



Cisco Support Community Expert Series Webcast


Cisco Data Center Overlays with focus on VXLAN

Vishal Mehta, CCIE Data Center, SP, and R&S

October 20, 2015

Upcoming Events


<https://supportforums.cisco.com/expert-corner/events>


**Cisco Support Community**
[Community Directory](#) [Expert Corner](#) [Community Corner](#) [Solutions](#)
Home / [Expert Corner](#) / [Events](#) Language: [English](#)

Events

[Ask the Experts](#) [Webcasts](#)

Cisco experts engage in discussions with you, our members, on specific networking issues. Each event runs for a two-week period.

**Ask the Expert: Cisco Data Center Overlays with Focus on VXLAN**
Welcome to this Cisco Support Community Ask the Expert conversation. This is an opportunity to learn and any ask questions aboutCisco Data Center Overlays with focus on VXLAN (Virtual Extensible Local Area Network) to Vishal Mehta and Pranav Doshi ...
Ends October 30, 2015
[Join the Discussion >](#)

**Ask the Expert: Switch and IOS Architecture and Unexpected Reboots on all Cisco Catalyst**
This session will provide an opportunity to learn and ask questions about Cisco Catalyst Switches IOS architecture, and how to troubleshoot any unexpected reboots and other errors on switches. Ask questions from Monday, October 5 to Friday, ...
Ended October 16, 2015
[Read the Q&A >](#)

Become an Event Top Contributor

Participate in Live Interactive Technical Events and much more
<http://bit.ly/1jll93B>



Top Contributors

Recognition Program | VIPs | Spotlight Awards | Hall of Fame | **Events Top Contributors** | Expert Interviews

Cisco Designated VIPs

The Cisco Designated VIP program recognizes the top external individual contributors in Cisco's online communities, including the Cisco Support Community (CSC), Cisco Learning Network (CLN) and the Cisco Developers Network (CDN). Cisco Designated VIPs are recognized by their peers for their expertise and tireless contributions, and their abundant participation is vital to community success. With this program, Cisco formally recognizes the positive, valuable influence our top individual members exert on the communities overall.

To learn more, please visit our [FAQ](#)

CISCO DESIGNATED VIP < 2015 2014 2013 2012 2011 >

 Aman Soi 2015 IP Telephony	 Anthony Holloway 2015 Contact Center
 Ayodeji otadipo Okantawon 2015 IP Telephony	 Carlo Poggiarelli 2015 IP Telephony
 Chris Deren 2015 IP Telephony	 Dan Lukes 2015 Small Business, Voice
 Gergely Szabo 2015 Contact Center	 John Blakley 2015 LAN, WAN

Experts Bureau

Use the Cisco Experts Bureau to find, connect, and follow recognized Subject Matter Experts and the programs they participate in regularly. The Experts Bureau comprises Cisco employees as well as Partners and Customers who have contributed to, or been selected for knowledge sharing programs on the Cisco Support Community, such as Webcasts, Ask the Expert Events, Facebook Forums, Tech-Talks, Meetups, and Blogs.

If you have interest in participating, apply online through this [simple form](#). After applying, a member of the Cisco Support Community team will be in contact with additional details.

Rate Content



Encourage and acknowledge people who generously share their time and expertise

Now your ratings on documents, videos, and blogs count give points to the authors!!!

So, when you contribute and receive ratings you now get the points in your profile.

Help us to recognize the quality content in the community and make your searches easier. Rate content in the community.

<https://supportforums.cisco.com/blog/154746>

Cisco Support Community Expert Series Webcast

Vishal Mehta

CCIE Data Center SP and R&S #37139



Ask the Expert Event following the Webcast

Now through October 30th

<https://supportforums.cisco.com/discussion/12604376/ask-expert-cisco-data-center-overlays-focus-vxlan>



Join the discussion for these Ask The Expert Events:

<http://bit.ly/events-webinar>

Thank You For Joining Us Today!



If you would like a copy of the presentation slides, click the PDF file link in the chat box on the right or go to:

<https://supportforums.cisco.com/document/12675756/cisco-data-center-overlays-focus-vxlan-slides-webcast>





Submit Your Questions Now!

Use the Q & A panel to submit your questions and the panel of experts will respond.

Please take a moment to complete the survey at the end of the webcast

Polling Question 1

Are you planning to implement VXLAN in your network ?

Yes

No

Still Evaluating



Cisco Support Community Expert Series Webcast

Cisco Data Center Overlays with focus on VXLAN

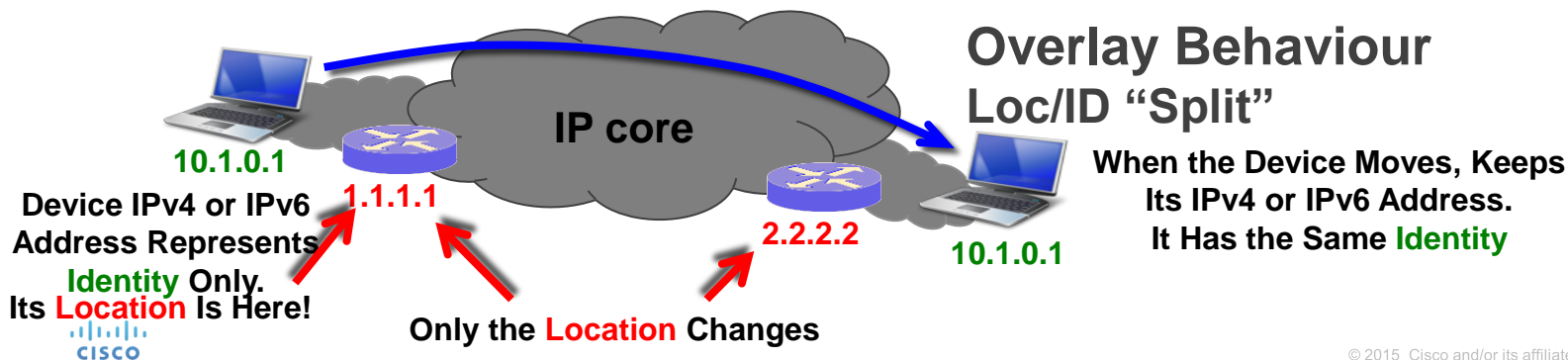
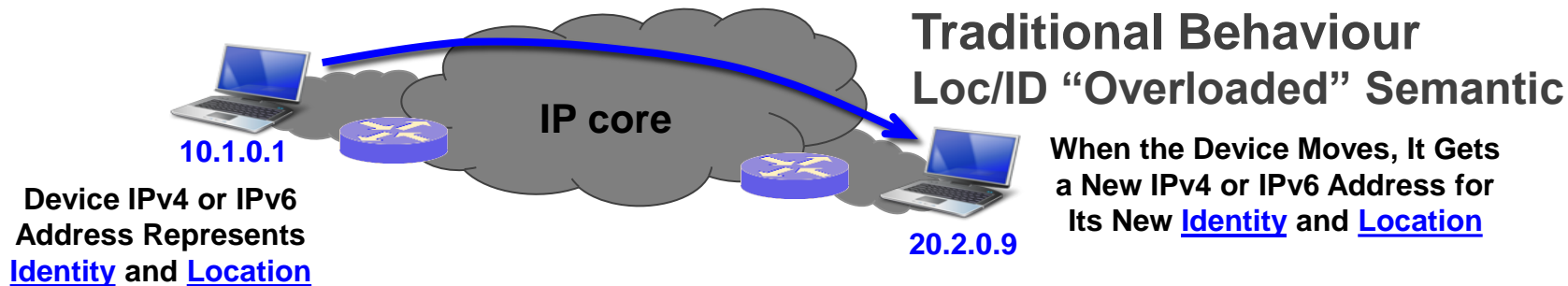
Vishal Mehta, CCIE Data Center, SP, and R&S

October 20, 2015

Agenda

- Overlays in Data Center
- Why VXLAN
- VXLAN Evolution
- Deployment Scenarios
- Comparison

Why Do We Need Overlays?



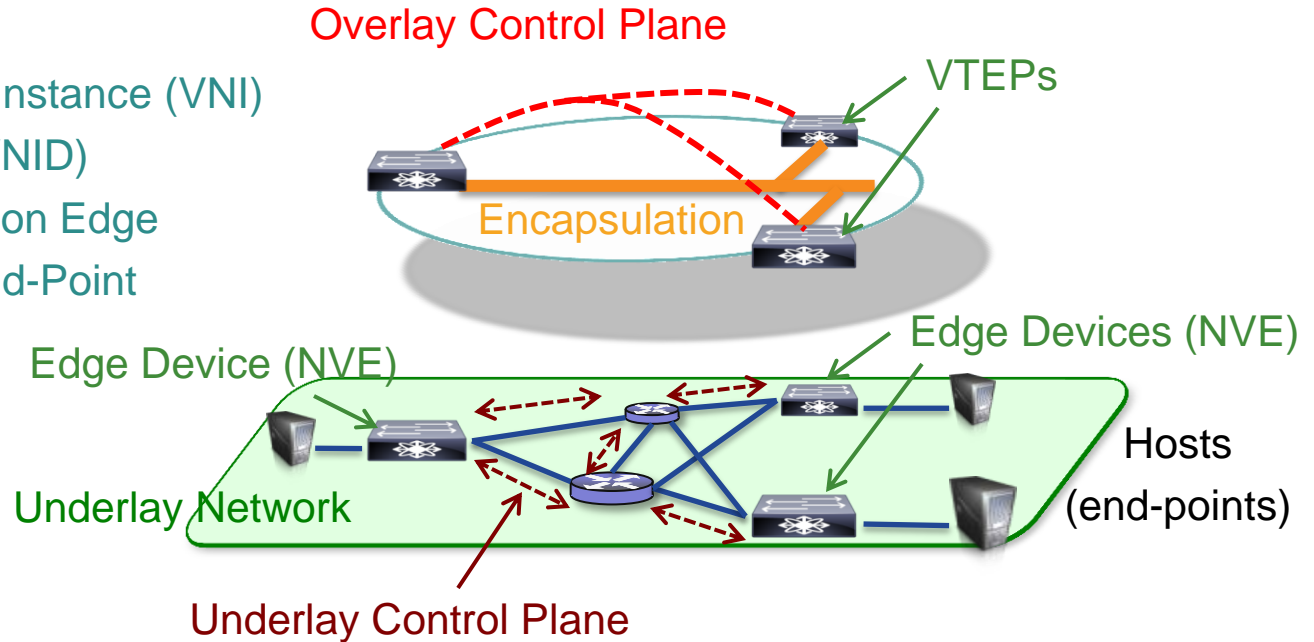
Overlay Taxonomy

Service = Virtual Network Instance (VNI)

Identifier = VN Identifier (VNID)

NVE = Network Virtualization Edge

VTEP = VXLAN Tunnel End-Point



Overlay Attributes

Service

Layer 2 Service

Layer 3 Service

Edge Device

Host Overlays

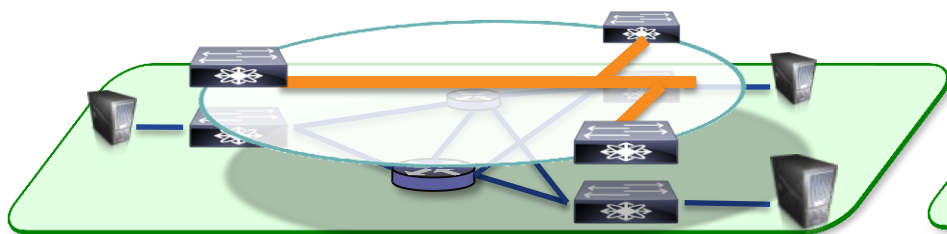
Network Overlays

Signalling

Data Plane
Learning

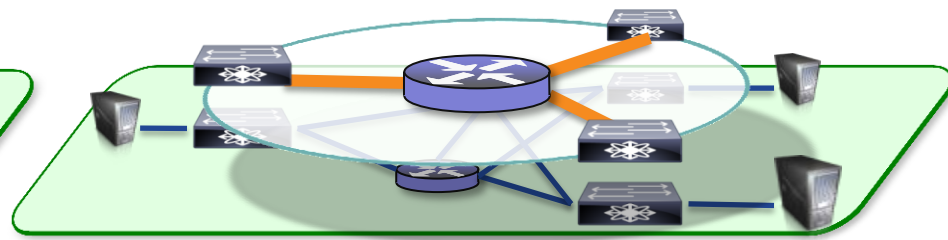
Control Plane
Learning

Types of Overlay Service



Layer 2 Overlays

- Emulate a LAN segment
- Transport Ethernet Frames (IP and non-IP)
- Single subnet mobility (L2 domain)
- Exposure to open L2 flooding
- Useful in emulating physical topologies



Layer 3 Overlays

- Abstract IP based connectivity
- Transport IP Packets
- Full mobility regardless of subnets
- Contain network related failures (floods)
- Useful in abstracting connectivity and policy

Hybrid L2/L3 Overlays offer the best of both domains

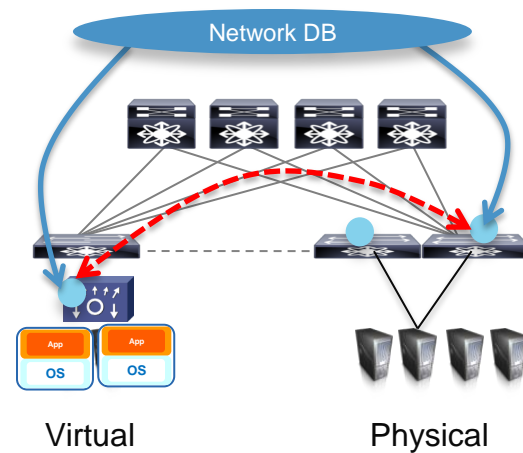
Overlay Edge Device & Data Plane Evolution

Service

Layer 2 Service
Layer 3 Service

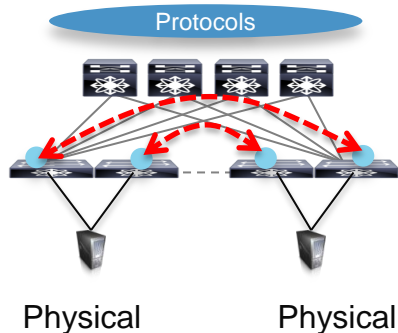
Edge Device

Host Overlays
Network Overlays



Overlay Network Evolution: Edge Devices

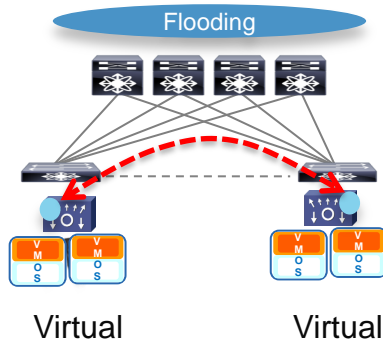
Network Overlays



- Router/switch end-points
- Protocols for resiliency/loops
- Traditional VPNs
- OTV, VPLS, LISP, FP

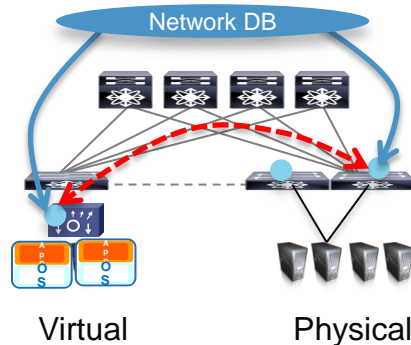


Host Overlays



- Virtual end-points only
- Single admin domain
- **VXLAN**, NVGRE, STT

Hybrid Overlays

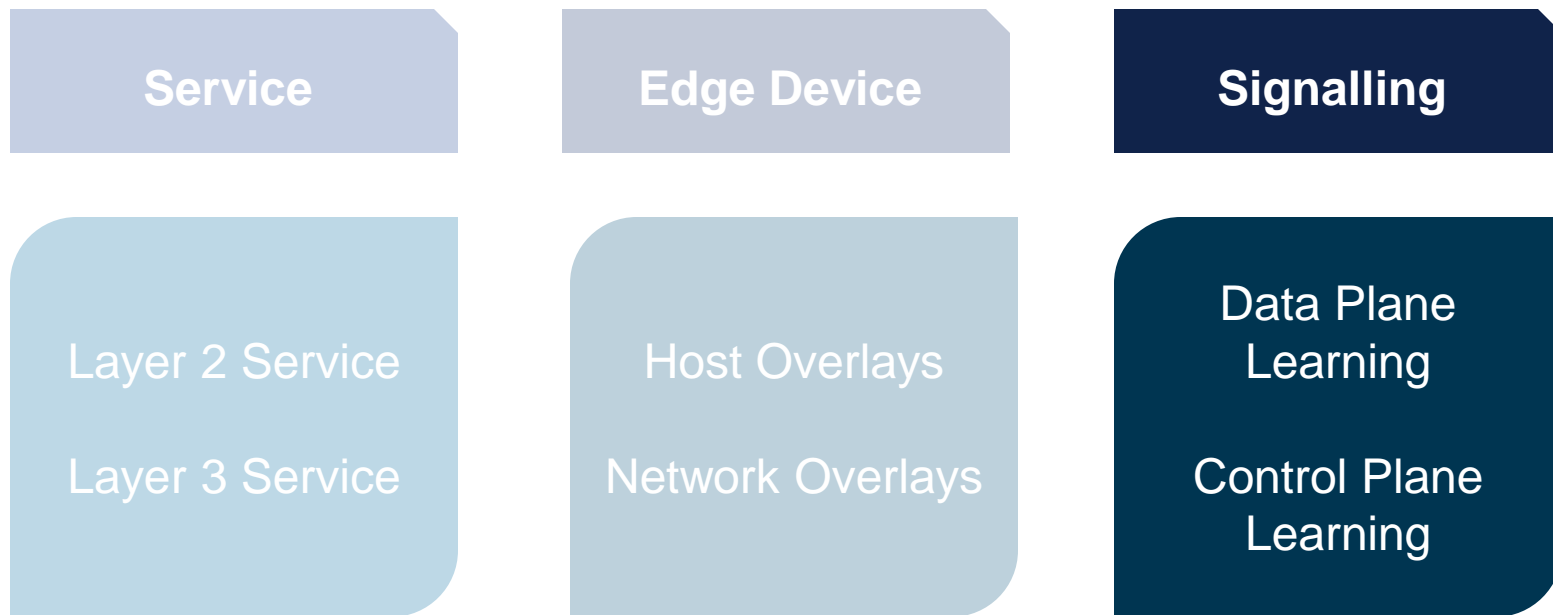


- Physical and Virtual - VXLAN
- Resiliency + Scale
- x-organization/federation
- Open Standards



Tunnel End-points

Overlay Signalling Evolution



Overlay Signalling

Data Plane

- Based on gleaning information from data plane events
 - Example: Source Learning on bridges, FabricPath, VXLAN (Multicast)
- Provides the following:
 - Address advertisement/mapping
 - Some tunnel management is possible
 - Does not provide Service Auto-discovery
- Requires a flood facility for data plane events to propagate:
 - Multicast tree
 - Unicast replication group at the head-end



Control Plane

- Provides:
 - Service Discovery
 - Address Advertising/Mapping
 - Tunnel Management
 - Extensions for multi-homing and advanced services can be provided

Protocol or Controller:

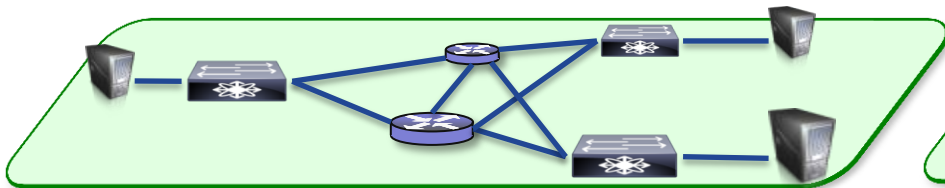
- Routing Protocol amongst Edge Devices
 - BGP, IS-IS, LISP
- Central database on a Controller
 - Distributed Virtual Switches (OVS, N1Kv/VSM)

Push or Pull:

- Push all information to all Edge Devices
 - BGP, IS-IS, Controllers
- Pull and cache on demand @ ED
 - LISP, DNS, Controllers

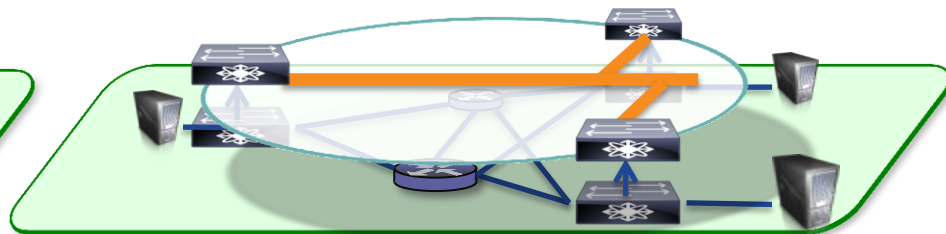
Modern DC Fabric

Seek well integrated best in class Overlays and Underlays



Robust Underlay/Fabric

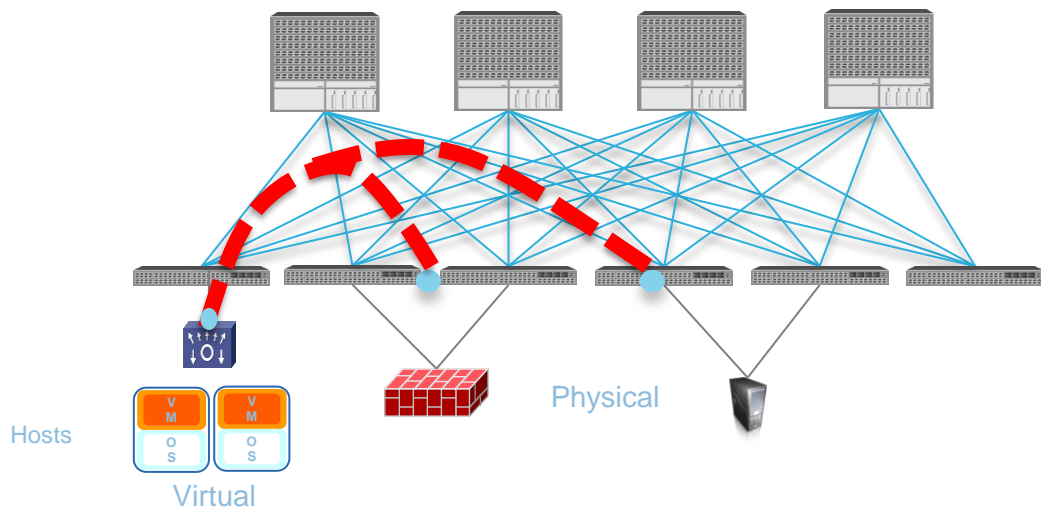
- High Capacity Resilient Fabric
- Intelligent Packet Handling
- Programmable & Manageable



Flexible Overlay Virtual Network

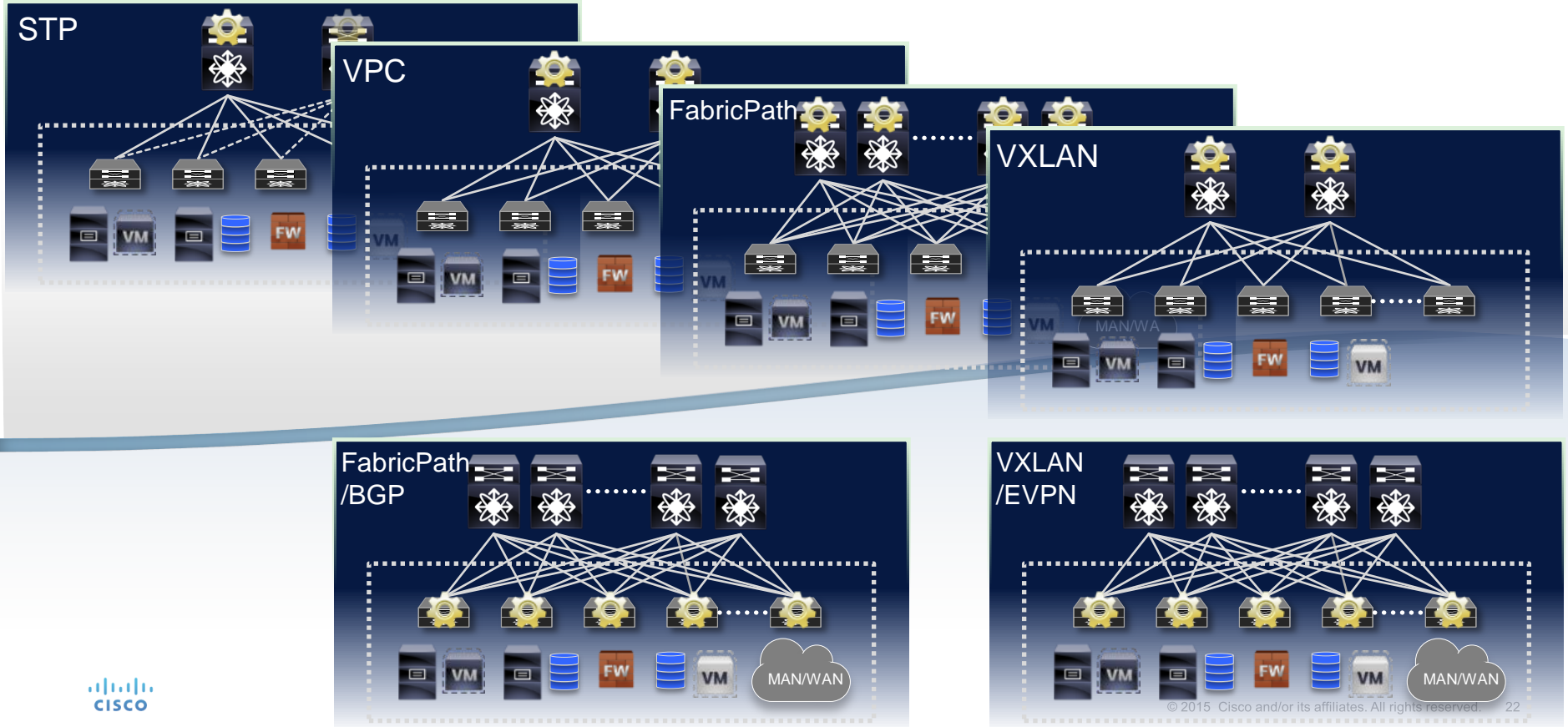
- Mobility – Track end-point attach at edges
- Scale – Reduce core state
 - Distribute and partition state to network edge
- Flexibility/Programmability
 - Reduced number of touch points

Trend: Flexible Data Center Fabrics



- Workload Mobility
- Workload Placement
- Segmentation
- Scale
- Automation & Programmability
- L2 + L3 Connectivity
- Physical + Virtual
- Open
- Network Virtualization

Data Center "Fabric" Journey



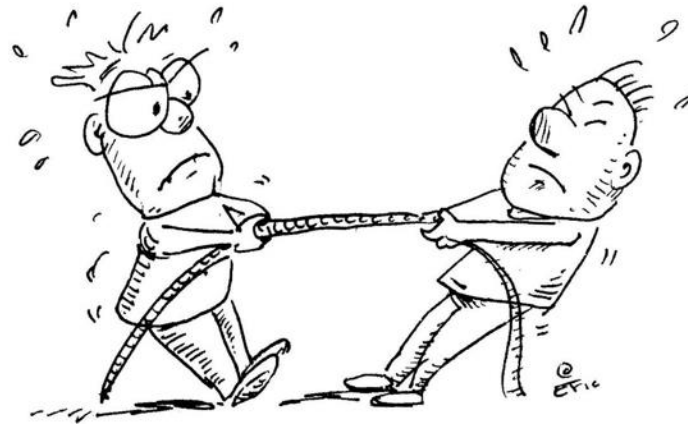
Which Encapsulation?

VXLAN

NVGRE

LISP

MPLS

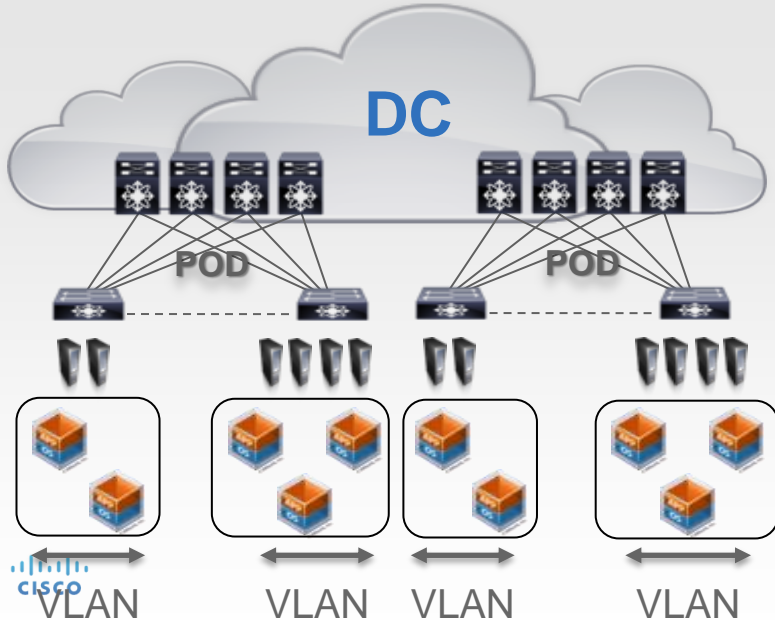


FabricPath

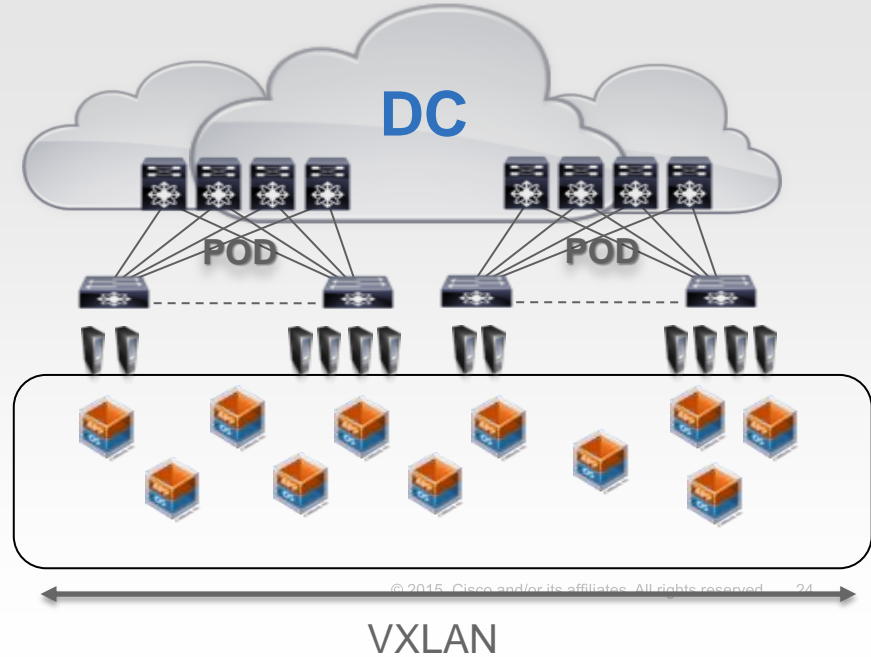
Why VXLAN

Standards based (VXLAN-RFC7348) Overlay with 16M identifiers
Leverages Layer-3 ECMP – all links forwarding
Integration of Physical and Virtual Nodes

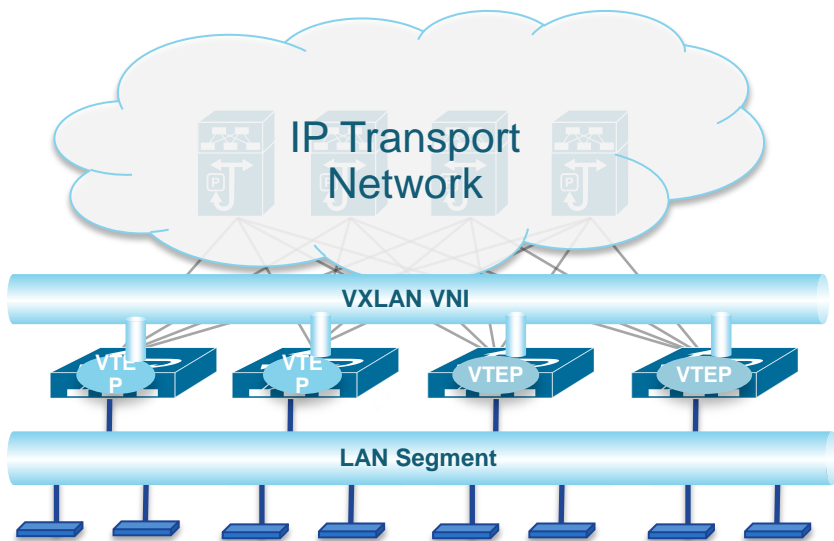
Limited Rack-wide VM Mobility



Virtual/Cloud Data Center



Network Virtualization with VXLAN



Underlay Network:

- IP routing – proven, stable, scalable
- Support any routing protocols --- OSPF, EIGRP, IS-IS, BGP, etc.
- ECMP – utilize all available network paths

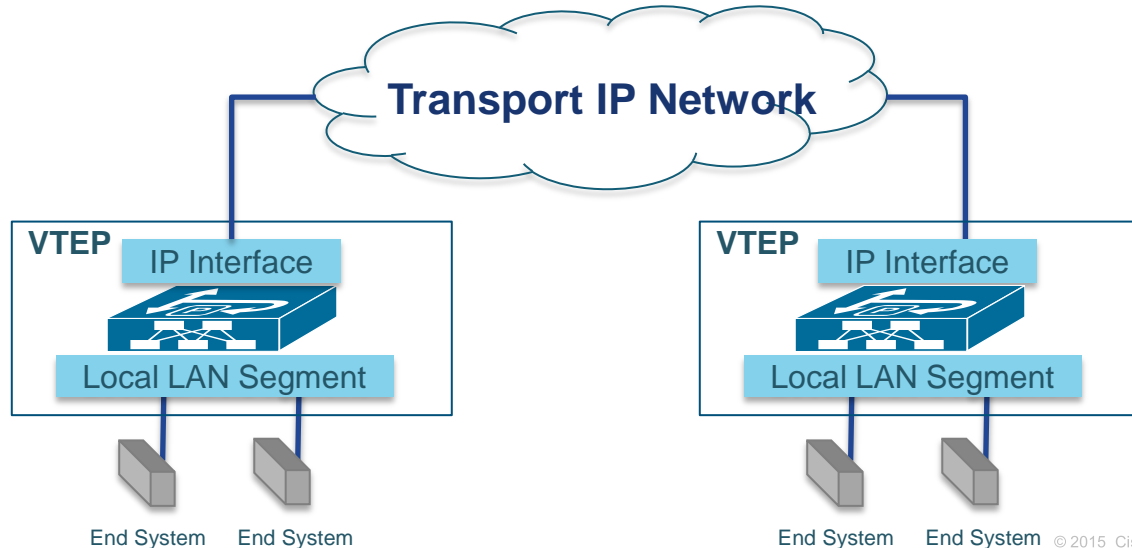
Overlay Network:

- Standards-based overlay
- Layer-2 extensibility and mobility
- Expanded Layer-2 name space
- Scalable network domain
- Multi-Tenancy

VXLAN VTEP

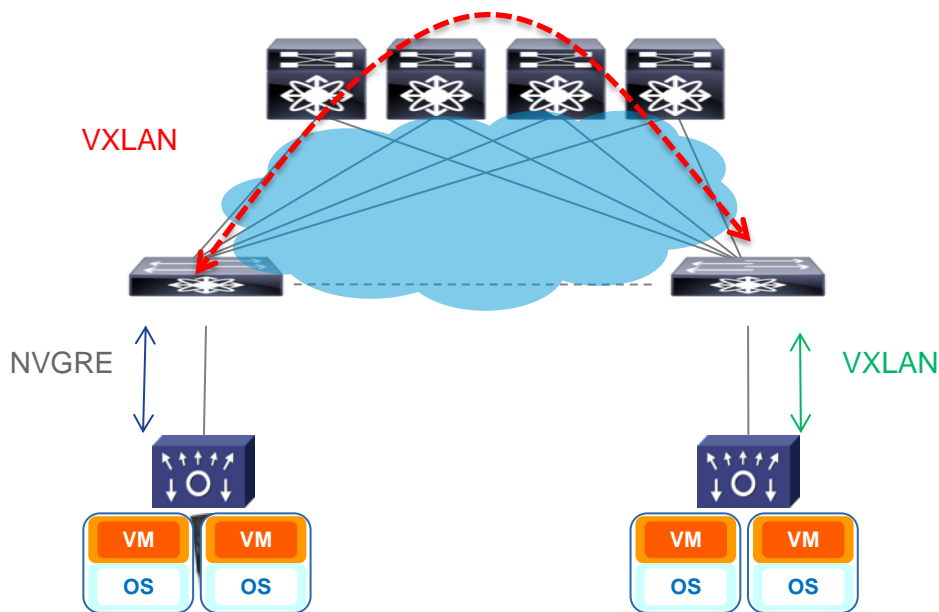
VXLAN terminates its tunnels on VTEPs (Virtual Tunnel End Point).

Each VTEP has two interfaces, one is to provide bridging function for local hosts, the other has an IP identification in the core network for VXLAN encapsulation/decapsulation.



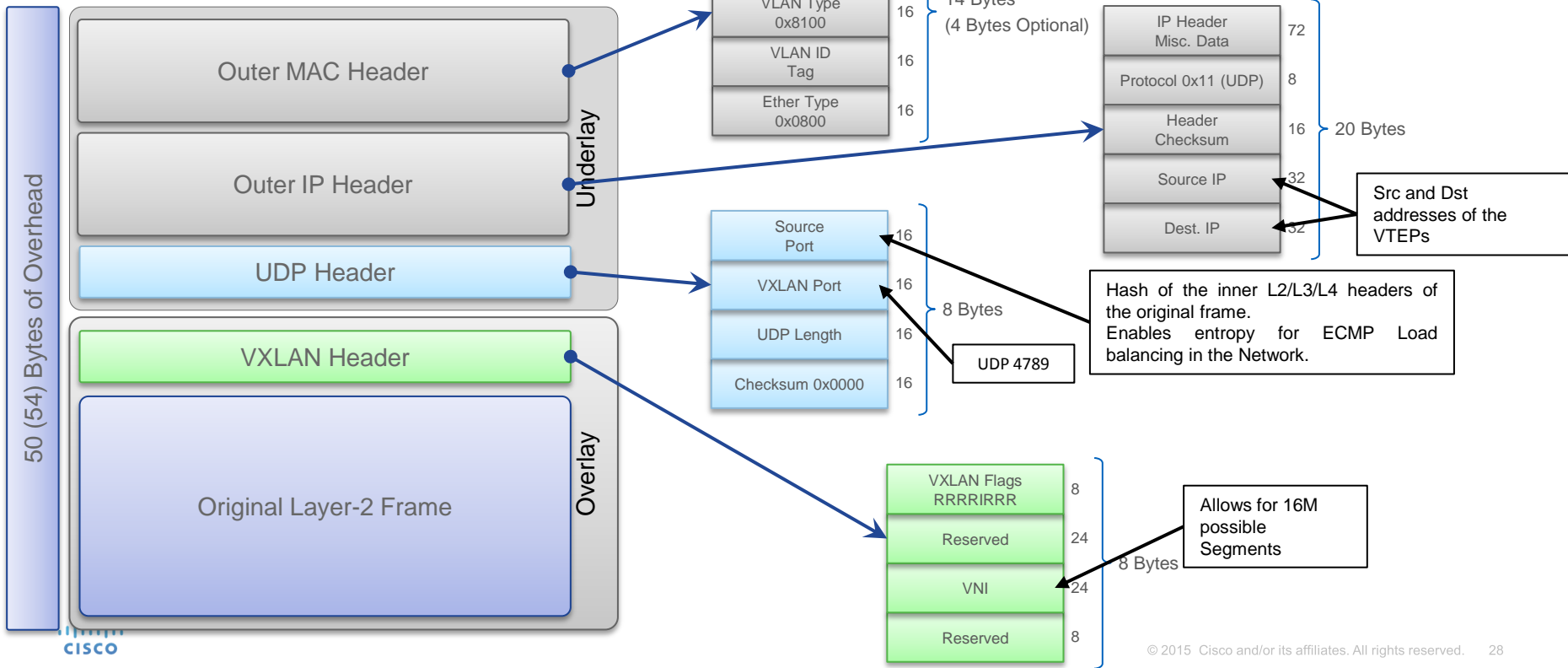
Normalization: The Encapsulation Doesn't Matter

- Intelligence in the Control Plane
- Capabilities Exchange in Control Plane (negotiate encapsulation)
- Normalize to common encapsulation
- Pervasive Multi-encap Gateways for optimal traffic patterns



VXLAN Frame Format

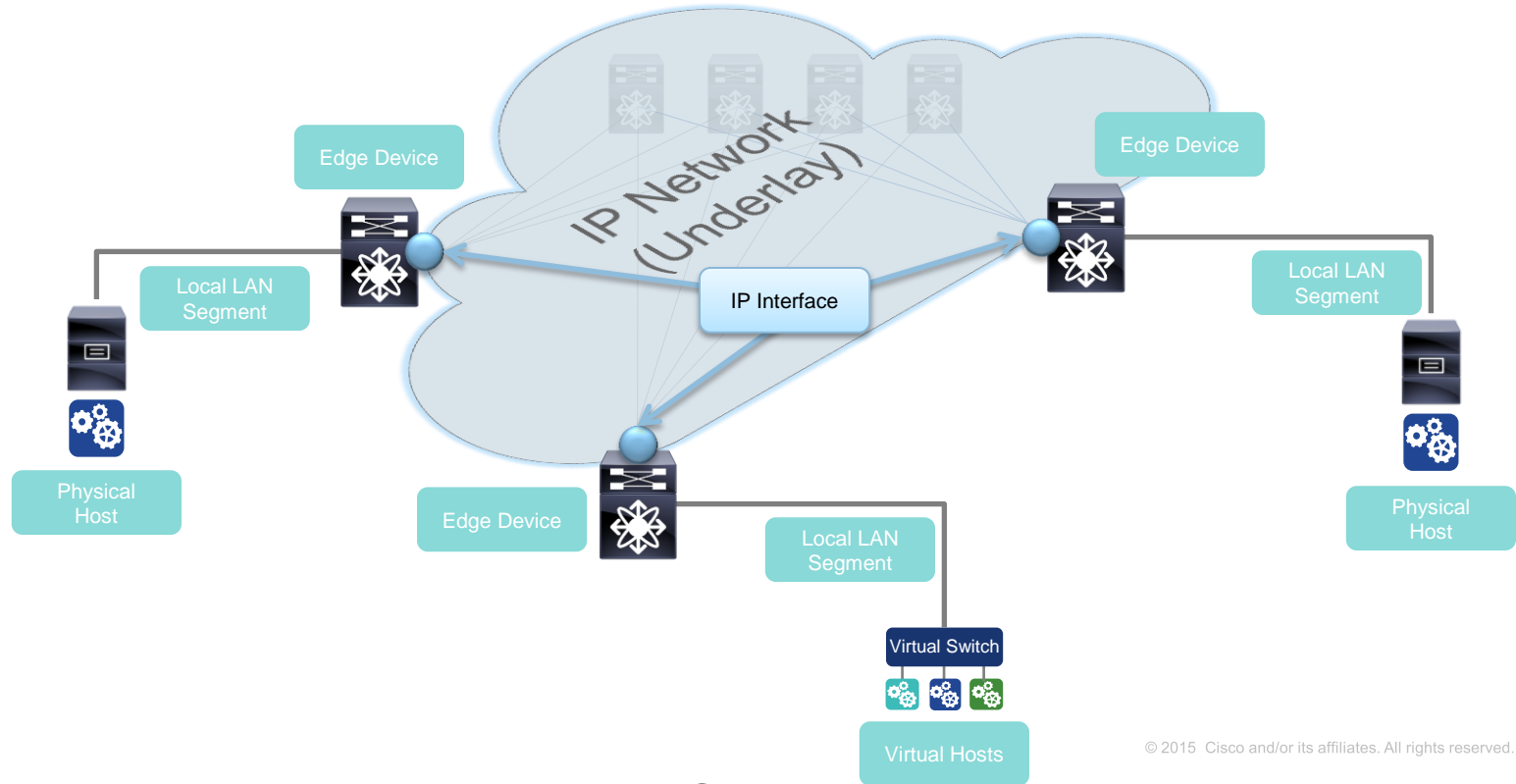
MAC-in-IP Encapsulation



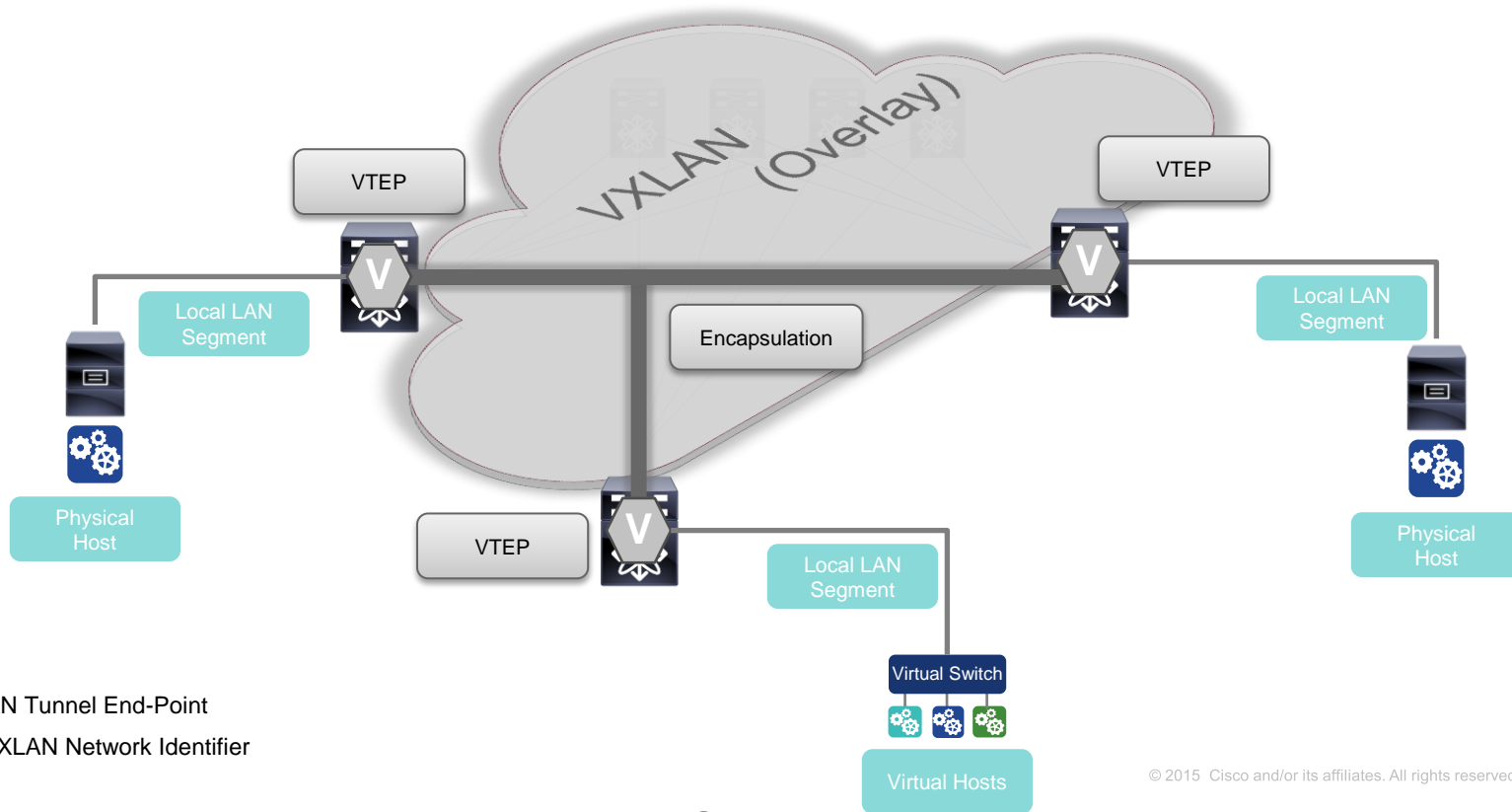
Terminology - Reference

- **Layer-2 VNI :**
 - VNI (VXLAN network identifier) carried in VXLAN packets bridged across VTEPs (VXLAN tunnel end point) . This VNI is configured per VLAN.
- **Layer-3 VNI:**
 - VNI carried in the VxLAN packets routed across VTEPs. This VNI is linked per Tenant VRF.
- **Anycast GW:**
 - All L3 VTEPs are configured with same mac and same subnet for host facing SVI.
- **VRF overlay VLAN:**
 - Every Tenant VRF will need a Vlan to be configured for VXLAN routing.
 - This VLAN is configured with L3-VNI.
- **VXLAN L2 Gateway:**
 - VTEP capable of switching VLAN->VXLAN, VXLAN->VLAN packets with in same VNI.
- **VXLAN L3 Gateway:**
 - VTEP capable of routing packets across different VNIs.

VXLAN Overview (1)



VXLAN Overview (2)



VTEP – VXLAN Tunnel End-Point

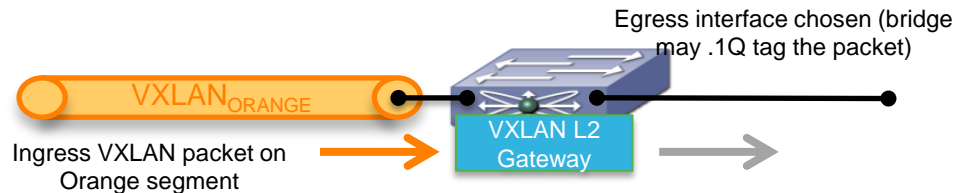
VNI/VNID – VXLAN Network Identifier



VXLAN L2 and L3 Gateways

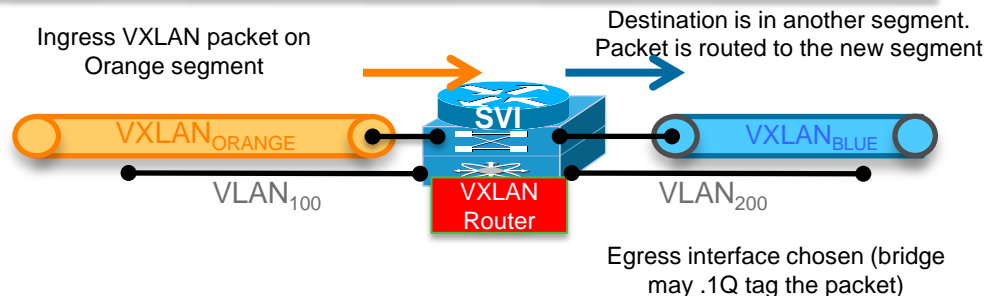
Connecting VXLAN to the broader network

L2 Gateway: VXLAN to VLAN Bridging



L3 Gateway: VXLAN to X Routing

- VXLAN
- VLAN



Cisco VXLAN Portfolio

Cisco VXLAN Solutions

Scale

Secure
Multi-tenancy

Workload
Mobility

Workload
Anywhere



ASR1000
CSR1000



Nexus
1000



Nexus
3100



Nexus
5600



Nexus
7000



Nexus
9000



ASR9000

L2 Gateway

L3 Gateway

BGP EVPN
Control Plane

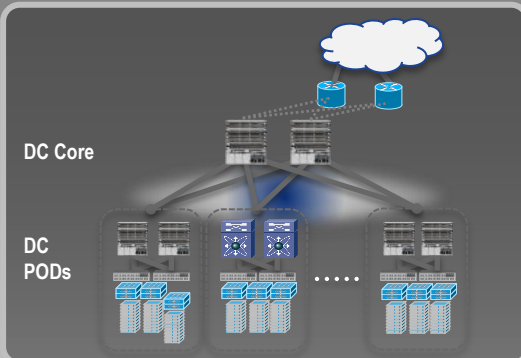
Anycast
Gateway

Head End
Replication



Common Building Blocks

EXISTING 3-TIER DESIGNS



Existing 2-Tier & 3-Tier Designs

VPC

FEX

PROGRAMMABLE SDN OVERLAY MODEL

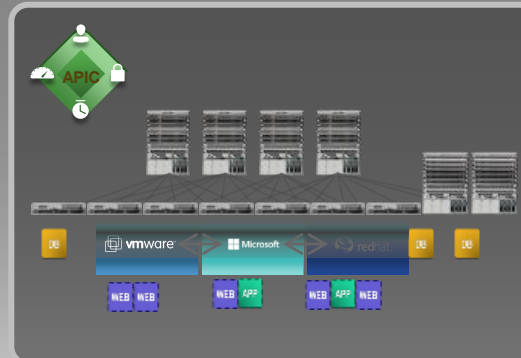


VXLAN Bridging & Routing

Integrated Network Virtualization

SDN Controllers

APPLICATION PROFILES & POLICIES



Application Centric Infrastructure

Policy Model

Automation

Nexus 3000, 5600, 7000

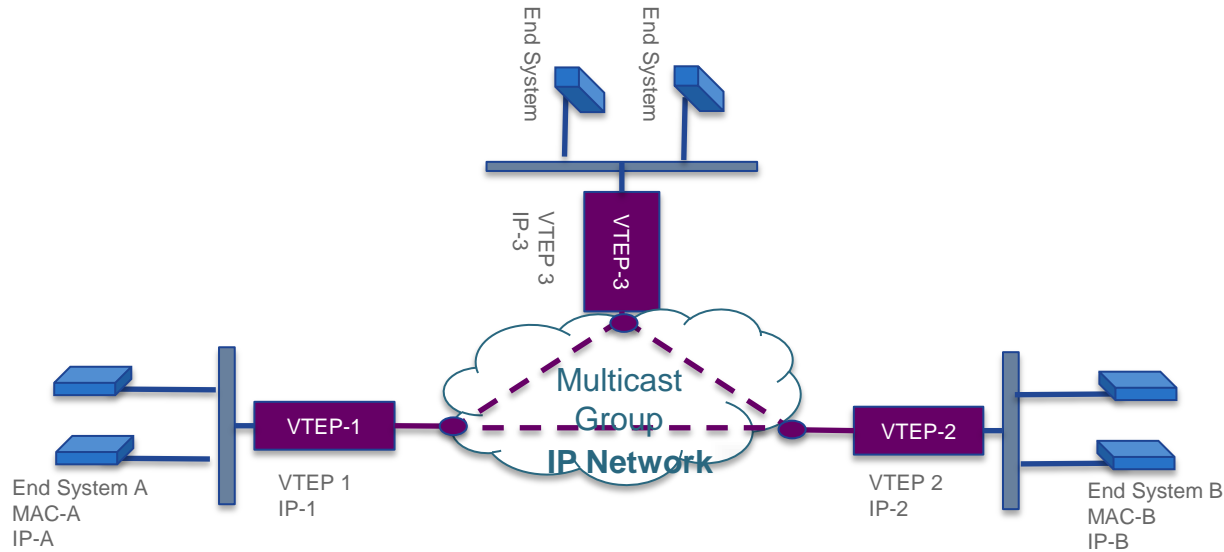
Nexus 9000

VXLAN Evolution

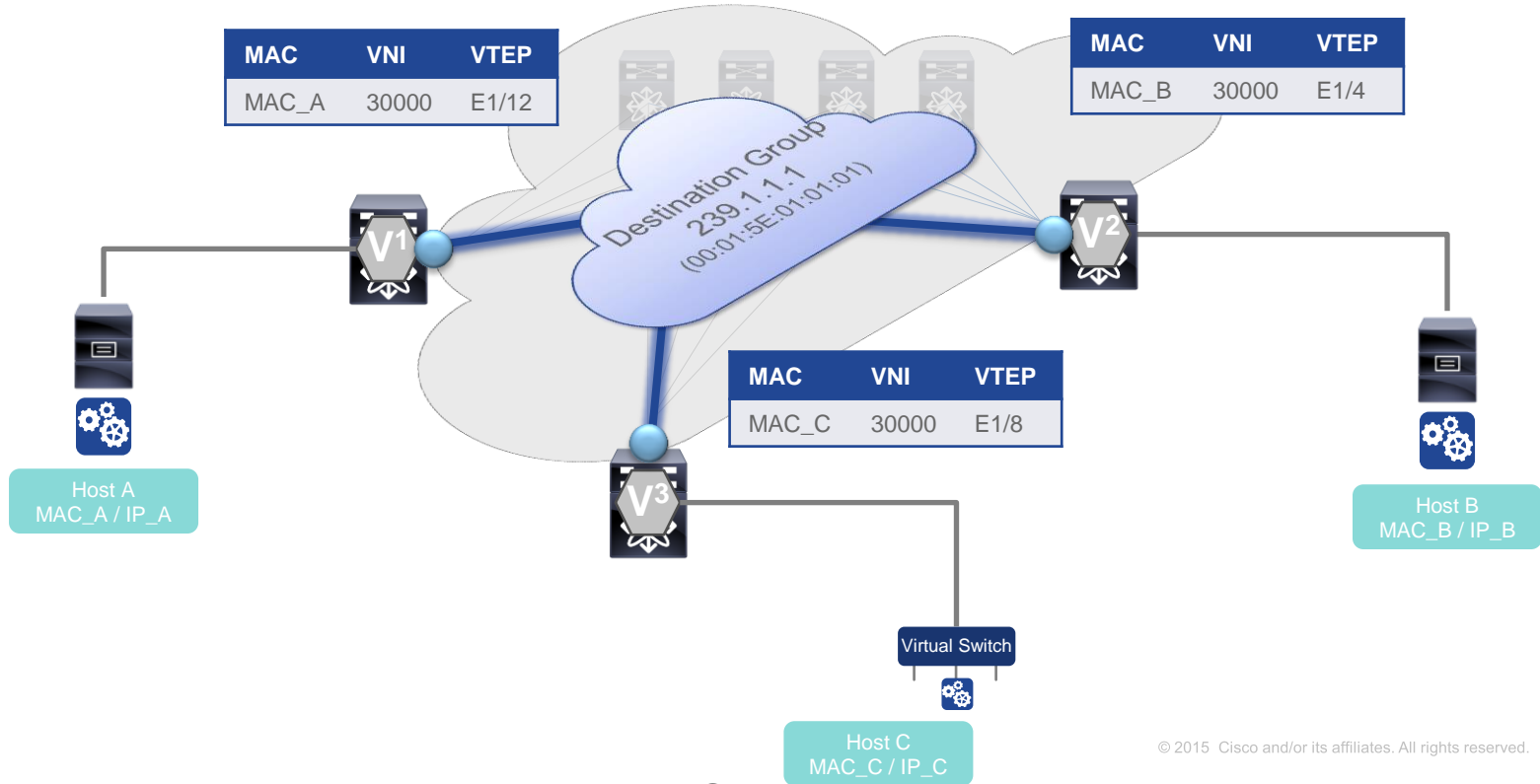
- Yesterday: VXLAN, yet another Overlay
 - Data-Plane only (Multicast based Flood & Learn)
- Today: VXLAN for the creation of scalable DC Fabrics – Intra-DC
 - Control-Plane, active VTEP discovery, Multicast and Unicast (Head-End Replication)
- Tomorrow: VXLAN for DCI – Inter-DC
 - DCI Enhancements (ARP caching/suppress, Multi-Homing, Failure Domain isolation, Loop Protection etc.)

Multicast-Based VxLAN

- No VXLAN control plane
- Data driven flood-&-learn
- Multicast transport for VXLAN BUM (Broadcast, Unknown Unicast and Multicast) traffic.

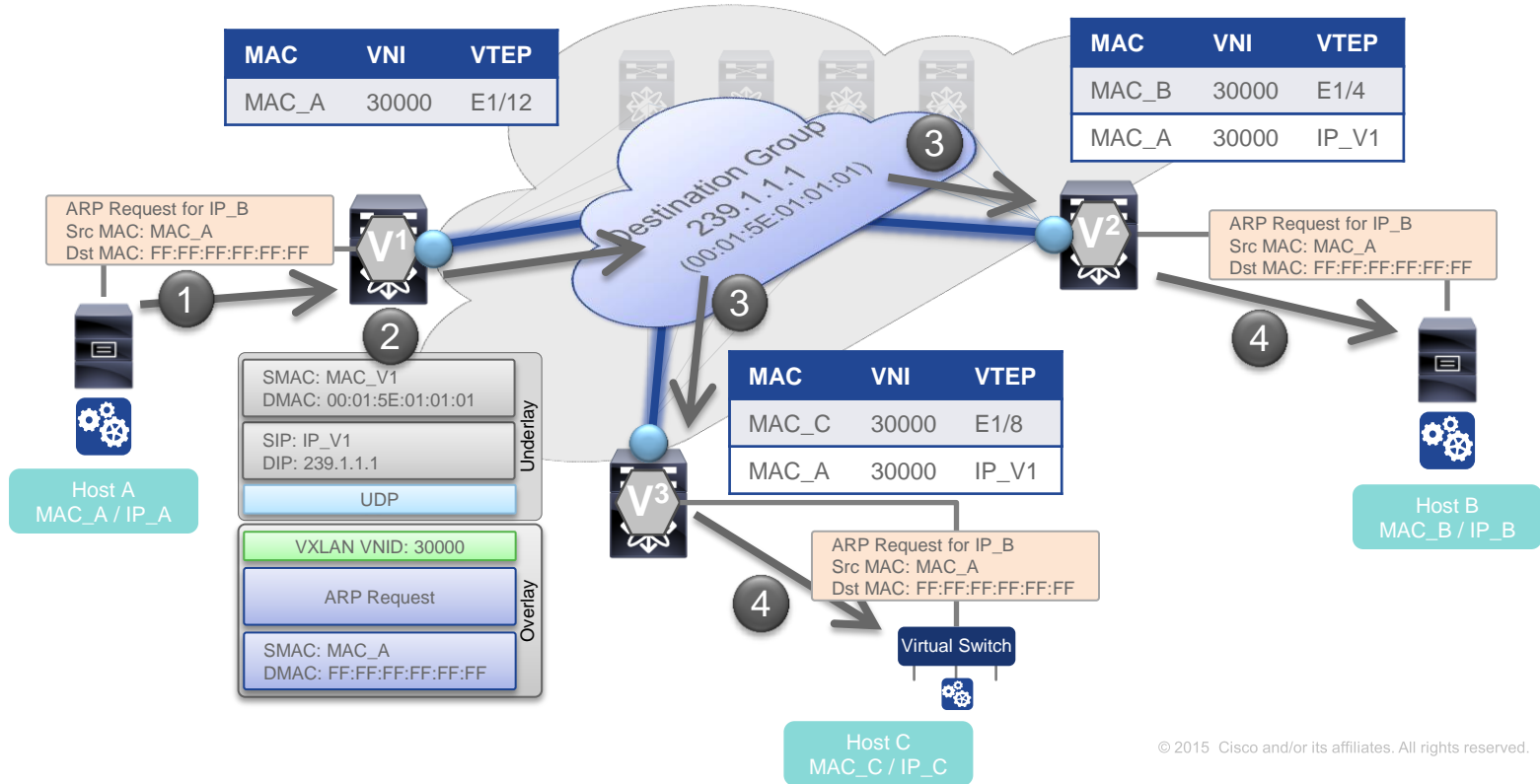


VXLAN Flood & Learn



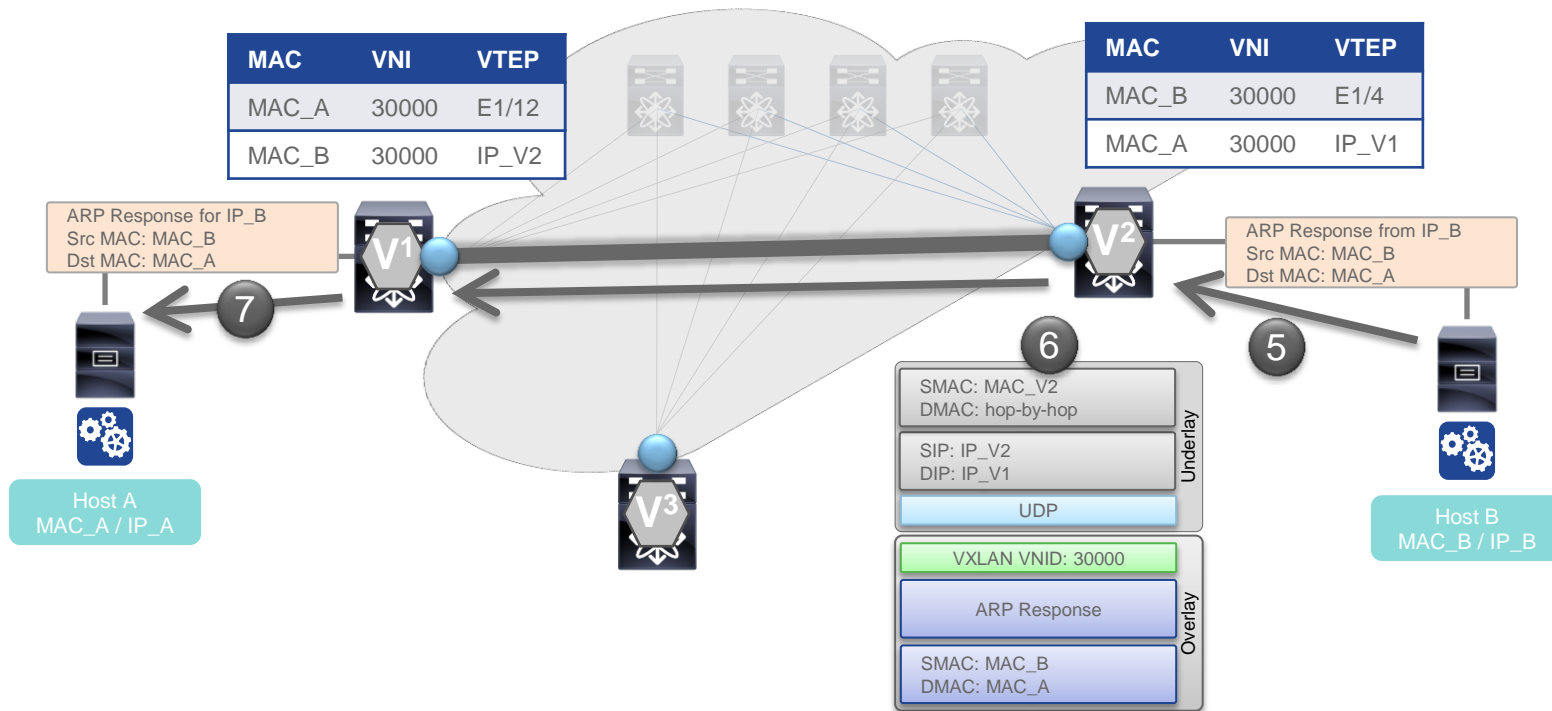
VTEP Peer Discovery & Address Learning (1)

VXLAN Flood & Learn



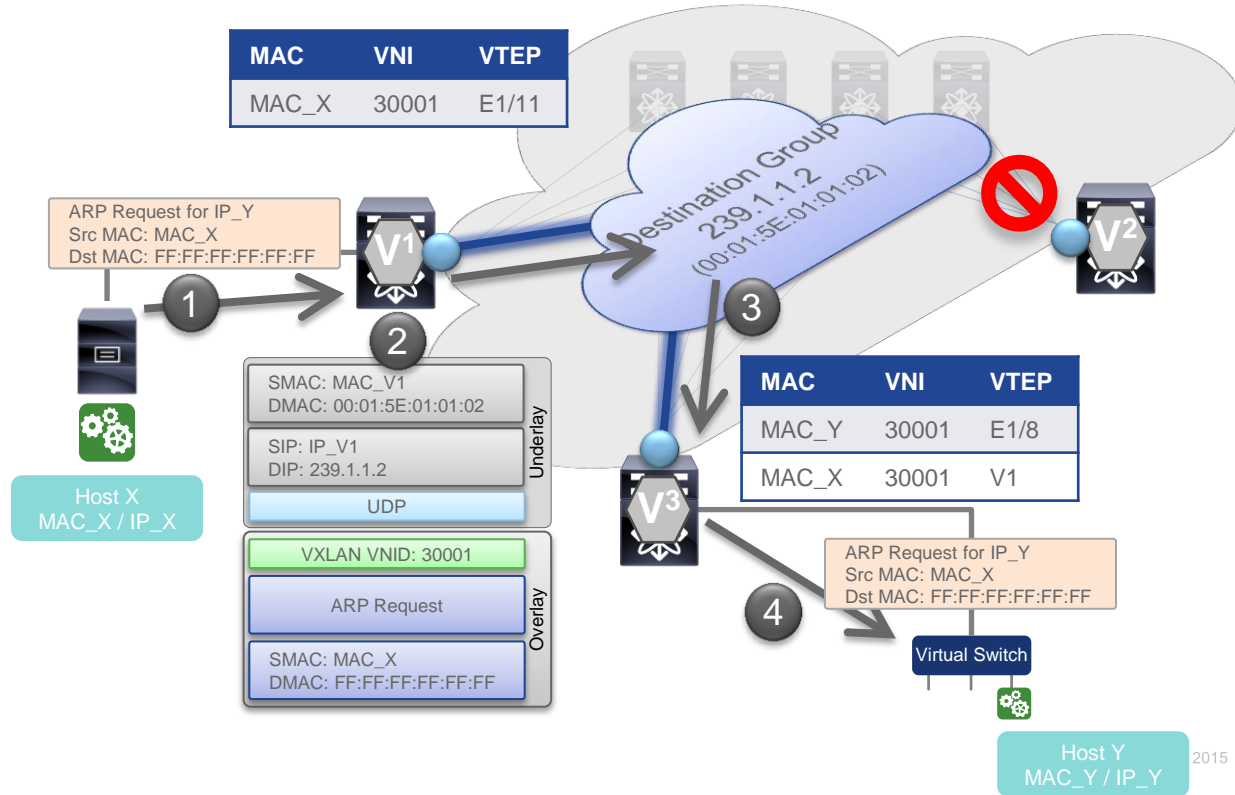
VTEP Peer Discovery & Address Learning (2)

VXLAN Flood & Learn



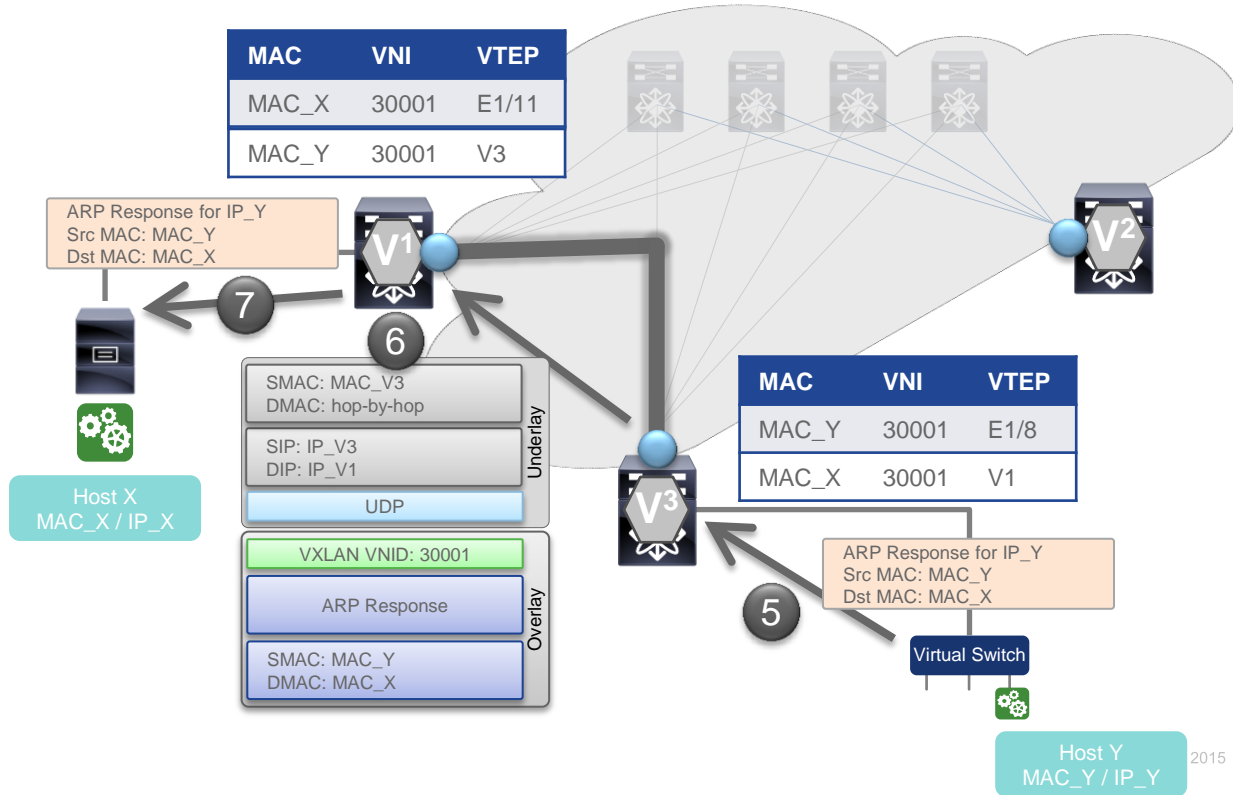
VTEP Peer Discovery & Address Learning (3)

VXLAN Flood & Learn



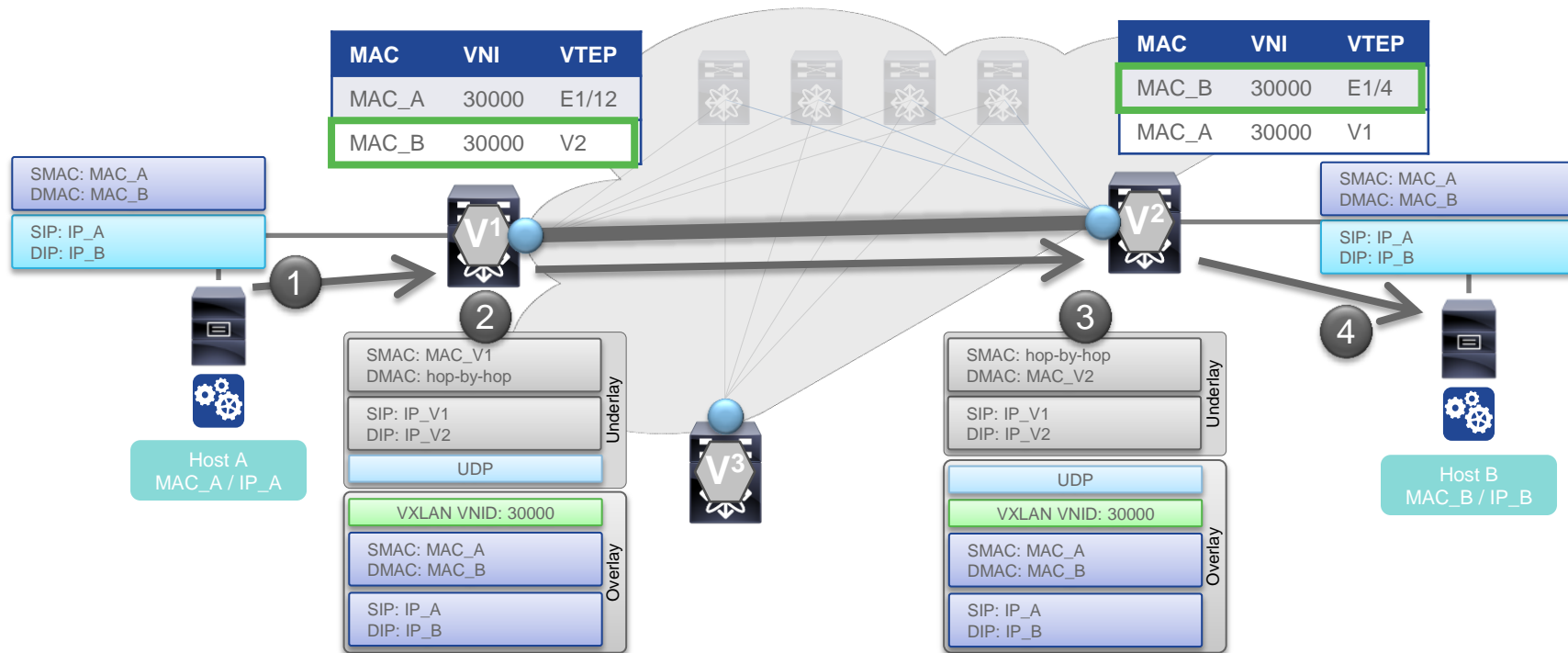
VTEP Peer Discovery & Address Learning (4)

VXLAN Flood & Learn



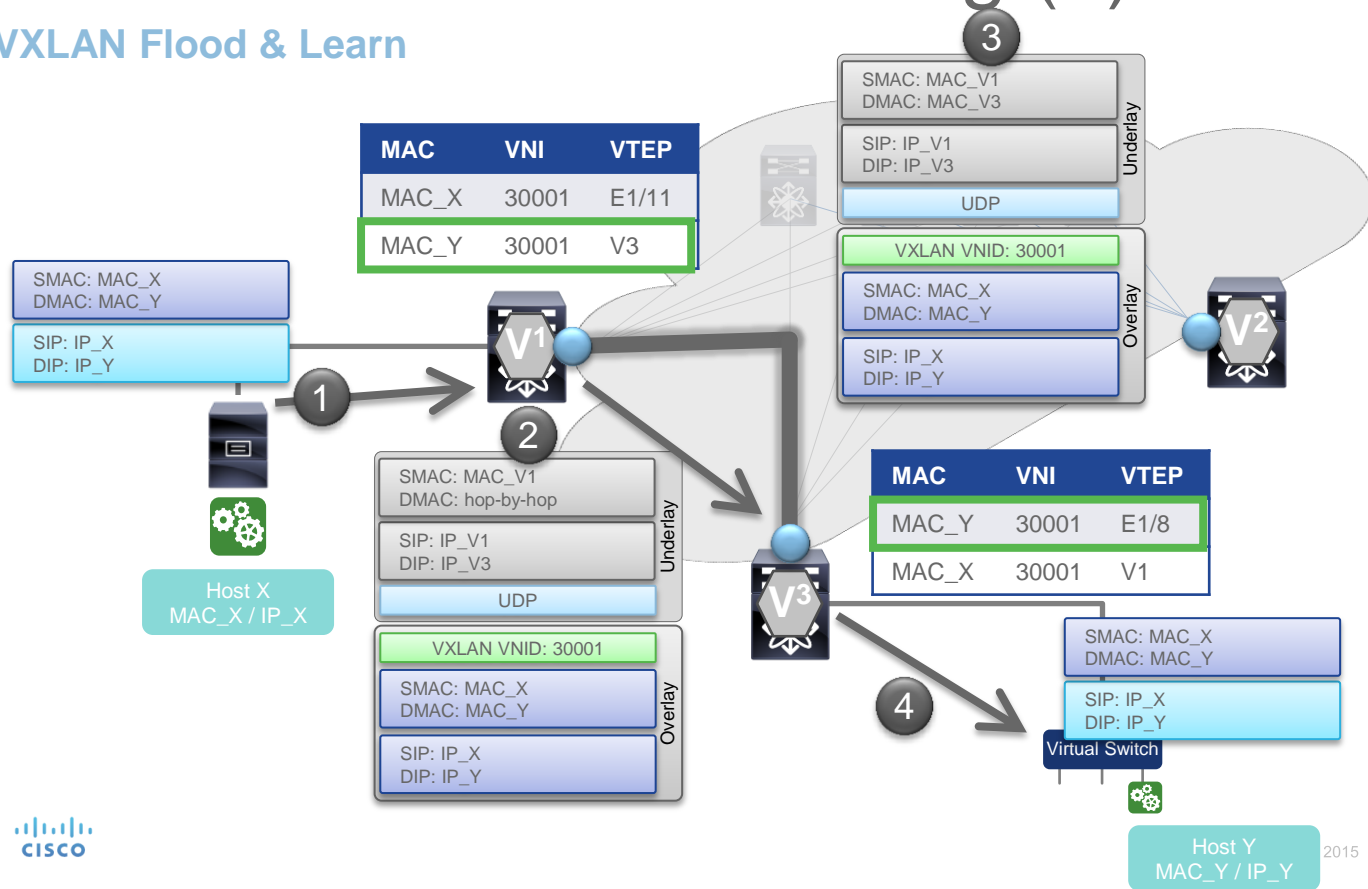
VXLAN Packet Forwarding (1)

VXLAN Flood & Learn



VXLAN Packet Forwarding (2)

VXLAN Flood & Learn



VXLAN Evolution

Multicast Independent*

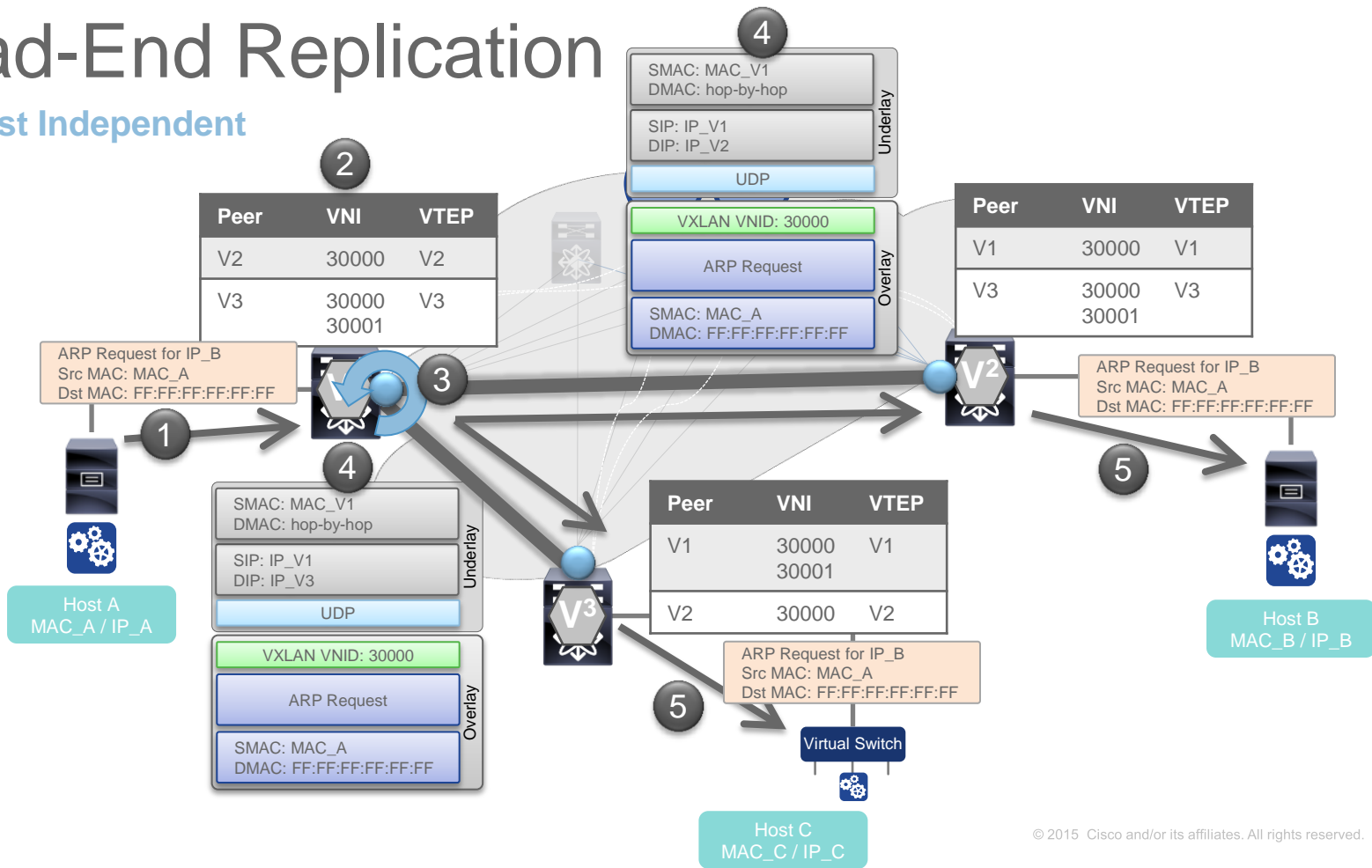
- Overlay Control-Plane provides dynamic VTEP discovery
- Head-End Replication enables Unicast-only mode (aka ingress Replication)

- Leveraging the Control-Plane to avoid Flood & Learn VTEP discovery (pro-active learning)
- Head-End Replication to relax the requirement for Multicast in the Underlay

*Multicast Independence requires the usage of the Overlay Control-Plane or static configuration

Head-End Replication

Multicast Independent



Polling Question 2

Is the thought of using Layer4 BGP protocol for DC switching a scary one ? 😊

Yes

No (I'm BGP Expert)

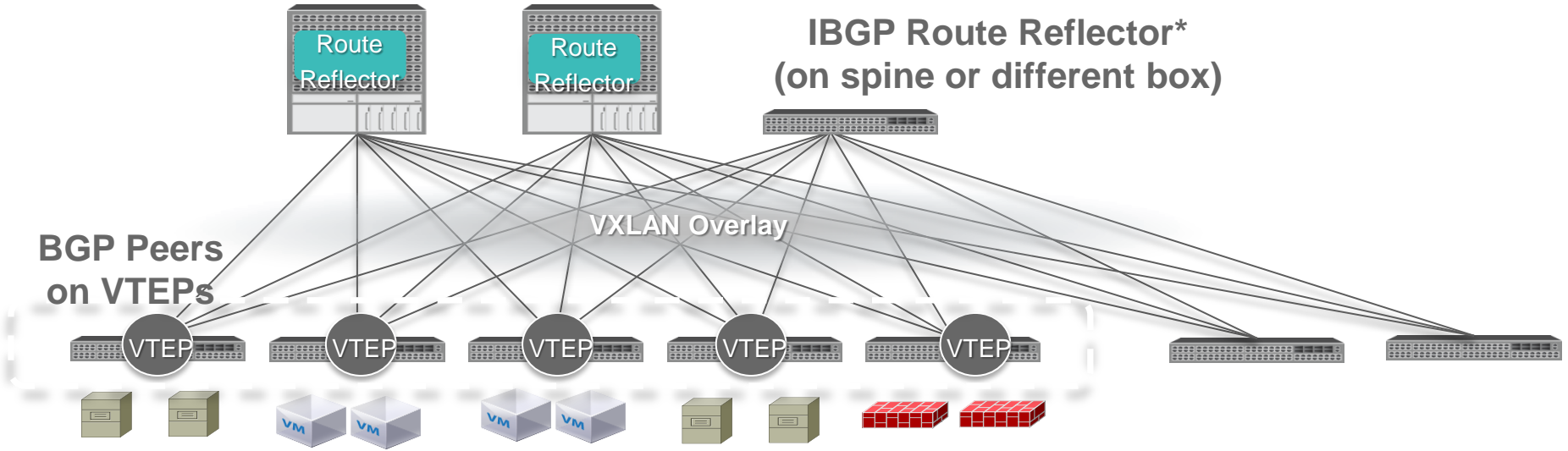
VXLAN Evolution

Protocol Learning

- Workload MAC and IP Addresses learnt by VXLAN Edge Devices (NVEs)
- Advertises Layer-2 and Layer-3 Address-to-VTEP Association (Overlay Control-Plane)
- Flood Prevention
- Optimized ARP forwarding

- Multi-Protocol BGP (MP-BGP) based Control-Plane using EVPN NLRI (Network Layer Reachability Information)
- Make Forwarding decisions at VTEPs for Layer-2 (MAC) and Layer-3 (IP); Integrated Route/Bridge (IRB)
- Reduce Flooding
- Reduce impact of ARP on the Network
- Standards Based (IETF draft)

BGP-EVPN VXLAN



Use Multi-Protocol BGP with EVPN Address family for :

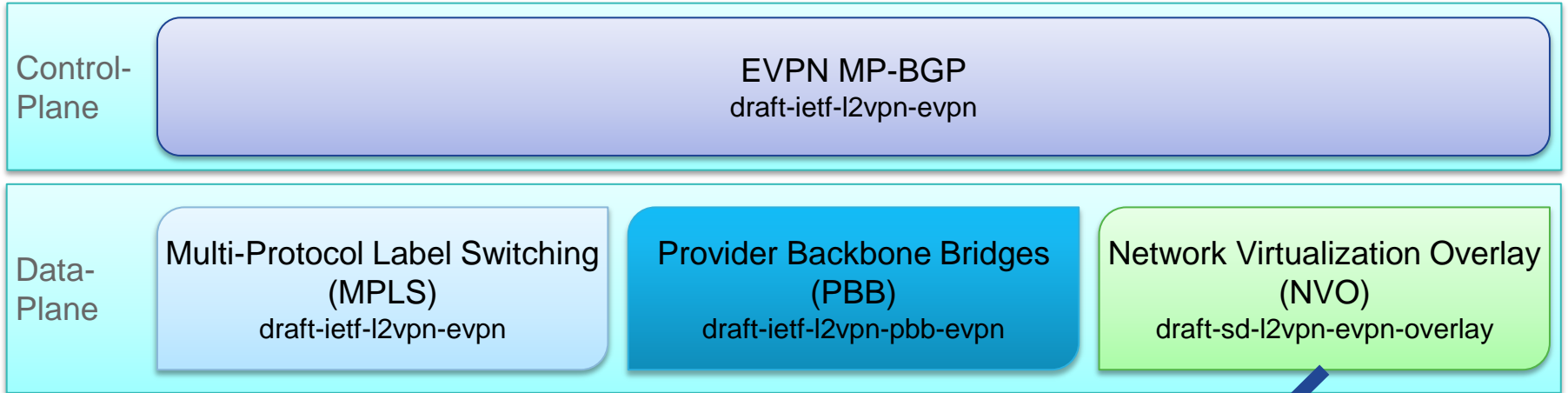
Tunnel Endpoints
Location

Host Reachability Information

- Mac Address
- IP address

EVPN – Ethernet VPN

VXLAN Evolution



- EVPN over NVO Tunnels (VXLAN, NVGRE, MPLSoE) for Data Center Fabric encapsulations
- Provides Layer-2 and Layer-3 Overlays over simple IP Networks

EVPN Solution Advantages

Early ARP
Termination

Suppresses flooding for Unknown Unicast ARP

Security

Authenticate Tunnel Endpoints

Distributed
Anycast Gateway

Seamless and Optimal vm-mobility

Ingress
Replication

Unicast Alternative to Multicast underlay

Active/Active
Multipathing

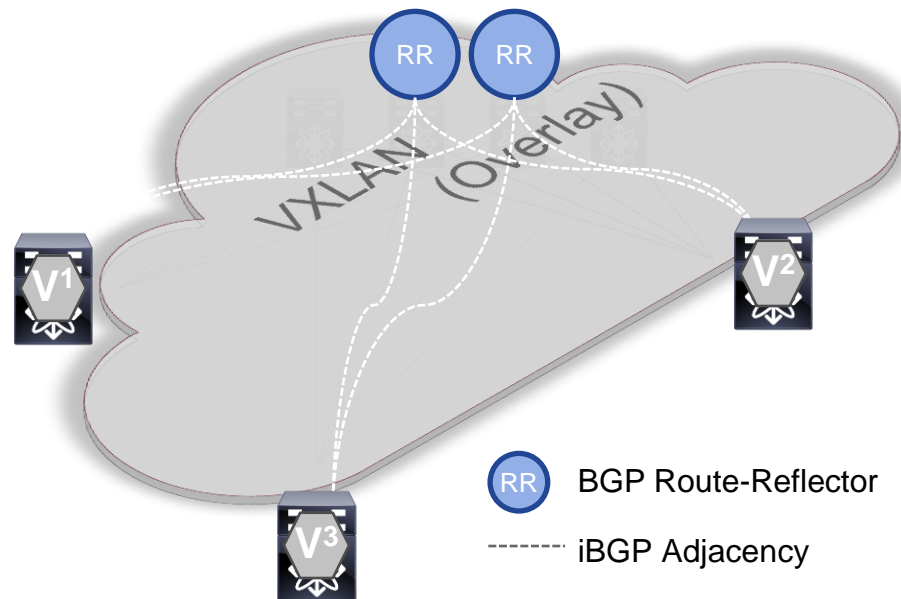
Active/Active and Resilient Multipathing
using vPC on Nexus



Host and Subnet Route Distribution

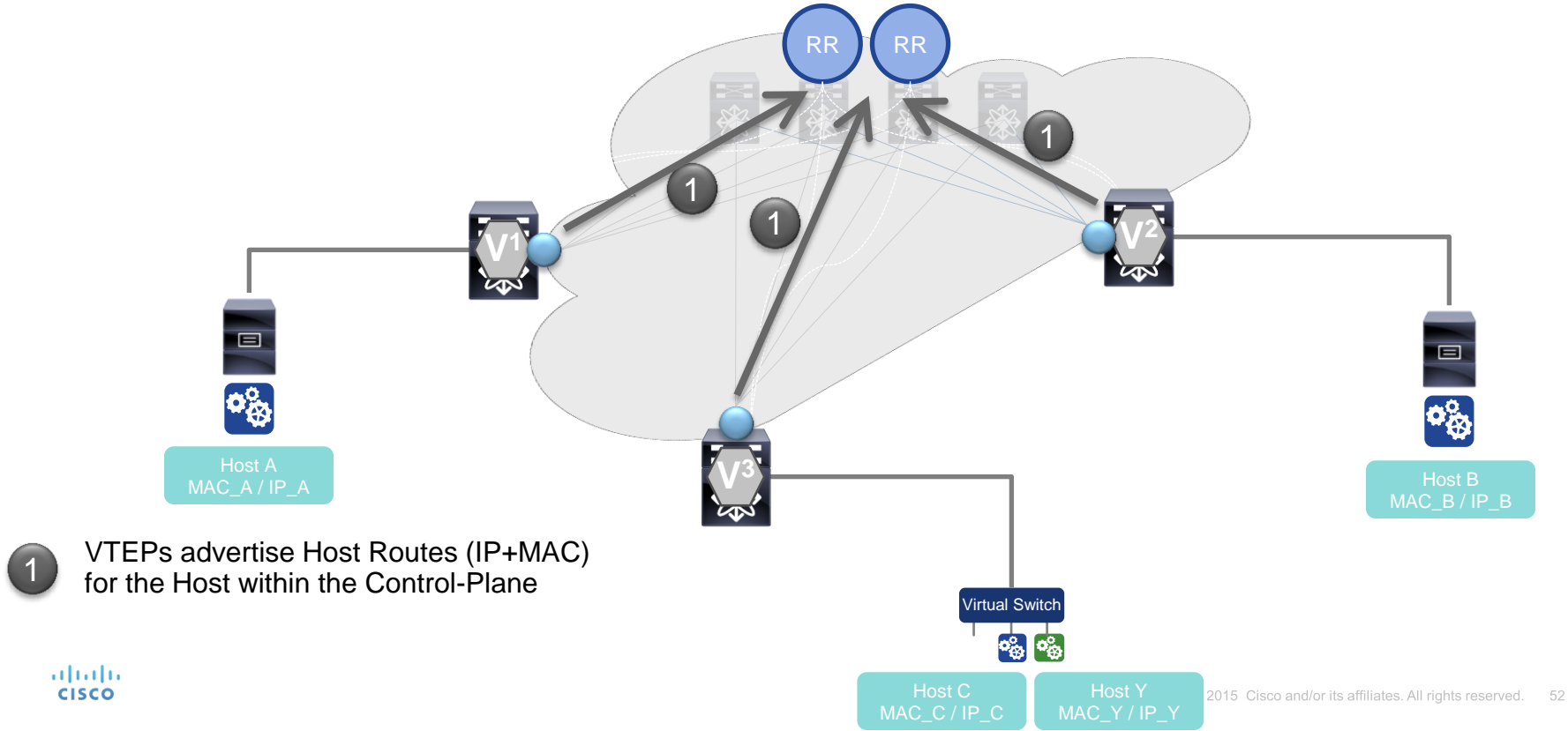
VXLAN/EVPN

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



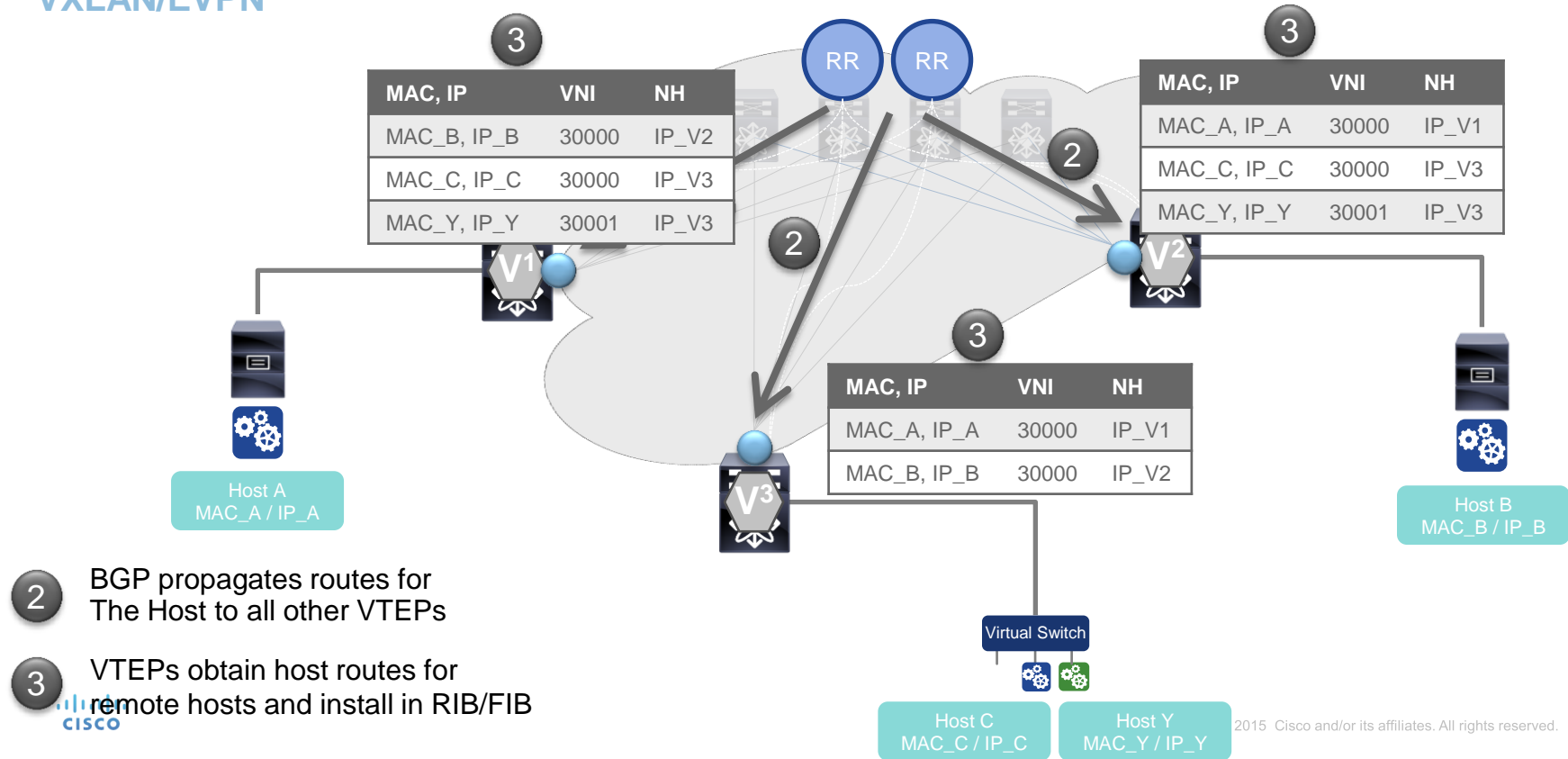
Protocol Learning & Distribution (1)

VXLAN/EVPN



Protocol Learning & Distribution (2)

VXLAN/EVPN



Host Advertisement

VXLAN/EVPN

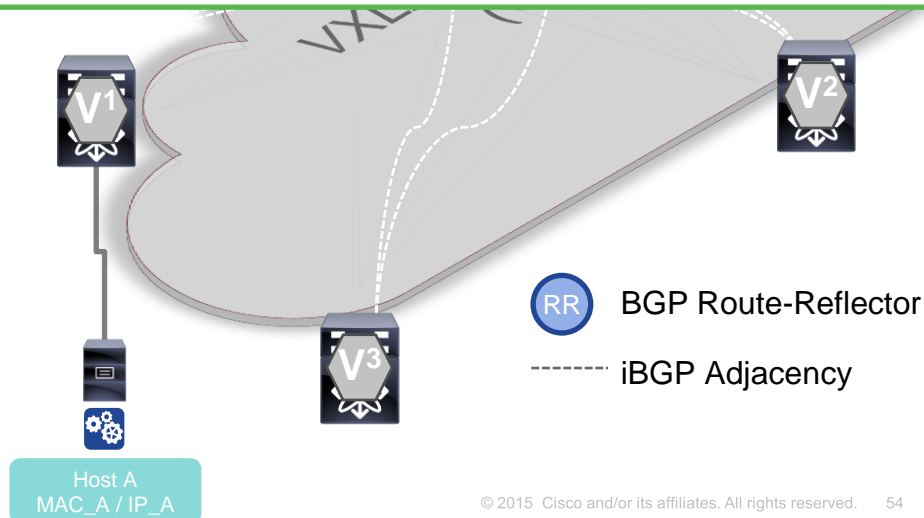
1. Host Attaches
2. VTEP V1 advertises Host A MAC (+IP) through BGP RR
3. Choice of Encapsulation is also advertised

MAC, IP	VNI (L2)	VNI (L3)	NH	Encap	Seq
MAC_A, IP_A	30000	50000	IP_V1	3:VXLAN	0

```

V1# sh bgp l2vpn evpn IP_A
BGP routing table information for VRF default, address family L2VPN EVPN
Route Distinguisher: 30000:V1
BGP routing table entry for [2]:[0]:[0]:[48]:[MAC_A]:[32]:[IP_A]/272, version 28838
Paths: (1 available, best #1)
Flags: (0x000202) on xmit-list, is not in l2rib/evpn

Advertised path-id 1
Path type: internal, path is valid, is best path, no labeled nexthop
AS-Path: NONE, path sourced internal to AS
IP_V1 (metric 3) from RR (RR)
Origin IGP, MED not set, localpref 100, weight 0
Received label 30000 50000
Extcommunity: RT:1000:30000 RT:1000:50000 ENCAP:3
Originator: IP_V1 Cluster list: RR
Remote Next-hop Attribute: IP_V1
encapsulation VXLAN VNID 50000 MAC MAC_V1
    
```



Host Moves

VXLAN/EVPN

1. Host Moves to V3
2. V3 detects Host A and advertises it with Seq #1
3. V1 sees more recent route and withdraws its advertisement

MAC, IP	VNI (L2)	VNI (L3)	NH	Encap	Seq
MAC_A, IP_A	30000	50000	IP_V3	3:VXLAN	1

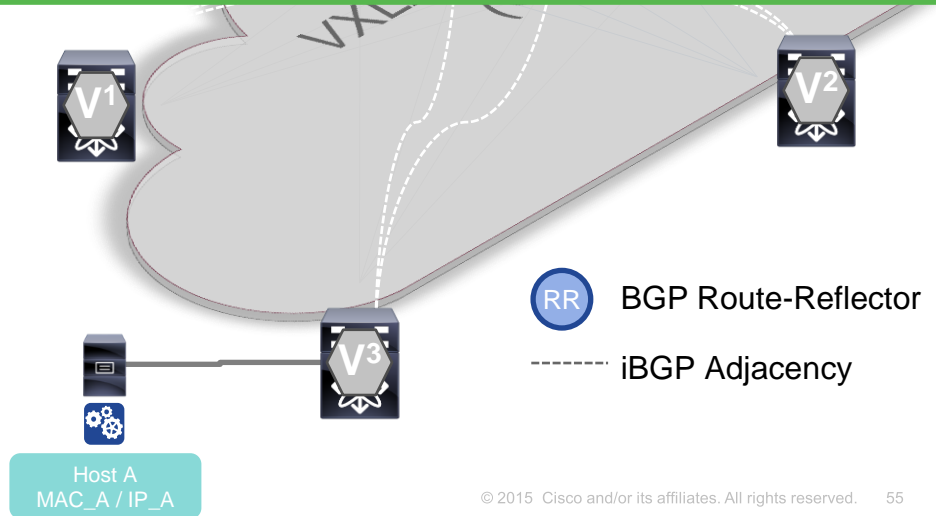
```

V1# sh bgp l2vpn evpn IP_A
BGP routing table information for VRF default, address family L2VPN EVPN
Route Distinguisher: 30000:V3
BGP routing table entry for [2]:[0]:[0]:[48]:[MAC_A]:[32]:[IP_A]/272, version 28839
Paths: (1 available, best #1)
Flags: (0x000202) on xmit-list, is not in I2rib/evpn

  Advertised path-id 1
  Path type: internal, path is valid, is best path, no labeled nexthop
  AS-Path: NONE, path sourced internal to AS
  IP_V3 (metric 3) from RR (RR)
  Origin IGP, MED not set, localpref 100, weight 0
  Received label 30000 50000
  Extcommunity: RT:1000:30000 RT:1000:50000 ENCAP:3
  Originator: IP_V3 Cluster list: RR
  Remote Next-hop Attribute: IP_V3
  encapsulation VXLAN VNID 50000 MAC MAC_V3
    
```

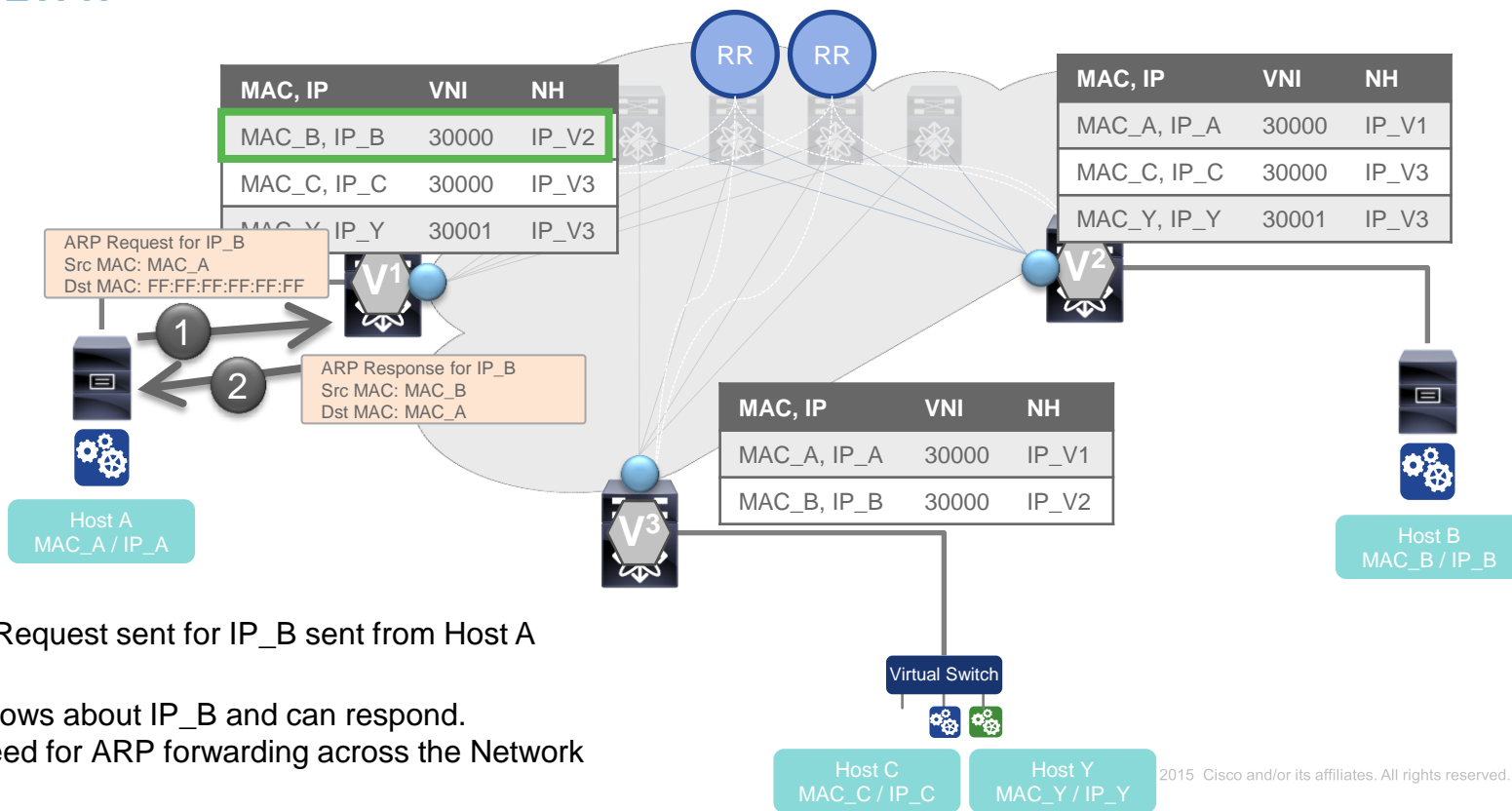
48, MAC, 32, IP

ENCAP:3 = VXLAN



ARP Suppression

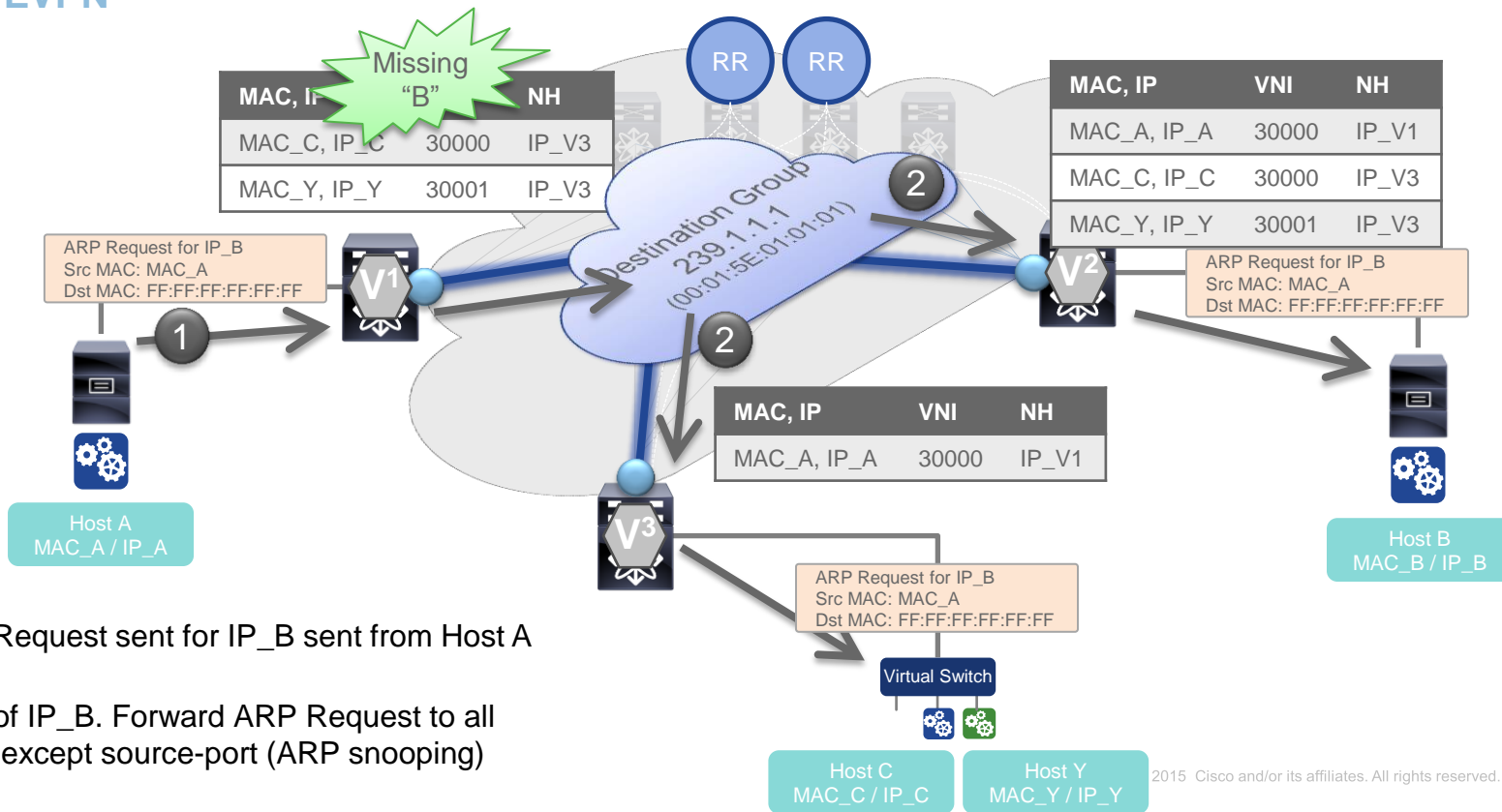
VXLAN/EVPN



- 1 ARP Request sent for IP_B sent from Host A
- 2 V1 knows about IP_B and can respond.
No need for ARP forwarding across the Network

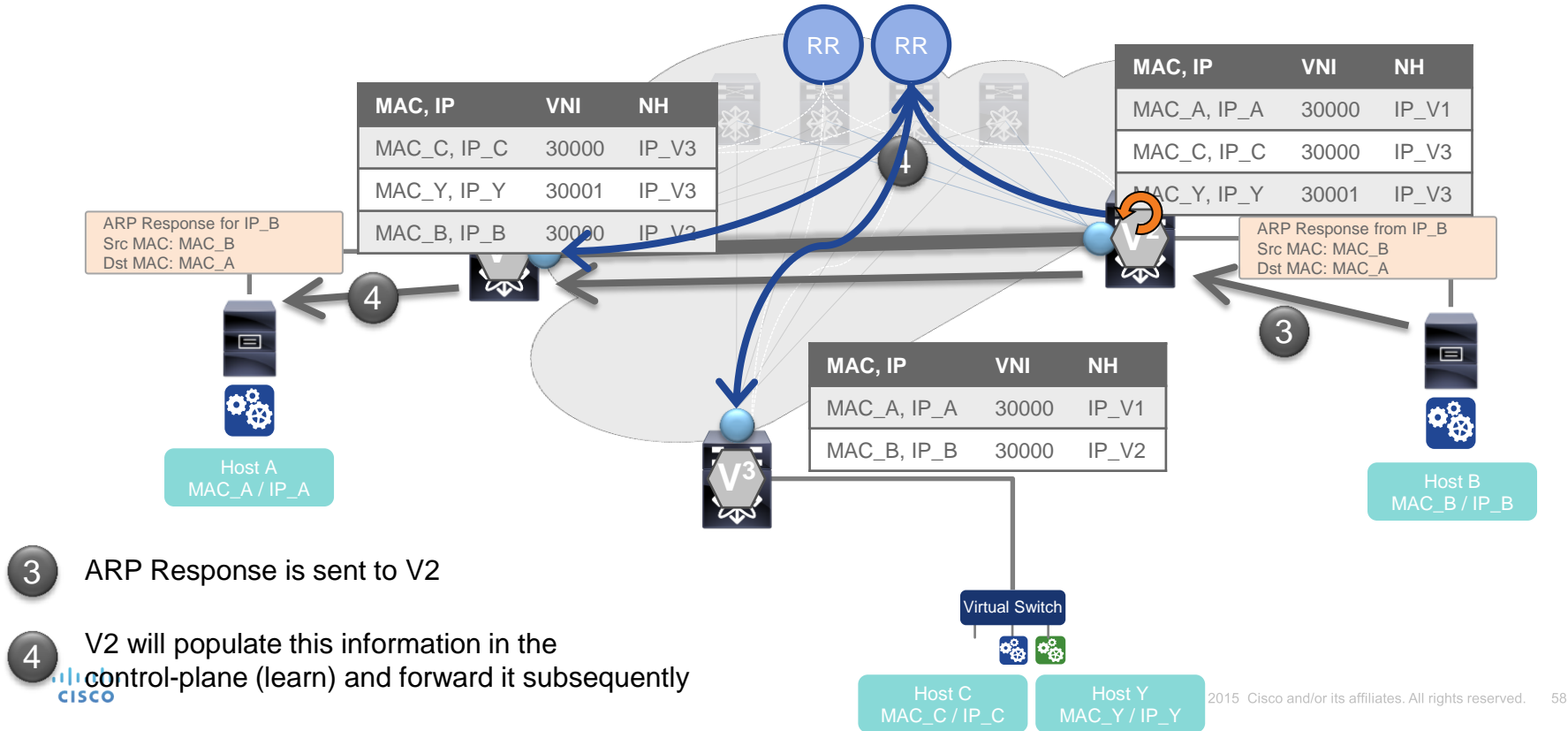
ARP Handling on Lookup “Miss” (1)

VXLAN/EVPN



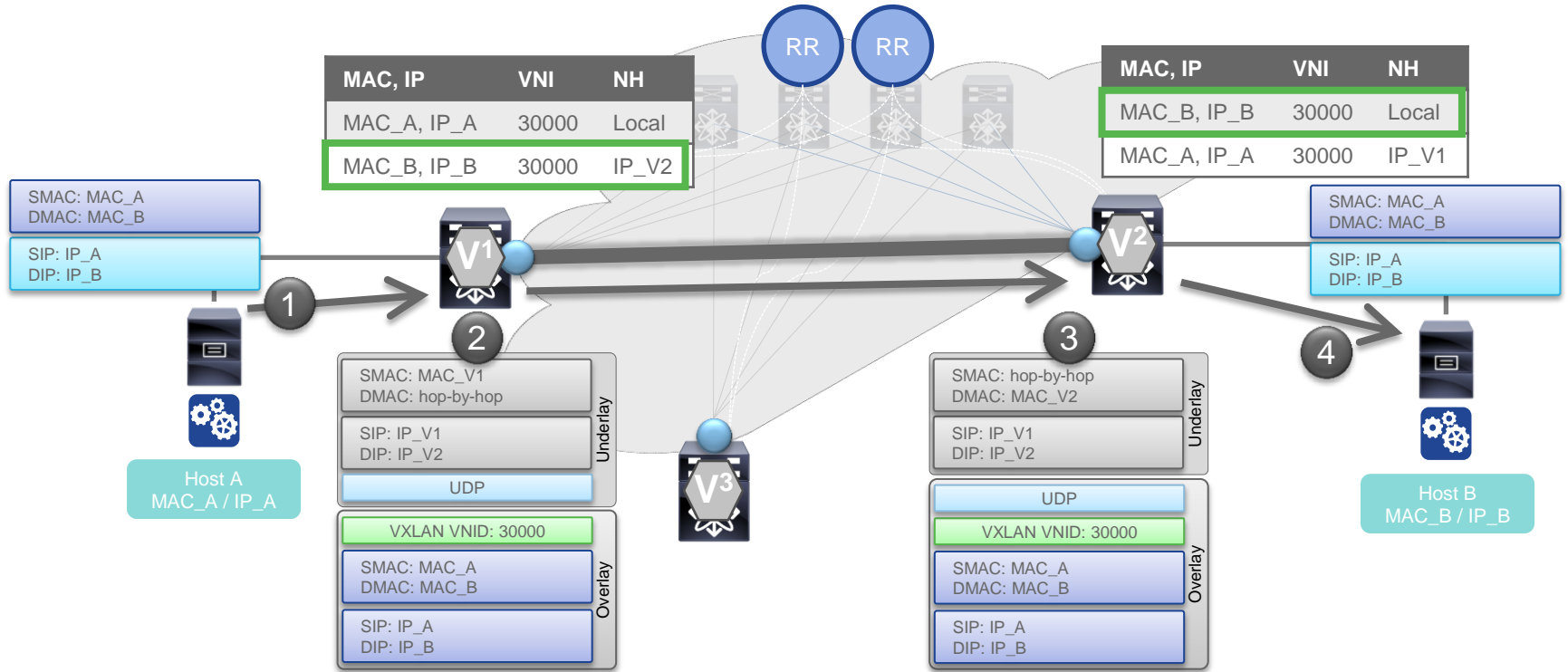
ARP Handling on Lookup “Miss” (2)

VXLAN/EVPN



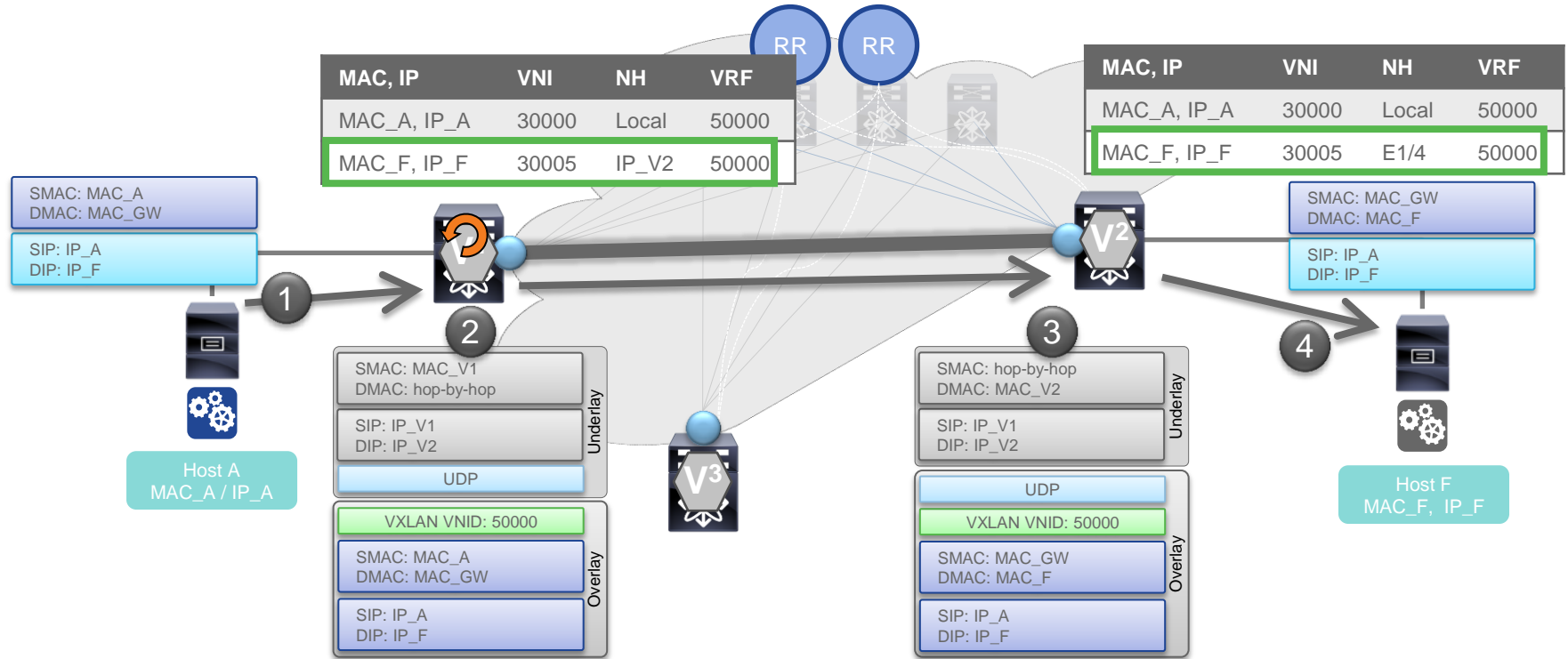
Packet Forwarding (Bridge)

VXLAN/EVPN



Packet Forwarding (Route)

VXLAN/EVPN



EVPN Control Plane Advantages

A multi-tenant fabric solution with host-based forwarding

- Industry standard protocol for multi-vendor interoperability
- Build-in multi-tenancy support
 - Leverage MP-BGP to deliver VXLAN with L3VPN characteristics
- Truly scalable with protocol-driven learning
 - Host MAC/IP address advertisement through EVPN MP-BGP
- Fast convergence upon host movements or network failures
 - MP-BGP protocol driven re-learning and convergence
 - Upon host movement, the new VTEP will send out a BGP update to advertise the new location of the host

EVPN Control Plane Advantages (Cont'ed)

A multi-tenant fabric solution with host-based forwarding

- Optimal traffic forwarding supporting host mobility
 - Anycast IP gateway for optimal forwarding for host generated traffic
 - No need for hair-pinning to reach the IP gateway
- ARP suppression
 - Minimize ARP flooding in overlay
- Head-end Replication with dynamically learned remote-VTEP list
 - Head-end replication enables multicast-free underlay network
 - Dynamically learned remote-VTEP list minimizes the operational overhead of head-end replication
- VTEP peer authentication via MP-BGP authentication
- Added security to prevent rogue VTEPs or VTEP spoofing

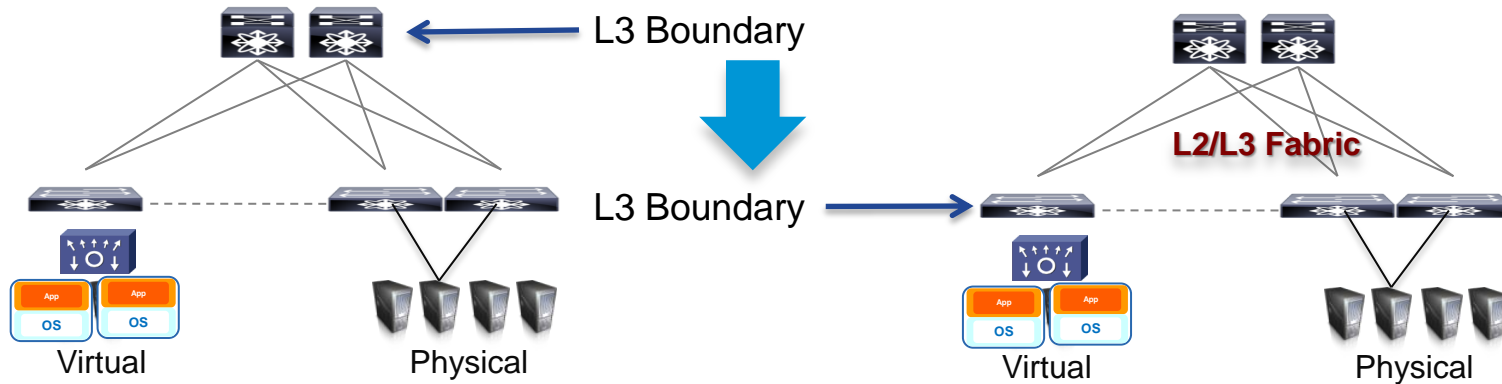
VXLAN Evolution

IP Services

- VXLAN Routing
- Distributed Anycast Gateway (requires Overlay Control-Plane)
- Multi-Tenancy

- Forward based on MAC or IP address learnt via Control-Plane (MP-BGP EVPN)
- Make routing decisions at VTEPs
- Scale and Multipathing (ECMP)
- Leverage Layer-3 Gateway capabilities along with Protocol Information
- LISP-ish / LISP-like approach for Host/IP Mobility
 - Location (VTEP), Identifier (MAC, IP of End-Host)

Distributed Gateway Function in L3 Overlays



Traditional L2 - centralised L2/L3 boundary

- Always bridge, route only at an aggregation point
- Large amounts of state converge
- Scale problem for large# of L2 segments
- Traditional L2 and L2 overlays

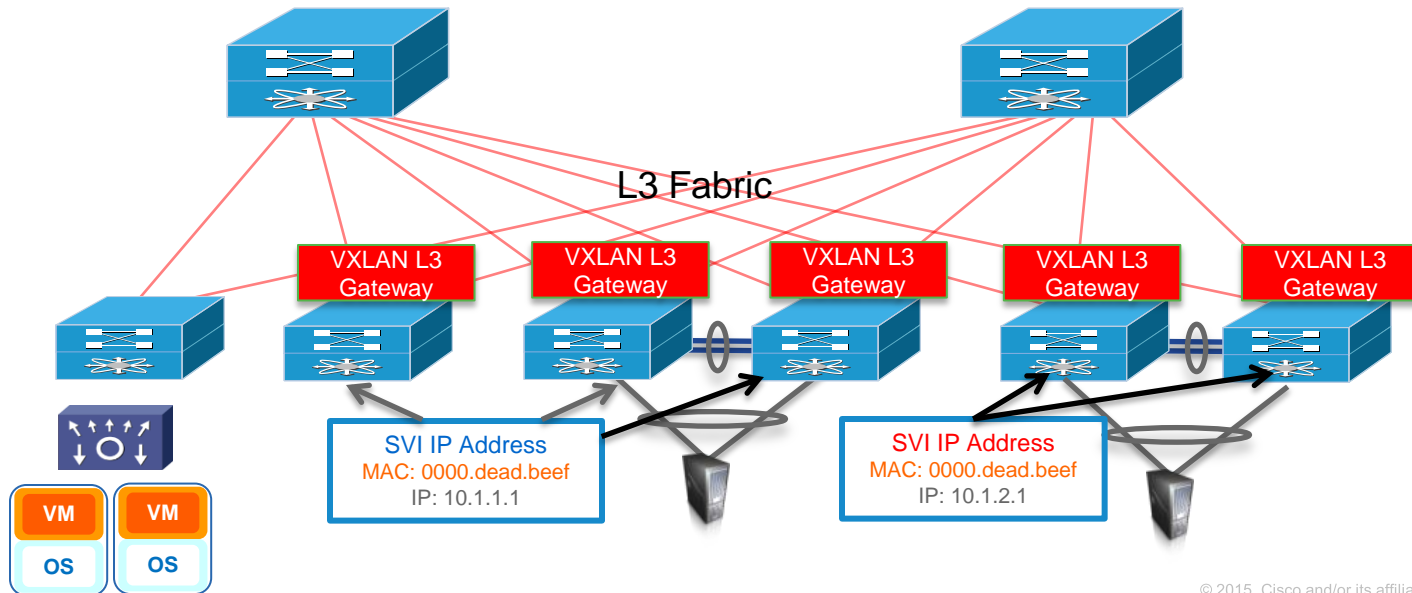
L2/L3 fabric (or overlay)

- Always route (at the leaves), bridge when necessary
- Distribute and disaggregate necessary state
- Optimal scalability
- Enhanced forwarding and L3 overlays

Distributed IP Anycast Gateway

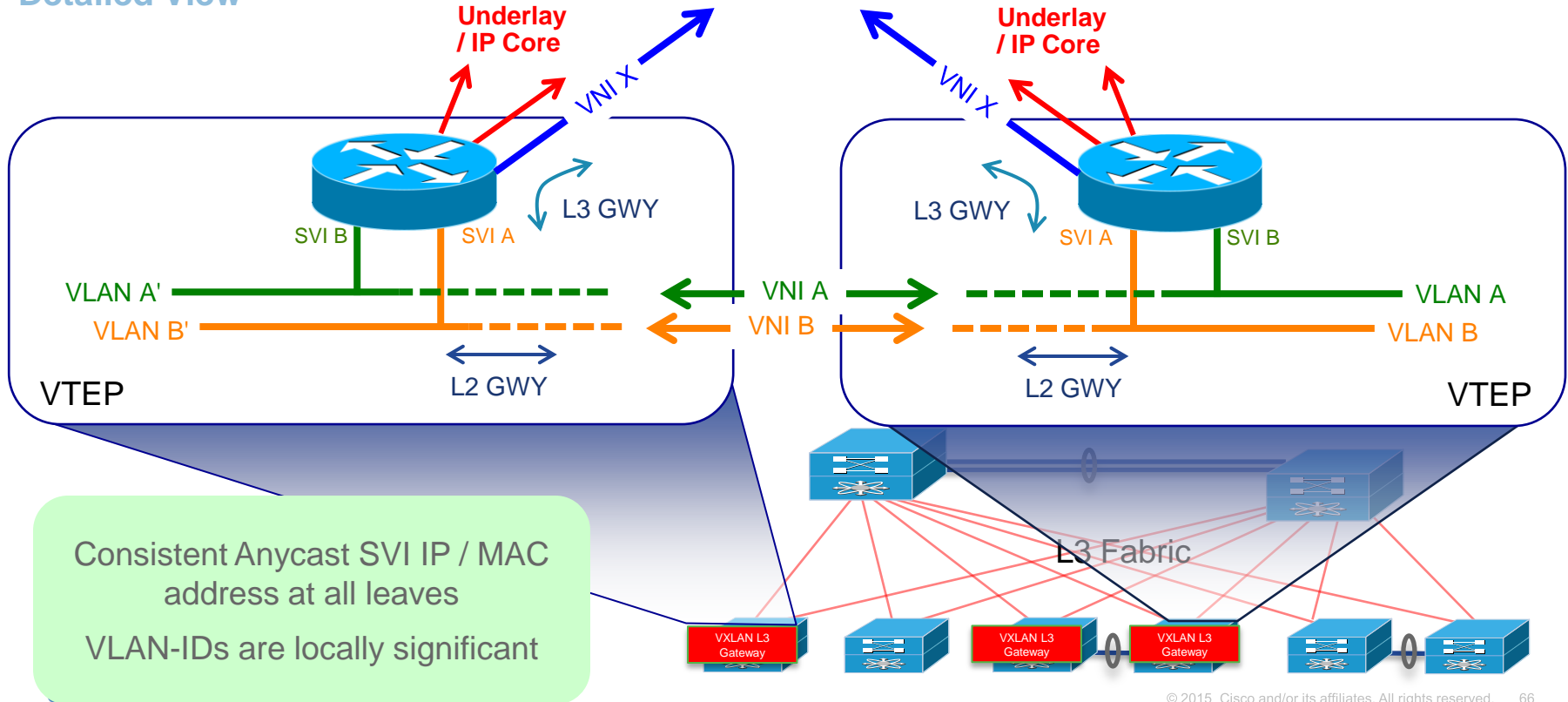


The same “Anycast” SVI IP/MAC is used at all VTEPs/ToRs
A host will always find its SVI anywhere it moves



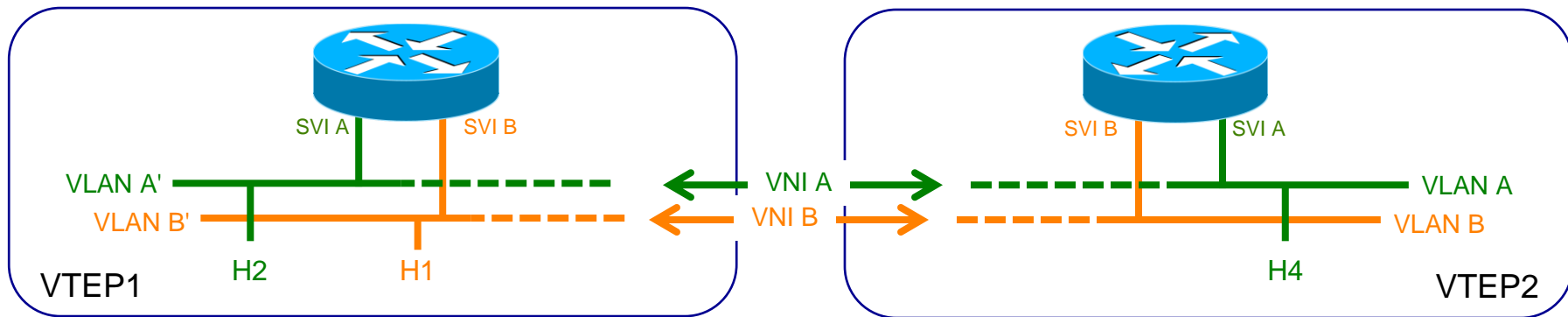
Distributed IP Anycast Gateway

Detailed View



VXLAN Bridging

802.1Q Tagged Traffic to VNI Mapping

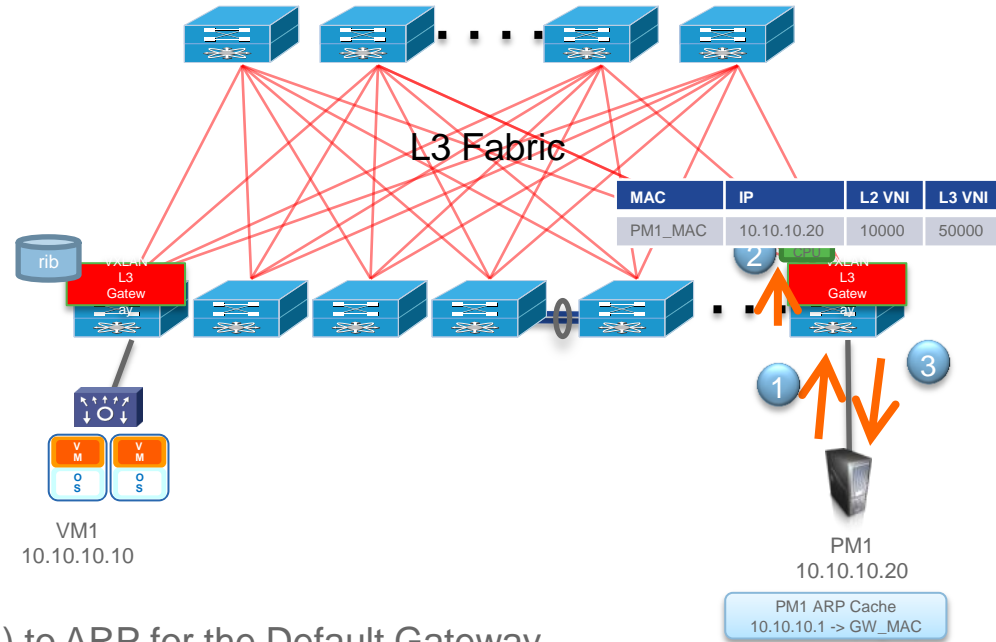


- VLANs are stretched over L2 VNIs
- VLANs (VLAN A) mapped to VNI (VNI A) at each VTEP: $VLAN A' \leftrightarrow VNI A \leftrightarrow VLAN A$
- Bridged traffic forwarded over the L2 VNIs

Distributed IP Anycast Gateway

Packet-Walk – IP Forwarding within the Same Subnet aka Bridging (ARP)

1. PM1 sends an ARP request for Default Gateway –10.10.10.1
2. The ARP request is suppressed at TOR and punted to the Supervisor, where MAC and IP is learned and distributed
3. TOR response with Gateway MAC to PM1

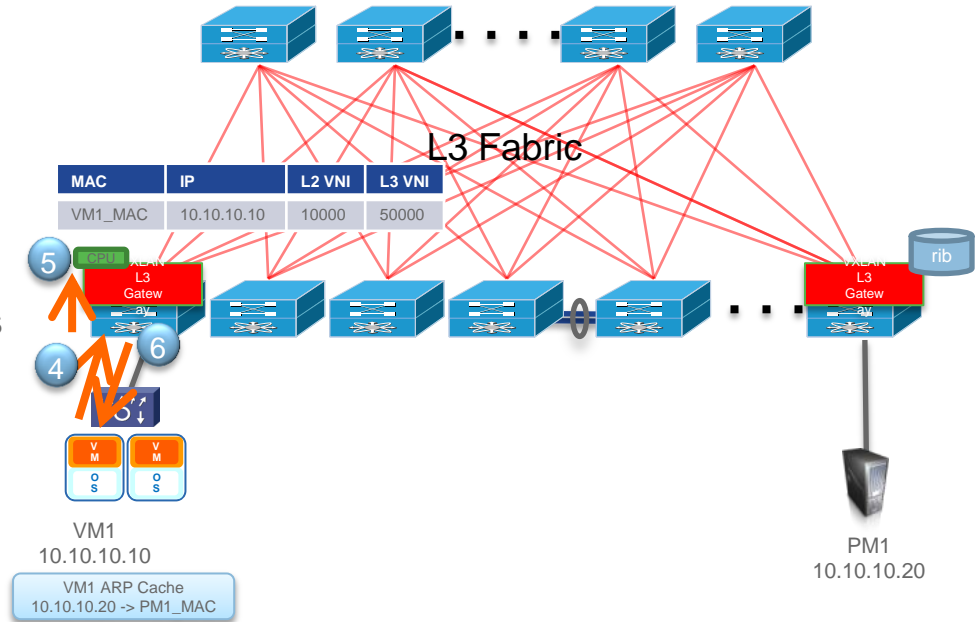


Standard behavior of End-Host (virtual or physical) to ARP for the Default Gateway

Distributed IP Anycast Gateway

Packet-Walk – IP Forwarding within the Same Subnet aka Bridging (ARP)

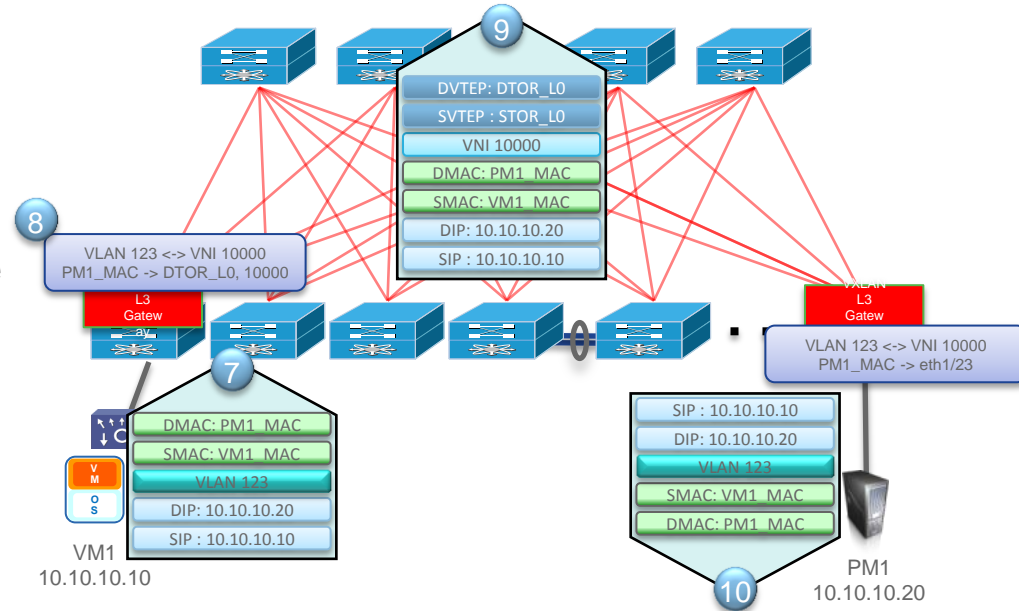
4. VM1 sends an ARP request for PM1 – 10.10.10.20
5. The ARP request is suppressed at TOR and punted to the Supervisor, where MAC and IP is learned and distributed
6. Assuming PM1 is known and a valid route does exist in the Unicast RIB, TOR responds to ARP with PM1 MAC as Source MAC. VM1 can build its ARP cache



Distributed IP Anycast Gateway

Packet-Walk – IP Forwarding within the Same Subnet aka Bridging (Data Packet)

7. VM1 generates a data packet with PM1_MAC as destination MAC
8. TOR receives the packet and performs Layer-2 lookup for the destination
9. TOR adds VXLAN-Header information (Destination VTEP, VNI, etc) and forwards the packet across the Layer-3 fabric, picking one of the equal cost paths available via the multiple Spines
10. The destination TOR receives the packet, strips off the VXLAN header and performs lookup and forwarding toward PM1

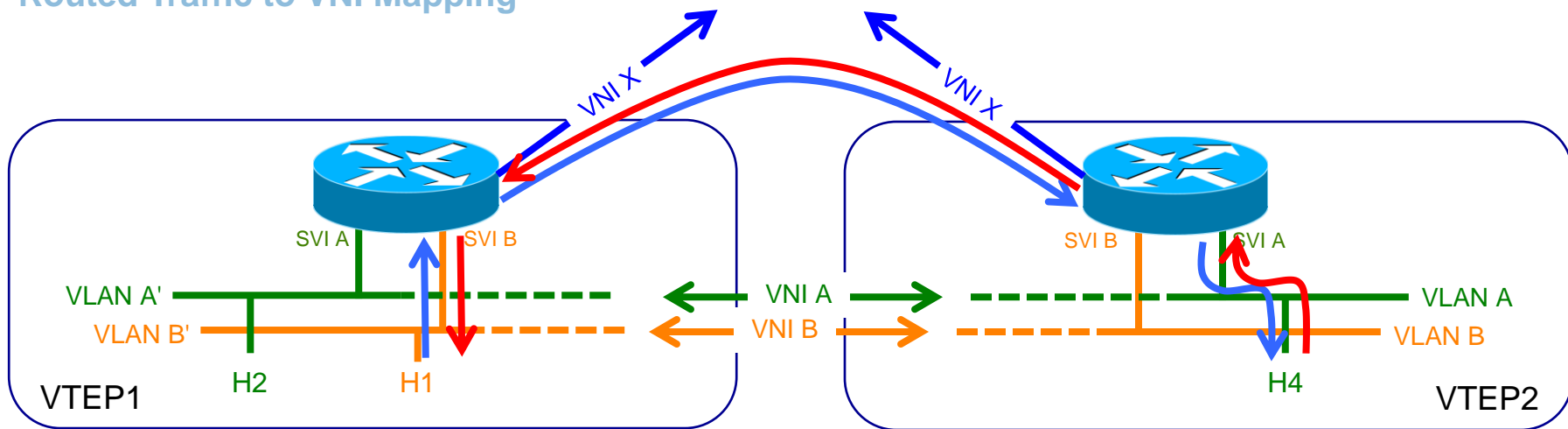


In case of VM1 is not known to PM1, PM1 would ARP for VM1. Destination TOR would Proxy for VM1.

No Silent-Host discovery problem.

VXLAN Routing

Routed Traffic to VNI Mapping



- A common VNI (**VNI X**) is provisioned amongst the different VTEPs to carry routed traffic
- Routed traffic between VTEPs will be encapsulated in **VNI X**
- Standard longest prefix match routing takes place:
 - Host routes for all known remote hosts are installed at every VTEP → Forward over **VNI X**

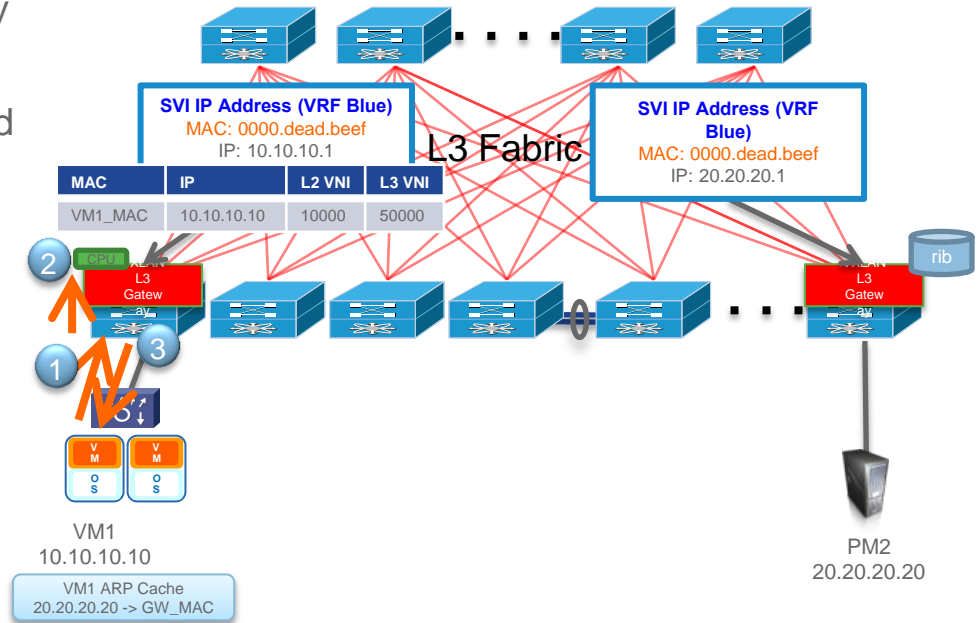


Local hosts are covered by directly connected prefix, a host route will not be present

Distributed IP Anycast Gateway

Packet-Walk – IP Forwarding within the Different Subnet aka Routing (ARP)

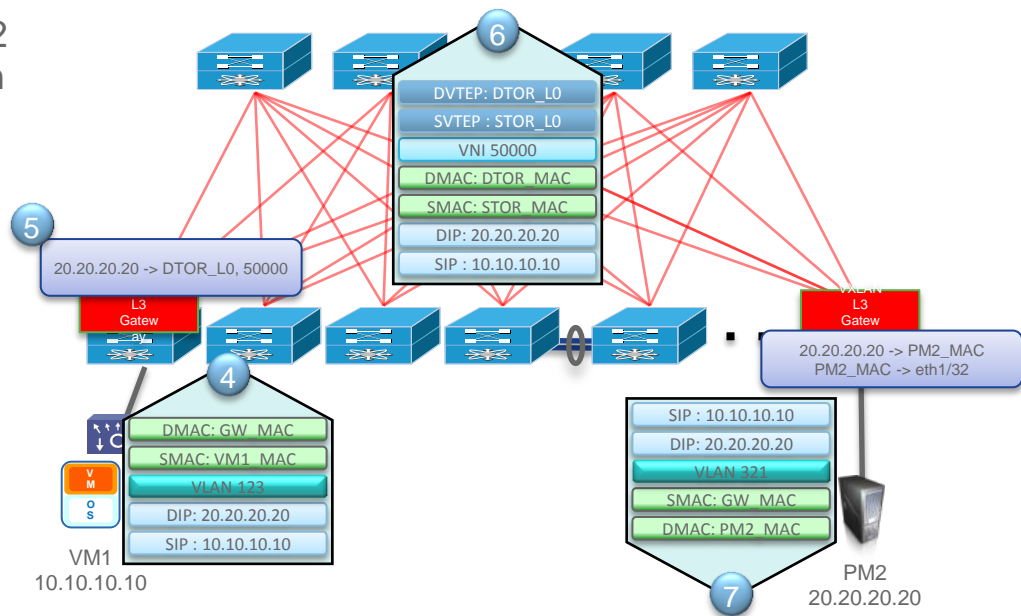
1. VM1 sends ARP request for Default Gateway –10.10.10.1
2. The ARP request will be received at TOR and punted to the Supervisor, where MAC and IP is learned and distributed
3. TOR acts as regular Default Gateway and sends ARP response with GW_MAC to VM1



Distributed IP Anycast Gateway

Packet-Walk – IP Forwarding within the Different Subnet aka Routing (Data Packet)

- VM1 generates a data packet destined to PM2 IP (20.20.20.20) with GW_MAC as destination MAC
- TOR receives the packet and performs Layer-3 lookup for the destination (known)
- TOR adds VXLAN-Header information (Destination VTEP, VNI, etc) and forwards the packet across the Layer-3 fabric, picking one of the equal cost paths available via the multiple Spines
- The destination TOR receives the packet, strips off the VXLAN header and performs lookup and forwarding toward PM2



VXLAN Evolution

Multicast Independent

- Head-end replication enables unicast-only mode
- Control Plane provides dynamic VTEP discovery

Protocol Learning prevents floods

- Workload MAC addresses learnt by VXLAN NVEs
- Advertise L2/L3 address-to-VTEP association information in a protocol

External Connectivity

- VXLAN HW Gateways to other encaps/networks
- VXLAN HW Gateway redundancy
- Enable hybrid overlays

IP Services

- VXLAN Routing
- Distributed IP Gateways

VXLAN Designs

VXLAN Design Considerations

VXLAN Mode:

- Flood-and-Learn
- With EVPN control Plane

BUM Traffic Handling:

- Multicast replication
- Unicast/ingress replication

Deployment Scenarios:

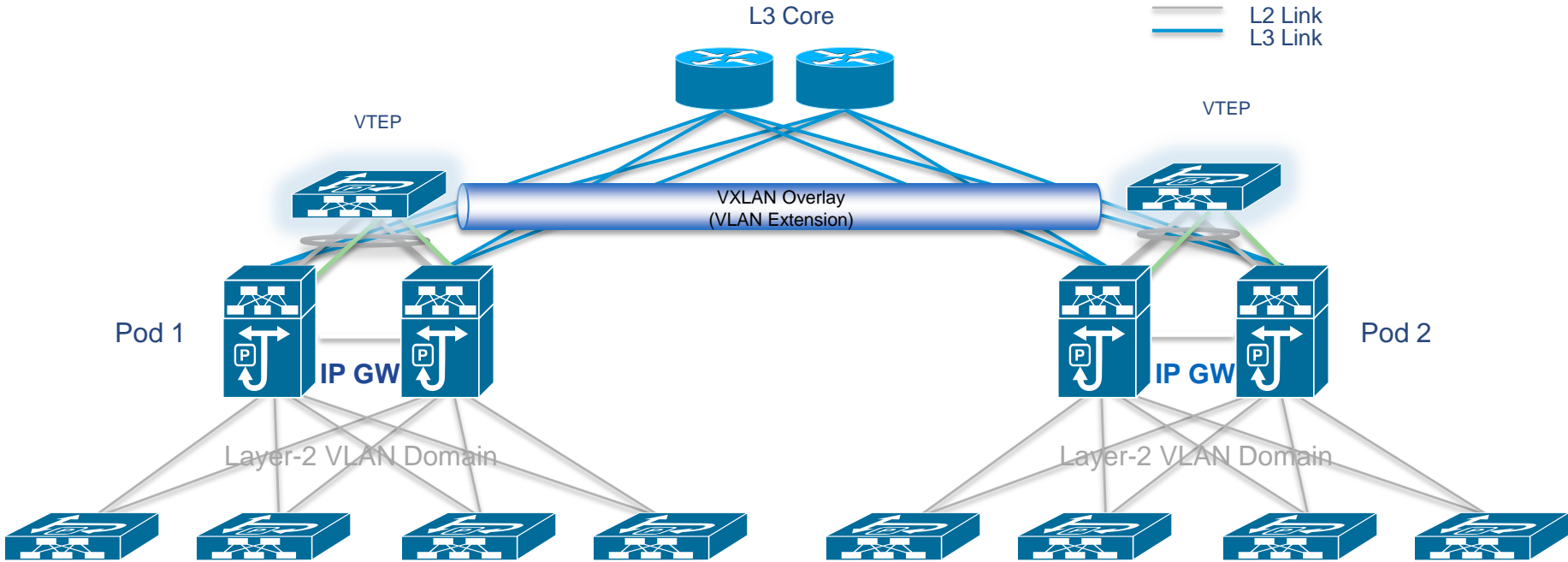
- Brown field vs green field
- Investment protection
- Multi-vendor environment?

Scalability:

- The number of VXLAN VNIs
- The number of VTEP peers
- The number of EVPN tenants
- The number of VXLAN Host IP routes
- The number of VXLAN Host MAC addresses
- The number of IPv4/IPv6 LPM routes
- The number of Ingress replication peers

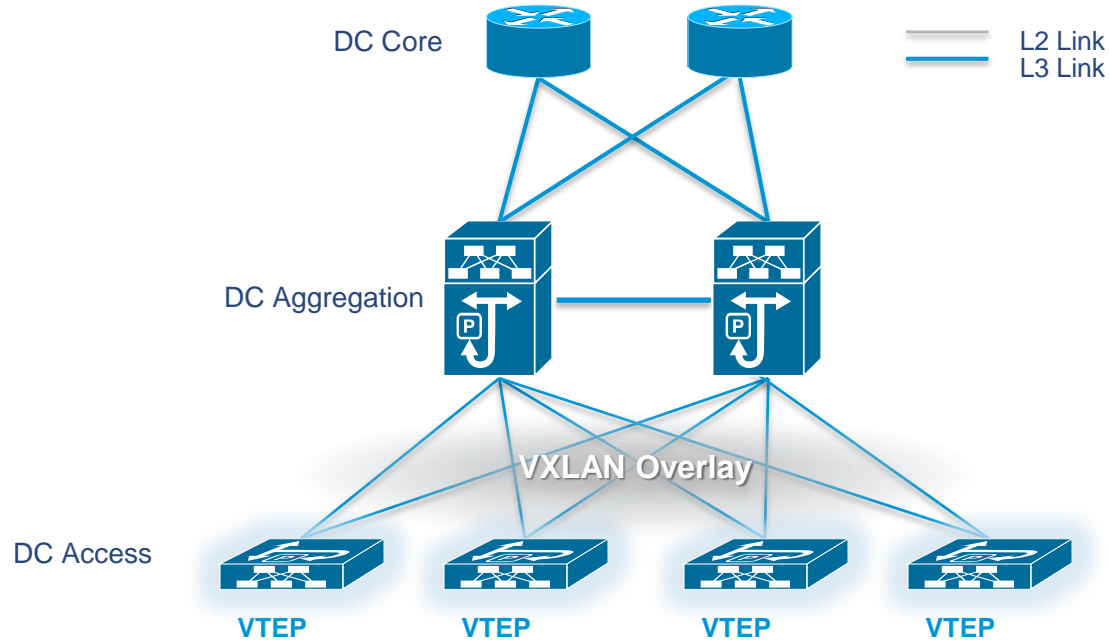
VXLAN Inter-PoD Extension

Brownfield: Connecting Two Data Center PODs



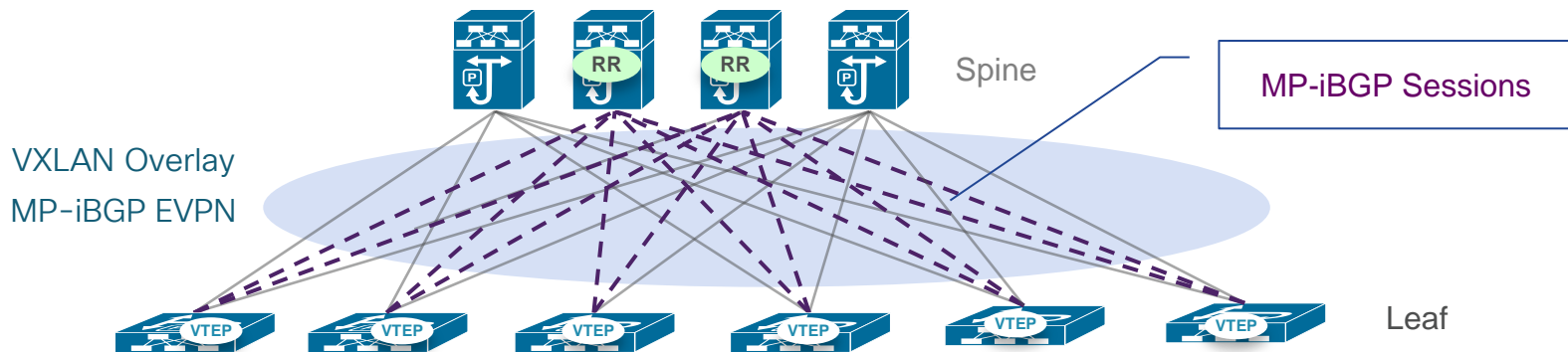
VXLAN in 3-Tier Network

Brownfield: Cross Layer 3 Boundaries



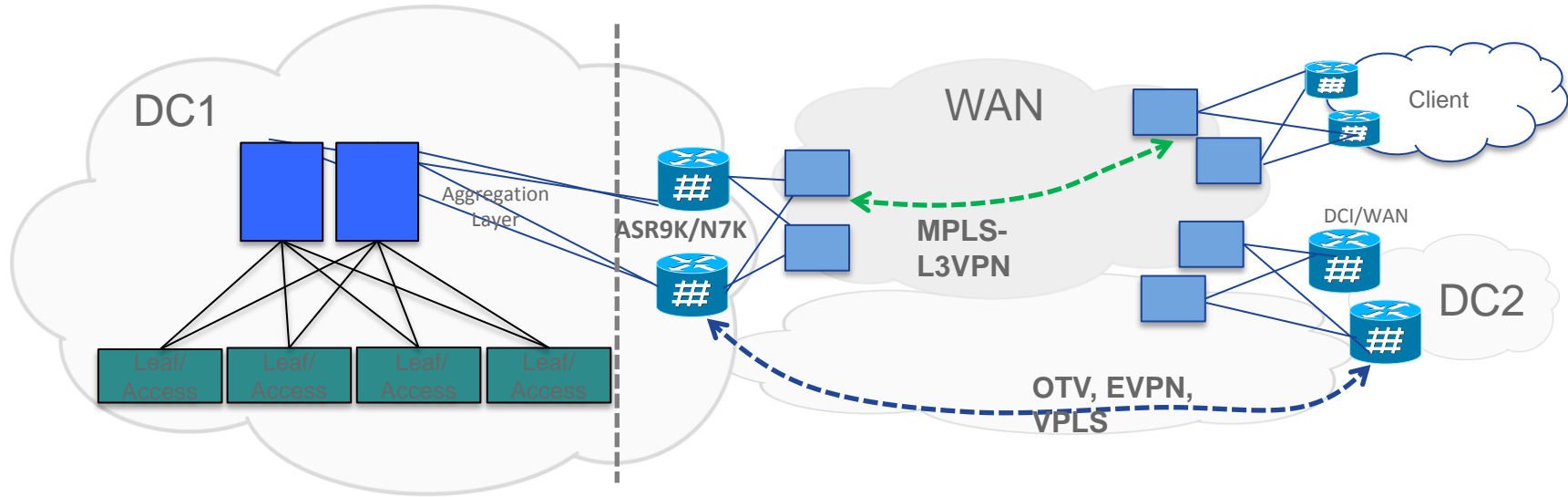
VXLAN Fabric Design with BGP EVPN

Greenfield: Multi-Tenancy with Mobility Support



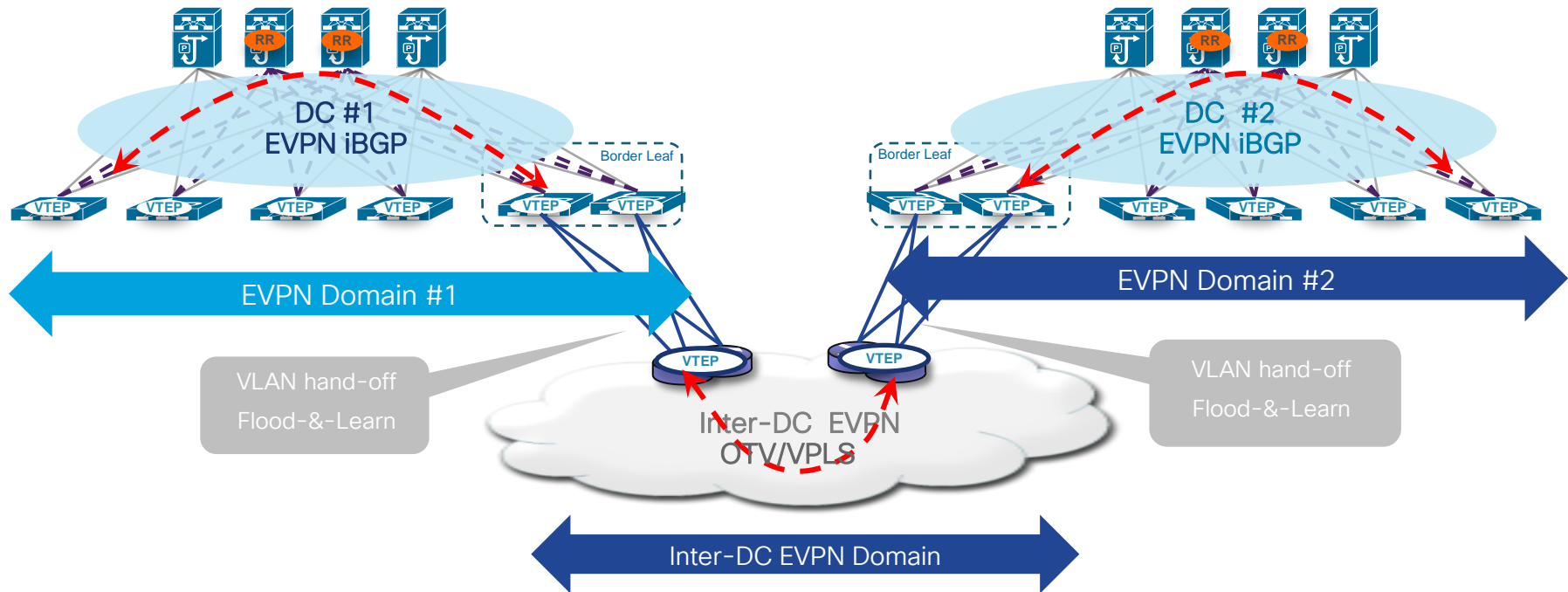
- Tunnel Endpoints are on leaf layer
- Spine nodes are iBGP Route Reflectors
- Supports Multi-tenancy with seamless Host Mobility

Integrate VXLAN with WAN

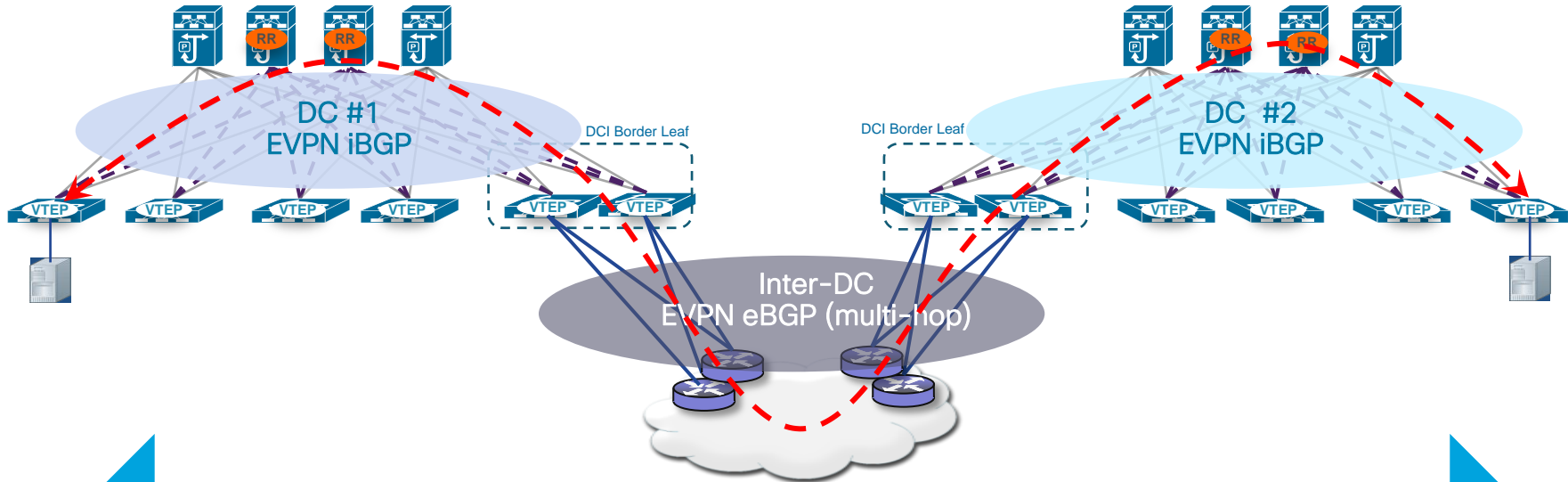


- For Disaster Recovery, High Availability
- Integrate EVPN/VXLAN to MPLS-L3VPN or LISP

Data Center Interconnectivity with VXLAN EVPN (Option A)



DCI with VXLAN EVPN (Option B)



One EVPN Administrative Domain
Stretched Across Two Data Centers

VXLAN: Flood-&-Learn vs EVPN Control Plane

	Flood-&-Learn	EVPN Control Plane
Overlay Services	L2+L3	L2+L3
Underlay Network	IP network with ECMP	IP network with ECMP
Encapsulation	MAC in UDP	MAC in UDP
Peer Discovery	Data-driven flood-&-learn	MP-BGP
Peer Authentication	Not available	MP-BGP
Host Route Learning	Local hosts: Data-driven flood-&-learn Remote hosts: Data-driven flood-&-learn	Local Host: Data-driven Remote host: MP-BGP
Host Route Distribution	No route distribution.	MP-BGP
L2/L3 Unicast Forwarding	Unicast encap	Unicast encap
BUM Traffic forwarding	Multicast replication Unicast/Ingress replication	Multicast replication Unicast/Ingress replication

Polling Question 3

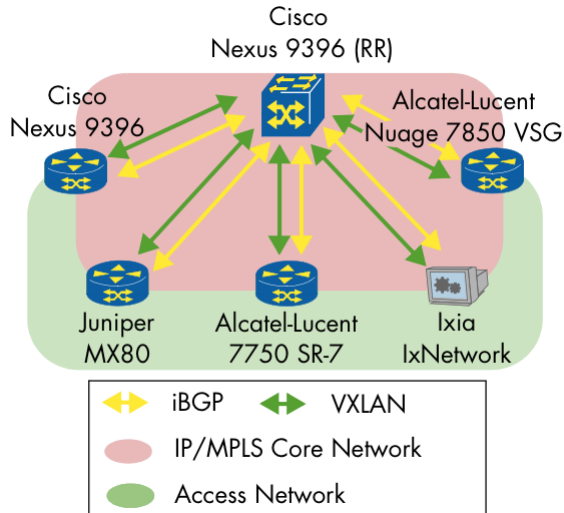
Since VXLAN w/BGP-EVPN is standard based, is multi-vendor integration a possibility ?

Yes

No

VXLAN/EVPN - Interoperability & Feasibility

- VXLAN/EVPN interoperability demonstrated during MPLS/SDN World Congress in Paris
- Participating Vendors are Cisco, Juniper, Alcatel Lucent & Ixia



Independently Tested at EANTC with public available Whitepaper

http://www.eantc.de/showcases/mpls_sdn_2015/intro.html

Resources

Resources

- **VXLAN Overview: Cisco Nexus 9000 Series Switches**
<http://www.cisco.com/c/en/us/products/collateral/switches/nexus-9000-series-switches/white-paper-c11-729383.html>
- **VXLAN Network with MP-BGP EVPN Control Plane**
<http://www.cisco.com/c/en/us/products/collateral/switches/nexus-9000-series-switches/guide-c07-734107.html>
- **Fundamentals of VXLAN**
http://www.cisco.com/c/en/us/products/collateral/switches/nexus-7000-series-switches/vidoe_fundamentals_vxlan.html
- **Digging Deeper into VXLAN, Part 1**
<http://blogs.cisco.com/datacenter/digging-deeper-into-vxlan>
- **Virtual Extensible LAN (VXLAN) Best Practices**
<http://www.cisco.com/c/en/us/products/collateral/switches/nexus-5000-series-switches/white-paper-c11-733618.html>



Submit Your Questions Now!

Use the Q & A panel to submit your questions and our expert will respond

Ask the Expert Event following the Webcast

Now through October 30th

<https://supportforums.cisco.com/discussion/12604376/ask-expert-cisco-data-center-overlays-focus-vxlan>



Collaborate within our Social Media

Learn About Upcoming Events



Facebook- <http://bit.ly/csc-facebook>



Twitter- <http://bit.ly/csc-twitter>



You Tube <http://bit.ly/csc-youtube>



Google+ <http://bit.ly/csc-googleplus>



LinkedIn <http://bit.ly/csc-linked-in>



Instagram <http://bit.ly/csc-instagram>



Newsletter Subscription
<http://bit.ly/csc-newsletter>

Cisco has support communities in other languages!

If you speak Spanish, Portuguese, Japanese, Russian or Chinese we invite you to participate and collaborate in your language



Spanish

<https://supportforums.cisco.com/community/spanish>

Portuguese

<https://supportforums.cisco.com/community/portuguese>

Japanese

<https://supportforums.cisco.com/community/csc-japan>

Russian

<https://supportforums.cisco.com/community/russian>

Chinese

<http://www.csc-china.com.cn>



More IT Training Videos and Technical Seminars on the Cisco Learning Network

View Upcoming Sessions Schedule
<https://cisco.com/go/techseminars>



Please take a moment to complete the survey

Thank you for Your Time!

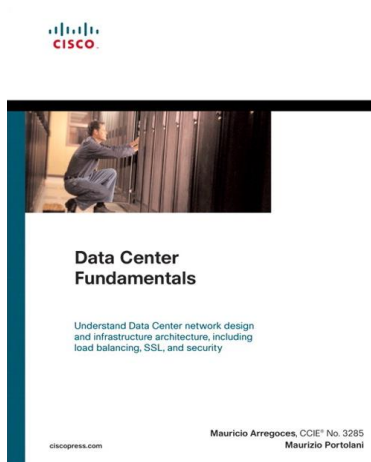
Thank you for participating!

- Redeem your 35% discount offer by entering code: CSC when checking out:

Visit Cisco Press at:

<http://bit.ly/csc-ciscopress-oct15>

Cisco Press





CISCO

TOMORROW starts here.