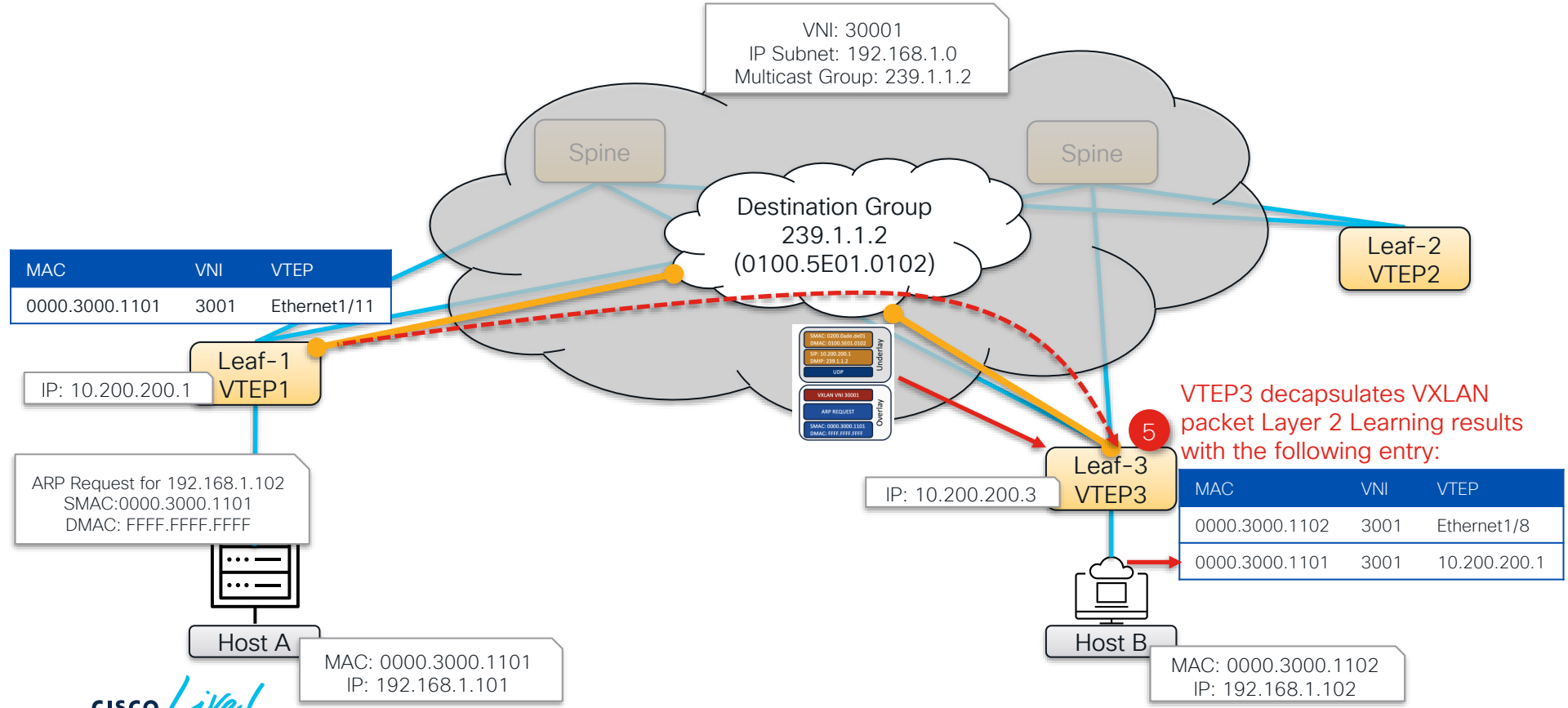
VXLAN Flood and Learn (F&L)
Host A to Host B Communication



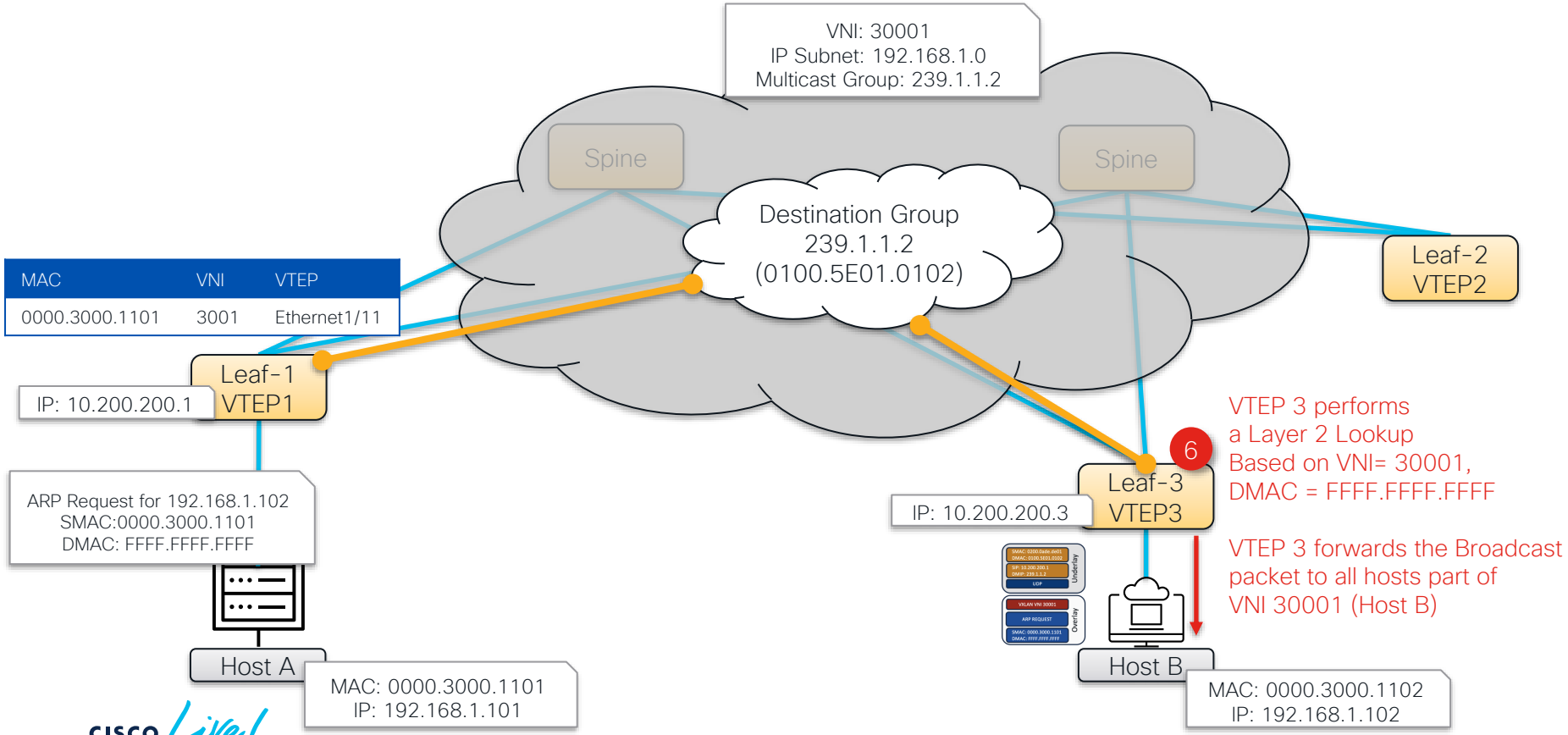
VXLAN Flood and Learn (F&L)
Host A to Host B Communication



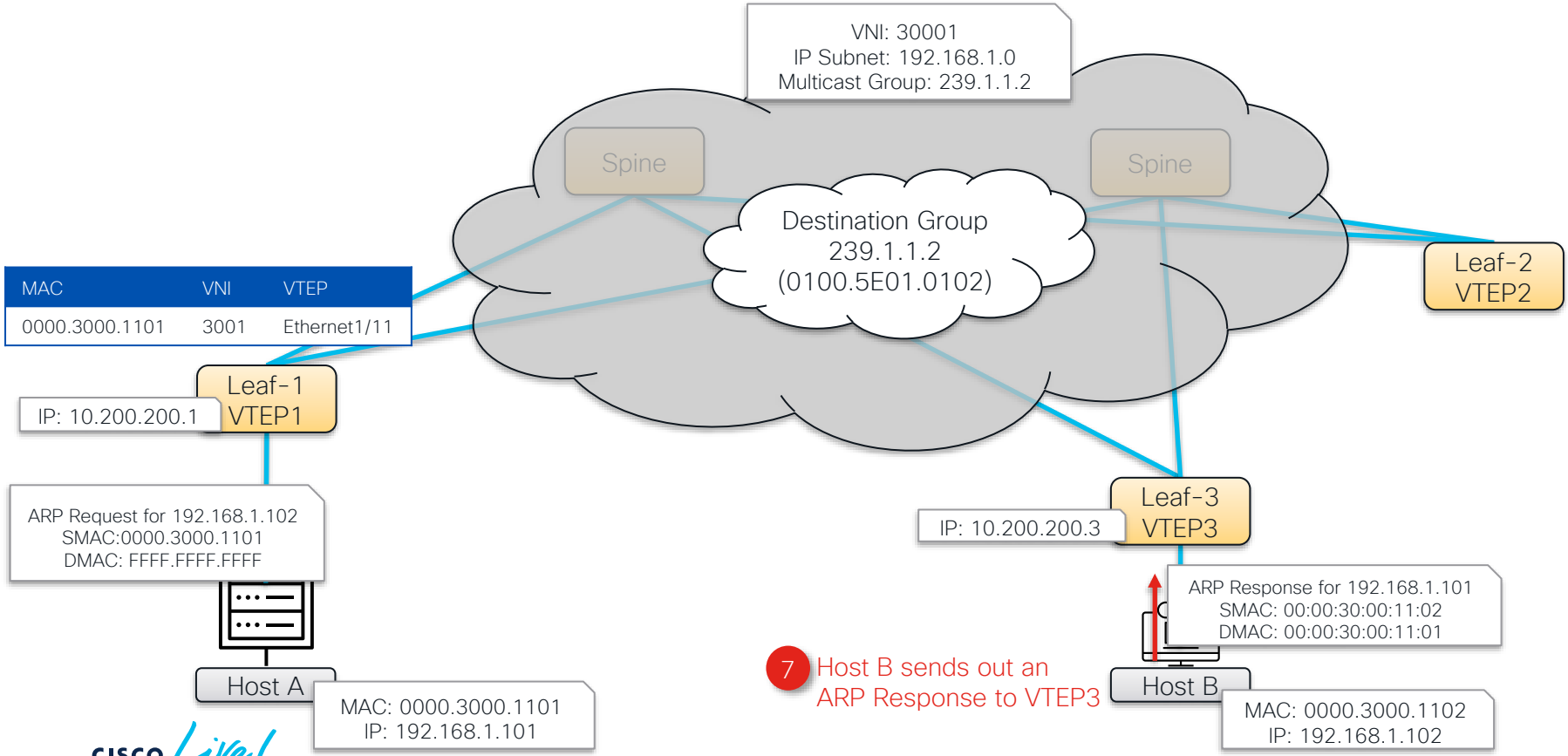
VXLAN Flood and Learn (F&L)
Host A to Host B Communication



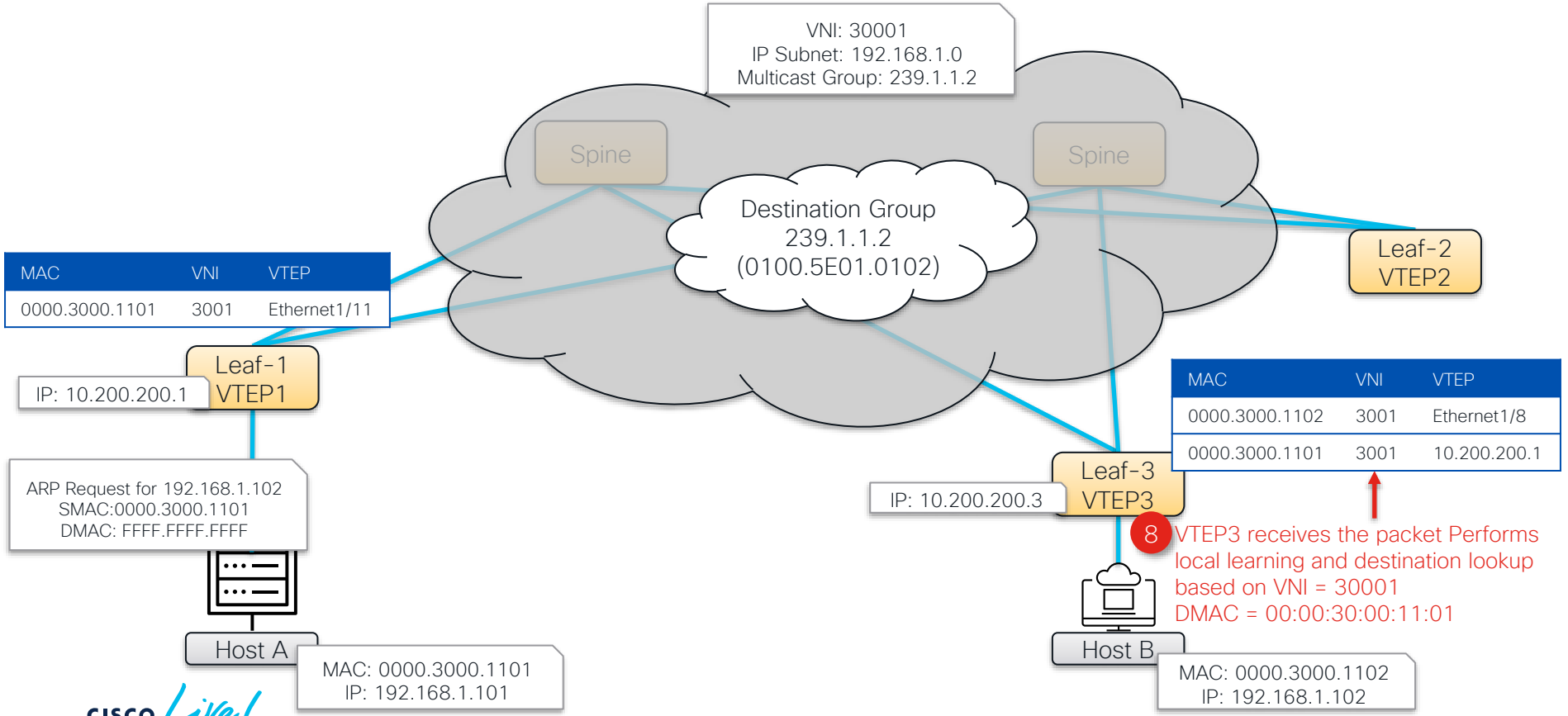
VXLAN Flood and Learn (F&L)
Host A to Host B Communication



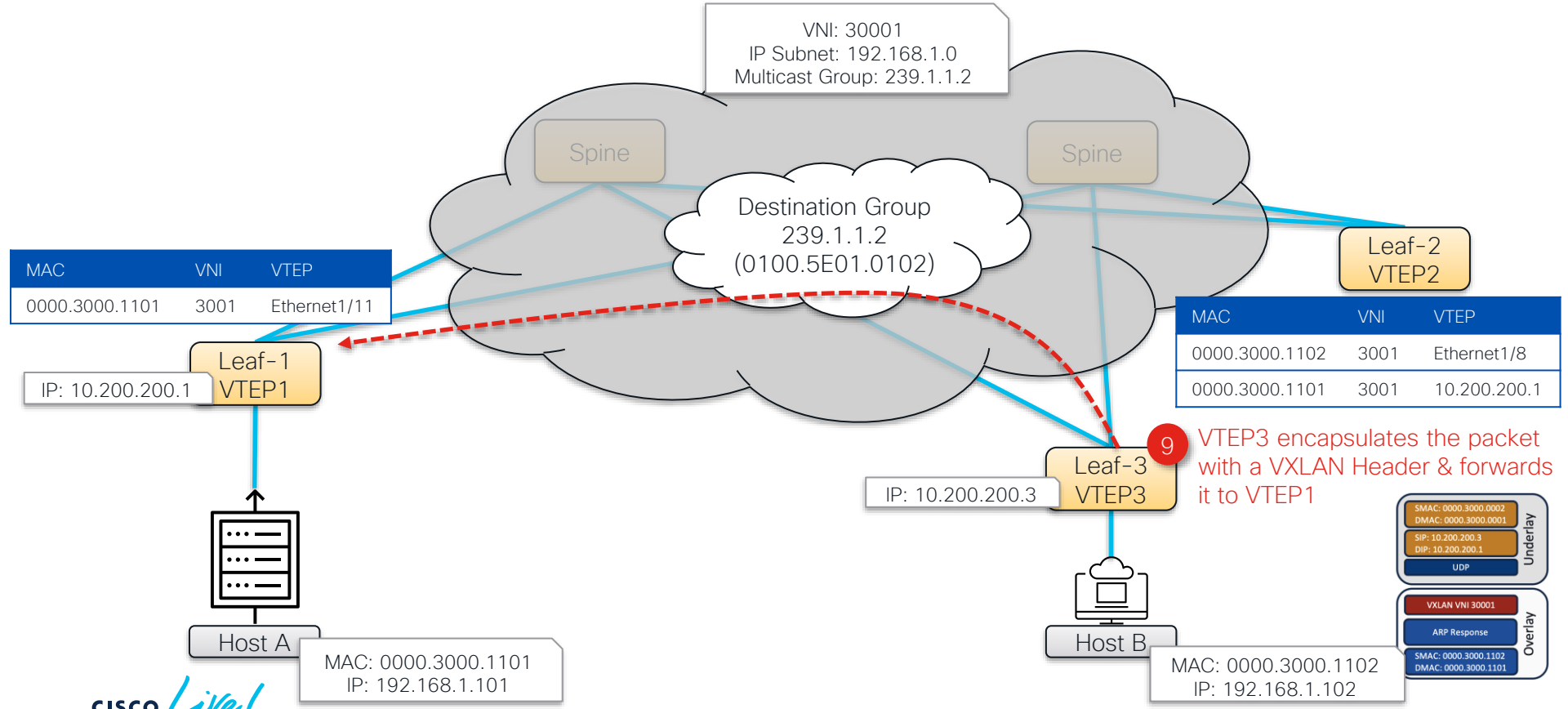
VXLAN Flood and Learn (F&L)
Host A to Host B Communication



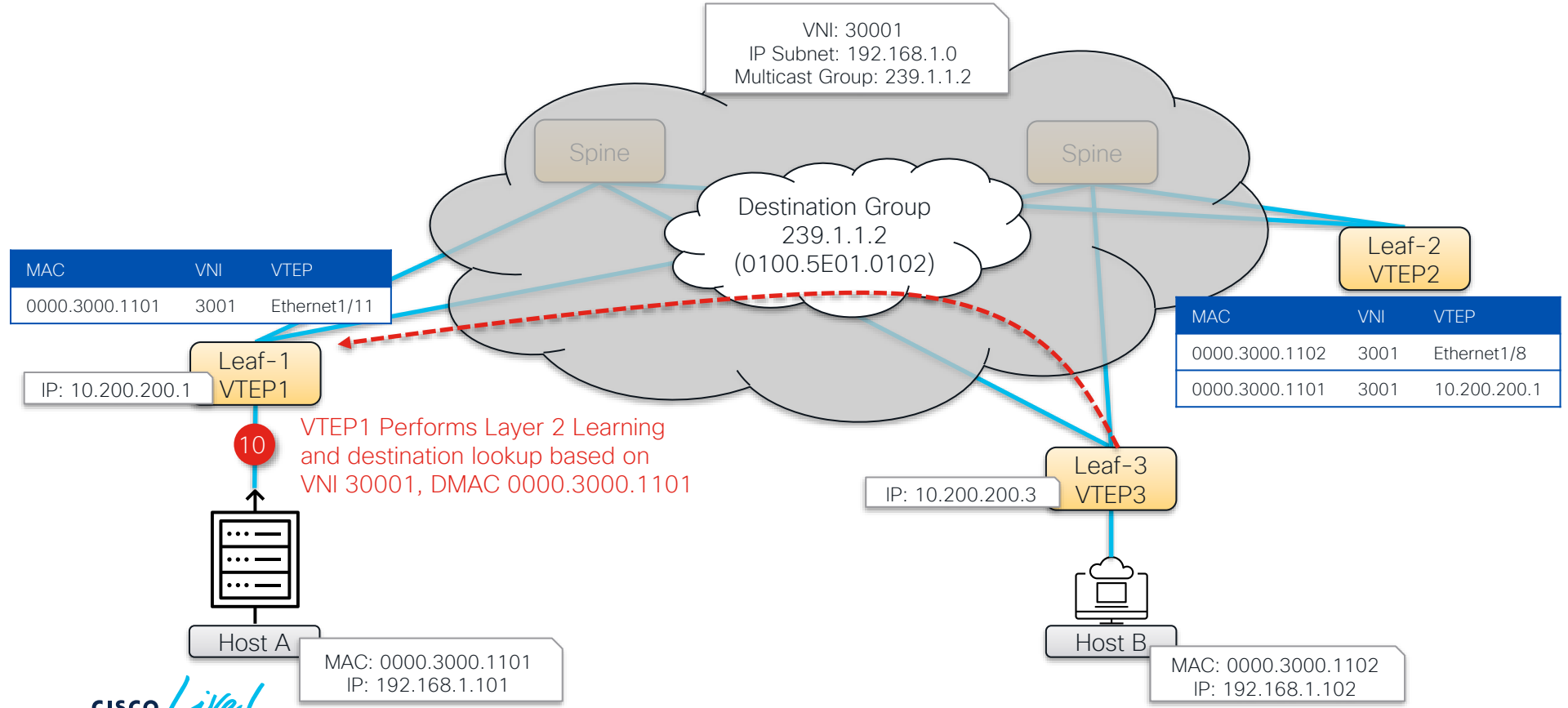
VXLAN Flood and Learn (F&L)
Host A to Host B Communication



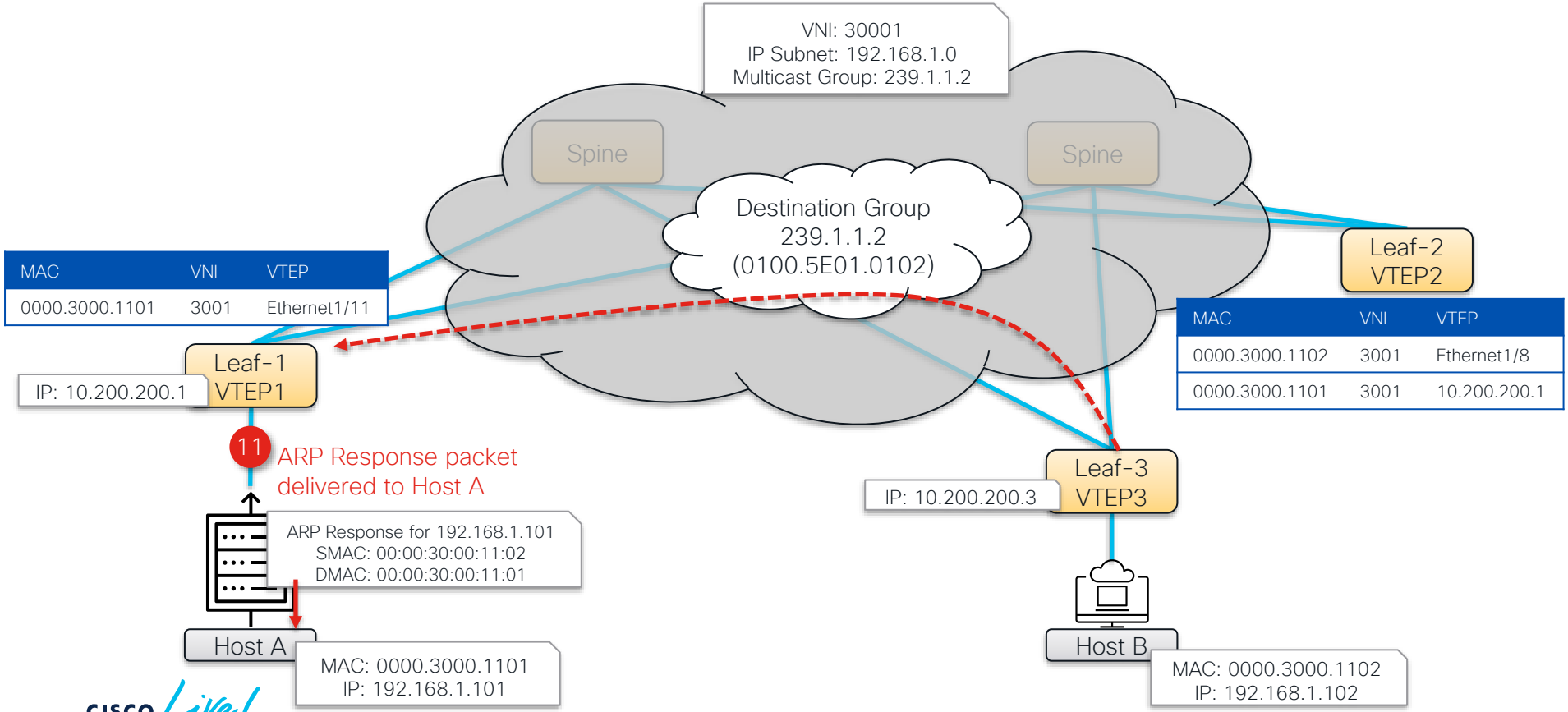
VXLAN Flood and Learn (F&L)
Host A to Host B Communication



VXLAN Flood and Learn (F&L)
Host A to Host B Communication

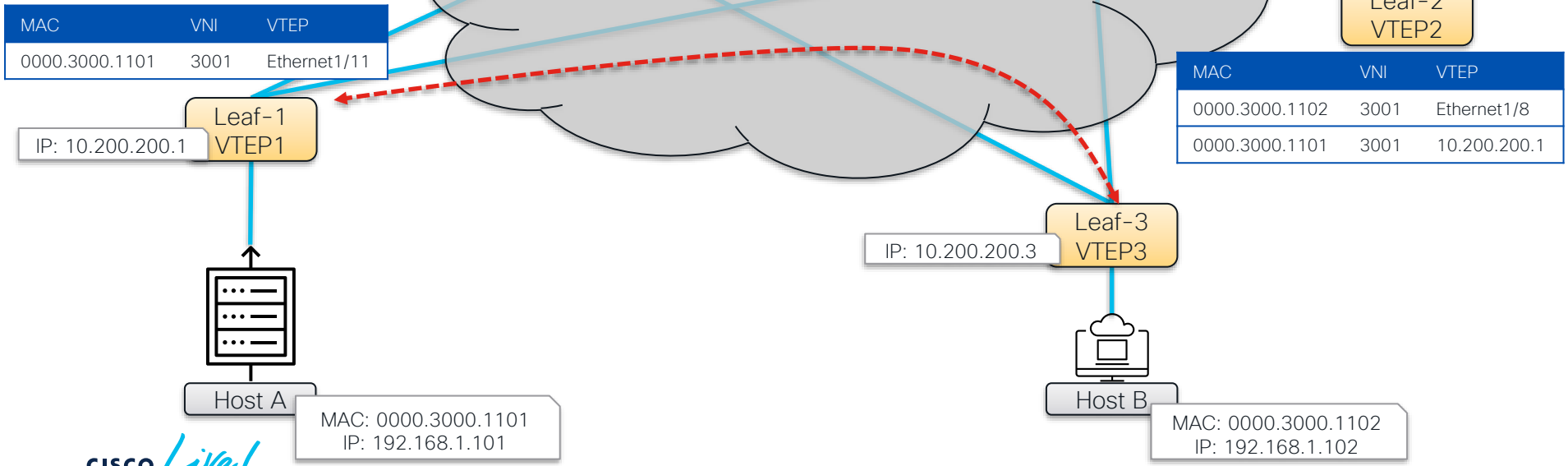


VXLAN Flood and Learn (F&L)
Host A to Host B Communication



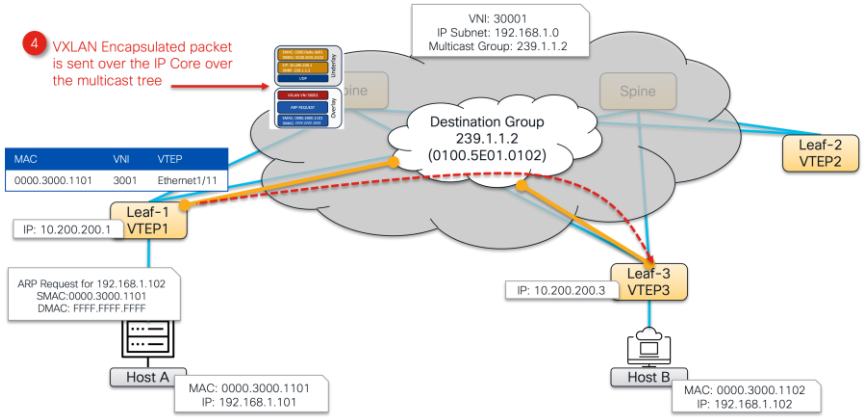
VXLAN Flood and Learn (F&L)
Host A to Host B Communication

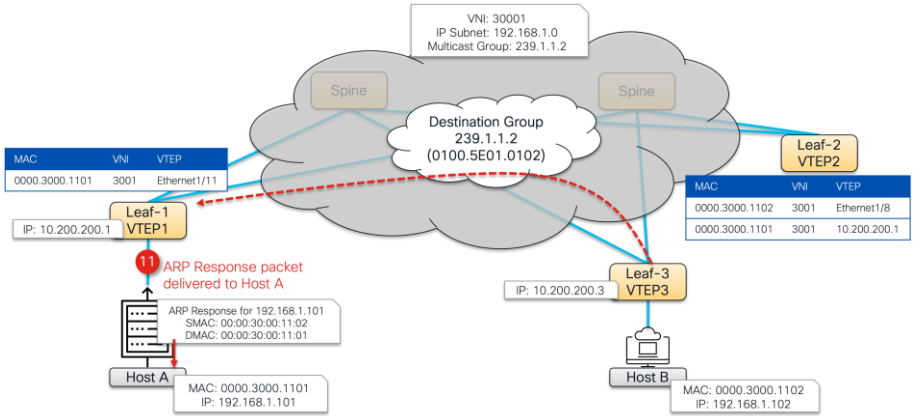
- Host MAC Information is populated at both VTEPs (1&3)
- Host A & B know of each other's MAC-IP Binding
- Subsequent traffic between Host A & Host B is unicast over VXLAN between VTEP1 & VTEP3



Flood & Learn Mechanism

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

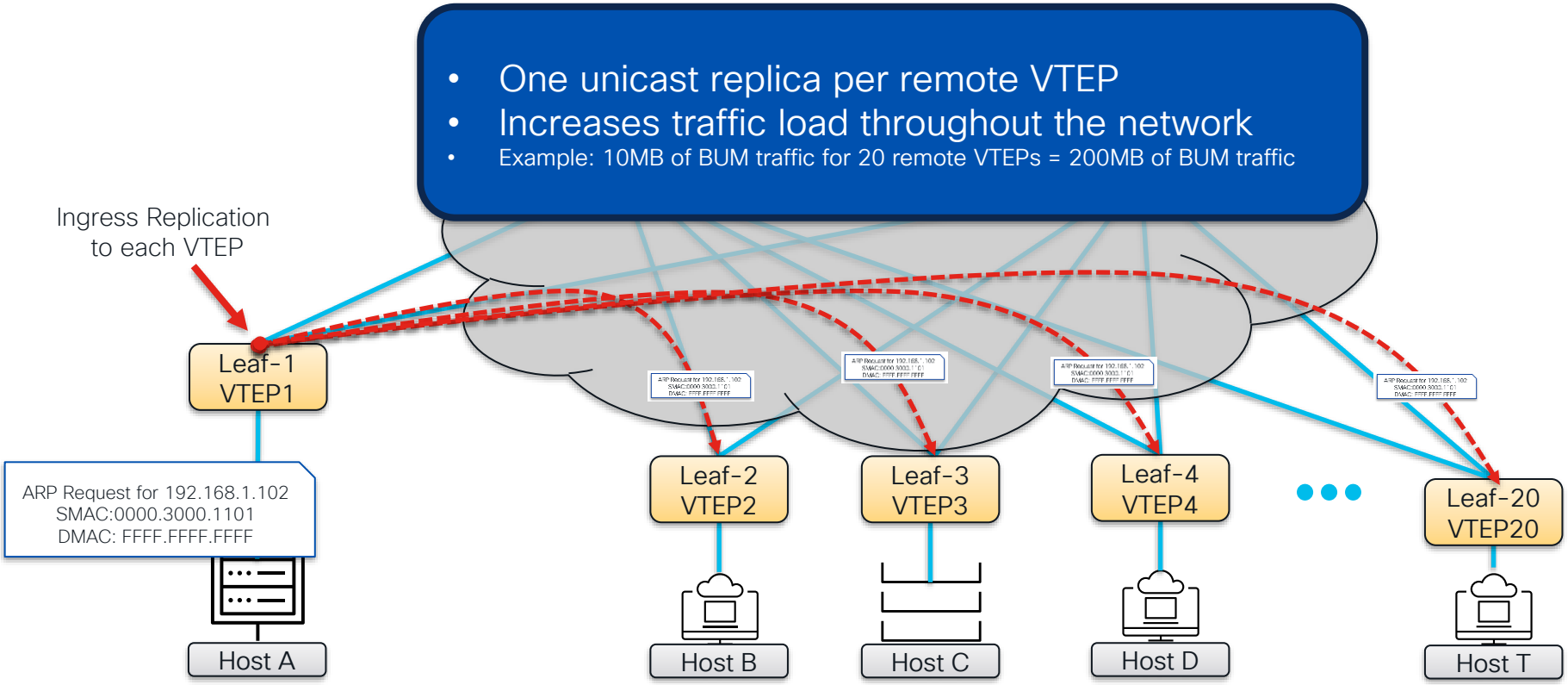




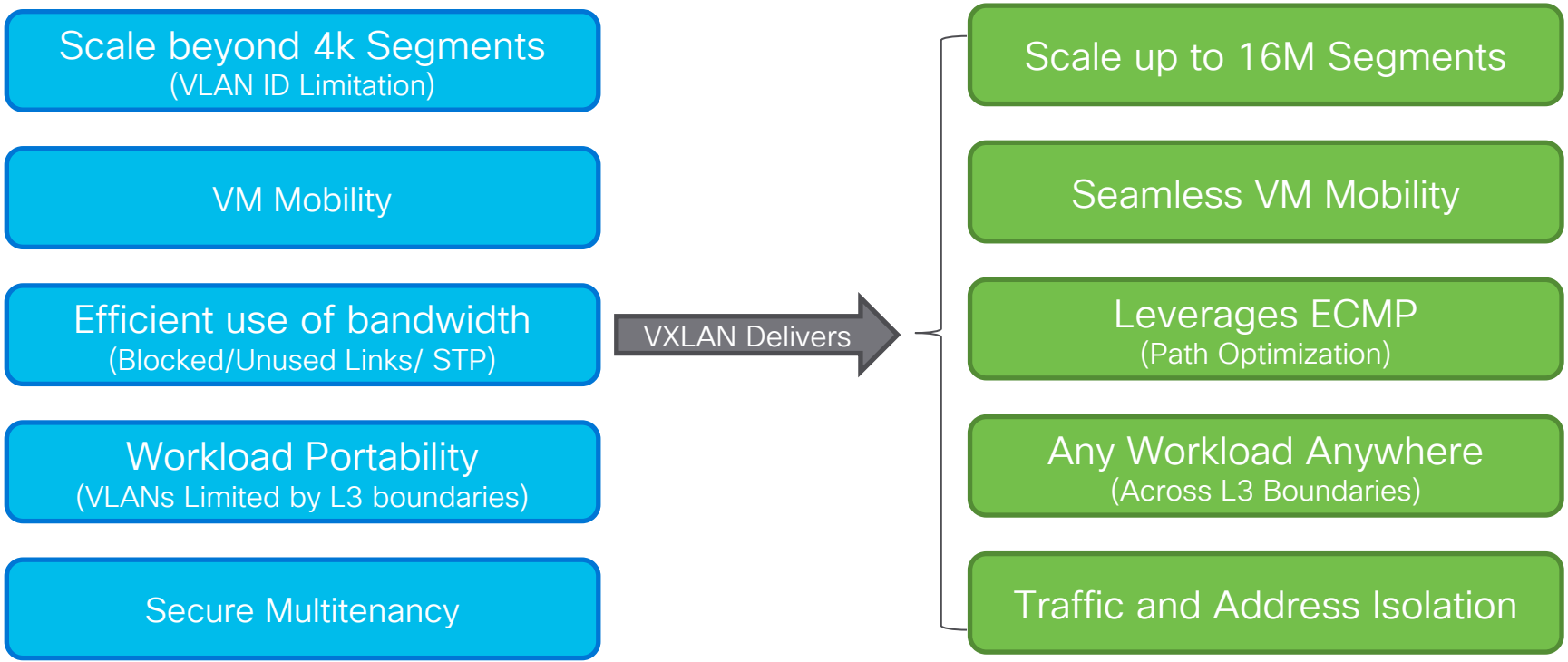
See hidden slides for packet walk

Flood & Learn Mechanism
Ingress Replication or Head-End Replication

- One unicast replica per remote VTEP
- Increases traffic load throughout the network
- Example: 10MB of BUM traffic for 20 remote VTEPs = 200MB of BUM traffic



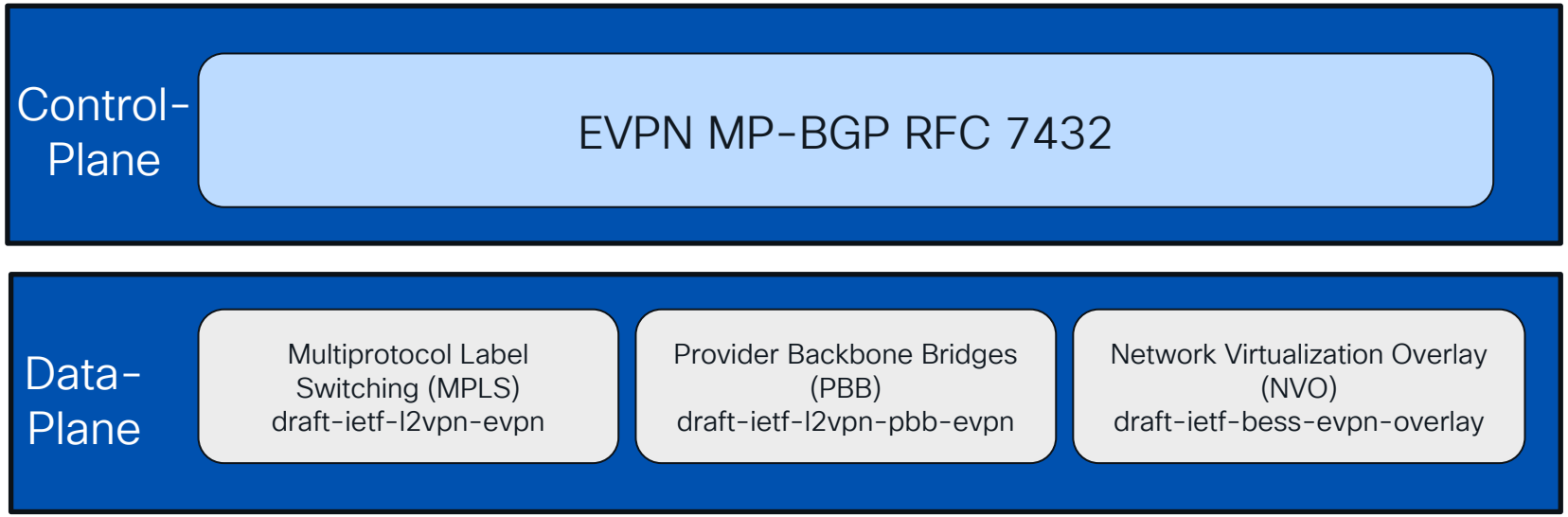
Why VXLAN? – How did we get here?



EVPN
Ethernet VPN

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

EVPN in the Data Center
IETF RFC/Draft for Control and Data Plane

For your reference

RFC / Draft	Title	Category
RFC 7348	Virtual Extensible Local Area Network	Data Plane
RFC 7432	BGP MPLS Based Ethernet VPNS	Control Plane
draft-ietf-bess-evpn-overlay	Network Virtualization Overlay Solution using EVPN	Control Plane
draft-ietf-bess-evpn-inter-subnet-forwarding	Integrated Routing and Bridging in EVPN	Control Plane
draft-ietf-bess-evpn-prefix-advertisement	IP Prefix Advertisement in EVPN	Control Plane
draft-tissa-nvo3-oam-fm	NVO3 Fault Management/OAM	Management Plane

BGP EVPN Route Types – RFC 7432

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

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

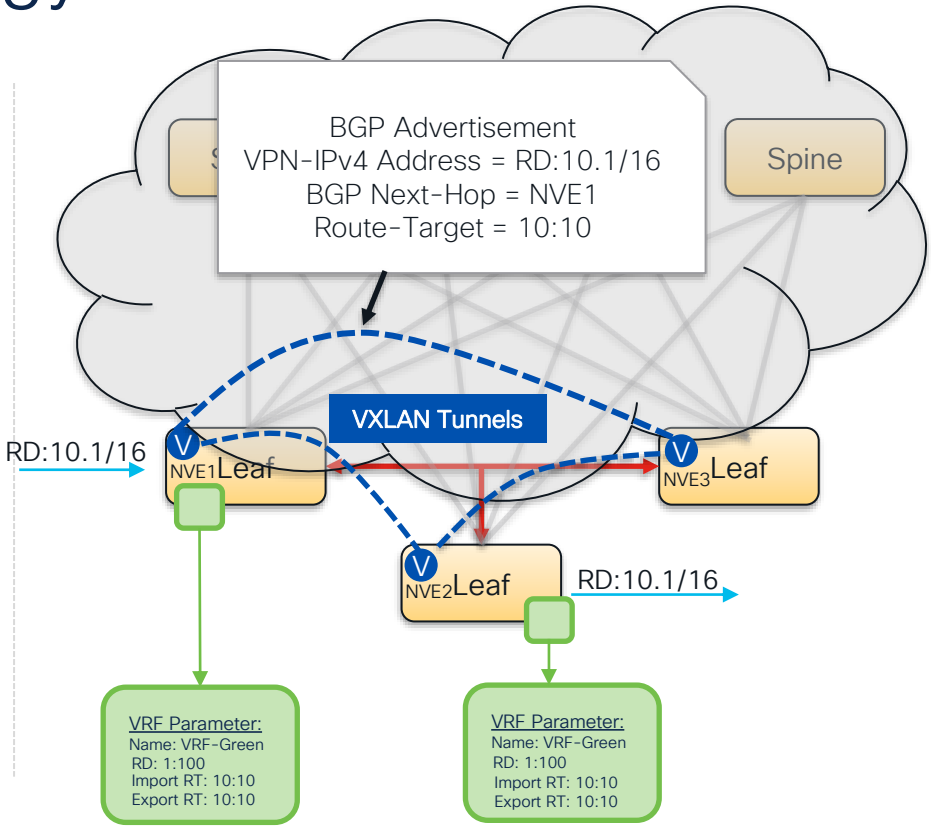
MP-BGP VPN Terminology

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



Control-Plane Comparison

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?!)

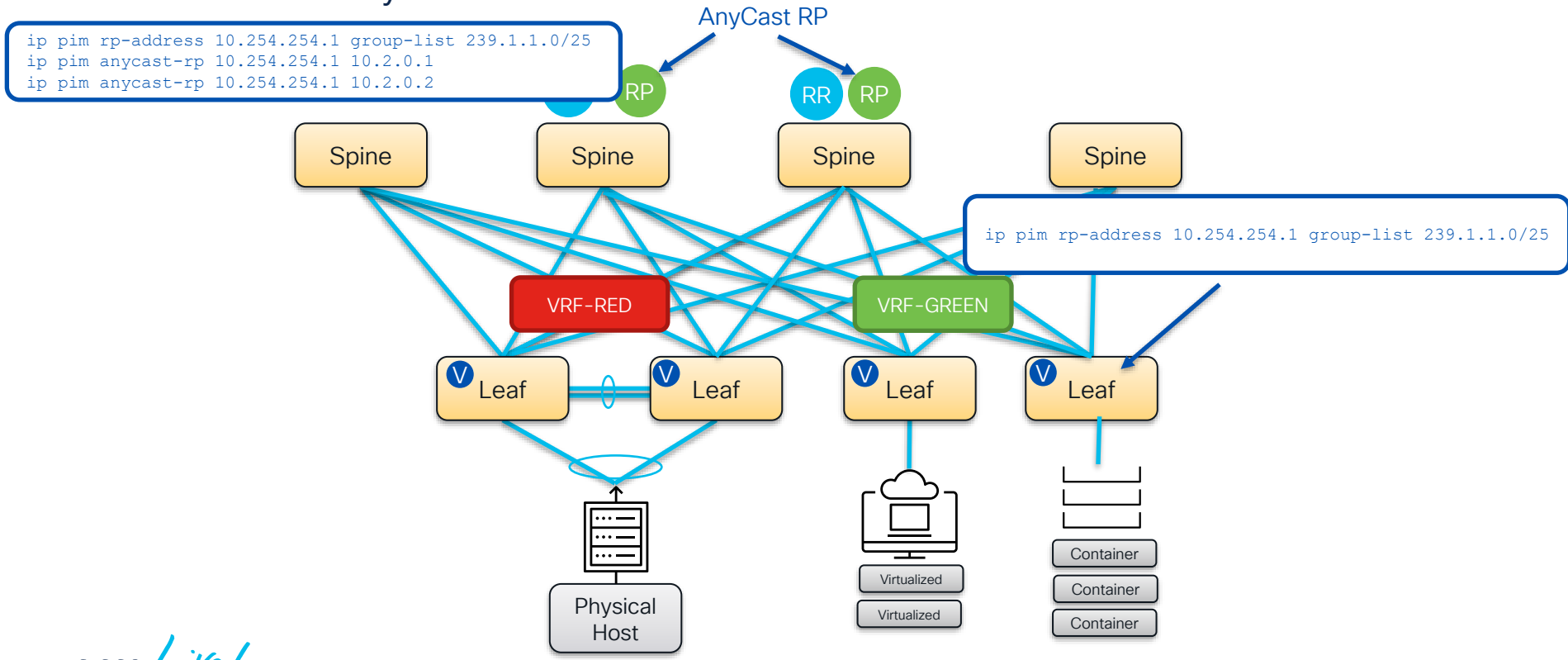
EVPN - Ethernet VPN (RFC8365)

- 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!

Configuring the EVPN VXLAN Fabric

VXLAN BGP EVPN

Network Underlay



VXLAN BGP EVPN

Network Underlay

RR Route Reflector

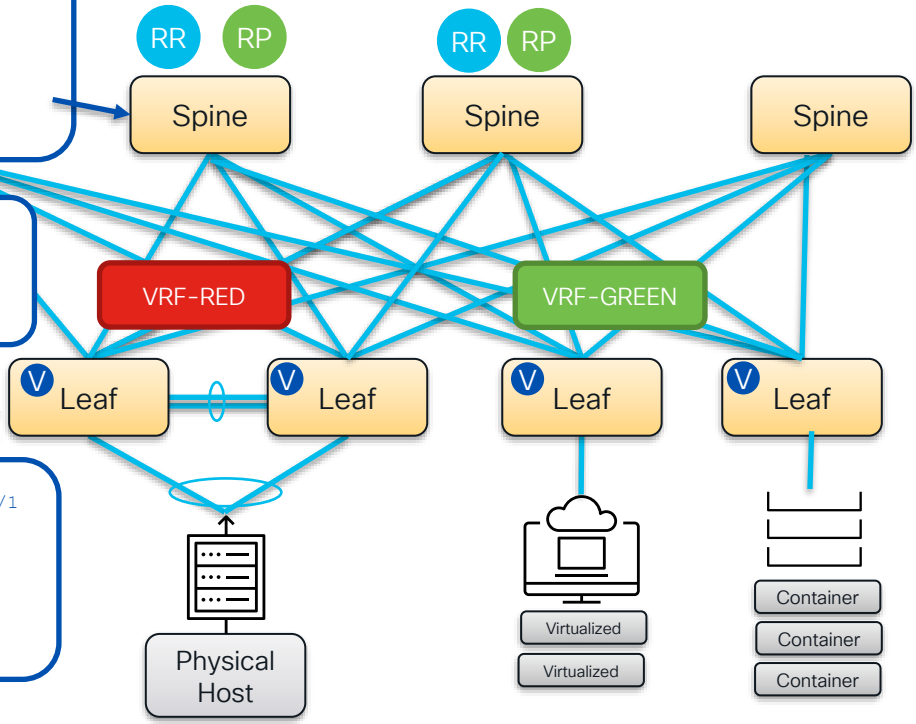
RP Rendezvous Point

```
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
```

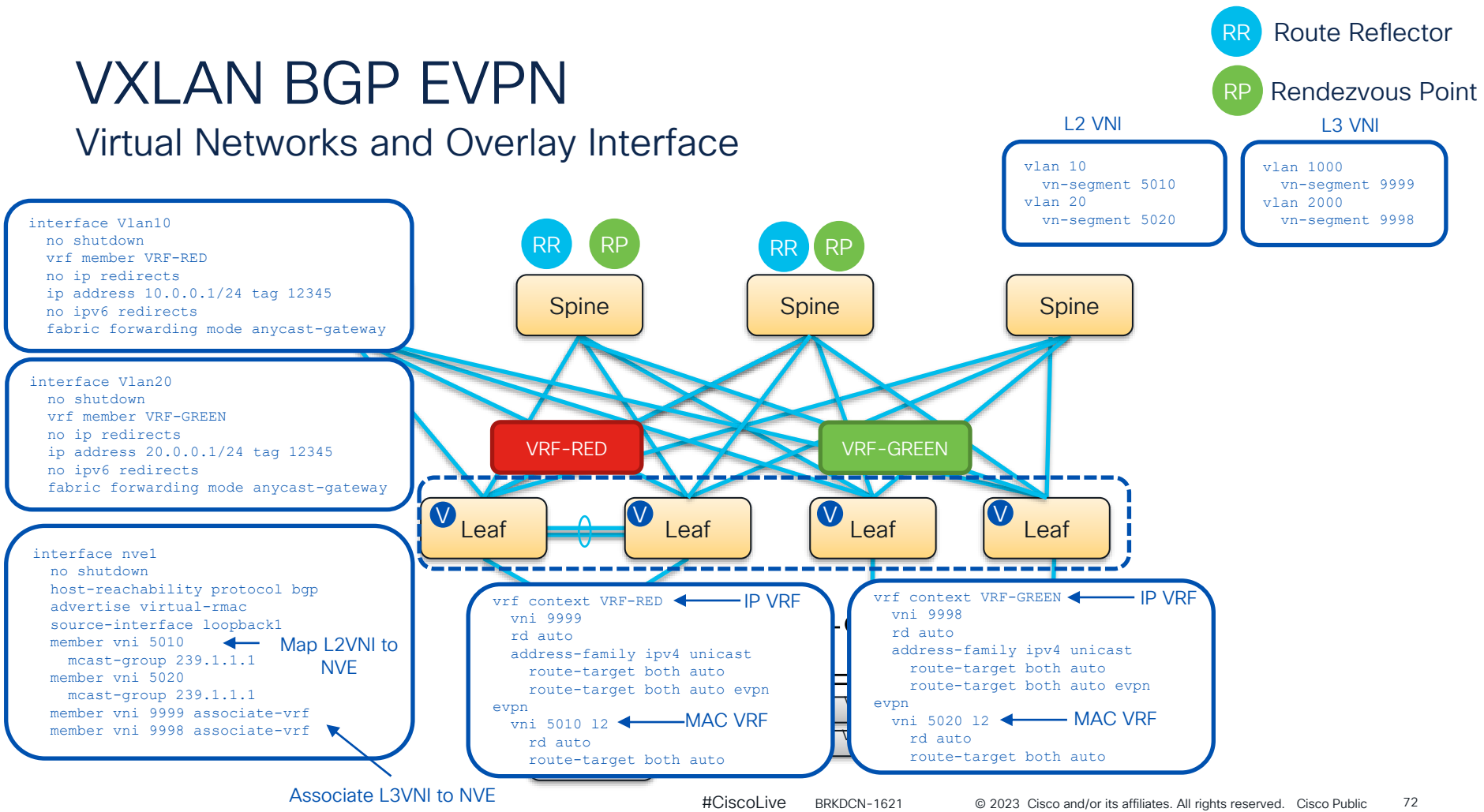
```
router ospf UNDERLAY
router-id 10.2.0.7
```

```
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

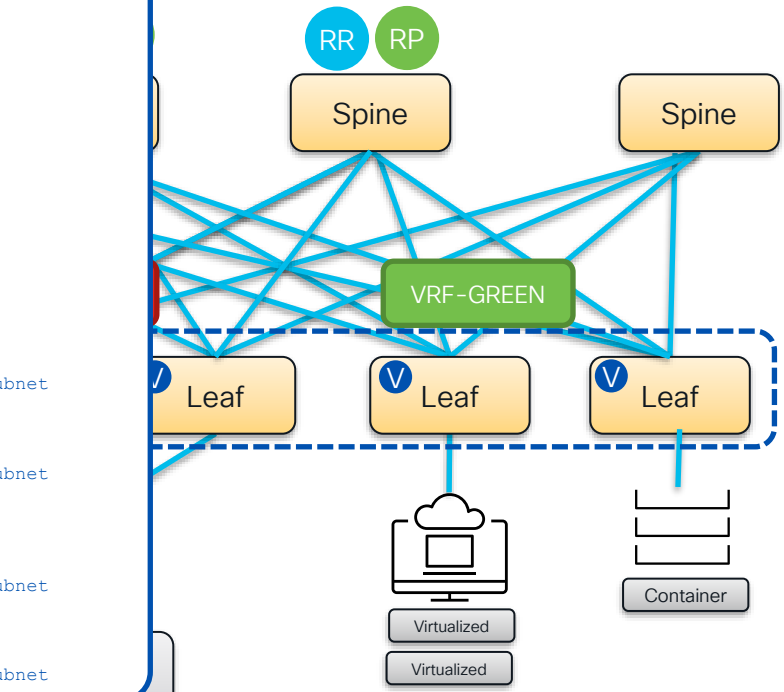
BGP on Leaf and Spine

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

VLAN to VNI Mapping

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

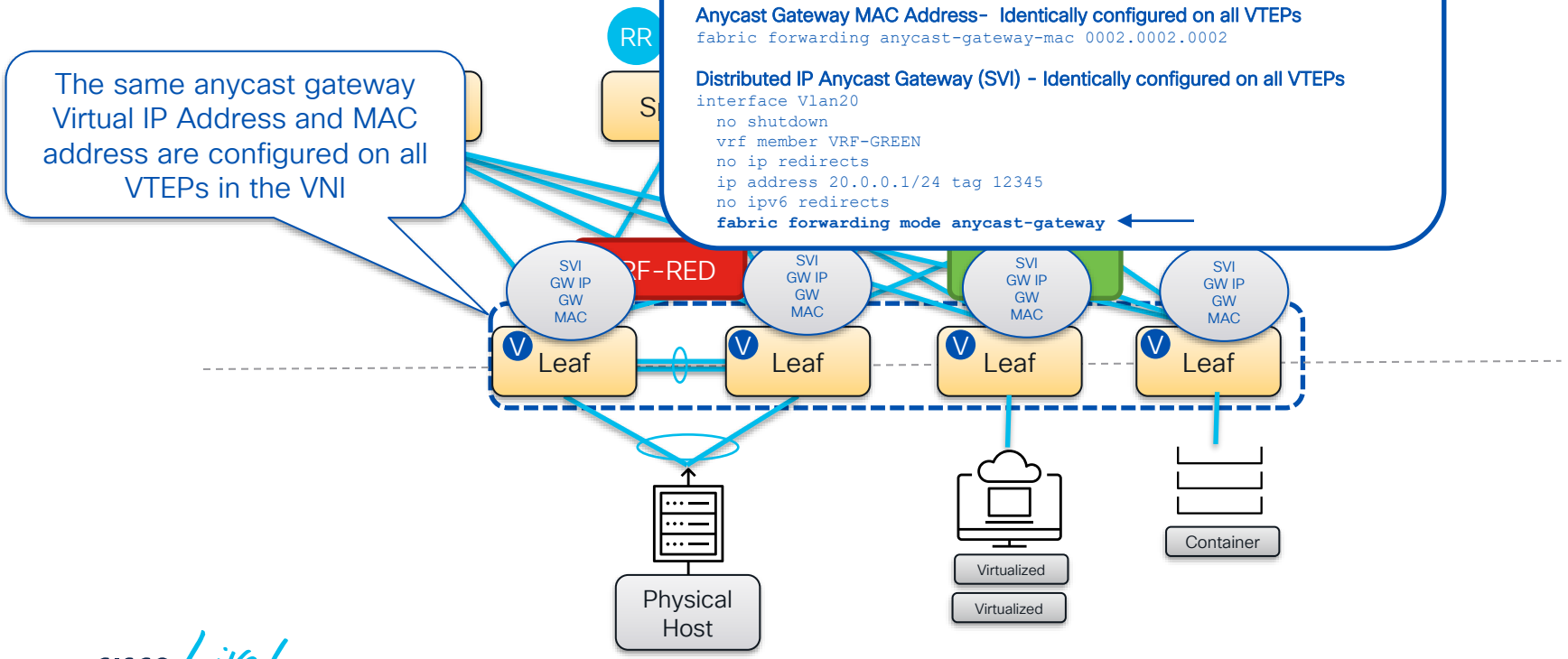
Anycast Gateway MAC Address- Identically configured on all VTEPs

```
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
```

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



Conclusion

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

Fill out your session surveys!

Attendees who fill out a minimum of four session surveys and the overall event survey will get **Cisco Live-branded socks** (while supplies last)!

Attendees will also earn 100 points in the **Cisco Live Challenge** for every survey completed.

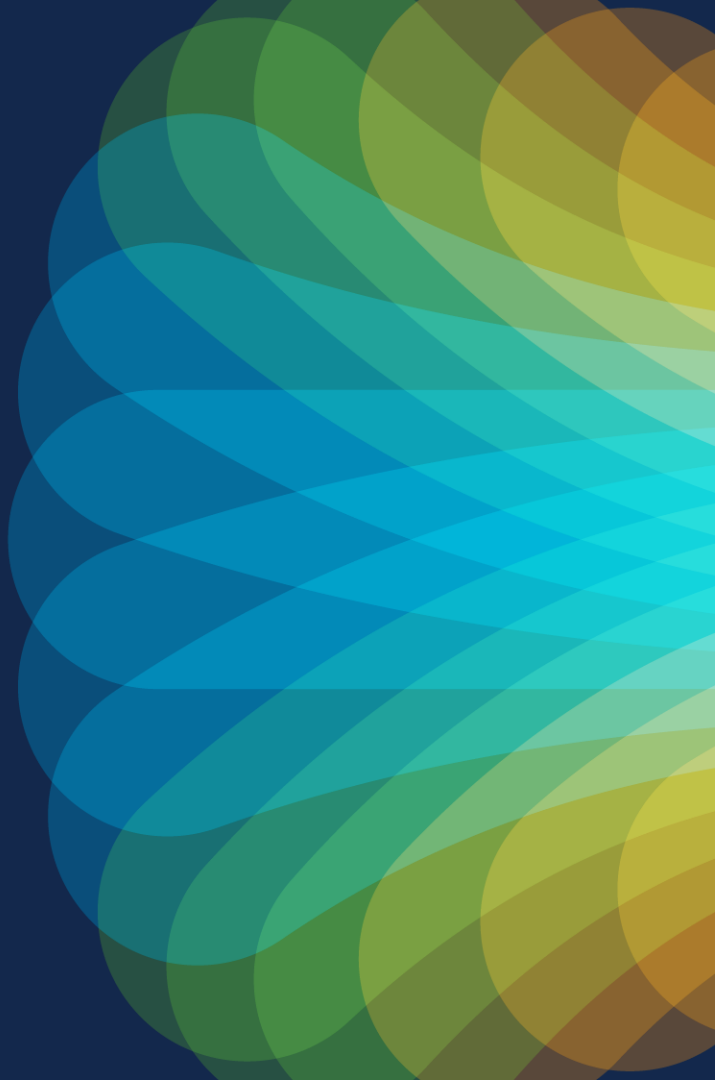
These points help you get on the leaderboard and increase your chances of winning daily and grand prizes

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

The bridge to possible

Thank you




cisco Live!
Let's go
#CiscoLive

Cisco Live
Challenge

Gamify your Cisco Live experience!

Get points for attending this session!

How:

1. Open the Cisco Events App.
2. Click on 'Cisco Live Challenge' in the side menu.
3. Click on View Your Badges at the top.
4. Click the + at the bottom of the screen and scan the QR code: