
cisco
You make possible

VXLAN BGP EVPN based Multi-Site

Lukas Krattiger – Principal Engineer

BRKDCN-2035

cisco Live!
Barcelona | January 27-31, 2020

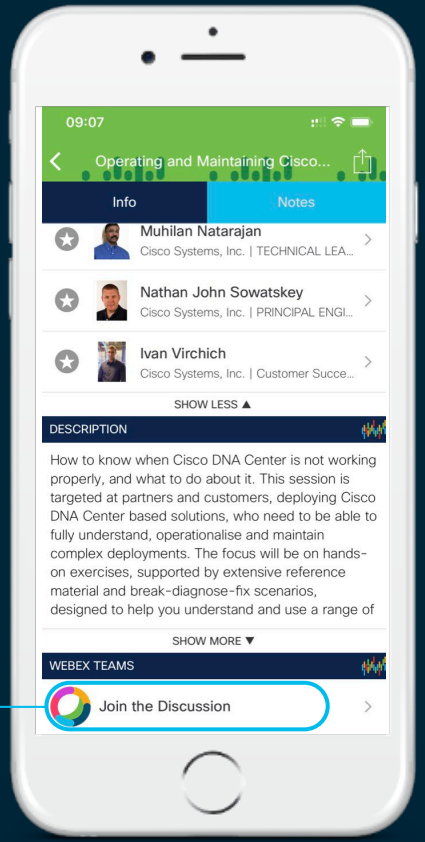
Cisco Webex Teams

Questions?

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

How

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



Agenda

- VXLAN EVPN and Data Center Interconnect (DCI) Evolution
- VXLAN Multi-Site Introduction
  - Functional Components and Use Cases
  - HW/SW Support and Scalability Values
  - Supported Topologies
- VXLAN Multi-Site Deep Dive
  - Border Gateway Deployment Considerations
  - Inter-Site BUM Traffic Handling
  - Control and Data Planes
  - Connectivity to the External Layer 3 Domain
  - Legacy Site Integration
  - Configuration Specifics (for your reference)
- Conclusions

VXLAN EVPN and Data Center Interconnect (DCI) Evolution

VXLAN Evolves as the Control Plane Evolves!

Before Yesterday

Yet Another Encapsulation

- Flood & Learn (Multicast-based)
- Data-Plane only

VXLAN Evolves as the Control Plane Evolves!

Before Yesterday

Yet Another Encapsulation

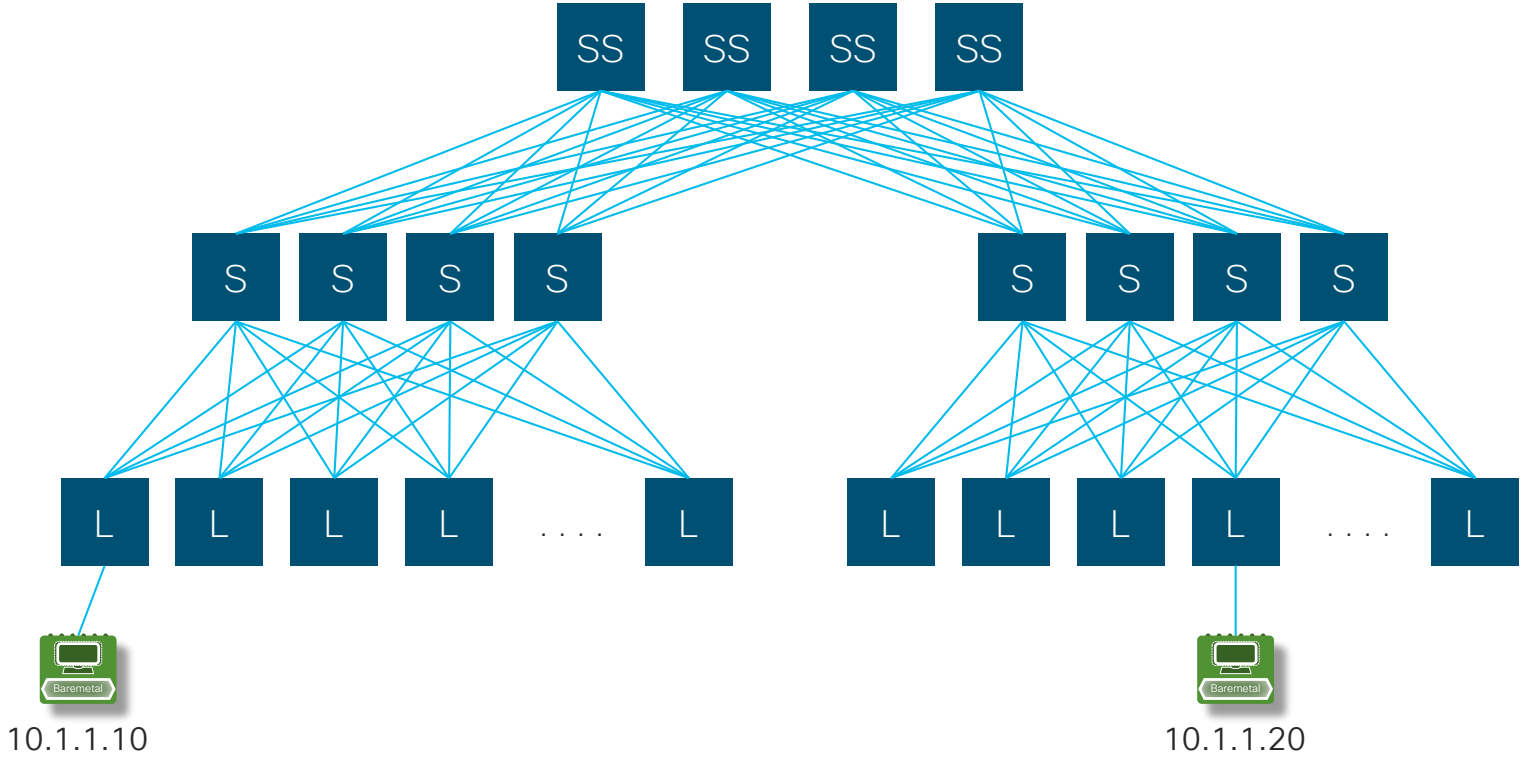
- Flood & Learn (Multicast-based)
- Data-Plane only

Yesterday

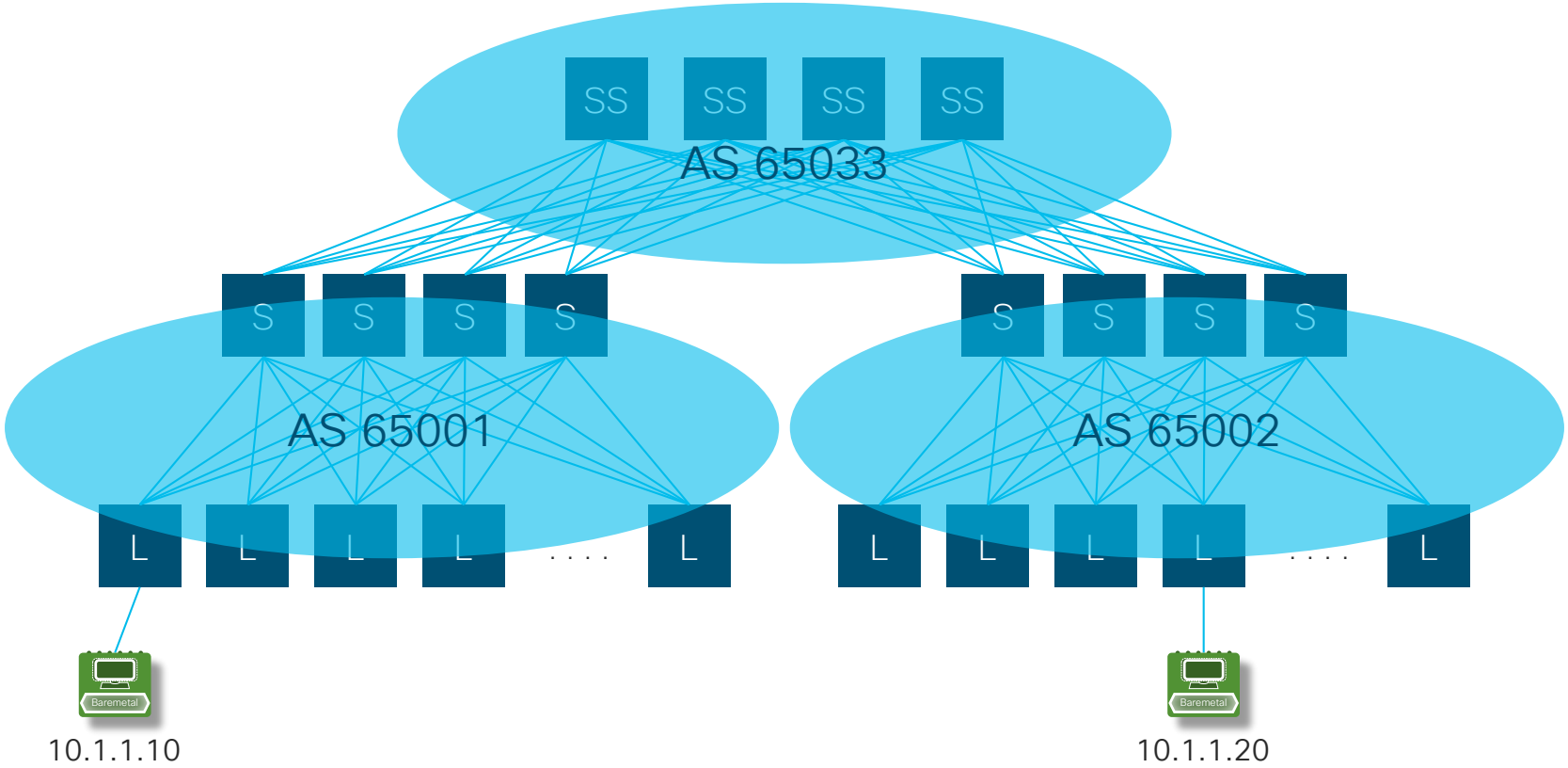
VXLAN for the Data Center – Intra-DC

- Control-Plane
- Active VTEP Discovery
- Multicast and Unicast

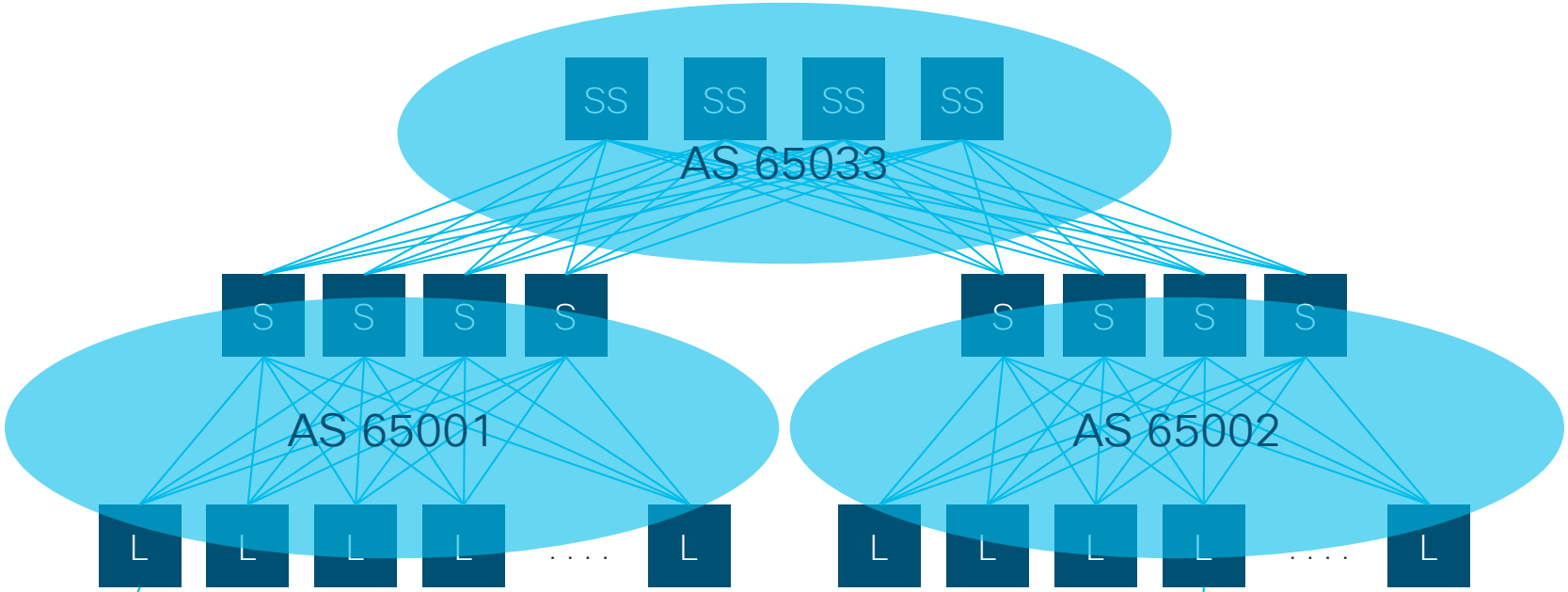
Traditional Overlay Forwarding



Traditional Overlay Forwarding



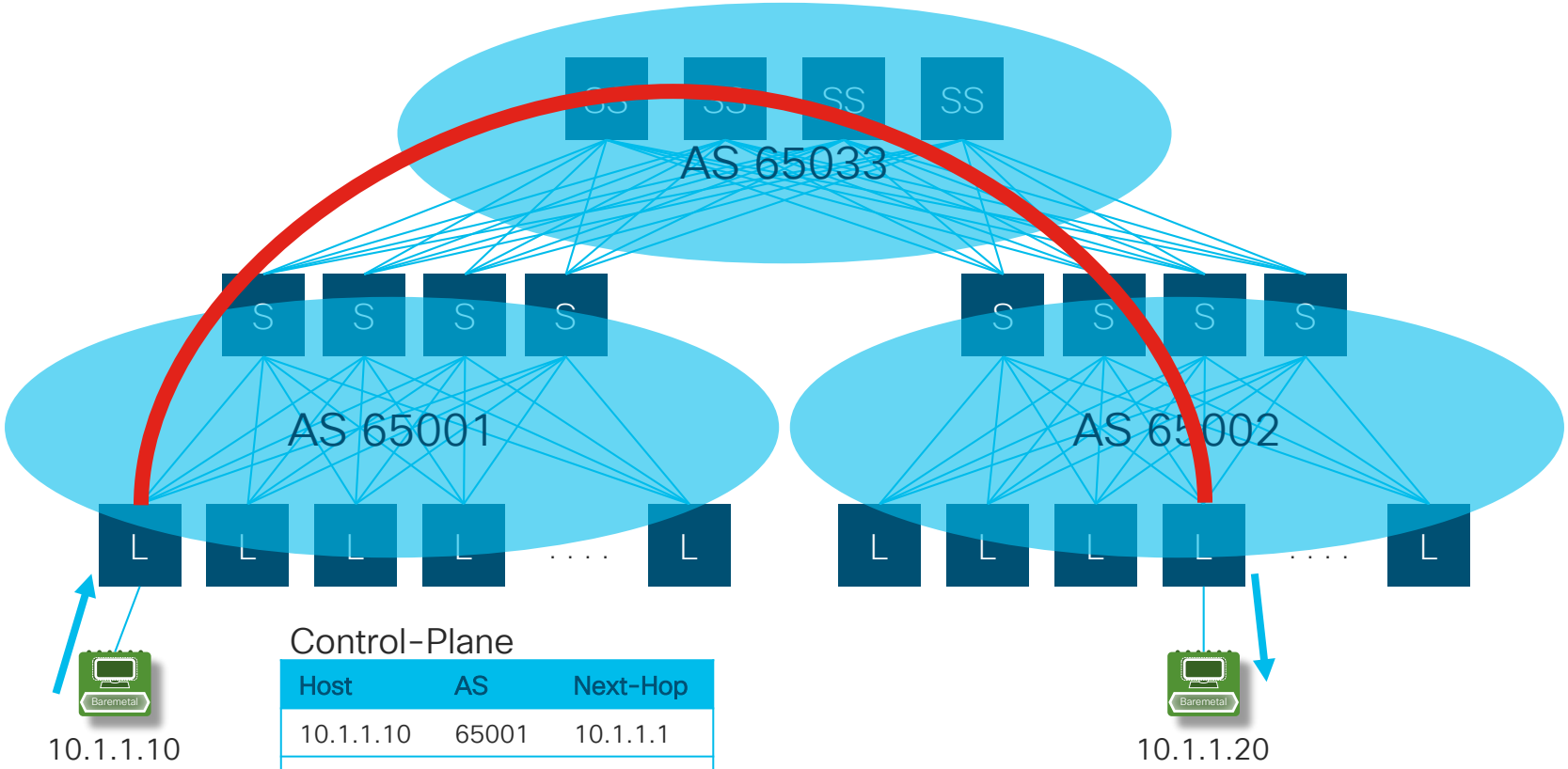
Traditional Overlay Forwarding



Control-Plane

Host	AS	Next-Hop
10.1.1.10	65001	10.1.1.1
10.1.1.20	65002	10.2.2.4

Traditional Overlay Forwarding

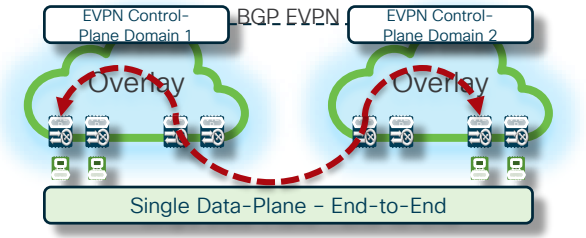


Control-Plane

Host	AS	Next-Hop
10.1.1.10	65001	10.1.1.1
10.1.1.20	65002	10.2.2.4

Inter-X Connectivity

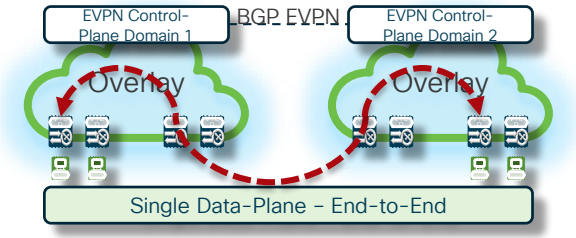
VXLAN Multi-Pod



- Single Fabric with End-to-End Encapsulation
- Build Hierarchy in the Underlay – Flatten it in the Overlay

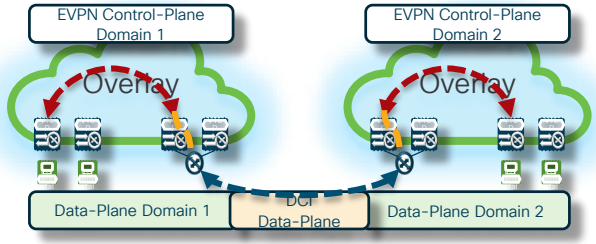
Inter-X Connectivity

VXLAN Multi-Pod



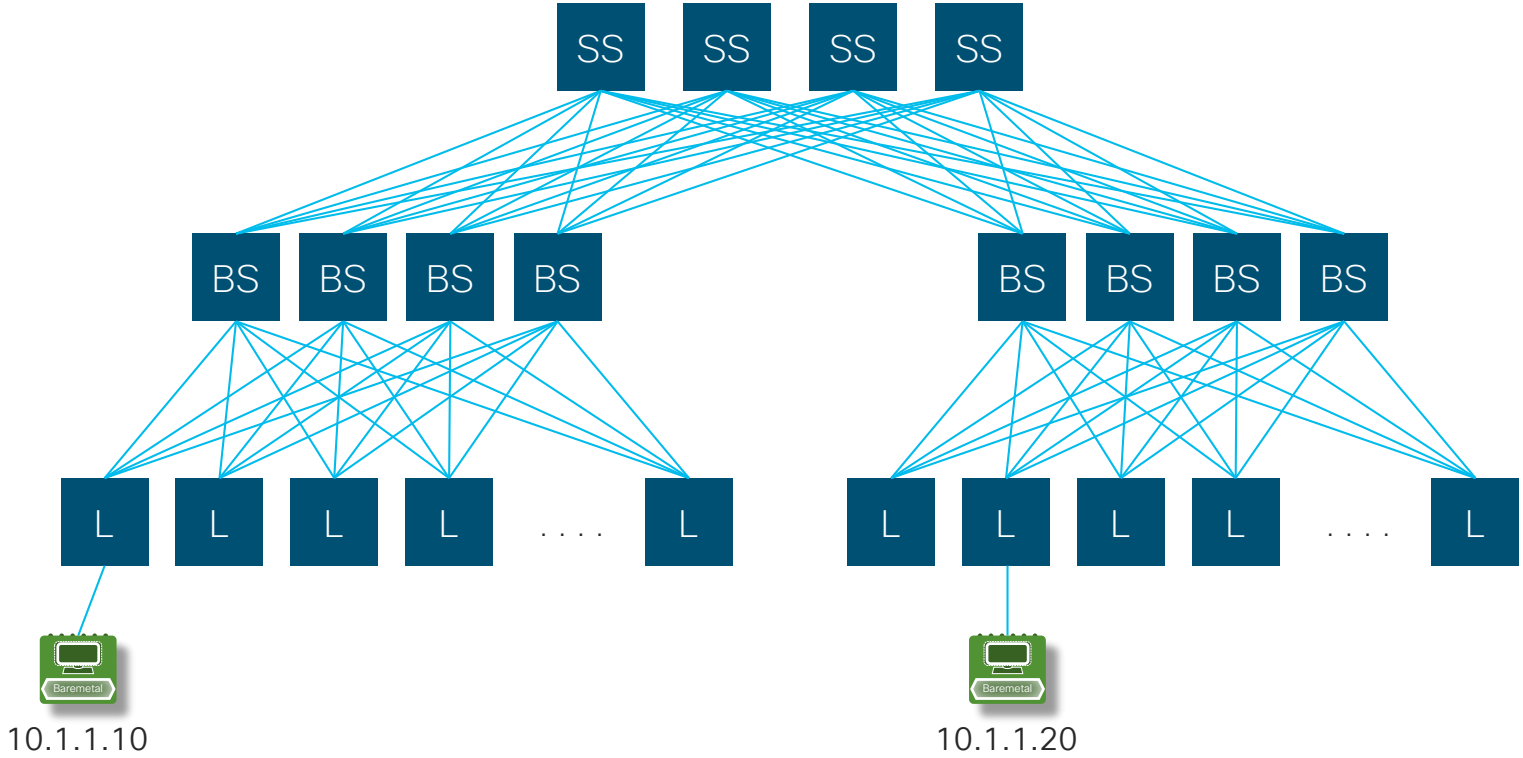
- Single Fabric with End-to-End Encapsulation
- Build Hierarchy in the Underlay – Flatten it in the Overlay

VXLAN Multi-Fabric

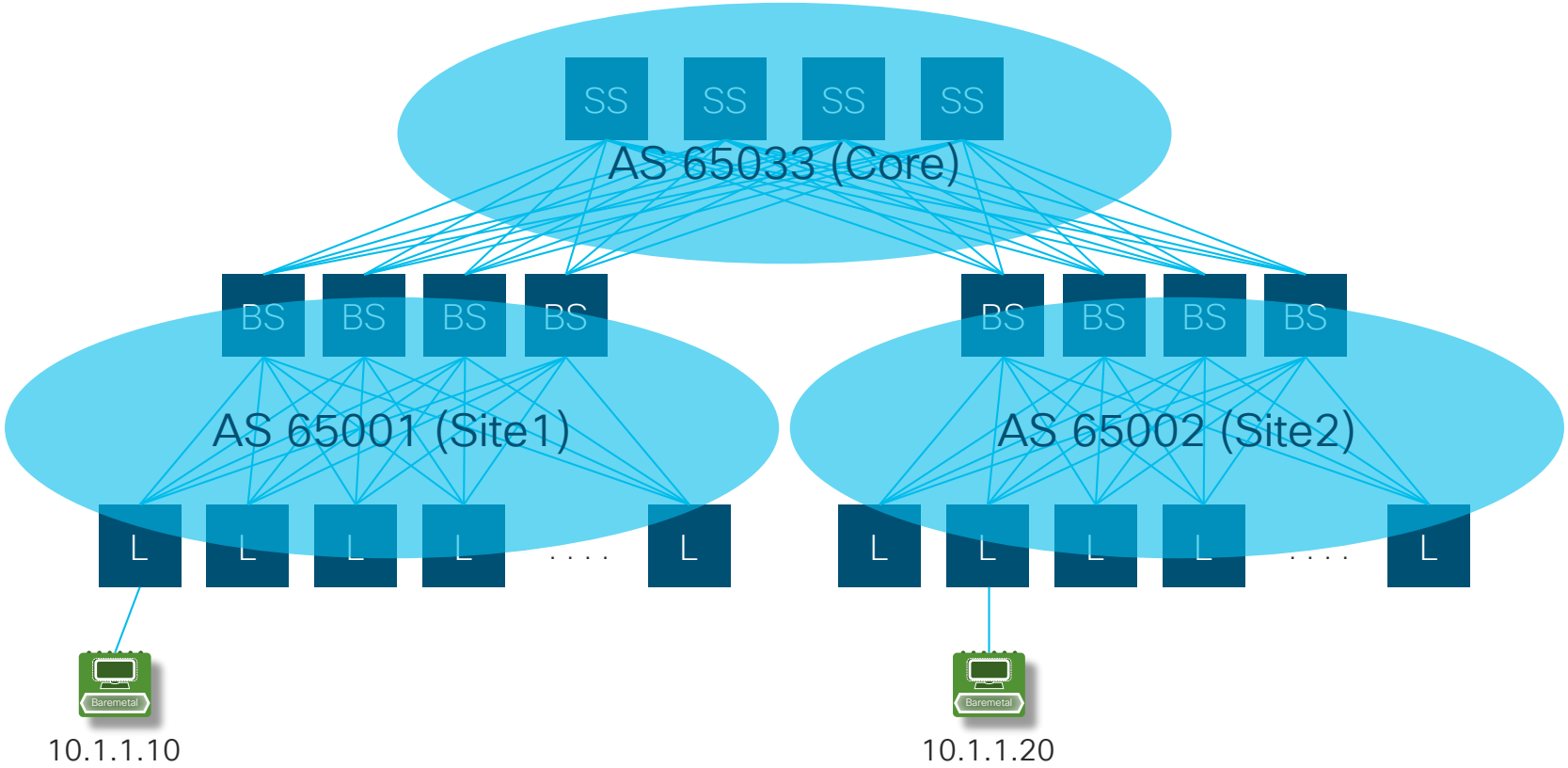


- Multiple Fabrics – Normalized through Ethernet
- Multiple Fabrics Interconnect using DCI (Layer 2 and Layer 3)

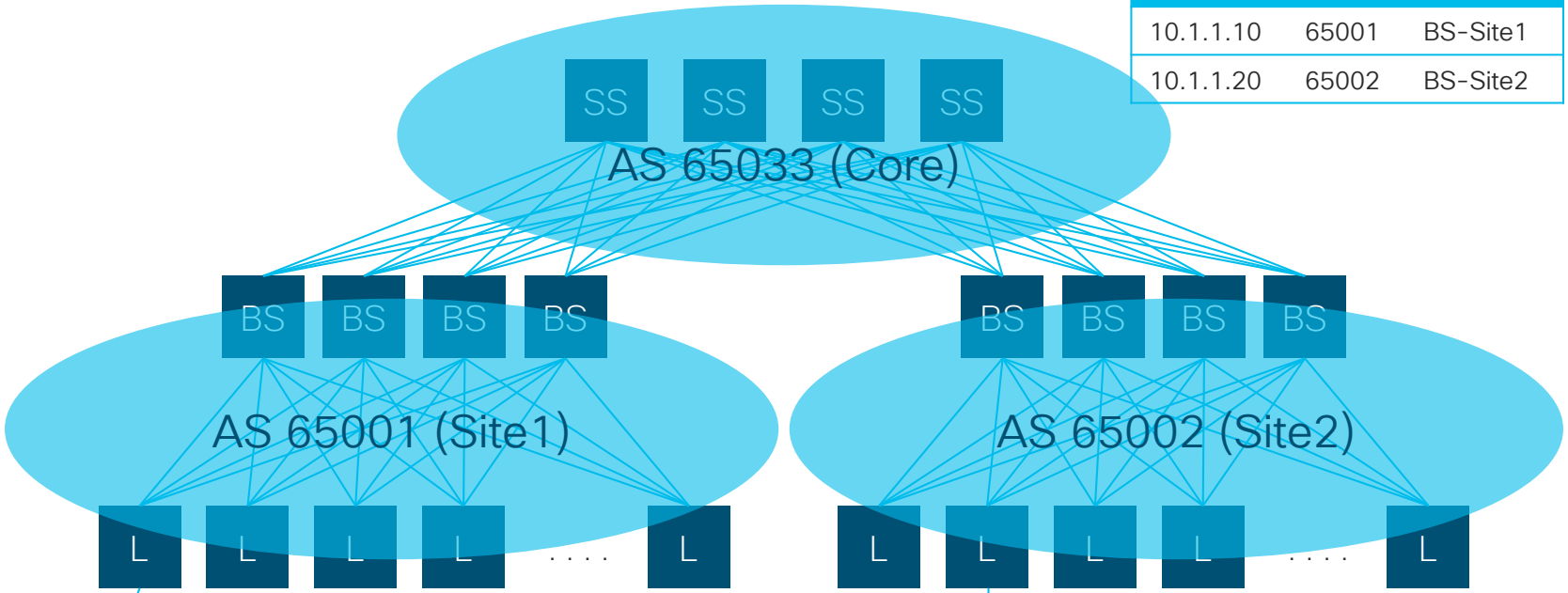
Network Routing Forwarding



Network Routing Forwarding



Network Routing Forwarding



Control-Plane (Core)

Host	AS	Next-Hop
10.1.1.10	65001	BS-Site1
10.1.1.20	65002	BS-Site2

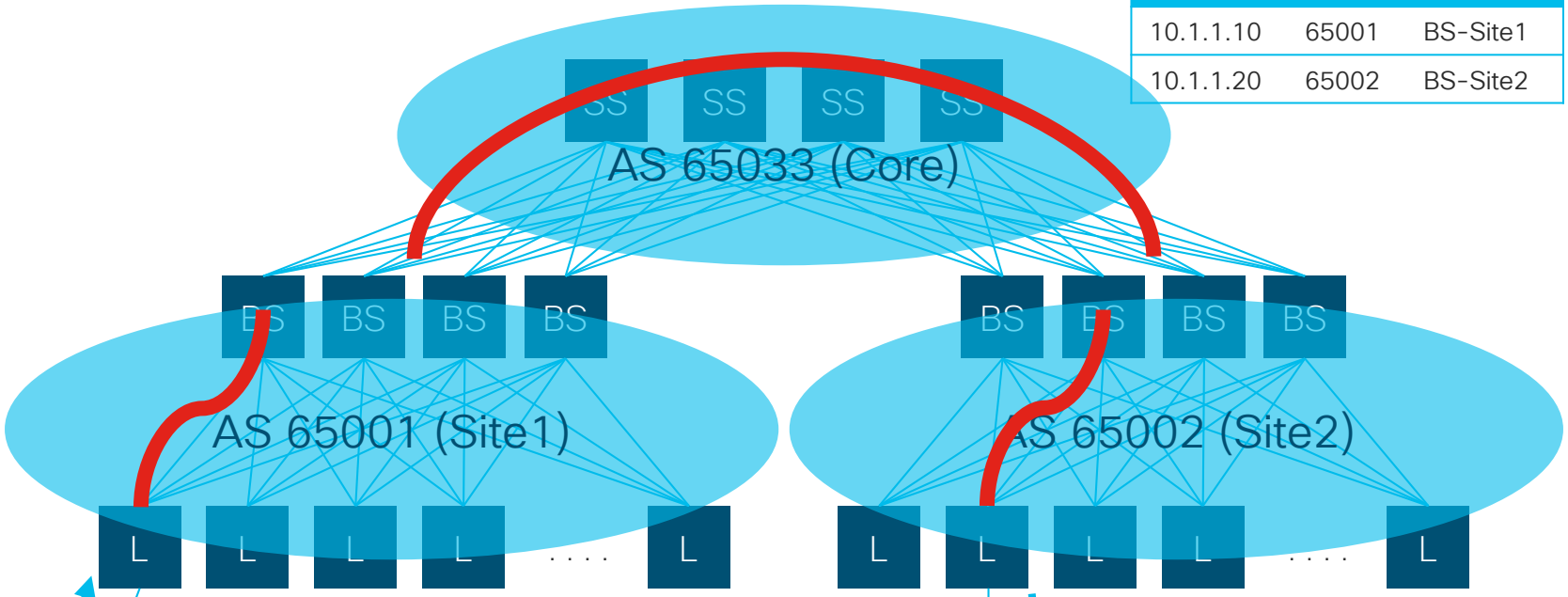
Control-Plane (Site1)

Host	AS	Next-Hop
10.1.1.10	65001	10.1.1.1
10.1.1.20	65002	BS-Site1

Control-Plane (Site2)

Host	AS	Next-Hop
10.1.1.10	65001	BS-Site2
10.1.1.20	65002	10.2.2.2

Network Routing Forwarding



Control-Plane (Core)

Host	AS	Next-Hop
10.1.1.10	65001	BS-Site1
10.1.1.20	65002	BS-Site2

Control-Plane (Site1)

Host	AS	Next-Hop
10.1.1.10	65001	10.1.1.1
10.1.1.20	65002	BS-Site1

Control-Plane (Site2)

Host	AS	Next-Hop
10.1.1.10	65001	BS-Site2
10.1.1.20	65002	10.2.2.2

VXLAN Evolves as the Control Plane Evolves!

Before Yesterday

Yet Another Encapsulation

- Flood & Learn (Multicast-based)
- Data-Plane only

Yesterday

VXLAN for the Data Center – Intra-DC

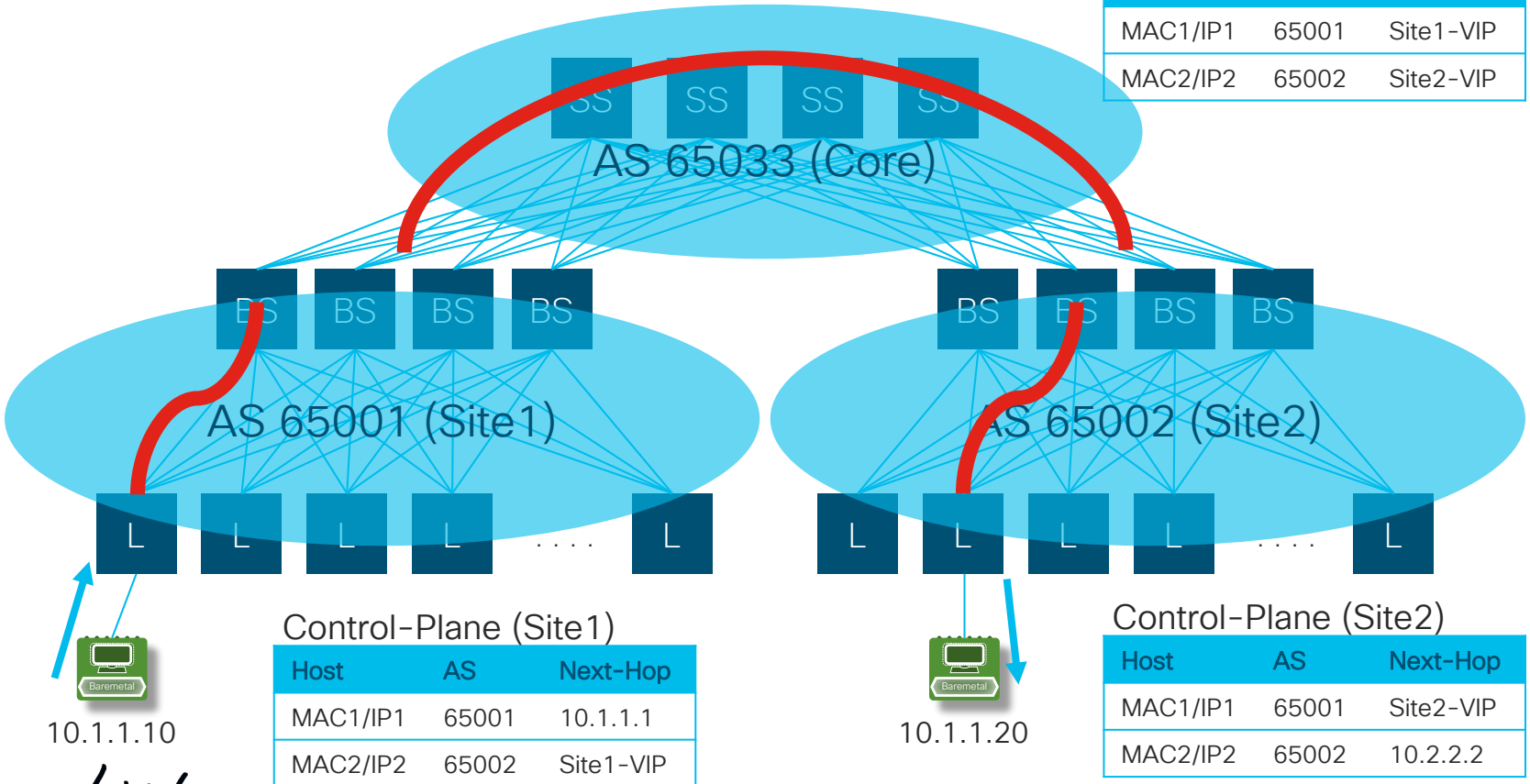
- Control-Plane
- Active VTEP Discovery
- Multicast and Unicast

Today

VXLAN for DCI – Inter-DC

- DCI Ready
- ARP/ND caching/suppress
- Multi-Homing
- Failure Domain Isolation
- Loop Protection

Multi-Site Overlay Forwarding



Control-Plane (Core)

Host	AS	Next-Hop
MAC1/IP1	65001	Site1-VIP
MAC2/IP2	65002	Site2-VIP

Control-Plane (Site1)

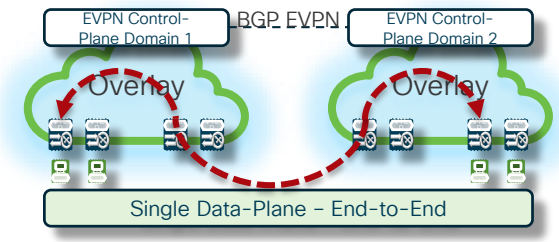
Host	AS	Next-Hop
MAC1/IP1	65001	10.1.1.1
MAC2/IP2	65002	Site1-VIP

Control-Plane (Site2)

Host	AS	Next-Hop
MAC1/IP1	65001	Site2-VIP
MAC2/IP2	65002	10.2.2.2

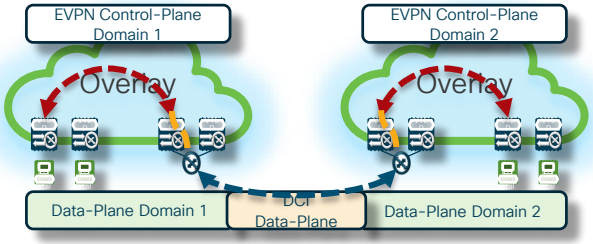
Inter-X Connectivity

VXLAN Multi-Pod



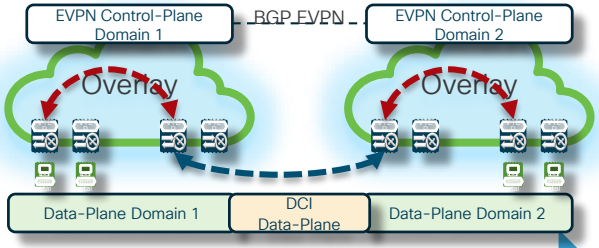
- Single Fabric with End-to-End Encapsulation
- Build Hierarchy in the Underlay – Flatten it in the Overlay

VXLAN Multi-Fabric



- Multiple Fabrics – Normalized through Ethernet
- Multiple Fabrics Interconnect using DCI (Layer 2 and Layer 3)

VXLAN Multi-Site



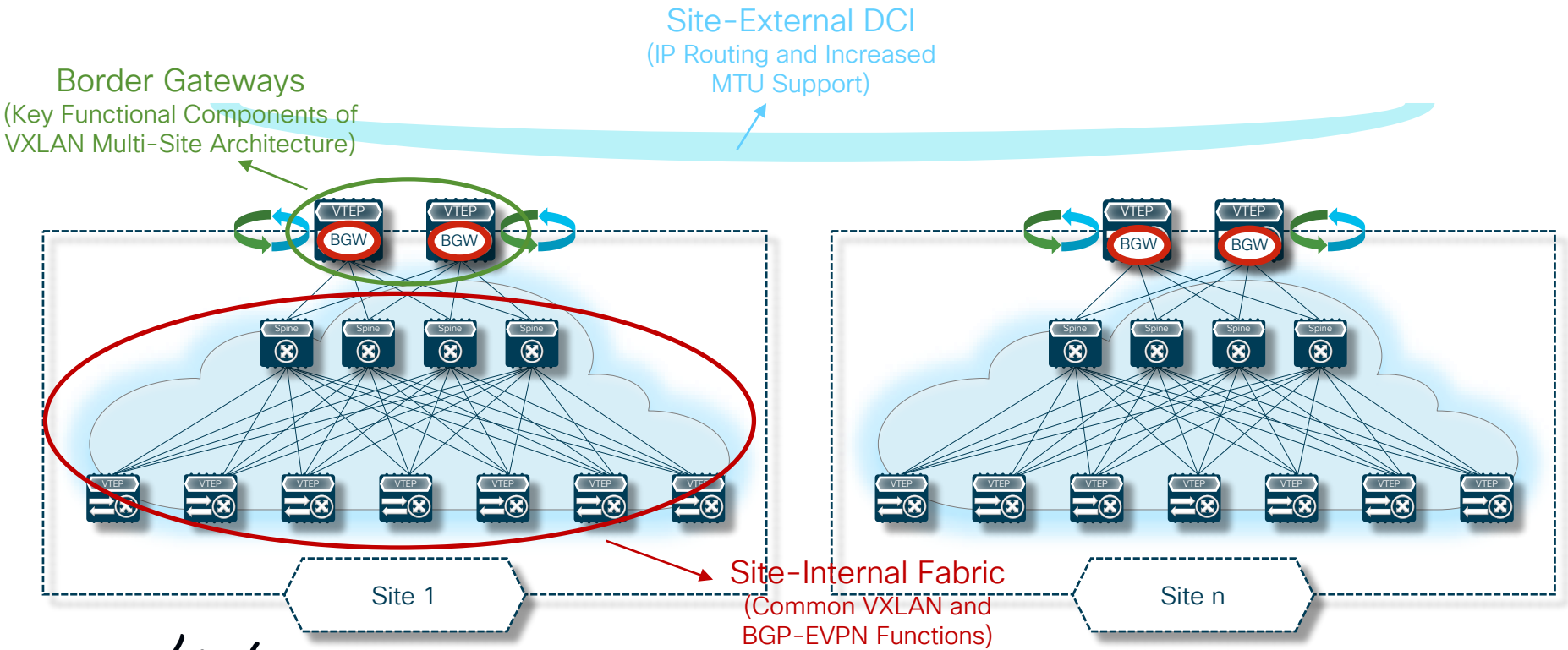
- Multiple Fabrics with Integrated DCI (DCI^2)
- Integrated DCI – Scaling within and between Fabrics

VXLAN Multi-Site Introduction

Functional Components and Use Cases

VXLAN Multi-Site
Functional Components

https://tools.ietf.org/html/draft-sharma-multi-site-evpn



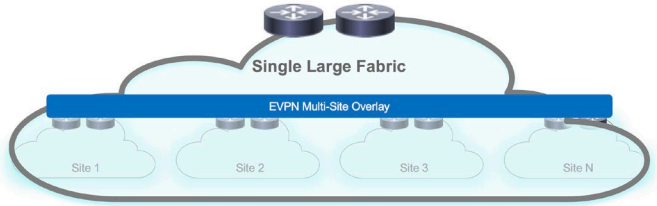
VXLAN Multi-Site Characteristics

- **Multiple** Overlay Domains – Interconnected & Controlled
- **Multiple** Overlay Control-Plane Domains – Interconnected & Controlled
- **Multiple** Underlay Domains - Isolated
- **Multiple** Replication Domains for BUM – Interconnected & Controlled
- **Multiple** VNI Administrative Domains – Phase 2

Underlay Isolation – Overlay Hierarchies

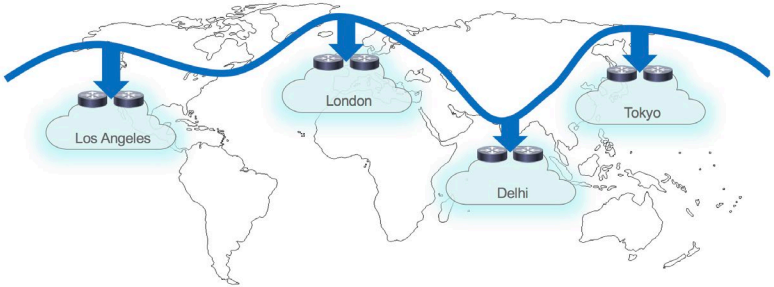
VXLAN Multi-Site

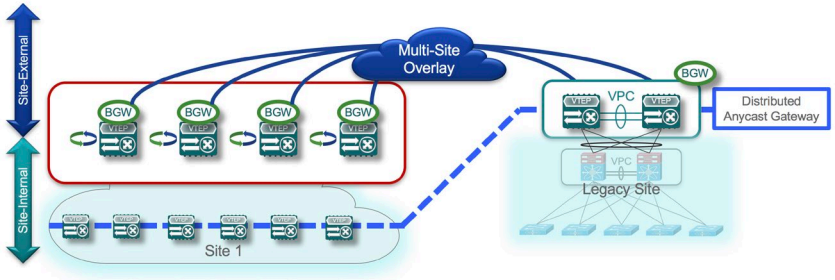
Main Use Cases



Scale-Up Model to Build a Large Intra-DC Network

Data Center Interconnect (DCI)

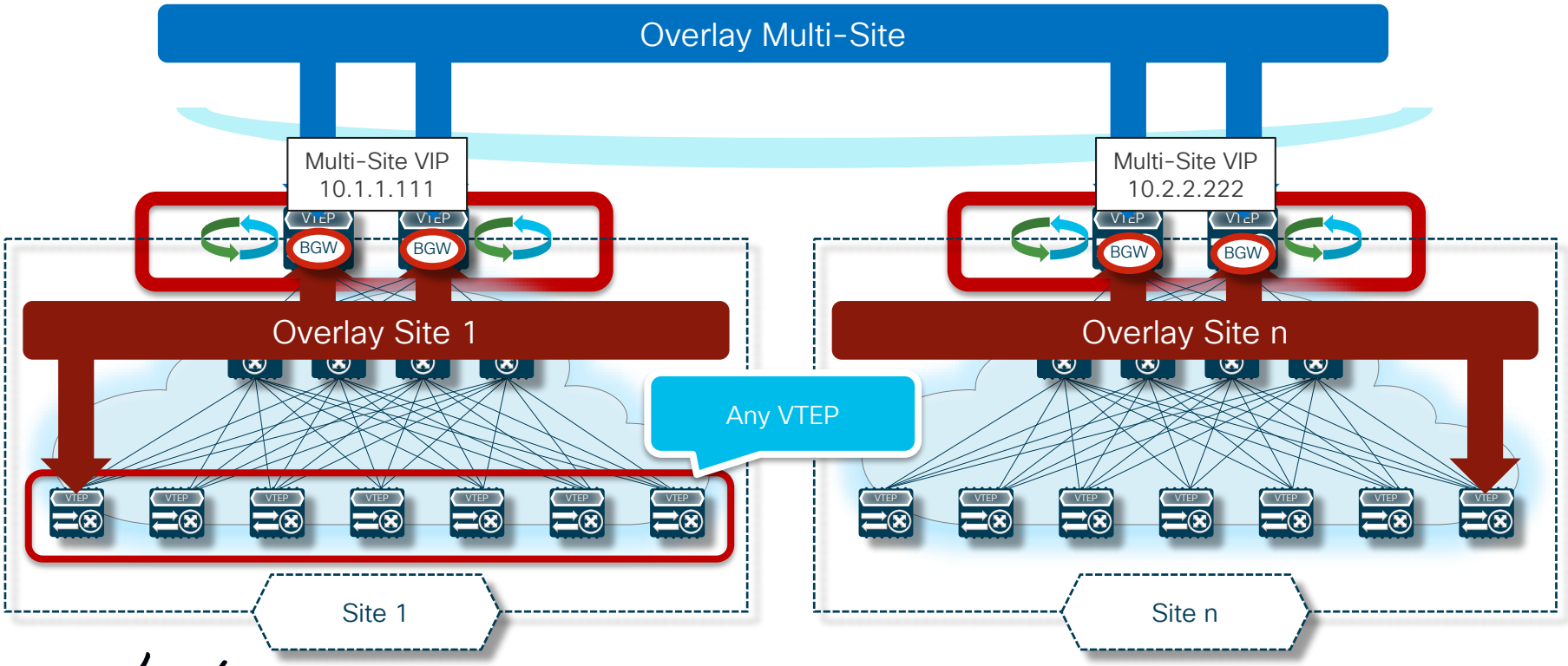




Integration with Legacy Networks (Coexistence and/or Migration)

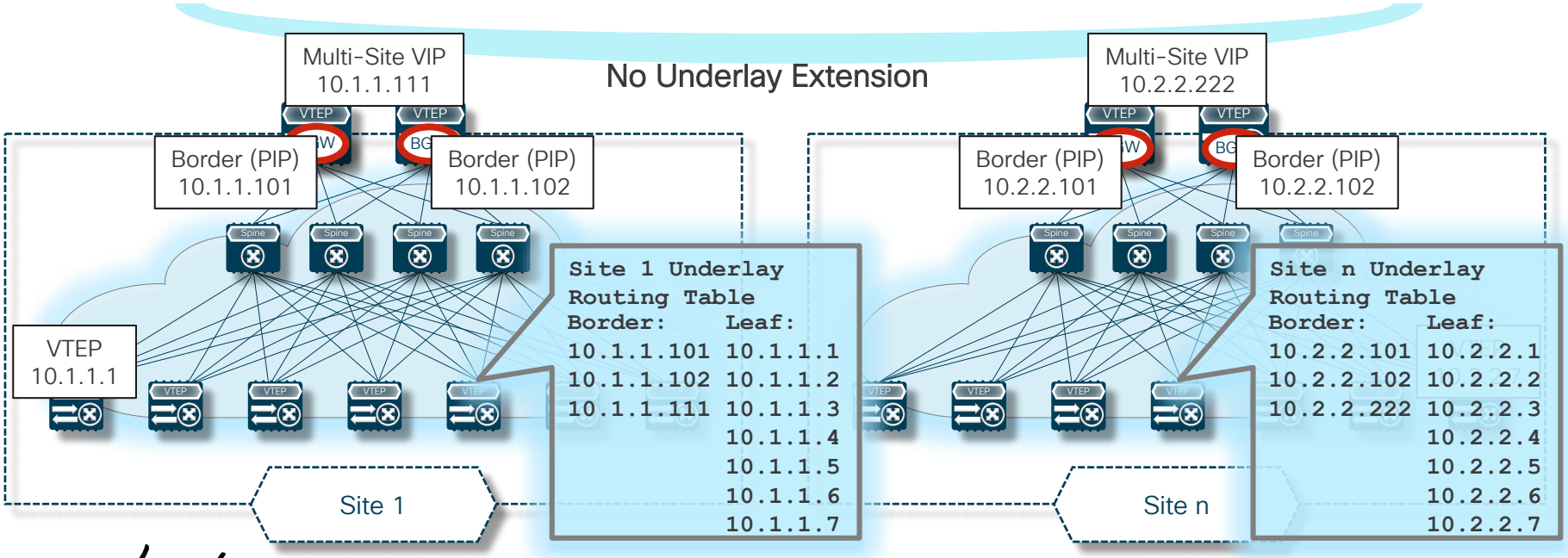
VXLAN Multi-Site

Introducing the Border Gateway



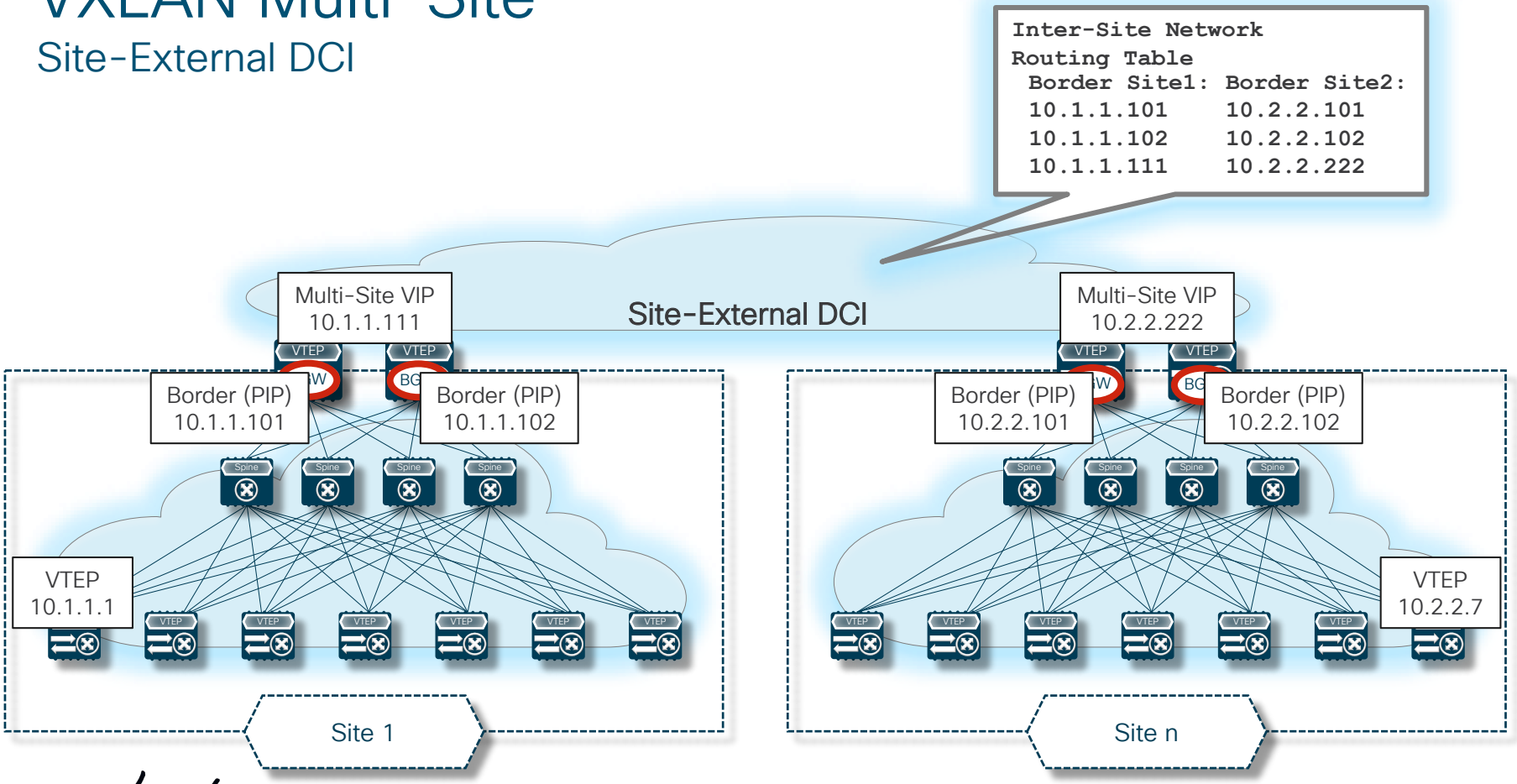
VXLAN Multi-Site

Underlay Isolation

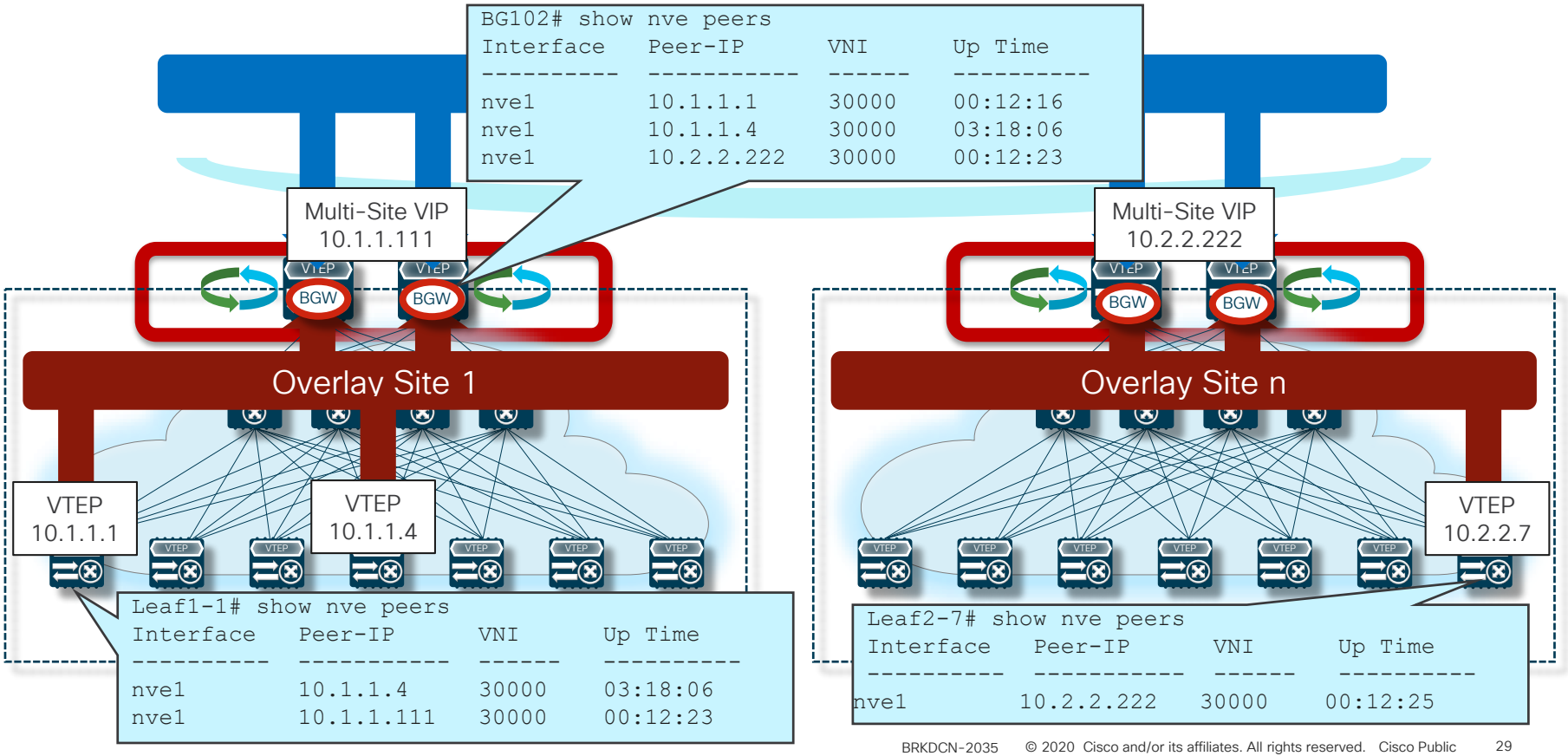


VXLAN Multi-Site

Site-External DCI



Multi-Site – VXLAN Tunnel Adjacencies



HW/SW Support and Scalability Values

VXLAN Multi-Site
HW/SW Support

- Minimum hardware and software requirements for Border Gateways

Item	Requirement
Cisco Nexus Hardware	• Cisco Nexus 9300 EX platform • Cisco Nexus 9300 FX platform • Cisco Nexus 9300 FX2 platform • Cisco Nexus 9364C platform • Cisco Nexus 9332C platform • Cisco Nexus 9500 platform with X9700-EX line card • Cisco Nexus 9500 platform with X9700-FX line card
Cisco Nexus Software	Cisco NX-OS Software Release 7.0(3)I7(1) or later

- The hardware and software requirements for the Site-Internal nodes of a VXLAN BGP EVPN site remain the same as those without the EVPN Multi-Site BGW

VXLAN Multi-Site

Scalability Values as of 9.2(3) Release

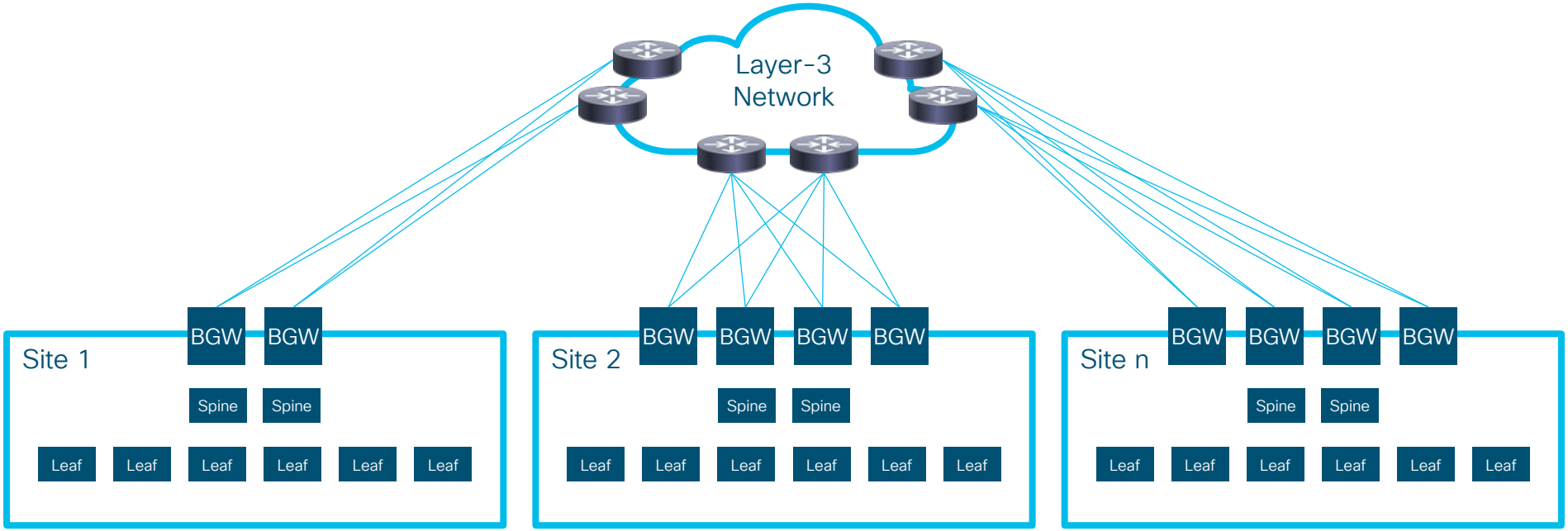
Multi-Site Scale	
Number of Sites	10
Number of BGWs per Site	4 (Anycast) or 2 (vPC)
VTEP per Site	256

Border Gateway (BGW) Scale	EX/FX/FX2	N9364C/N9332C
Number of Layer-2 VNI (VLAN)	2,000	
Number of Layer-3 VNI (VRF)	1,000	
MAC per BGW	90,000	64,000
IPv4 Host Routes per BGW*	~530,000	~60,000
IPv4 Network Routes per BGW*	~530,000	~8,000
IPv6 Host Routes per BGW*	~24,000	~7,000
IPv6 Network Routes per BGW*	~260,000	~2,000

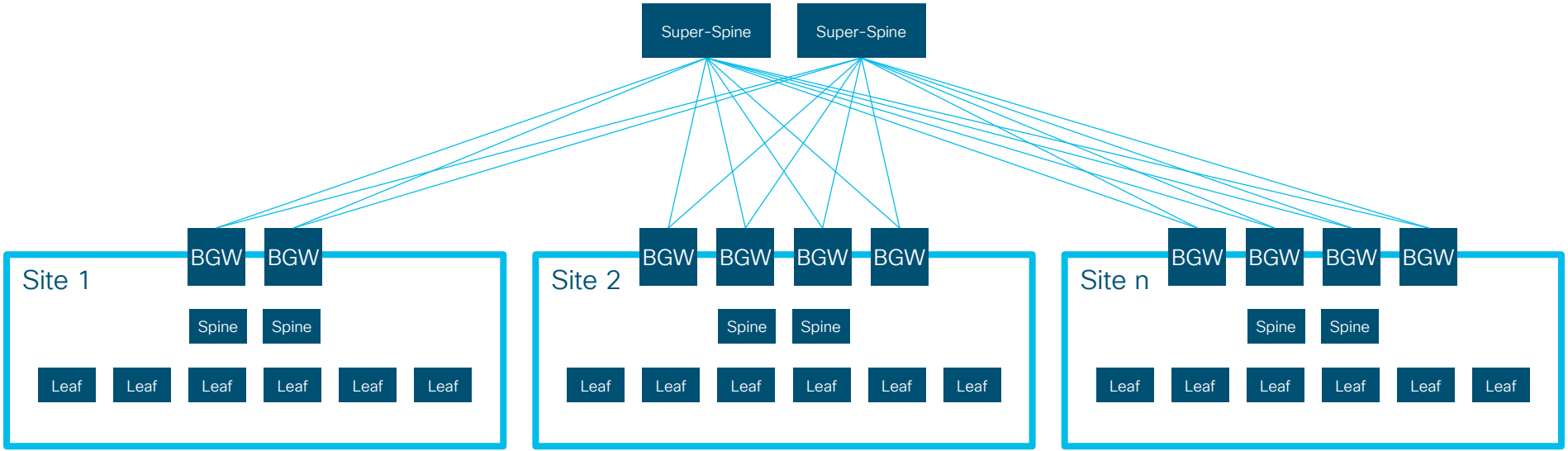
*The values provided in these tables focus on the scalability of one particular Route scale at a time

Supported Topologies

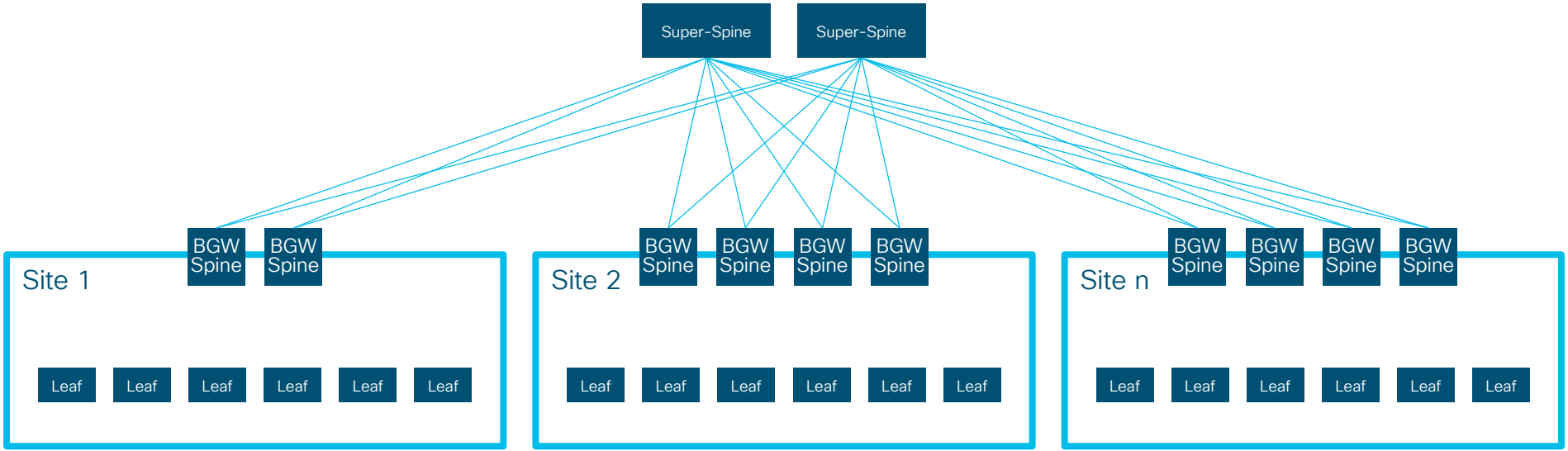
BGW-to-Cloud



BGWs between Spine and Super-Spine

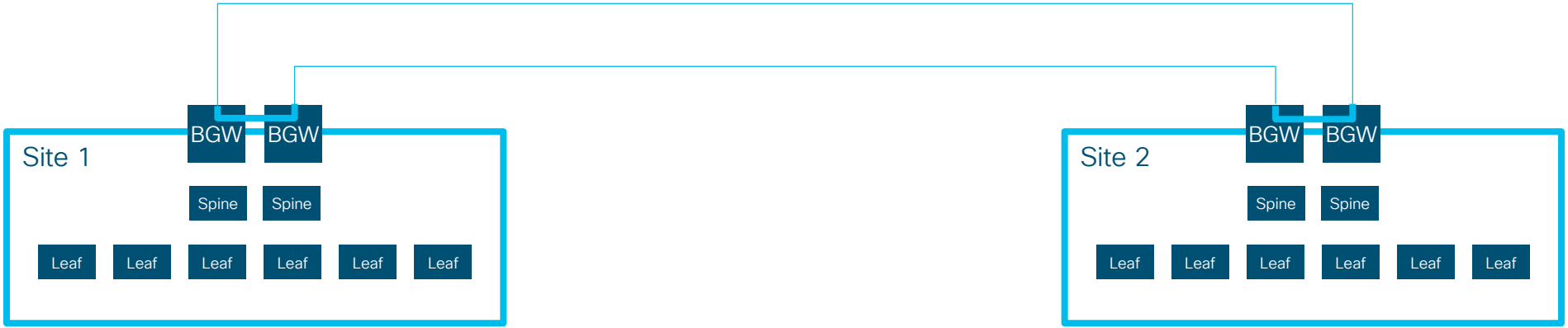


BGWs on Spine



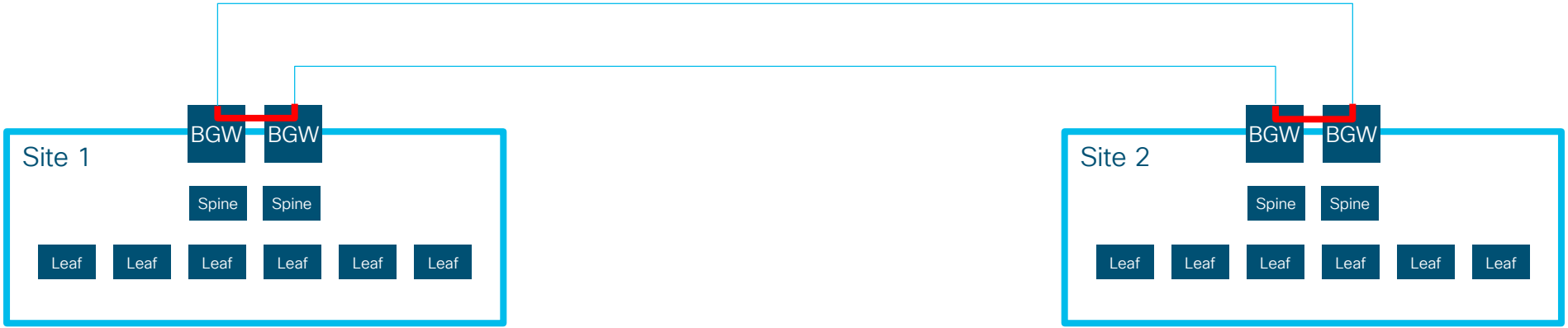
BGWs Back-to-Back

- Recommended to limit the back-to-back deployment to two sites
  - 2 Site topology can be fully automated using DCNM
  - Recommended to insert Layer-3 Core network with 3+ sites



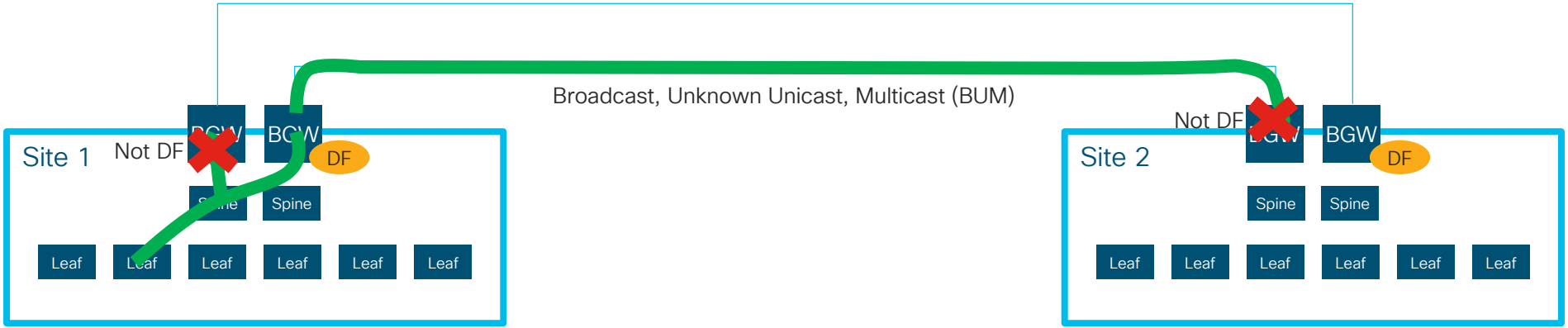
BGWs Back-to-Back

- Minimal Topology
  - Any to Any BGW Communication Required
  - BGW Local Link for Any to Any Reachability



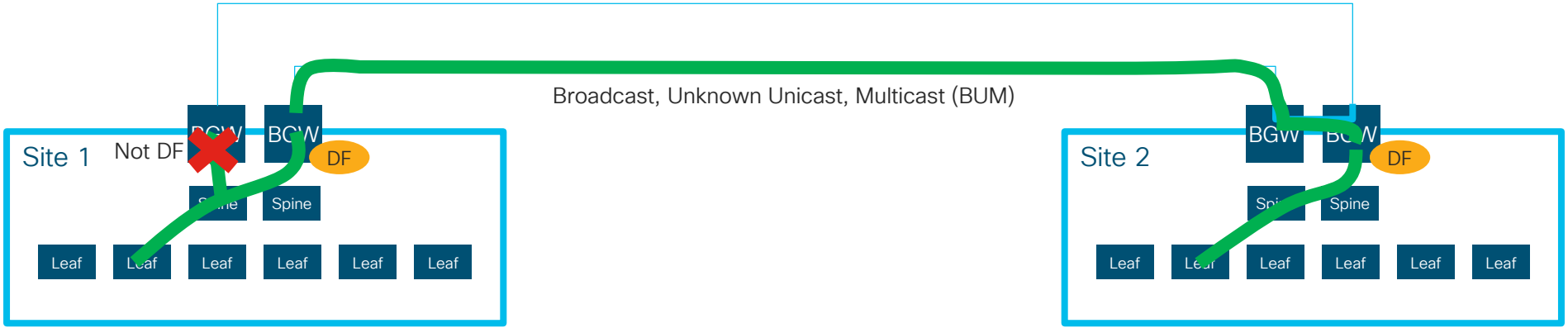
BGWs Back-to-Back

- Without Minimal Topology and Layer-2 Stretch



BGWs Back-to-Back

- With Minimal Topology and Layer-2 Stretch



VXLAN Multi-Site Deep Dive

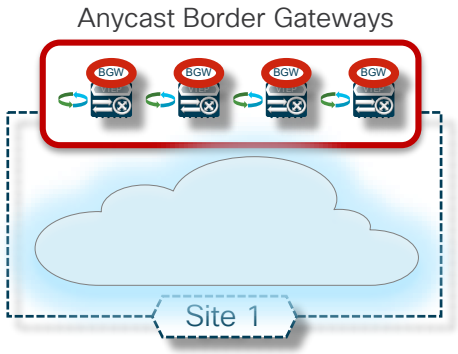
cisco Live!

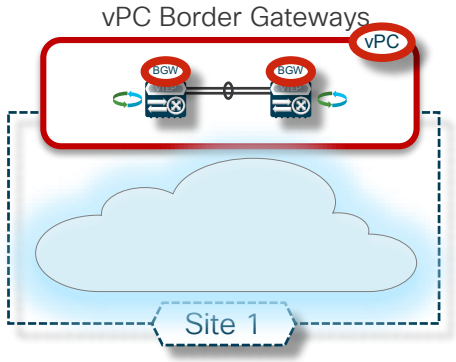
Border Gateway Deployment Considerations

VXLAN Multi-Site

Border Gateways Deployment Considerations

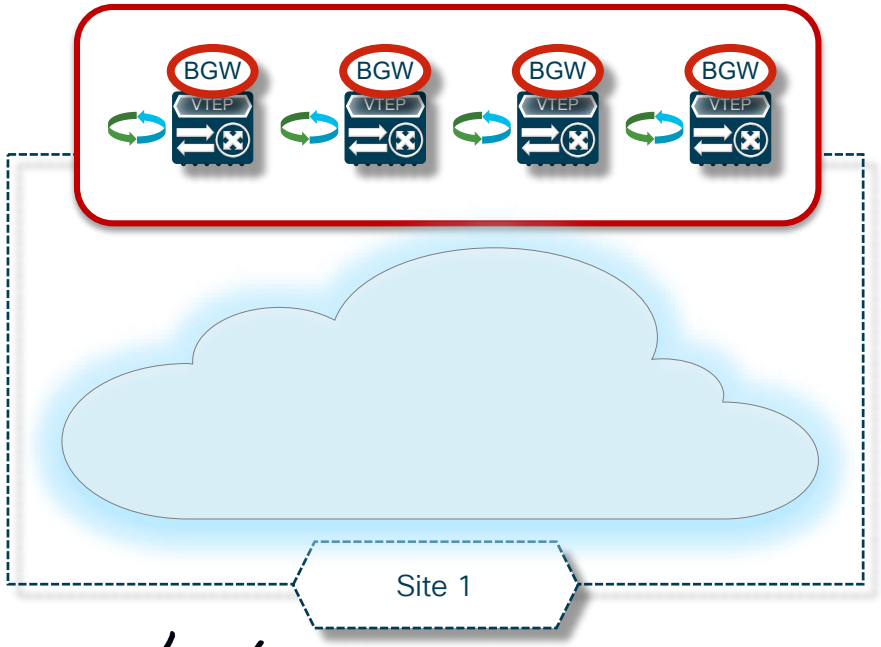
- Border Gateways used for two main functions:
  - Interconnecting each site to the Inter-Site network (for East-West traffic flows)
  - Connecting each site to the external Layer 3 domain (for North-South traffic flows)
  - May also be used to connect endpoints and/or network service nodes (FWs, ADCs)
- Possible deployment models:
  - Anycast Border Gateways
  - vPC Border Gateways
- BGW function enablement in the VXLAN EVPN fabric:
  - BGWs as leaf nodes
  - BGWs as spine nodes (Border-Spines)





Anycast Border Gateways

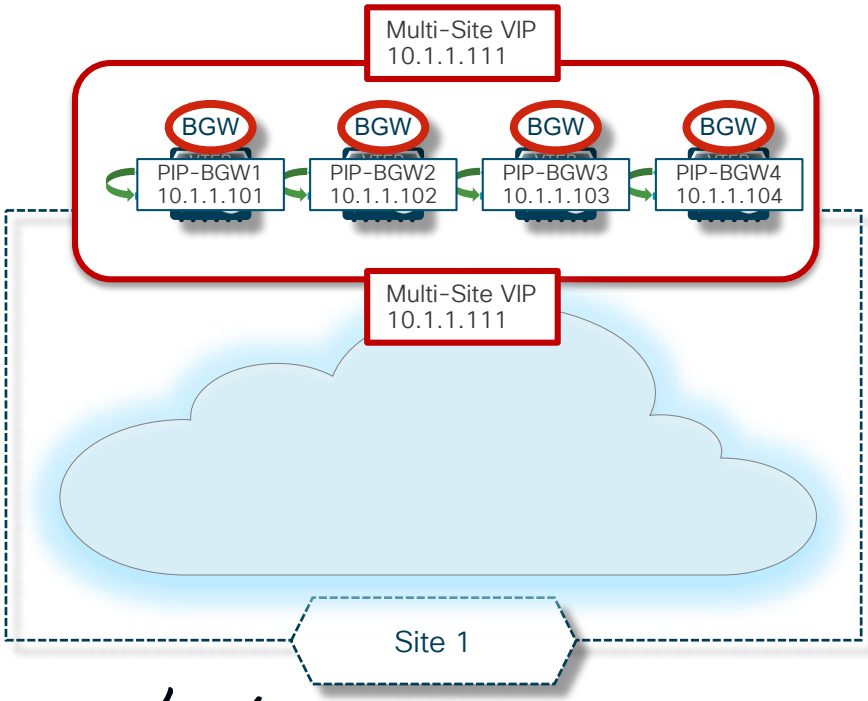
VXLAN Multi-Site
Anycast Border Gateway (1)



Anycast Border Gateway

- Up to 4 Border Gateways
- Border Gateway
  - Deploying at Leaf – 7.0(3)I7(1)
  - Deploying at Spine – 7.0(3)I7(2)

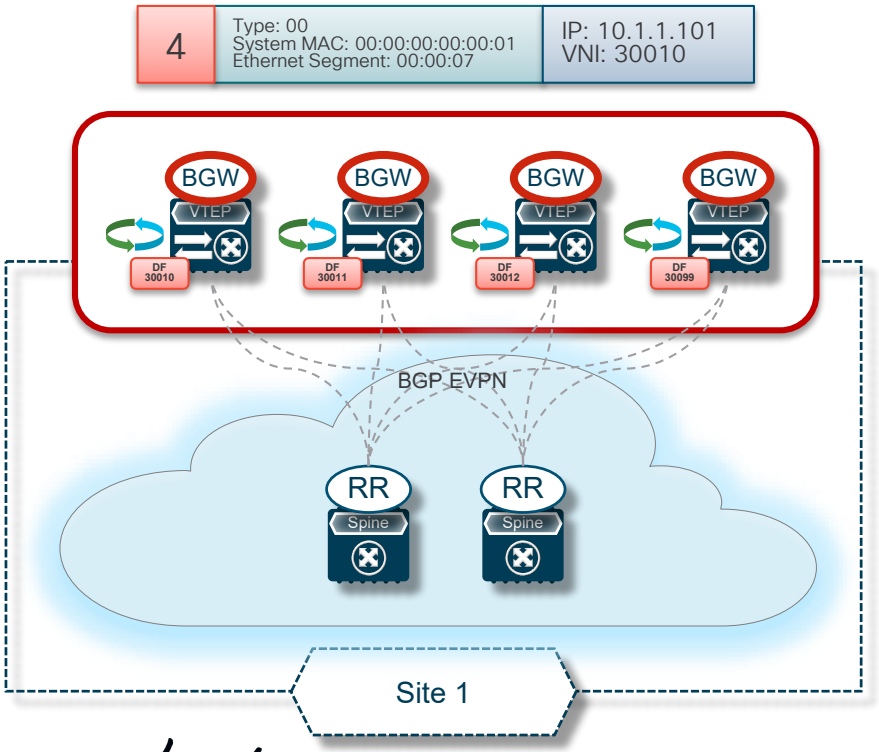
VXLAN Multi-Site
Anycast Border Gateway (2)



Anycast Border Gateway

- Common Multi-Site Virtual IP (Multi-Site VIP) across BGWs
  - Multi-Site VIP for communication between the Border Gateways in **different Sites**
  - Multi-Site VIP for communication between Border Gateways and Leaf nodes **within a Site**
- Individual Primary IP (PIP) per BGW
  - Used for Broadcast, Unknown Unicast and Multicast (BUM) replication
  - PIP for communication with Single-Homed endpoints (routed only), intra- and inter-Site

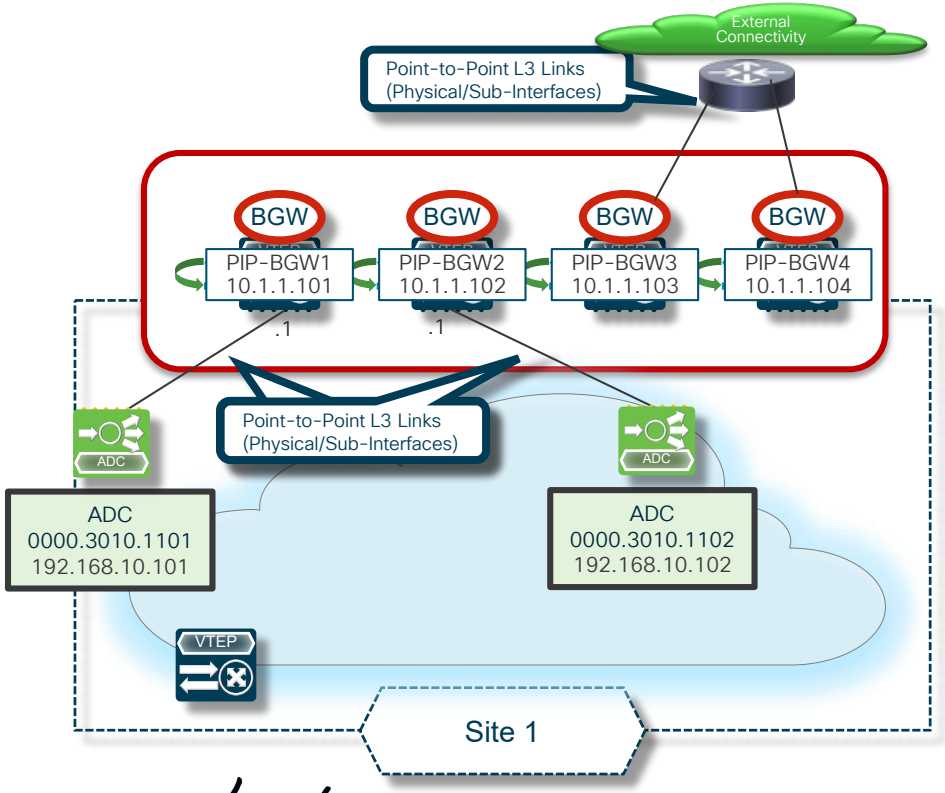
VXLAN Multi-Site
Anycast Border Gateway (3)



Anycast Border Gateway

- Per-VNI Designated Forwarder (DF) election
  - Each BGW can serve as DF for a single or a set of Layer-2 VNIs
  - DF election and assignment is automatic
- Using BGP EVPN Route Type 4 for DF election
  - Operator Managed Assignment (Type: 00)
  - Six Octet Site Identifier (System MAC: 00:00:00:00:00:01)
  - Multi-Site Discriminator (Ethernet-Segment: 00:00:07)
  - Originators IP Address (PIP): 10.1.1.101
  - Layer-2 VNI: 30010

VXLAN Multi-Site
Anycast Border Gateway (4)



Anycast Border Gateway

- Single-Homed End-Points only connected with L3 links
  - Services Appliance (i.e. Firewall, ADC etc.)
  - External routers
  - No SVI support on BGW nodes
- Advertised and Reachable through Individual Primary IP Address (PIP)
  - Intra-Site: Leaf nodes use PIP to reach the device connected to Border Gateways
  - Inter-Site: Remote Border Gateways use PIP to reach the device connected to Border Gateways

vPC Border Gateways

NXOS Release 9.2(1)

Anycast BGW vs. vPC Border Gateway

NXOS Release 9.2(1)

Anycast Border Gateway

- Up to 4 BGW
  - Shared Nothing
  - Simple Failure Scenarios
- Any Deployments
  - No End-Point or Network Services Connectivity on BGW
- Greenfield Deployments

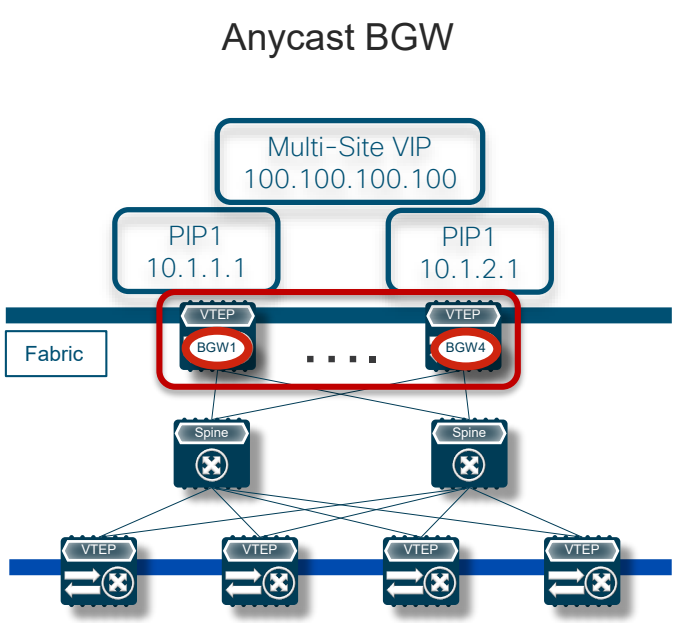
vPC Border Gateway

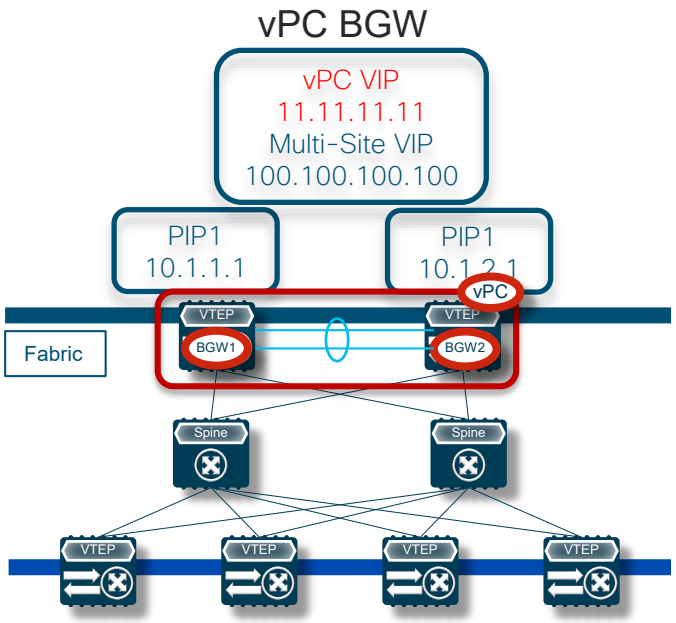
- 2 BGW with physical vPC Peer-Link
- Small Deployments
  - End-Point or Network Services Connectivity on BGW
- Migration Use-Cases (Brownfield)
  - Classic Ethernet/FabricPath to VXLAN EVPN

Multi-Site Border Gateway – Anycast vs. vP

NXOS Release 9.2(1)

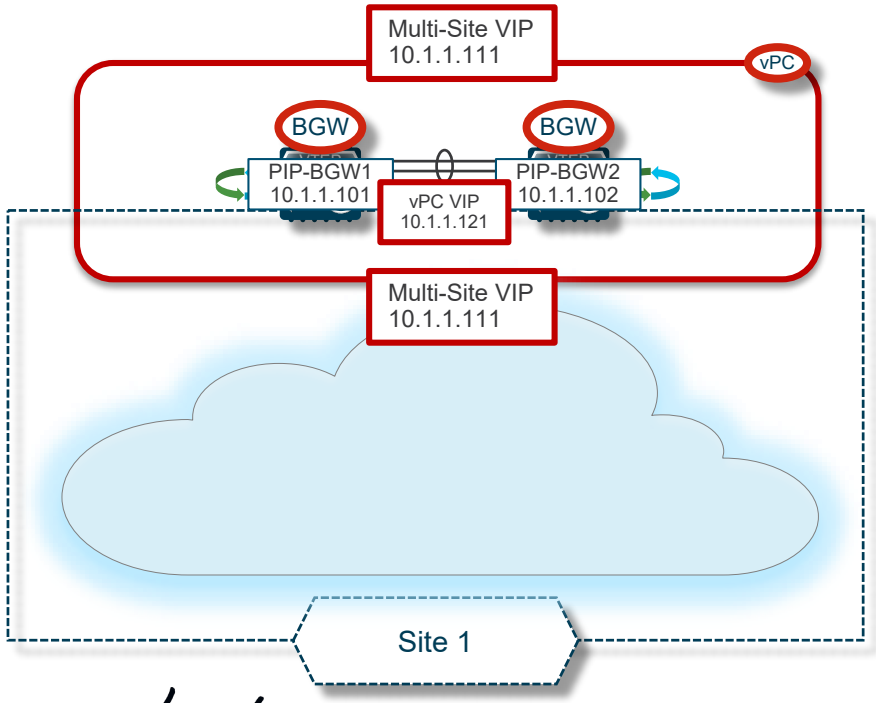
- Both Anycast and vPC Border Gateway needs to be configured with a common Multi-Site VIP address and an individual Primary IP (PIP) address
- vPC Border Gateways share a secondary IP address to be used as vPC virtual IP (vPC VIP)





VXLAN Multi-Site
vPC Border Gateway and Transit Traffic

NXOS Release 9.2(1)



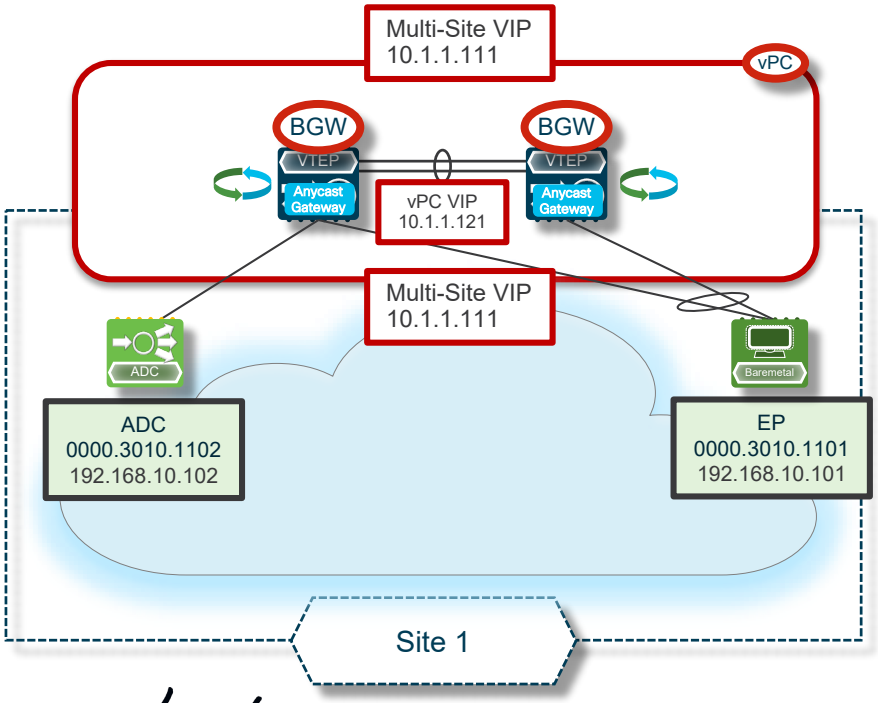
vPC Border Gateway

- Common Multi-Site Virtual IP (Multi-Site VIP) across BGWs
  - Multi-Site VIP for Inter-Site transit communication (transit)
- Common vPC Virtual IP (vPC VIP) across BGWs
  - Used by default for communication with external networks
  - Used for Broadcast, Unknown Unicast and Multicast (BUM) replication
- Individual Primary IP (PIP) per BGW
  - Used for communication with external networks when “advertised-pip” is configured

cisco Live!

VXLAN Multi-Site
vPC Border Gateway and Locally Attached End-Points

NXOS Release 9.2(1)

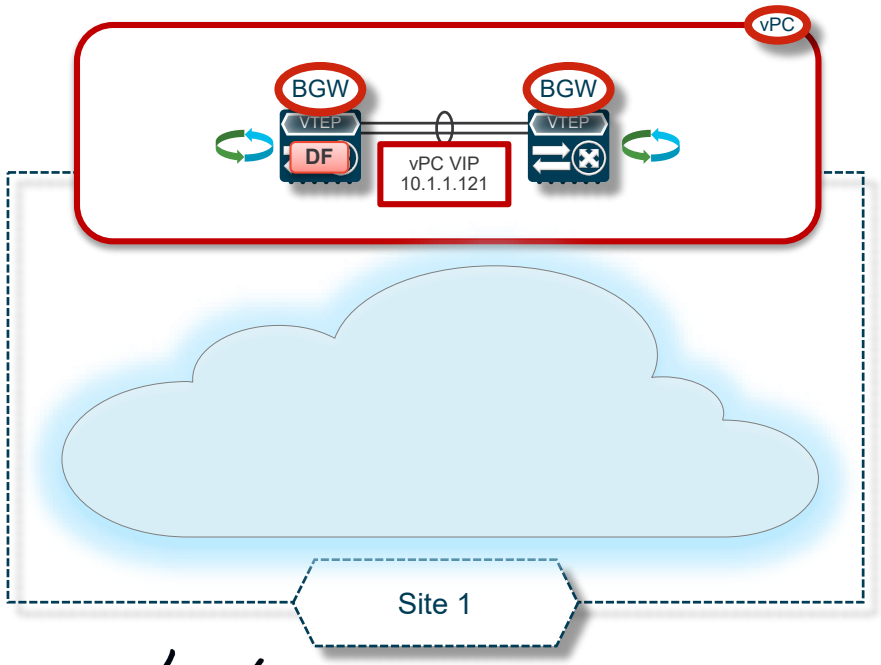


vPC Border Gateway

- Single- or Dual-Homed End-Points
  - Services Appliance (i.e. Firewall, ADC etc.)
  - Physical or Virtual Servers
  - Anycast Gateway function offered to the endpoints
- Advertised and Reachable through vPC Virtual IP Address (vPC VIP)
  - Intra-Site: Leaf nodes use vPC VIP to reach End-Points connected to Border Gateways
  - Inter-Site: Remote Border Gateways use vPC VIP to reach End-Points connected to Border Gateways
  - Traffic potentially traverses vPC Peer-Link

VXLAN Multi-Site
vPC Border Gateway and Designated BUM Forwarder

NXOS Release 9.2(1)



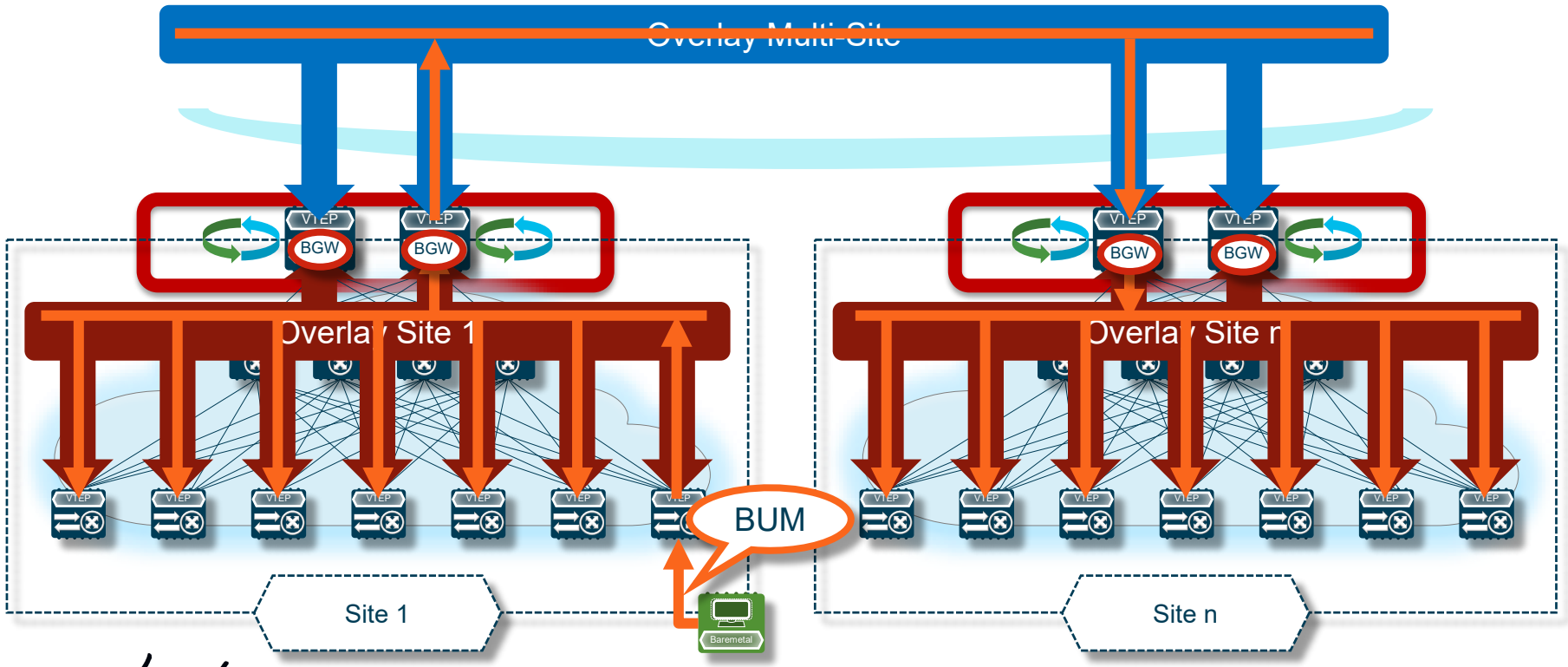
vPC Border Gateway

- vPC-based Designated Forwarder Election
- Per-Site Designated Forwarder (DF) election
  - Using same approach as in vPC
  - Best Path to Rendezvous-Point or vPC Primary Node
  - Same vPC node is elected DF for **all** the Layer-2 VNIs

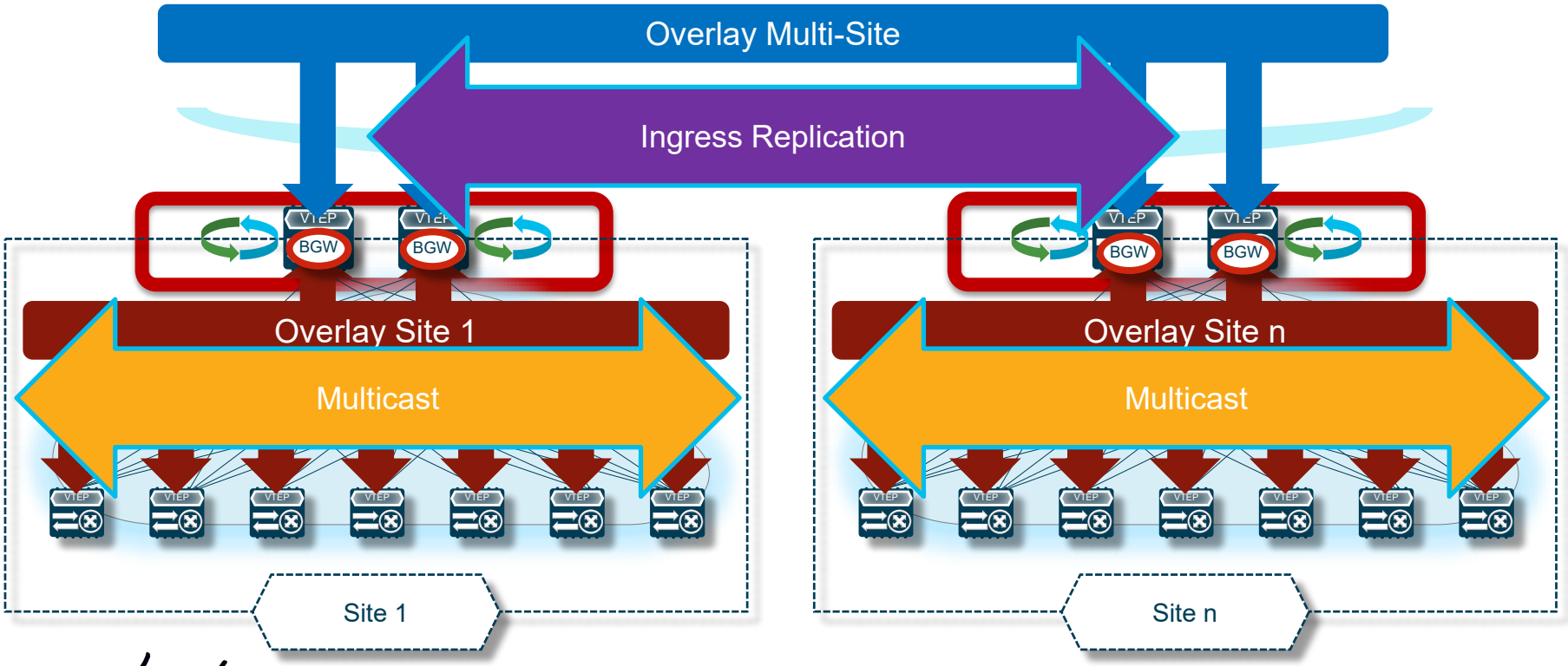
cisco Live!

Inter-Site BUM Traffic Handling

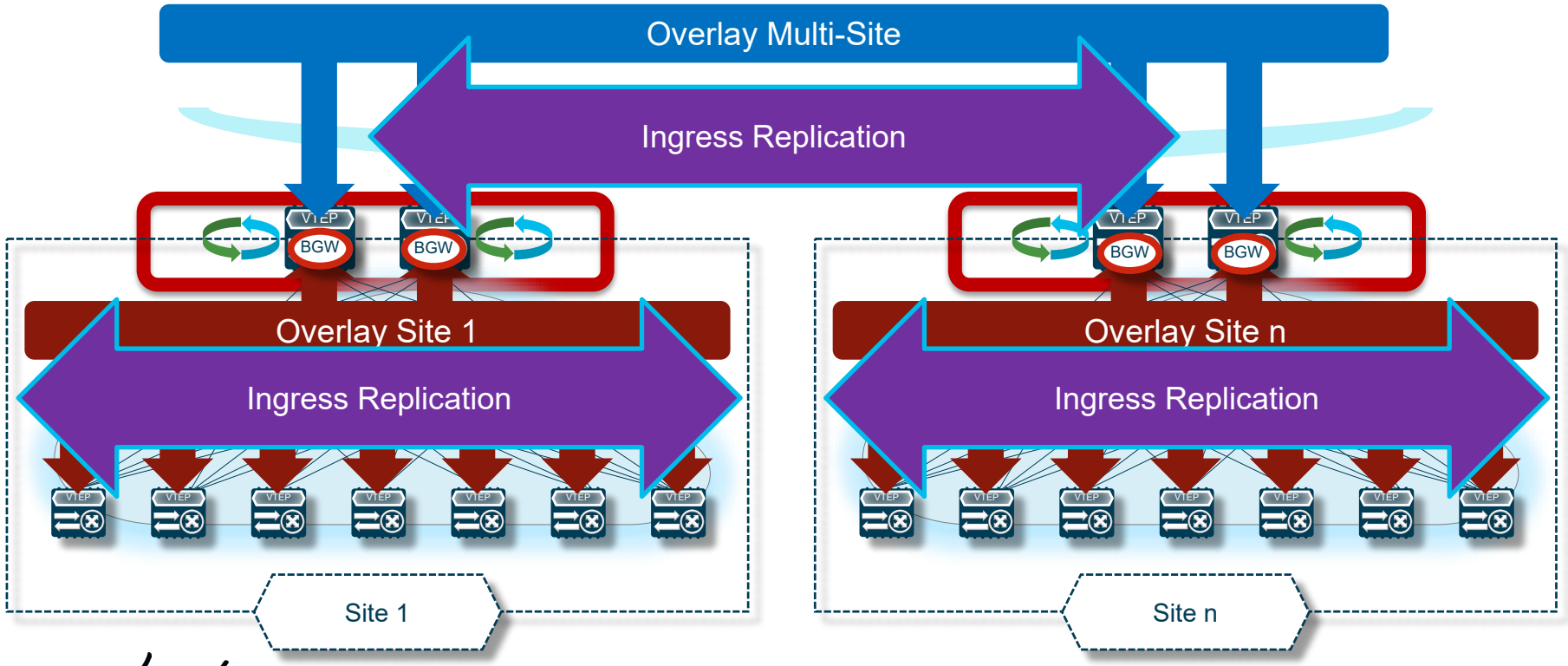
VXLAN Multi-Site
BUM Traffic Forwarding



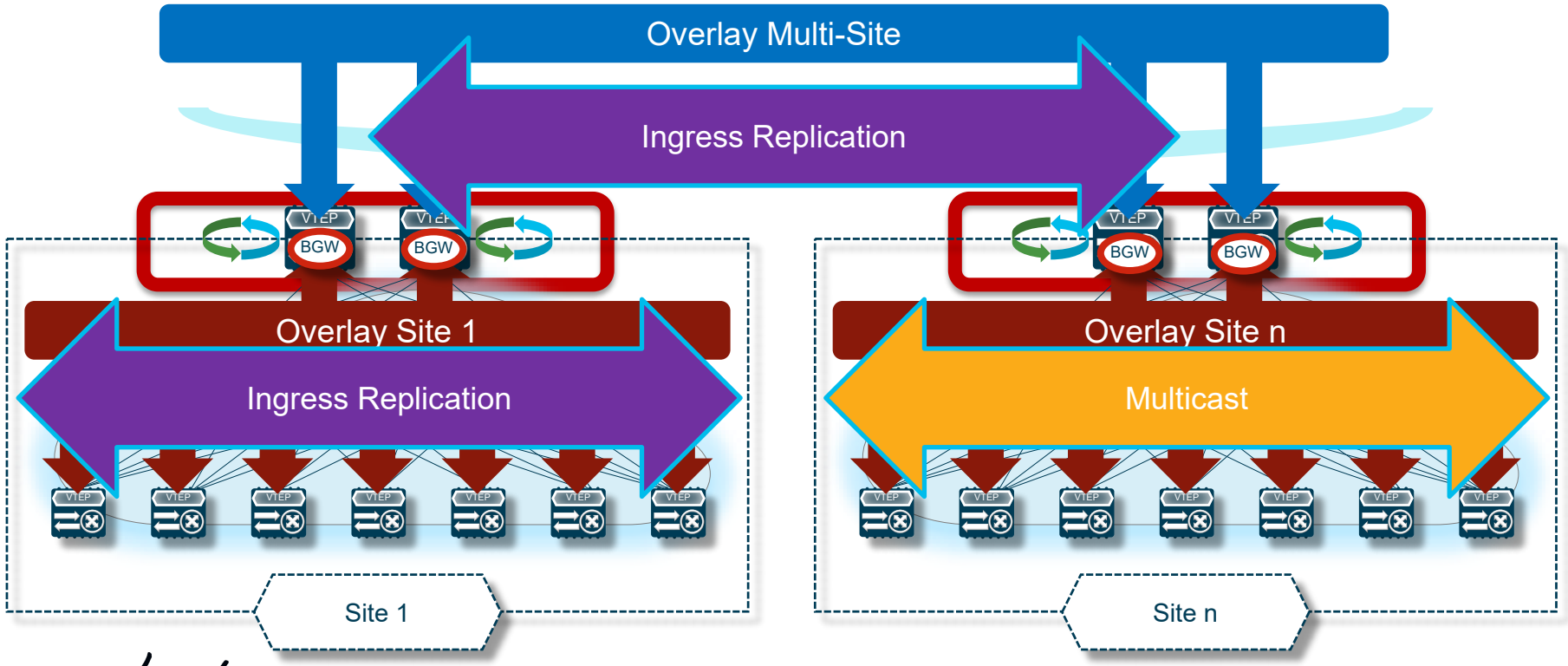
VXLAN Multi-Site
BUM Replication Modes (Multicast Intra-Site)



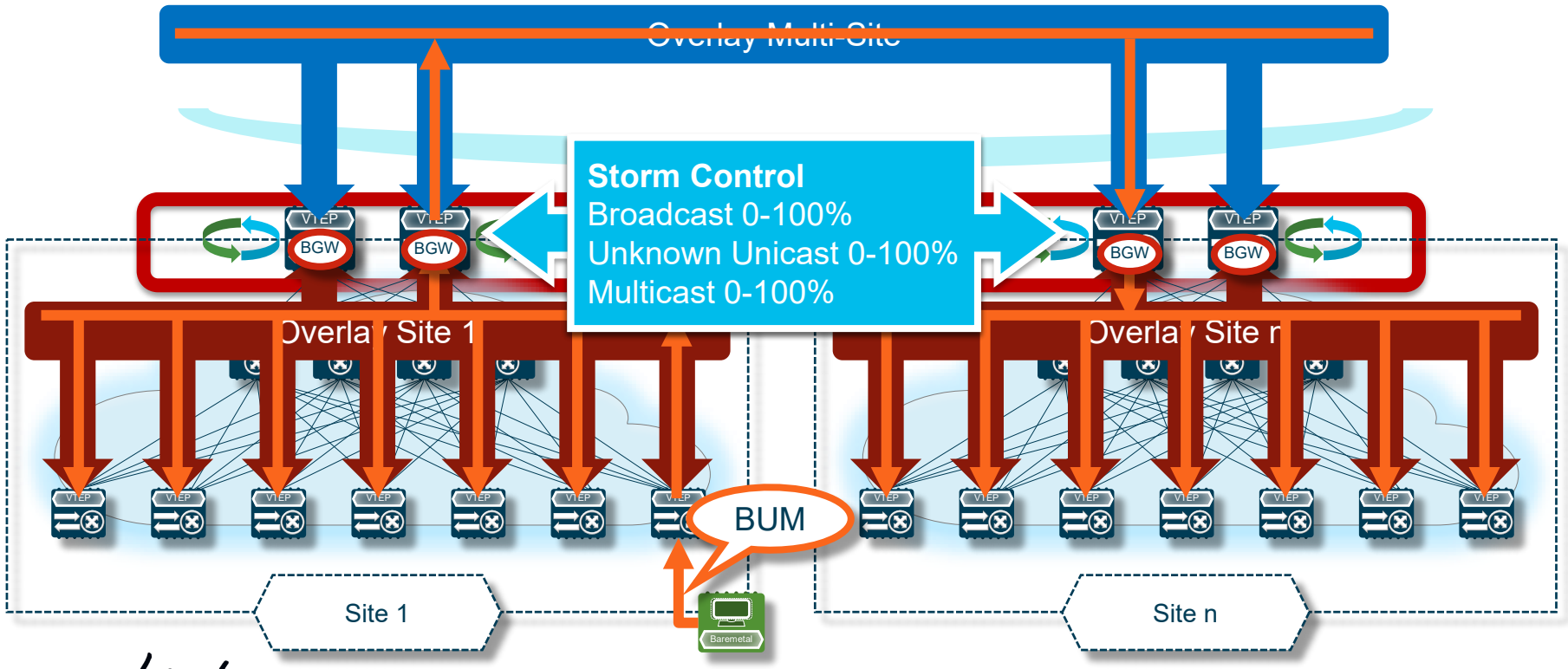
VXLAN Multi-Site
BUM Replication Modes (Ingress Replication Only)



VXLAN Multi-Site
BUM Replication Modes (Mixed Mode Intra-Site)



VXLAN Multi-Site
BUM Traffic Policing



Control and Data Planes

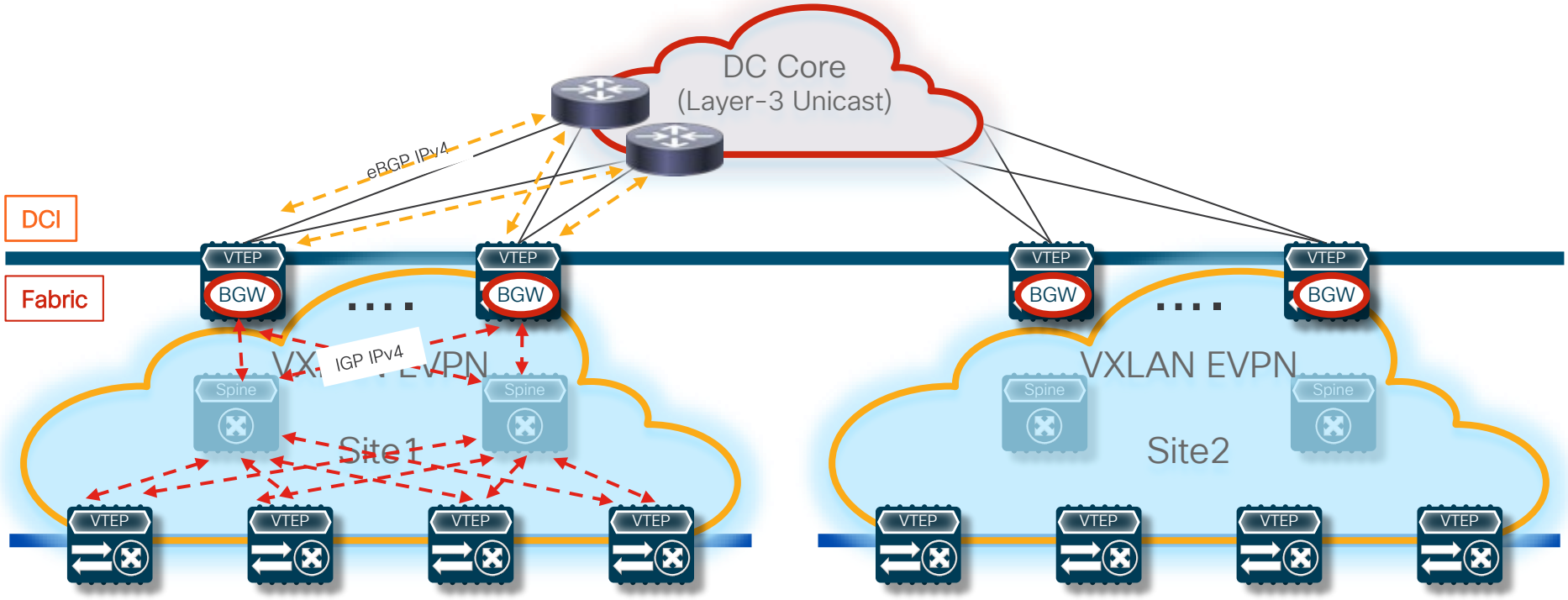
Multi-Site Control Plane

VXLAN Multi-Site
Control Plane Deployment Considerations

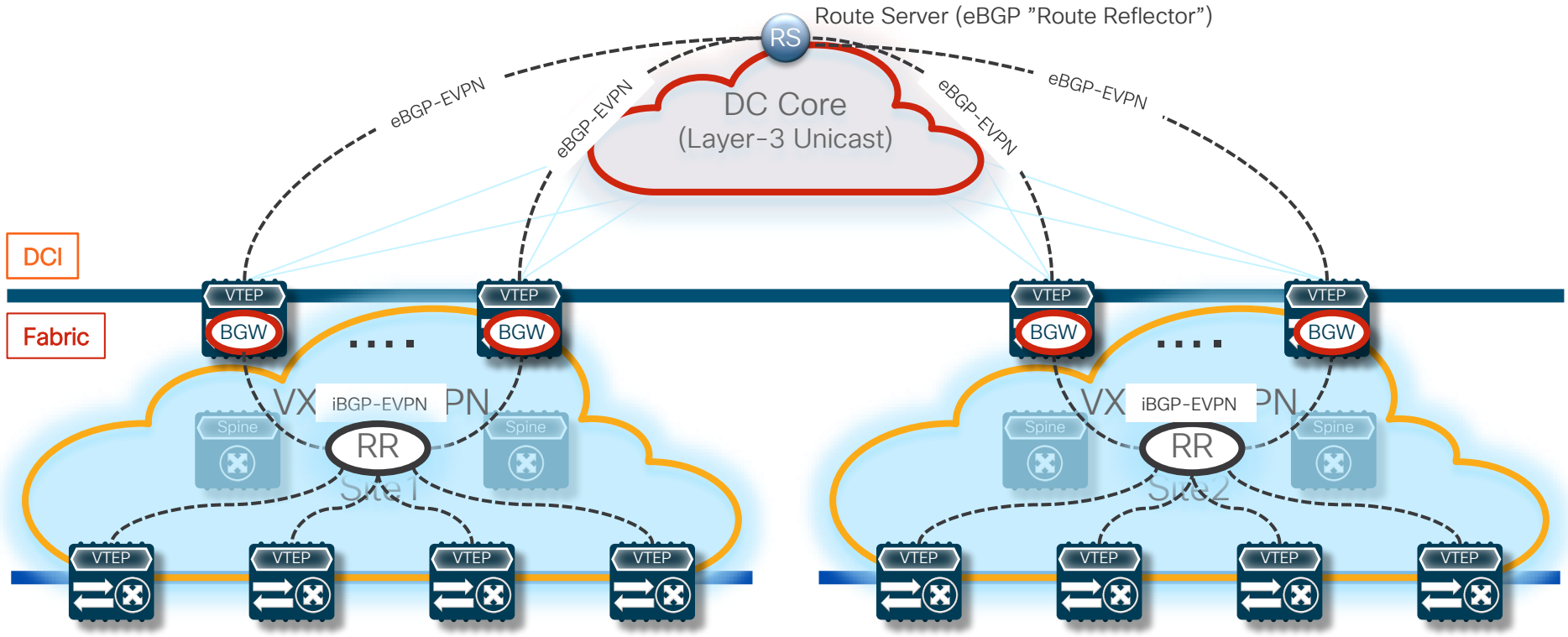
- MP-eBGP EVPN only inter-Sites
  - Next-hop behavior (VXLAN tunnel termination and reorigination) and loop protection (as-path attribute)
- Two main options for underlay and overlay control plane deployment
  1. I-E-I (Recommended)
     - Intra-Site: IGP (OSPF, IS-IS) as underlay CP, iBGP as overlay CP
     - Inter-Sites: eBGP for both underlay and overlay CPs
  2. E-E-E*
     - Intra-Site and Inter-Sites: eBGP for both underlay and overlay CPs
- Full mesh of MP-eBGP EVPN adjacencies across sites
  - Recommended to deploy a couple of **Route-Servers** with 3 or more sites
  - RS in a separate AS only perform control plane functions (“eBGP Route-Reflectors”, IETF RFC 7947)
  - RS functions: EVPN routes reflection, next-hop-unchanged, route-target rewrite

*For more information on why eBGP for both underlay and overlay CP is not a good idea:
https://learningnetwork.cisco.com/blogs/community_cafe/2017/10/17/the-magic-of-super-spines-and-rfc7938-with-overlays-guest-post

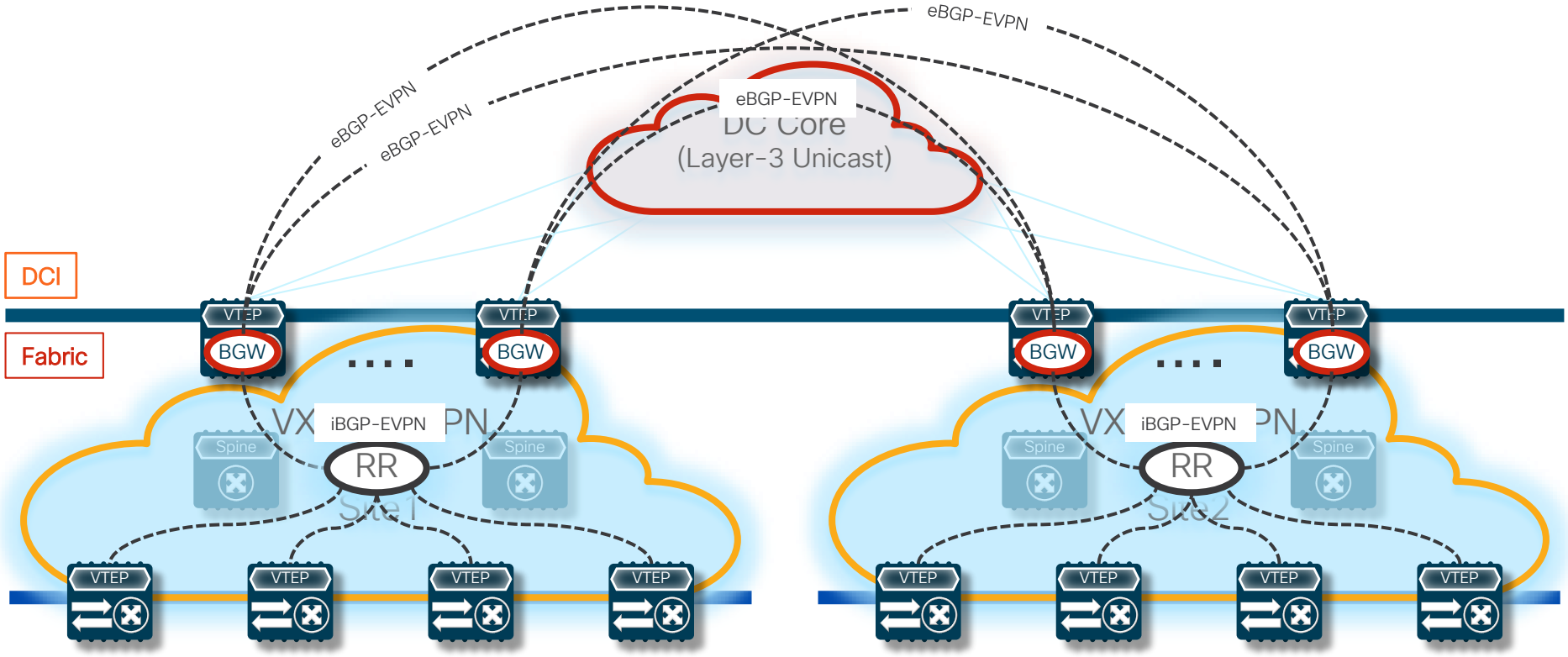
VXLAN Multi-Site
Underlay Control Plane



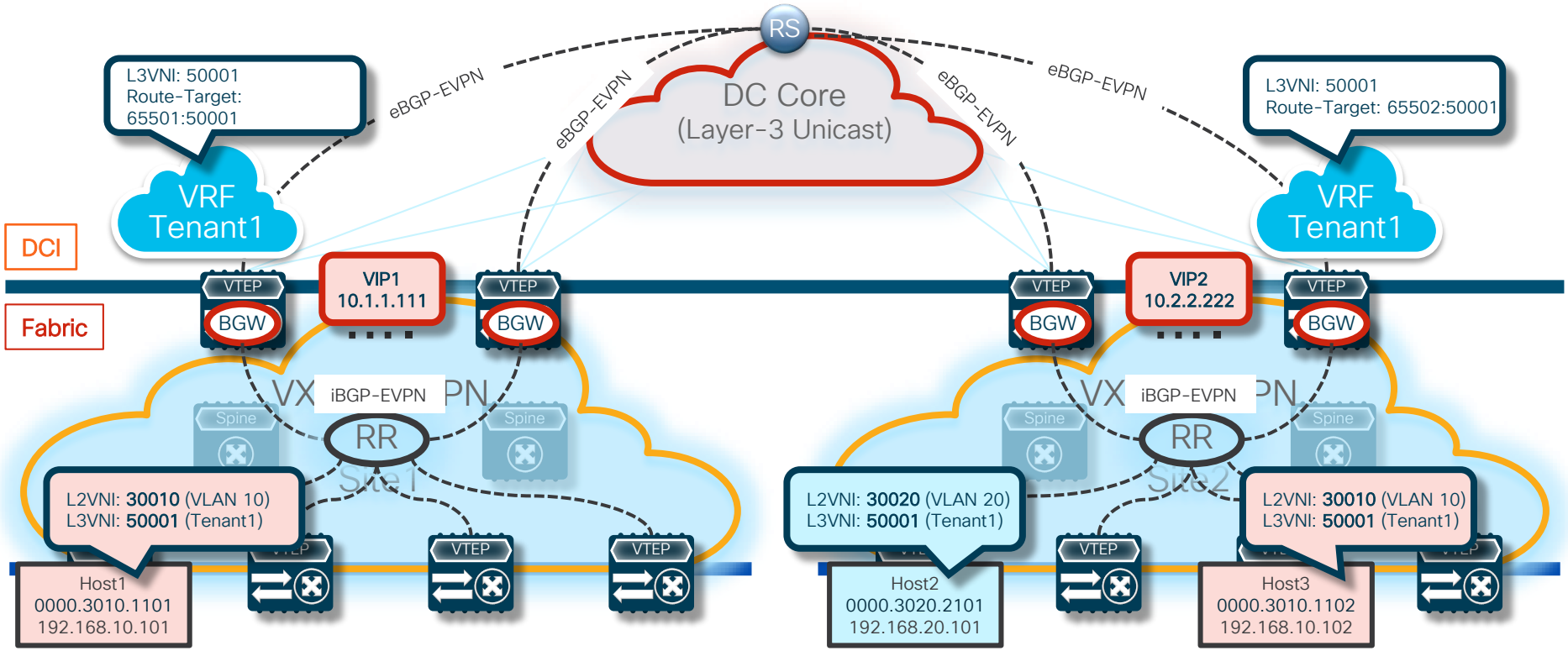
VXLAN Multi-Site
Overlay Control Plane (L3 Core)



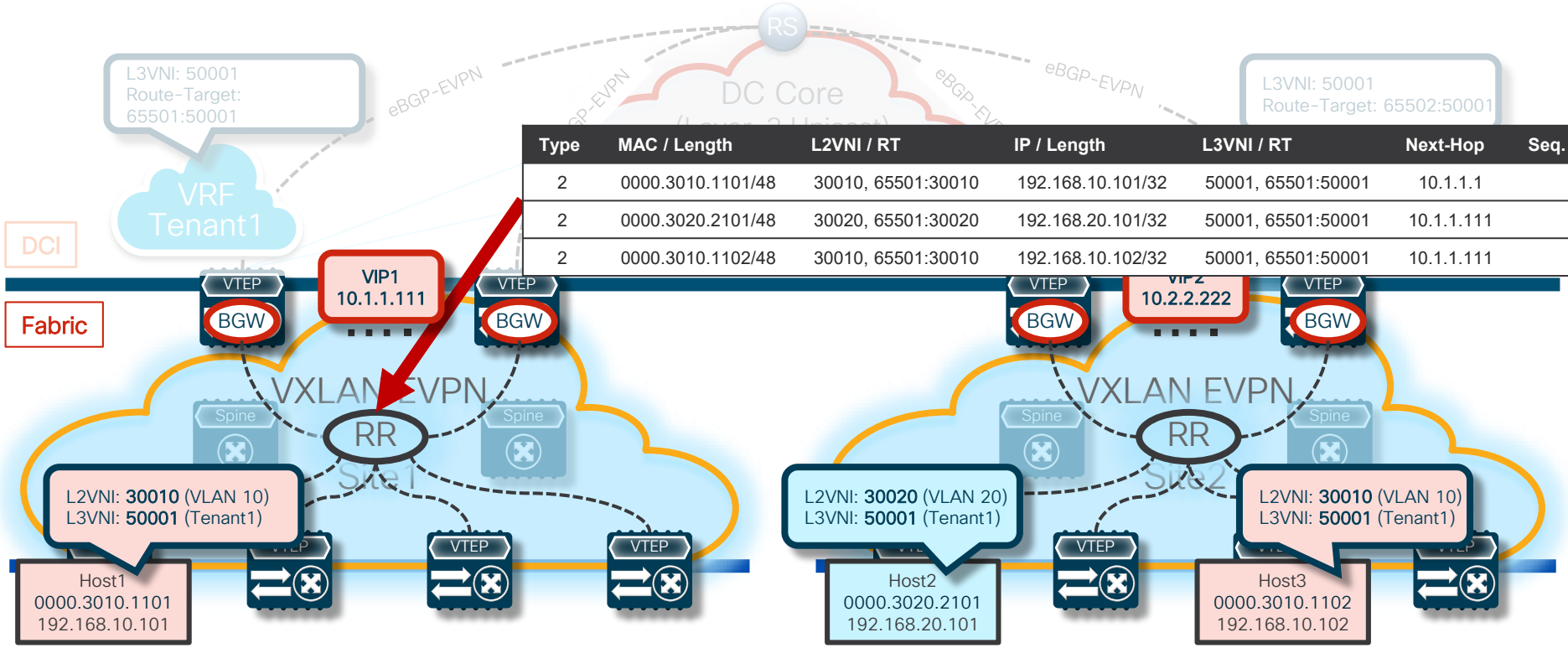
VXLAN Multi-Site
Overlay Control Plane (L3 Core, no RS)



VXLAN Multi-Site
Overlay Control Plane



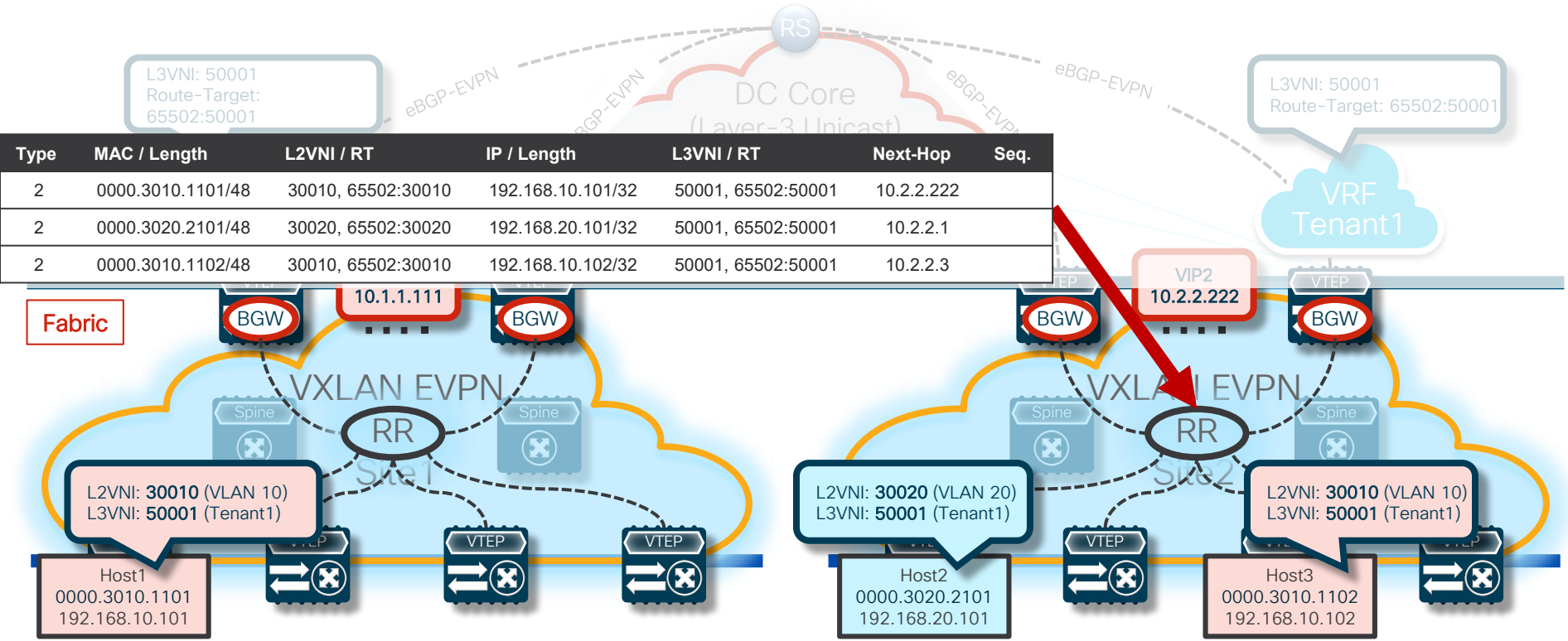
VXLAN Multi-Site
Overlay Control Plane (Site 1)



Type	MAC / Length	L2VNI / RT	IP / Length	L3VNI / RT	Next-Hop	Seq.
2	0000.3010.1101/48	30010, 65501:30010	192.168.10.101/32	50001, 65501:50001	10.1.1.1	
2	0000.3020.2101/48	30020, 65501:30020	192.168.20.101/32	50001, 65501:50001	10.1.1.111	
2	0000.3010.1102/48	30010, 65501:30010	192.168.10.102/32	50001, 65501:50001	10.1.1.111	

VXLAN Multi-Site
Overlay Control Plane (Site 2)

Type	MAC / Length	L2VNI / RT	IP / Length	L3VNI / RT	Next-Hop	Seq.
2	0000.3010.1101/48	30010, 65502:30010	192.168.10.101/32	50001, 65502:50001	10.2.2.222	
2	0000.3020.2101/48	30020, 65502:30020	192.168.20.101/32	50001, 65502:50001	10.2.2.1	
2	0000.3010.1102/48	30010, 65502:30010	192.168.10.102/32	50001, 65502:50001	10.2.2.3	

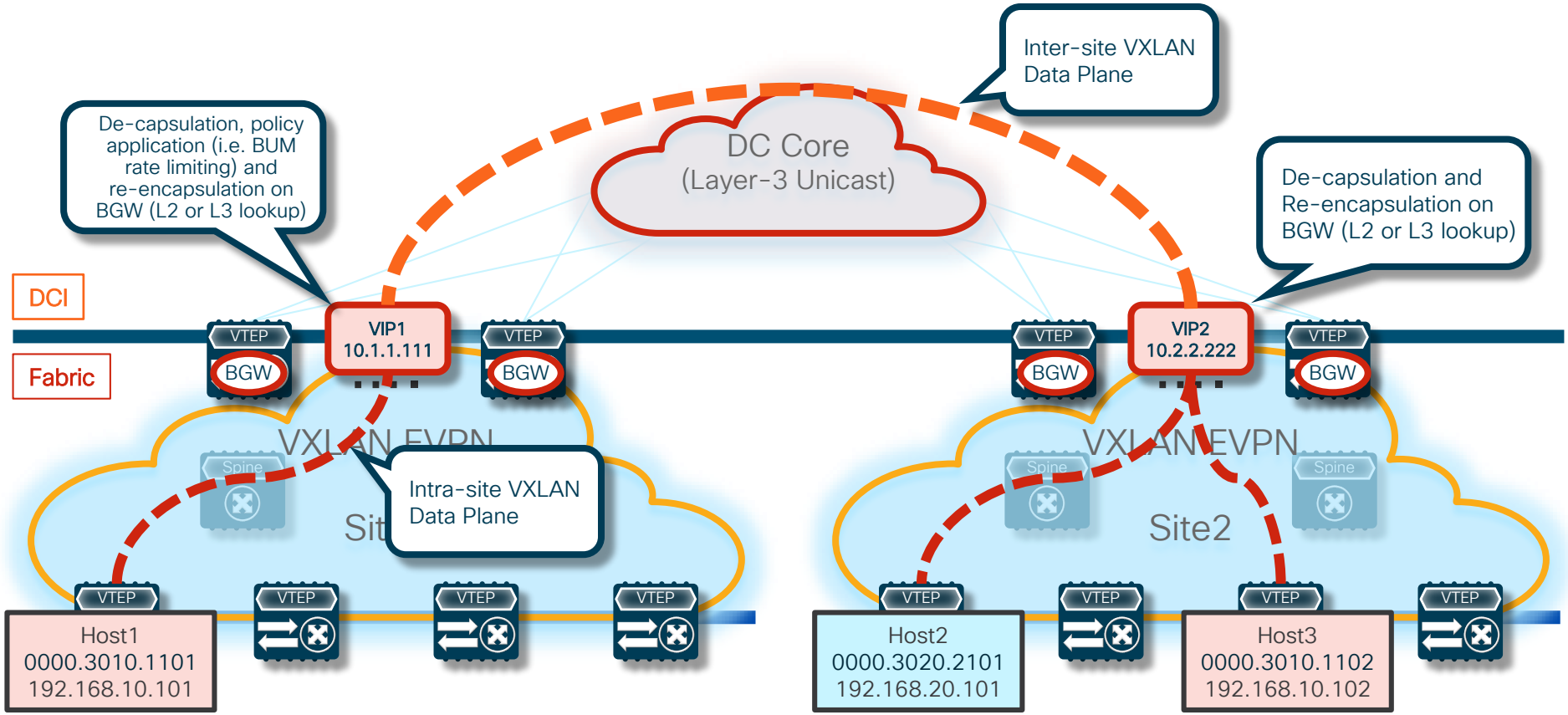


Multi-Site – Selective Advertisements

- The Multi-Site architecture provides granular control on how Layer-2 and Layer-3 communication is extended across sites
- Layer-2 and/or Layer-3 VNIs configured on the Border Gateways (BGW) control the Control-Plane advertisement towards DCI
- Enhances the overall scalability of the solution
  - Scale up the total number of End-Points supported across sites

Multi-Site Data Plane

VXLAN Multi-Site
Overlay Data Plane



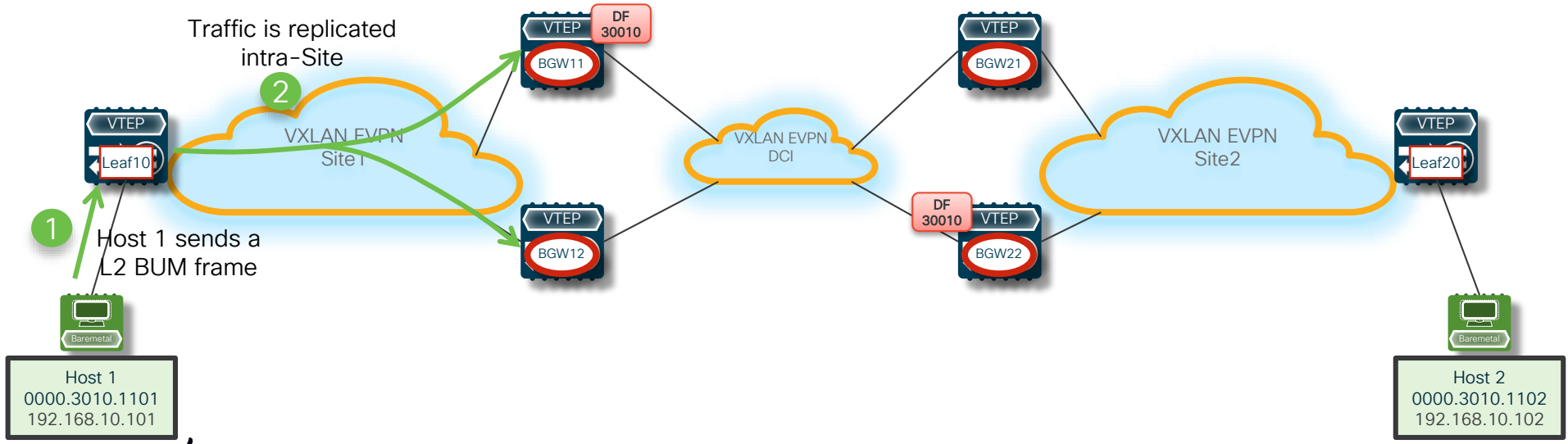
Multi-Site Packet Walk (BUM)

VXLAN Multi-Site Packet Walk

Layer 2 (BUM) – Site 1

BUM Forwarding

SIP	DIP	VXLAN	SMAC	DMAC	SIP	DIP	Payload
L10	DGROUP	30010	H1-MAC	ALL-F	H1-IP	ALL-255	

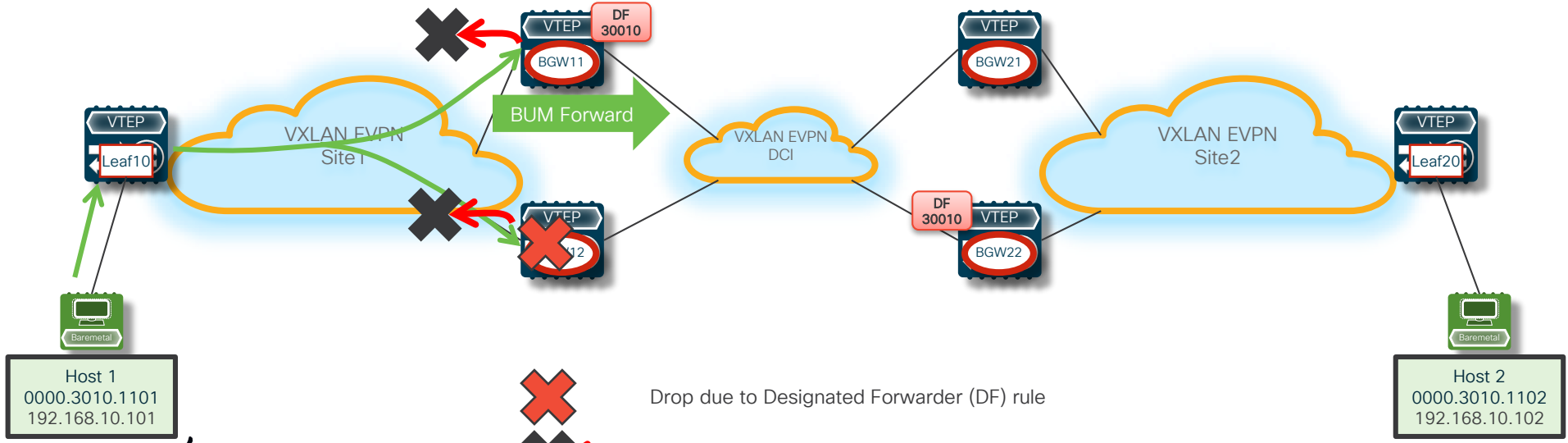


VXLAN Multi-Site Packet Walk

Layer 2 (DF and Split Horizon) – Site 1

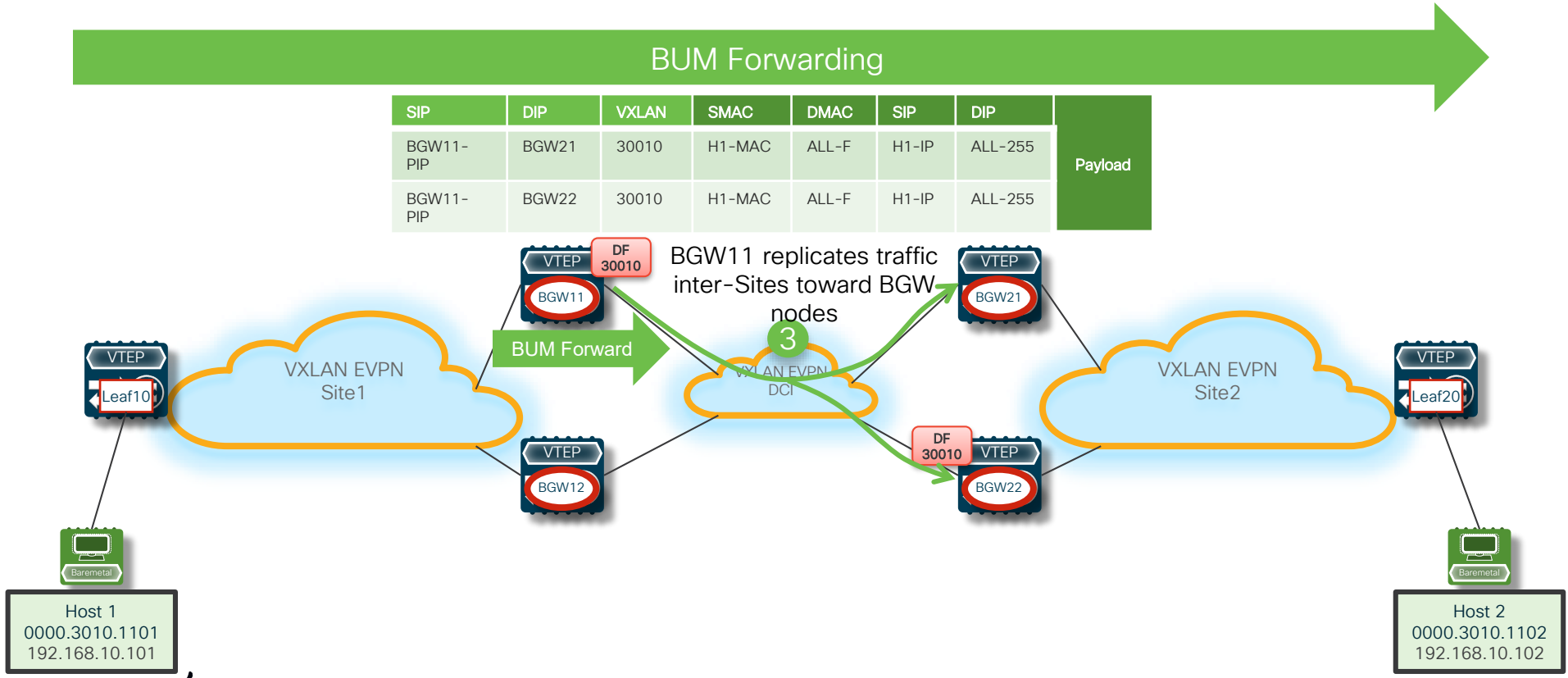
BUM Forwarding

SIP	DIP	VXLAN	SMAC	DMAC	SIP	DIP	Payload
L10	DGROUP	30010	H1-MAC	ALL-F	H1-IP	ALL-255	



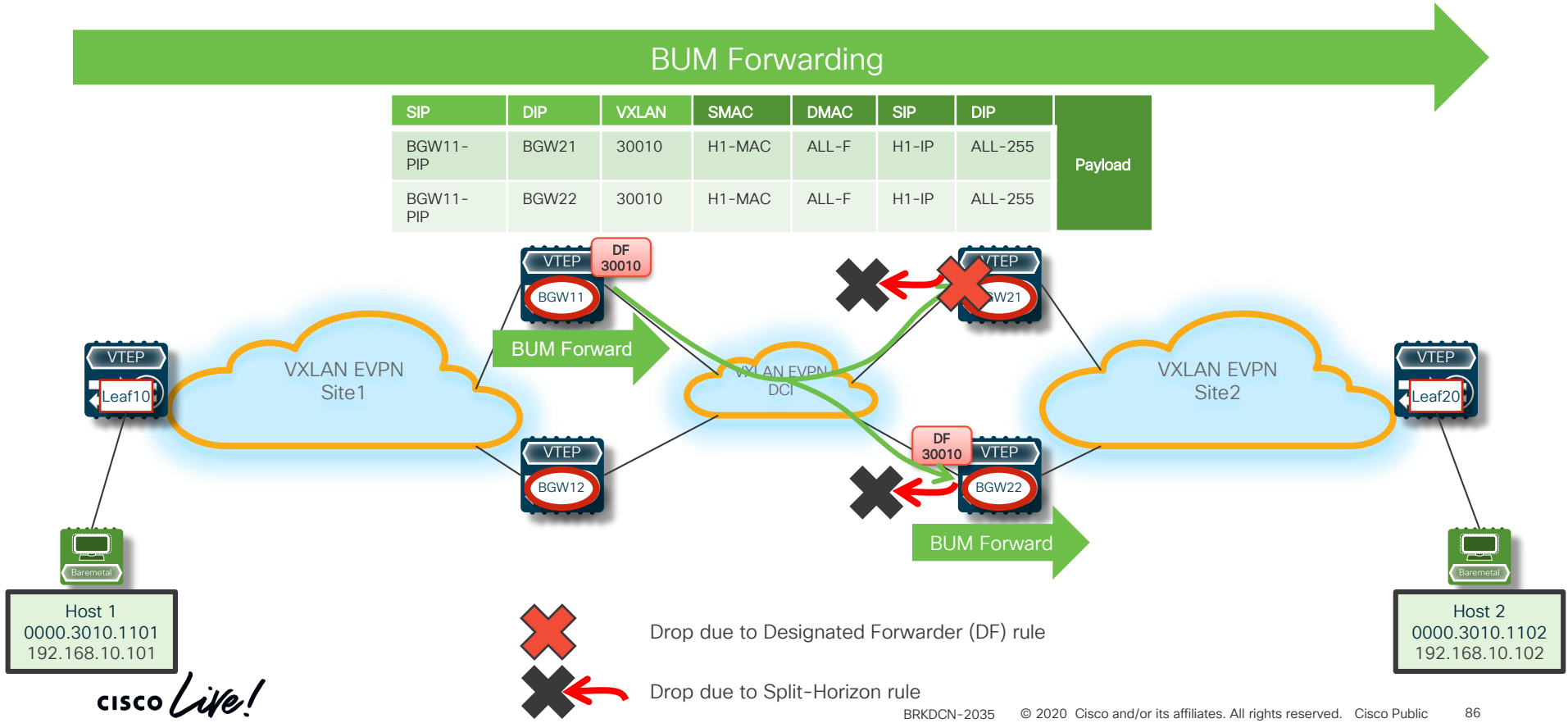
VXLAN Multi-Site Packet Walk

Layer 2 (BUM) – DCI



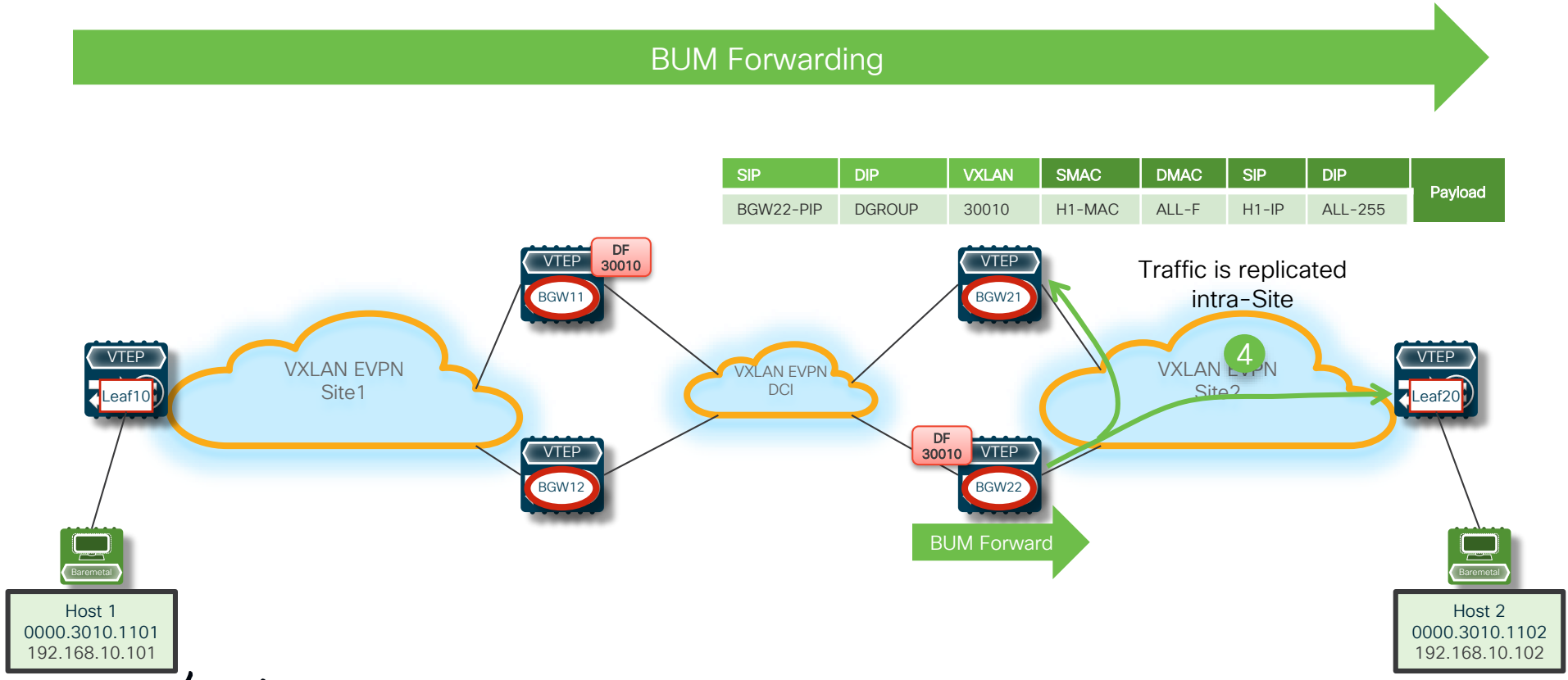
VXLAN Multi-Site Packet Walk

Layer 2 (DF and Split Horizon) – DCI



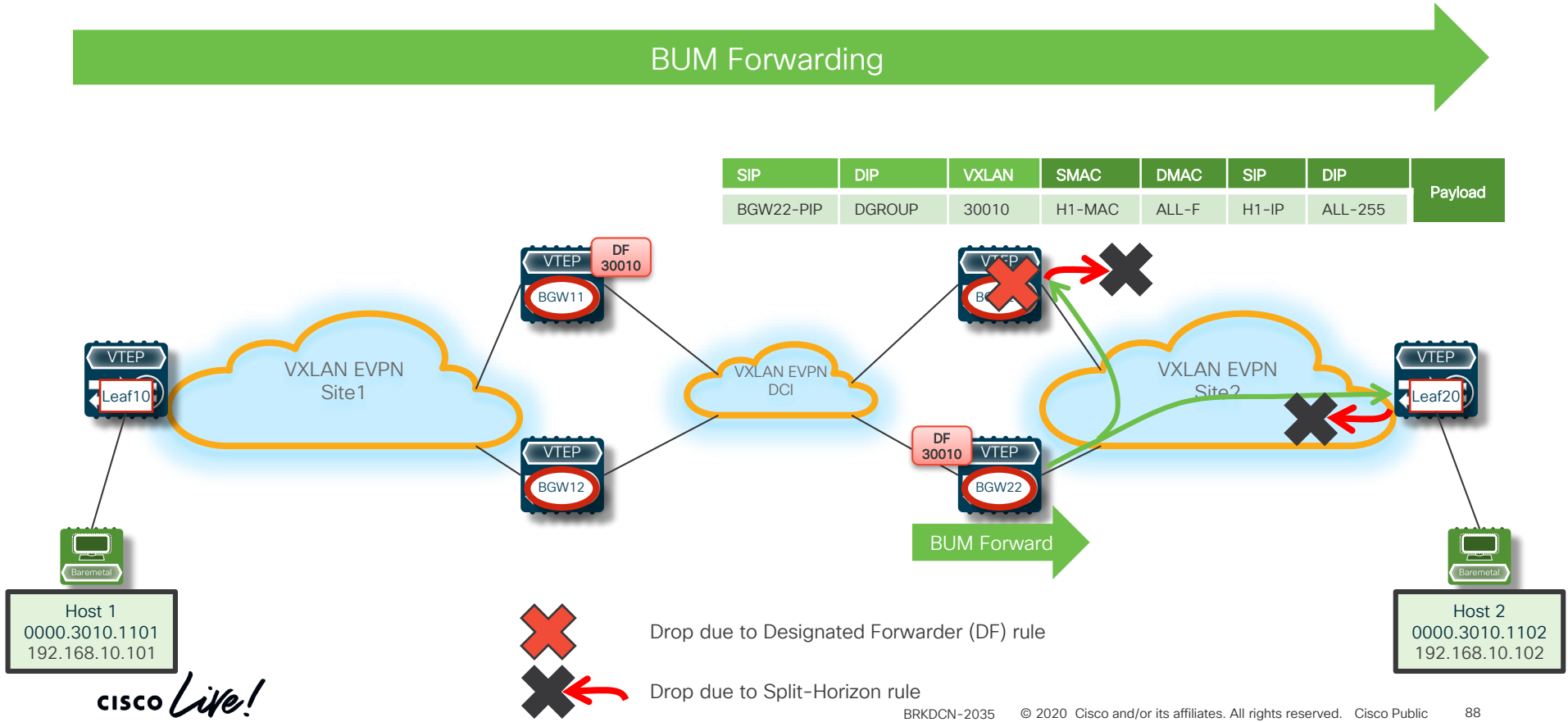
VXLAN Multi-Site Packet Walk

Layer 2 (BUM) – Site 2



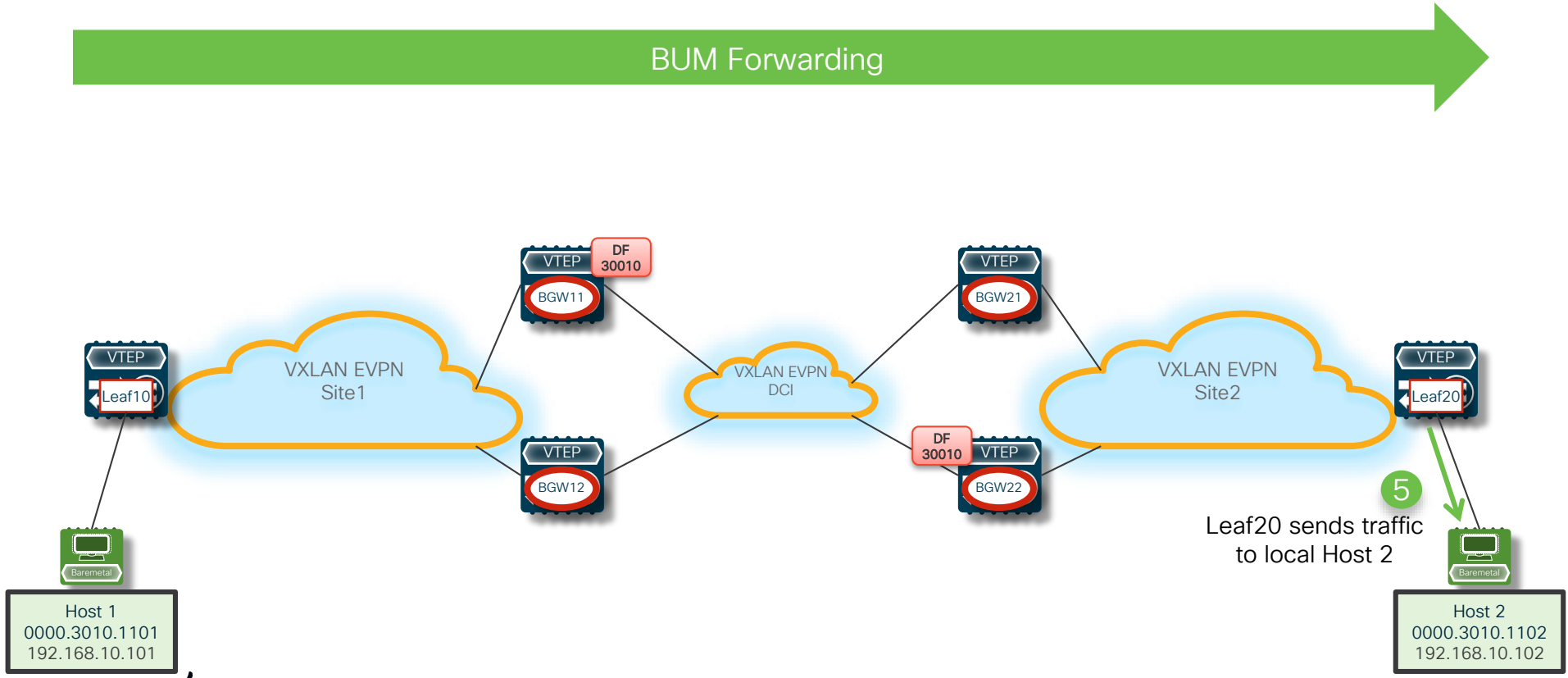
VXLAN Multi-Site Packet Walk

Layer 2 (DF and Split Horizon) – Site 2



VXLAN Multi-Site Packet Walk

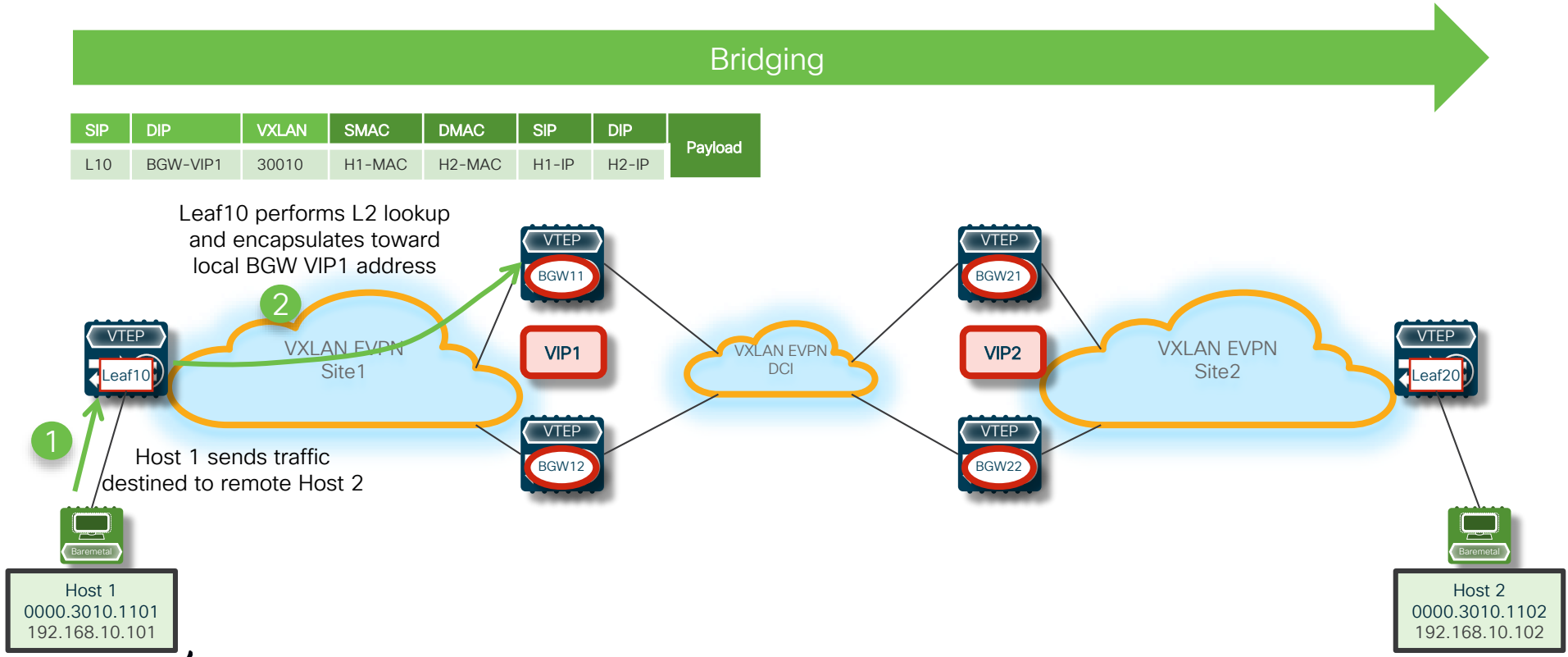
Layer 2 (BUM) – Site 2



Multi-Site Packet Walk (Bridging)

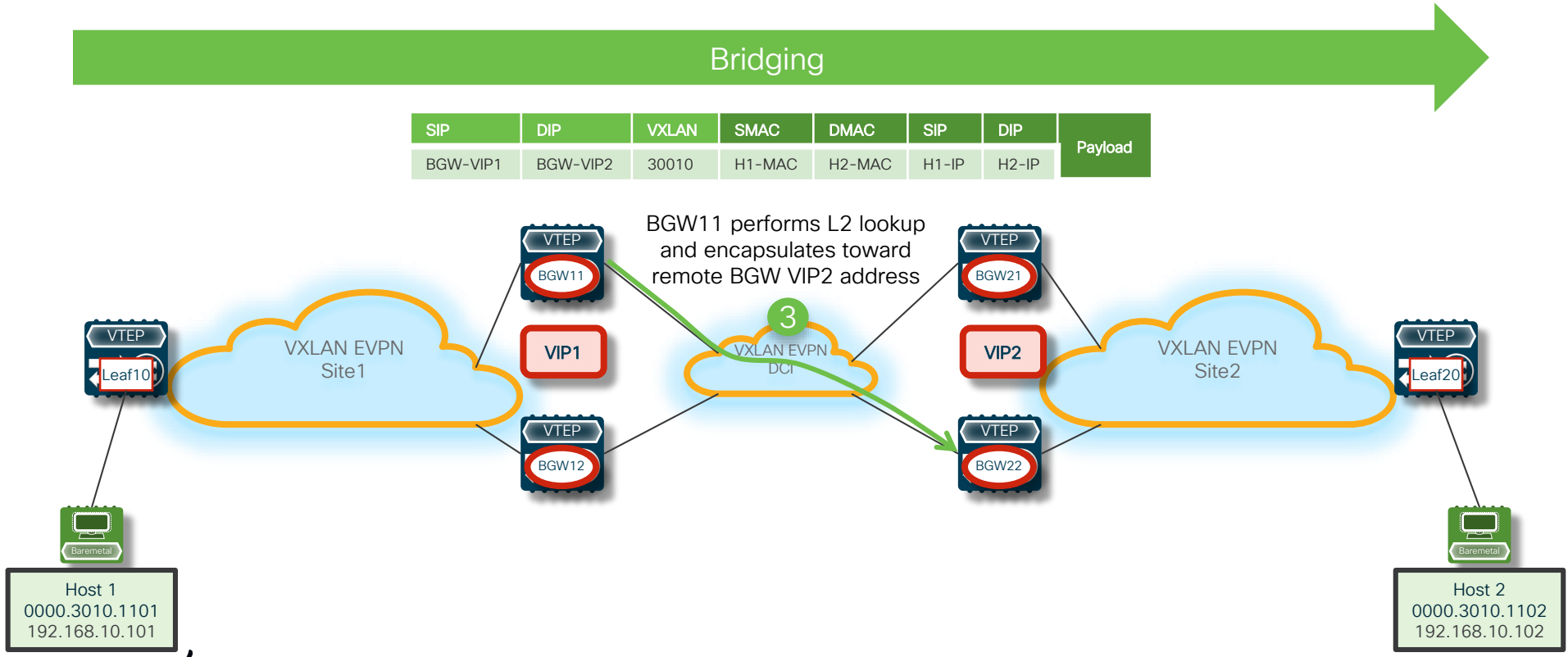
VXLAN Multi-Site Packet Walk

Layer 2 (Host 1 to Host 2) – Site 1



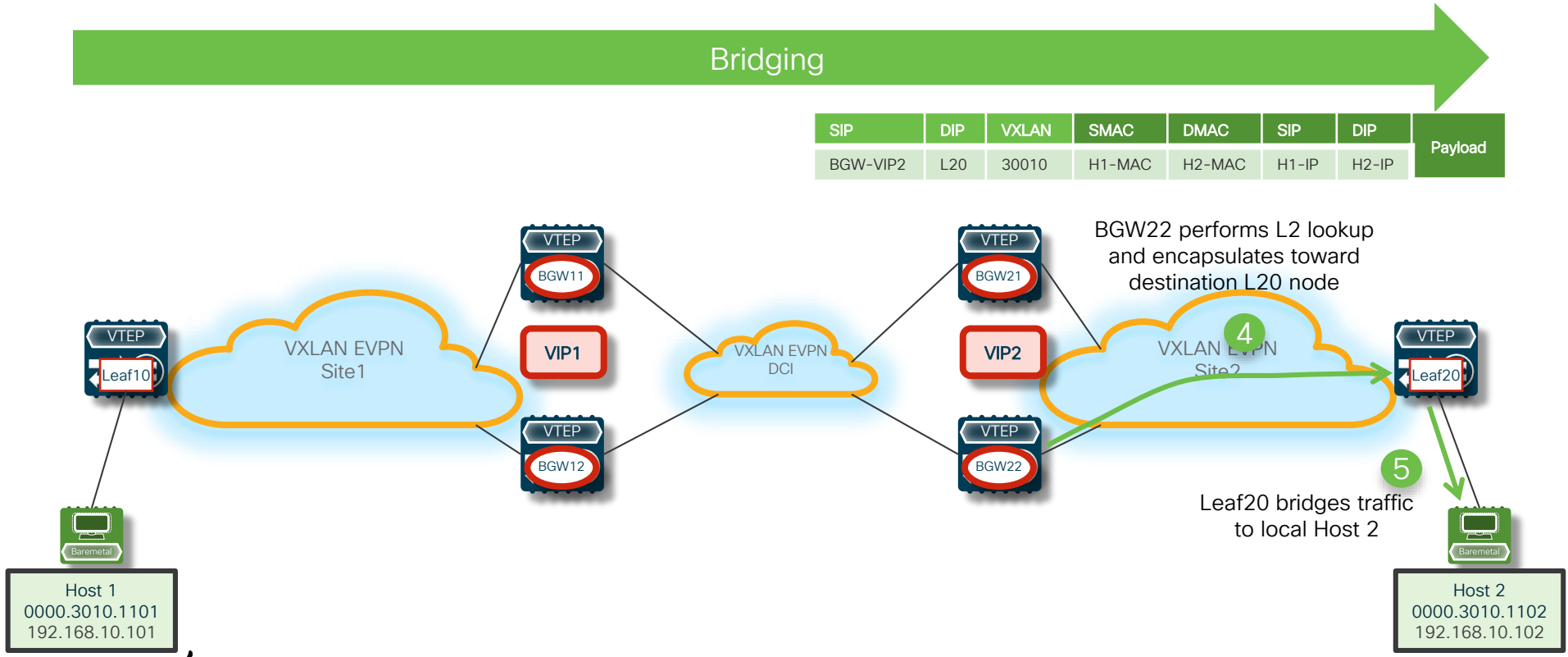
VXLAN Multi-Site Packet Walk

Layer 2 (Host 1 to Host 2) – DCI



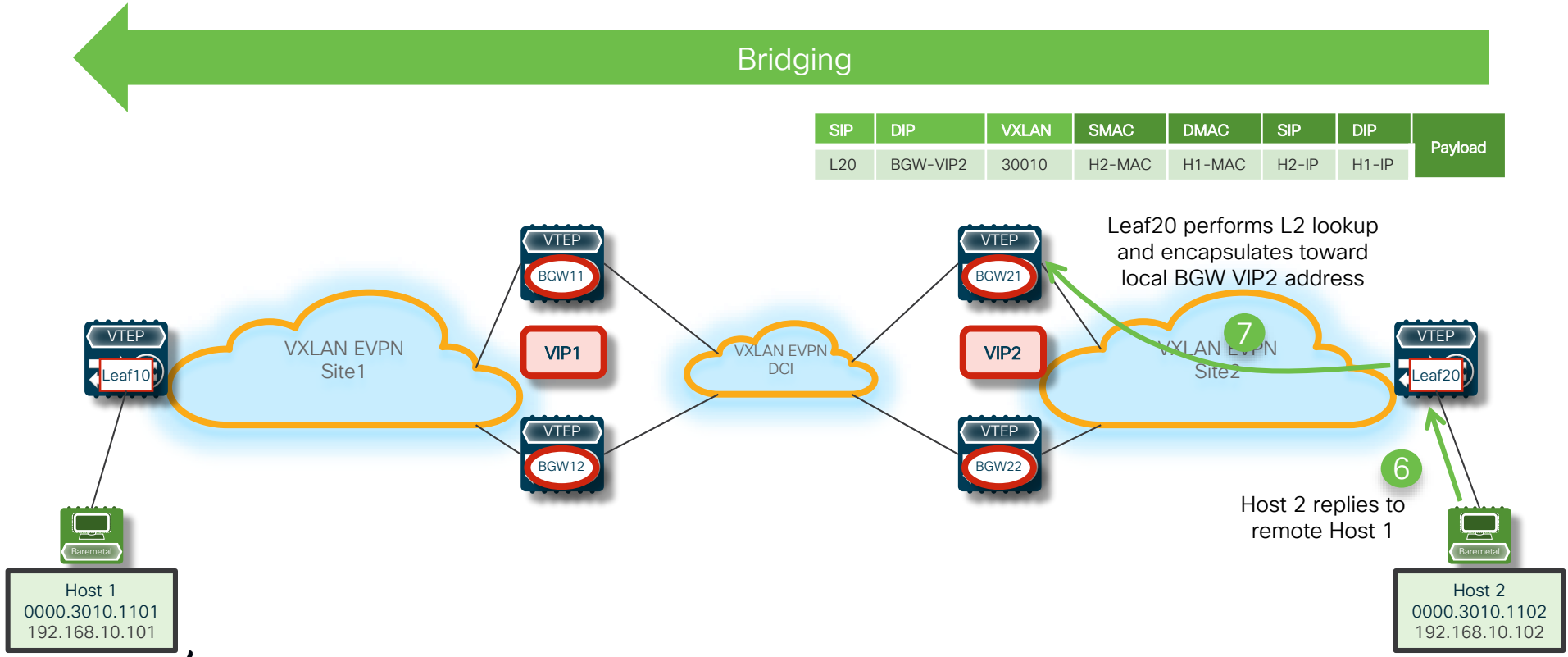
VXLAN Multi-Site Packet Walk

Layer 2 (Host 1 to Host 2) – Site 2



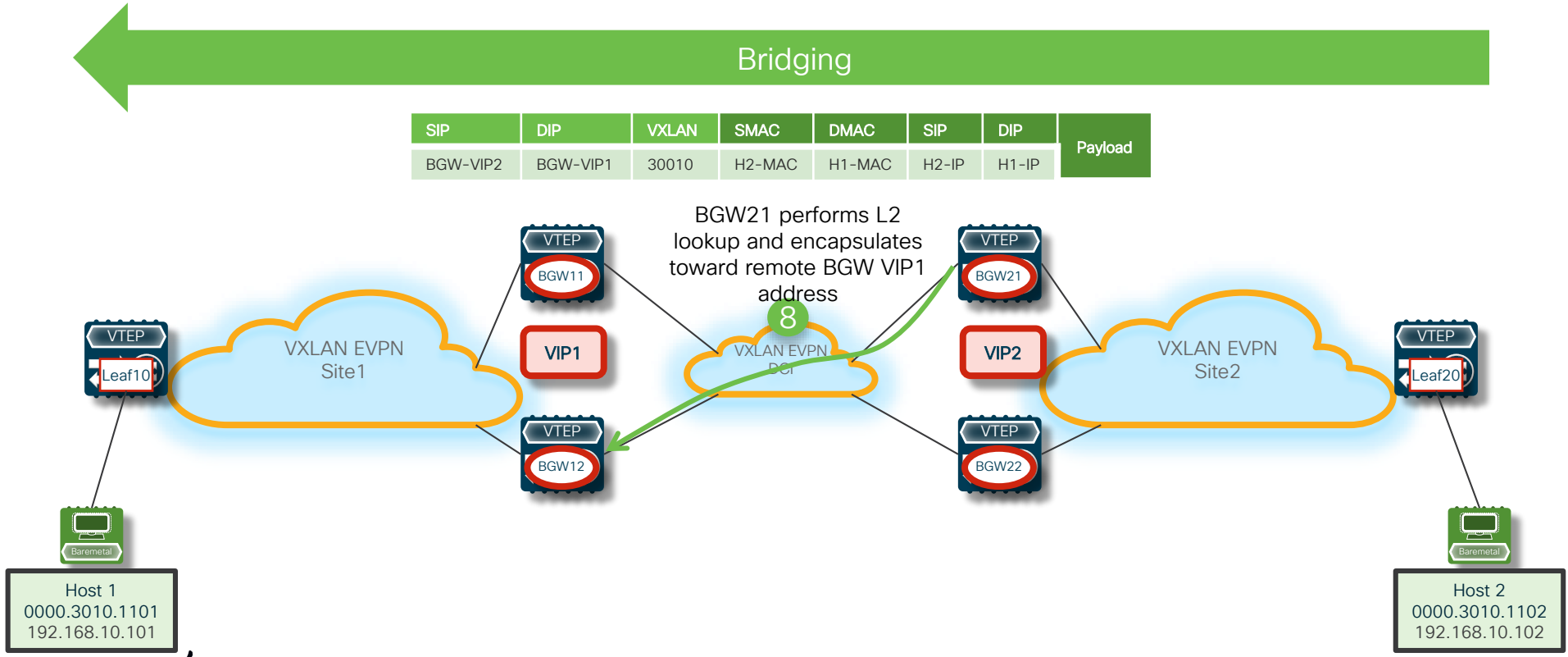
VXLAN Multi-Site Packet Walk

Layer 2 (Host 2 to Host 1) – Site 2



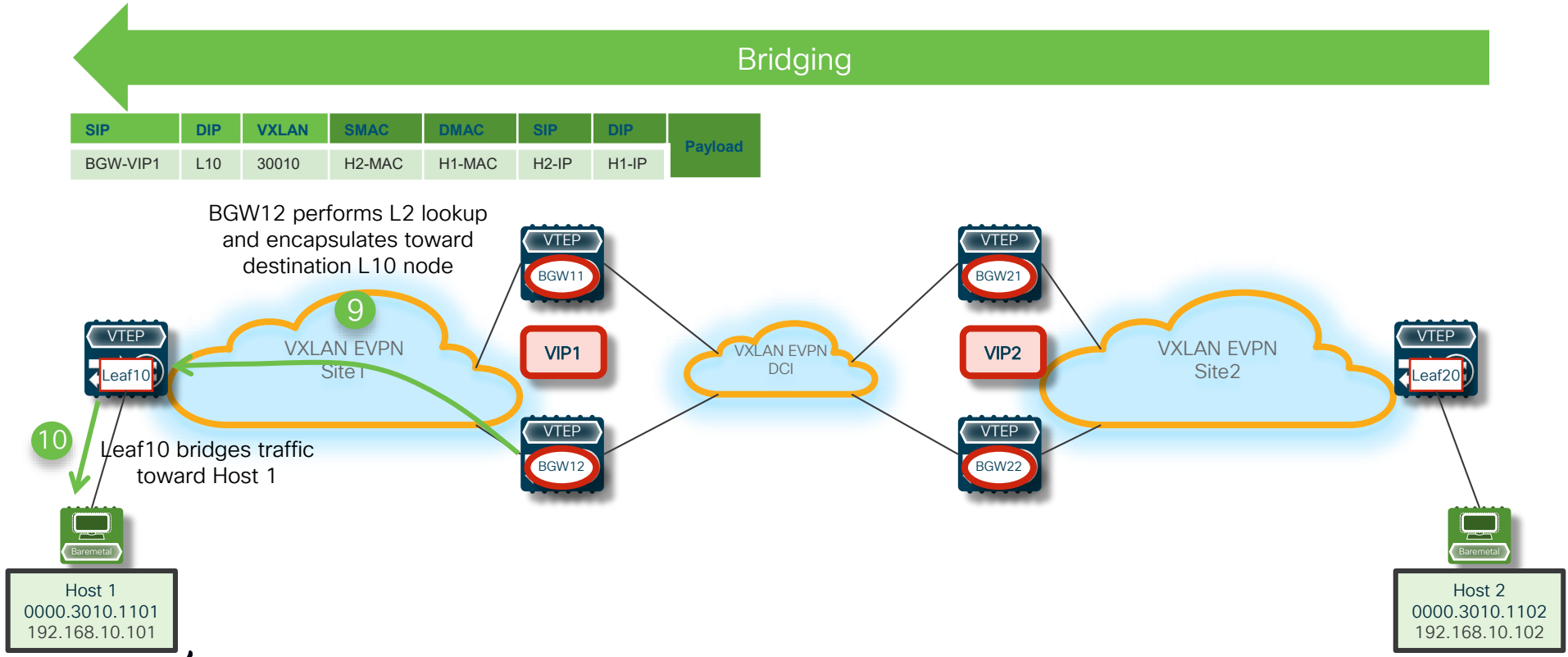
VXLAN Multi-Site Packet Walk

Layer 2 (Host 2 to Host 1) – DCI



VXLAN Multi-Site Packet Walk

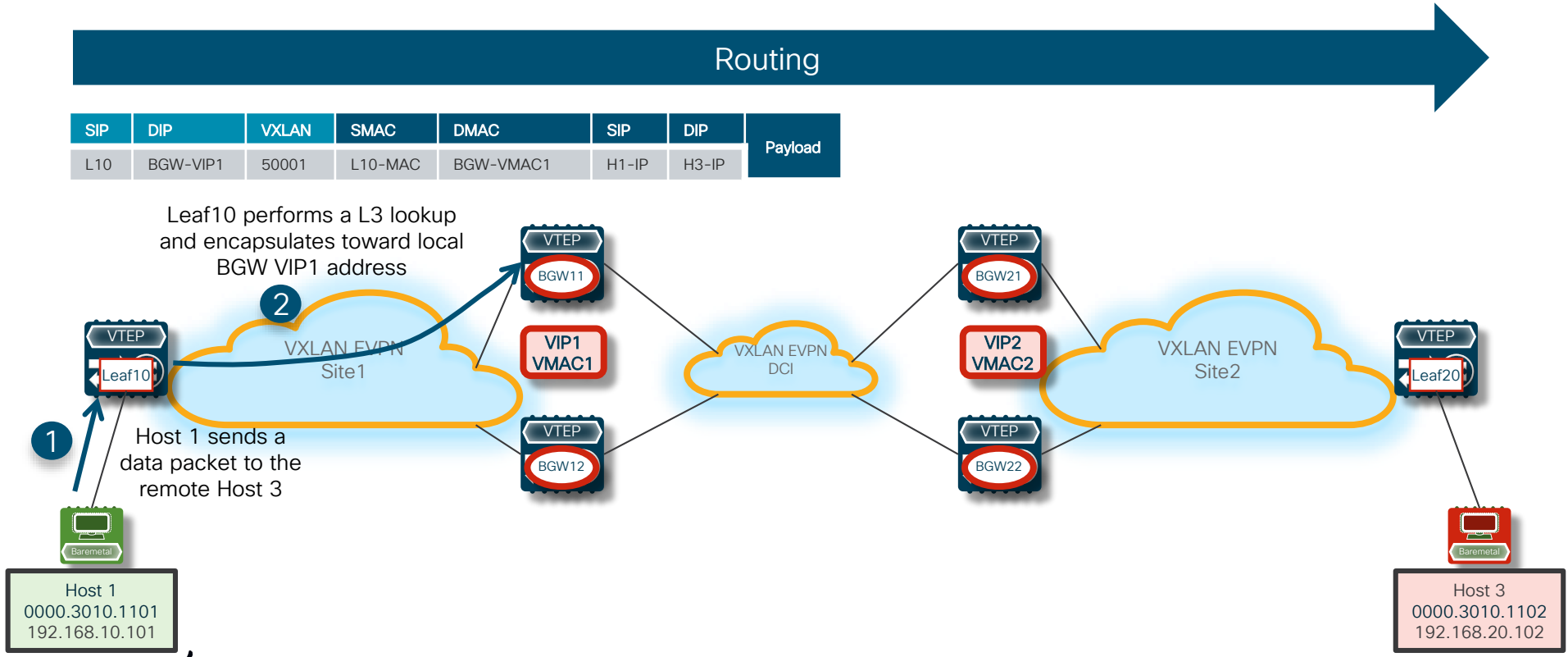
Layer 2 (Host 2 to Host 1) – Site 1



Multi-Site Packet Walk (Routing)

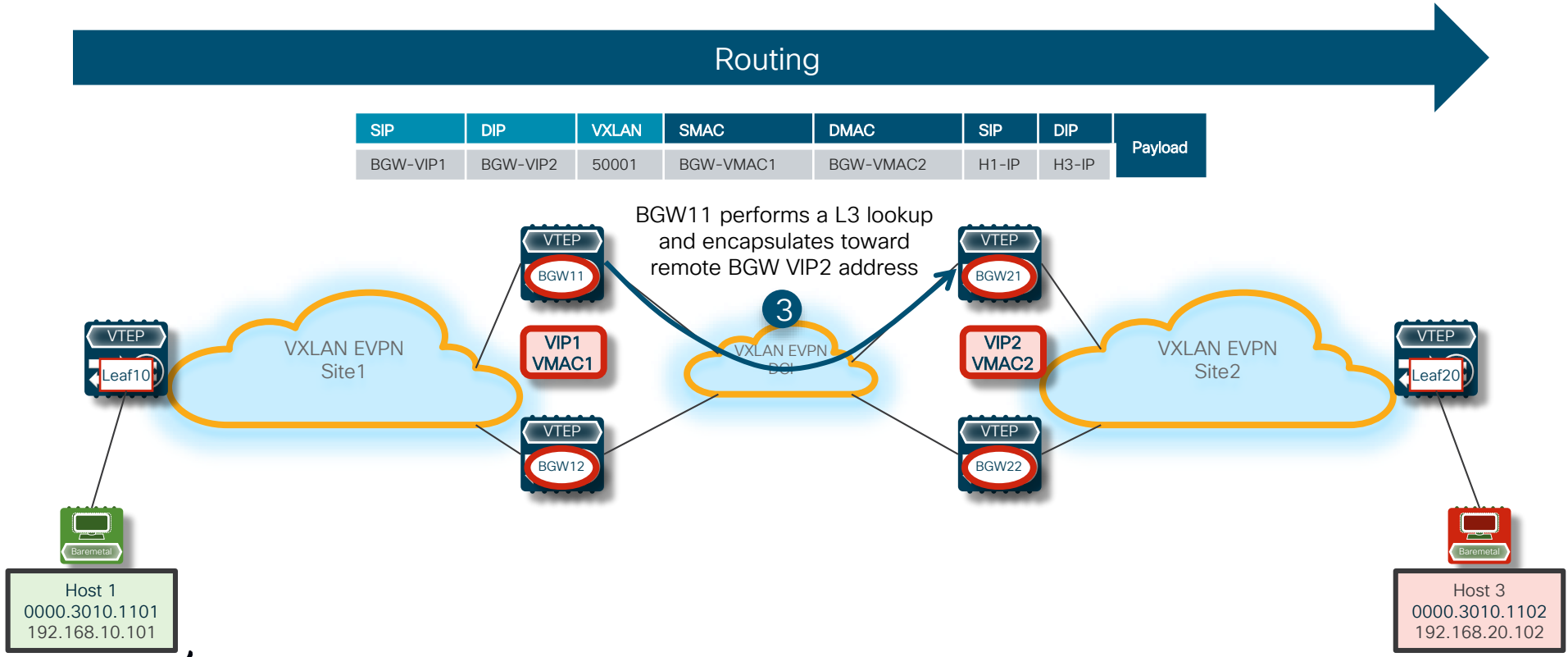
VXLAN Multi-Site Packet Walk

Layer 3 (Host 1 to Host 3) – Site 1



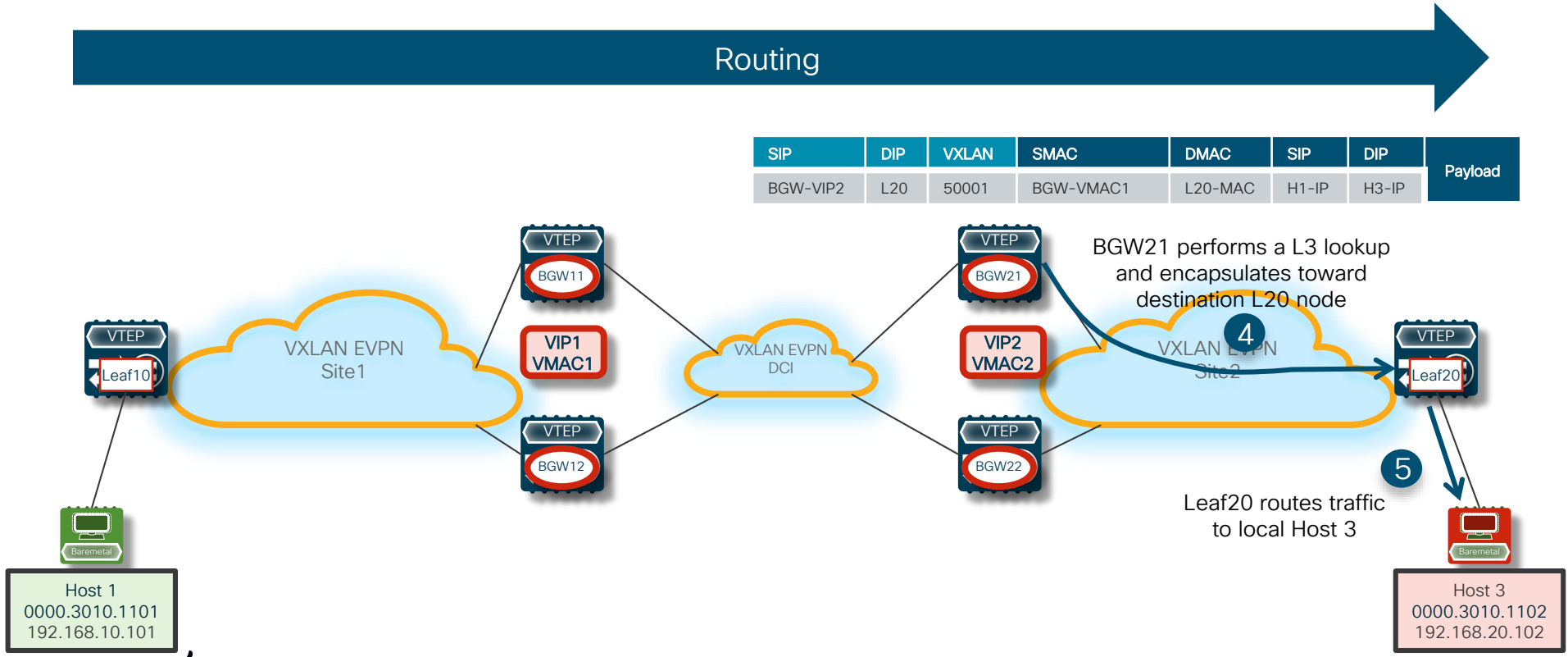
VXLAN Multi-Site Packet Walk

Layer 3 (Host 1 to Host 3) – DCI



VXLAN Multi-Site Packet Walk

Layer 3 (Host 1 to Host 3) – Site 2

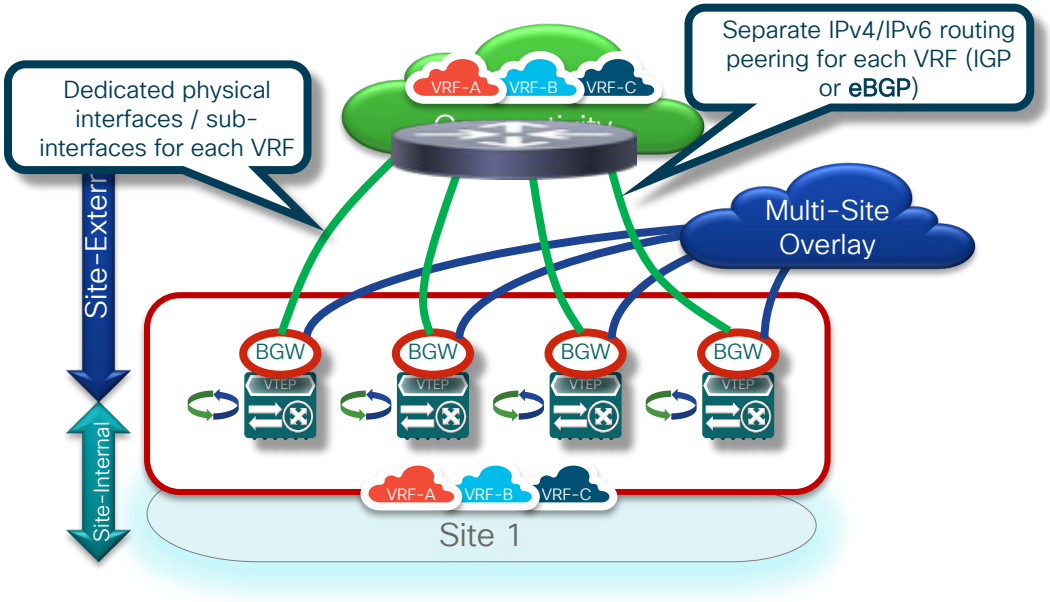


Connectivity to the External Layer 3 Domain

VXLAN Multi-Site
Connectivity to the External Layer 3 Domain

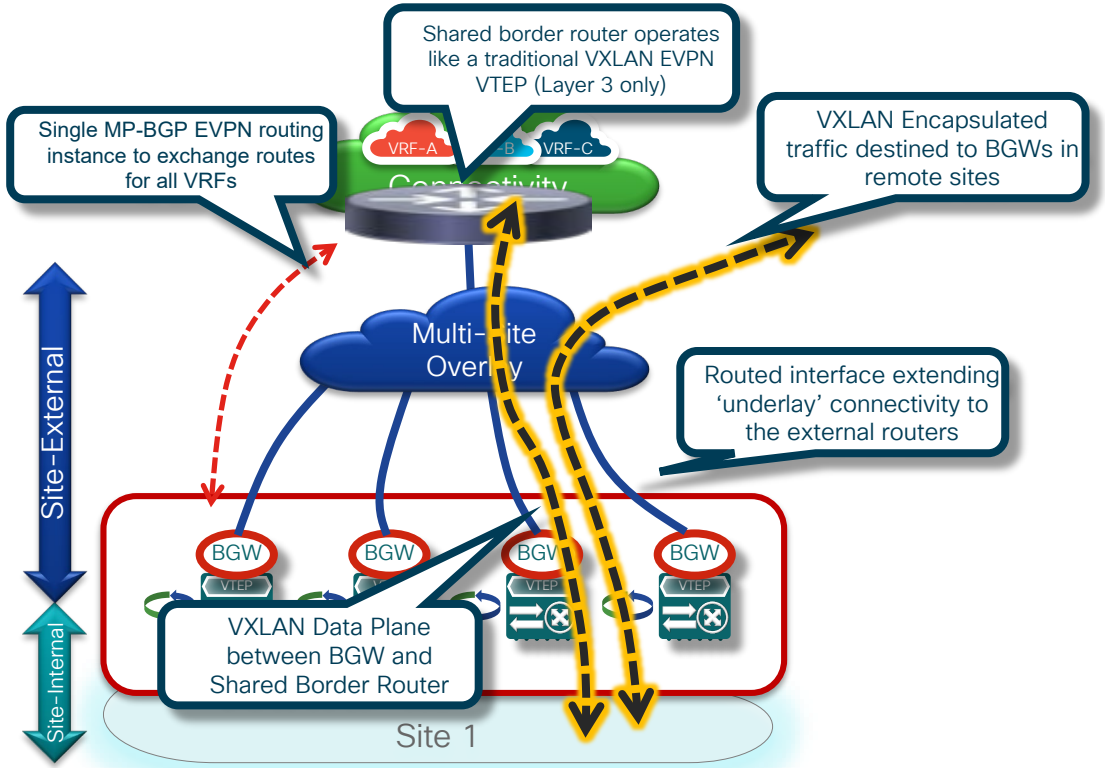
- Different connectivity models are supported
  - VRF-Lite peering with external WAN Edge routers
  - MP-BGP EVPN peering with external WAN Edge routers (Shared Border deployment model)
- Dedicated or shared pair of WAN Edge routers across sites
- The BGW nodes can also be used to provide Layer-3 external connectivity to each site

VXLAN Multi-Site
Border Gateways and VRF-Lite to External Routers



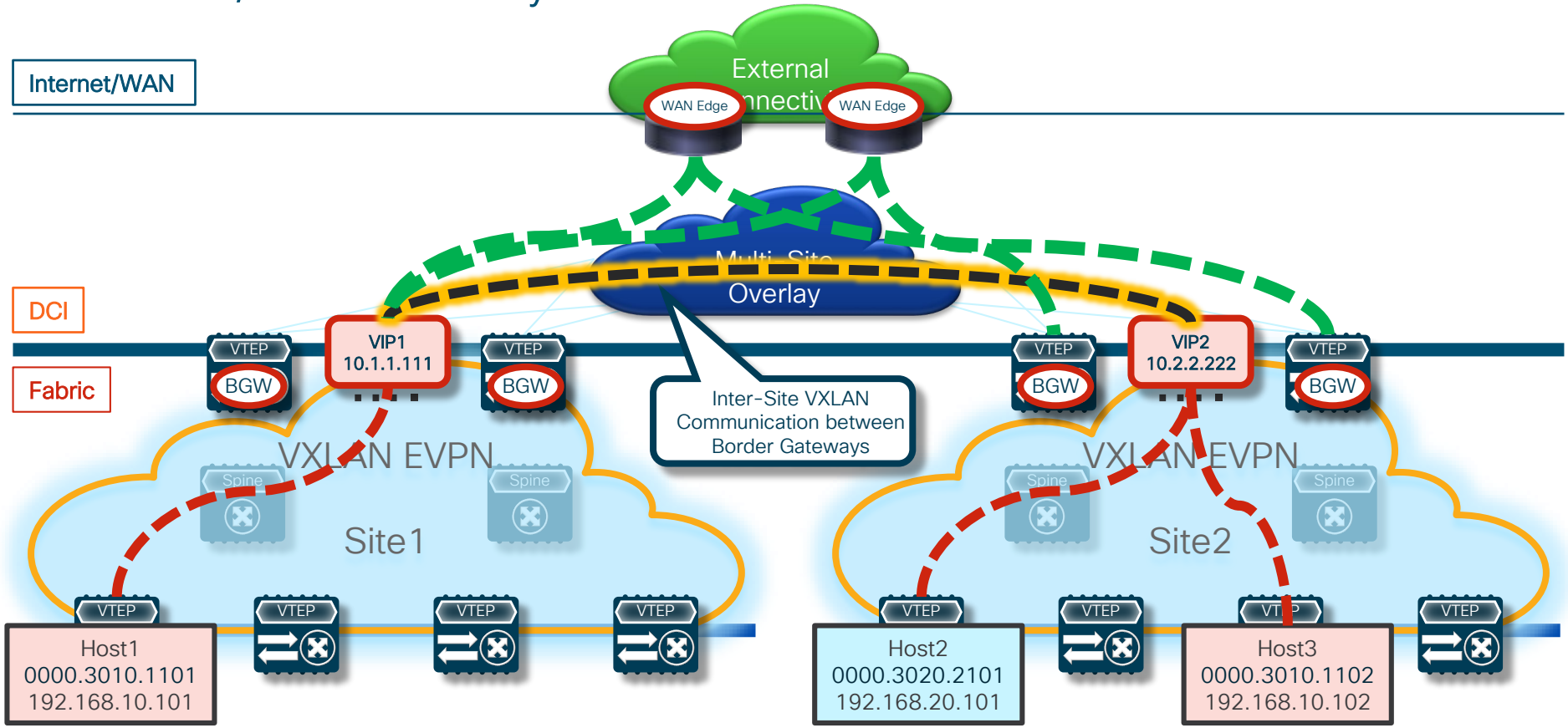
- Separate IPv4/IPv6 routing peering for each VRF established with the external routers on dedicated physical interfaces/sub-interfaces
- Must use separate interfaces for inter-site communication
  - No support for VXLAN encapsulated traffic on sub-interfaces

VXLAN Multi-Site
Border Gateway Connectivity to Shared Border Router

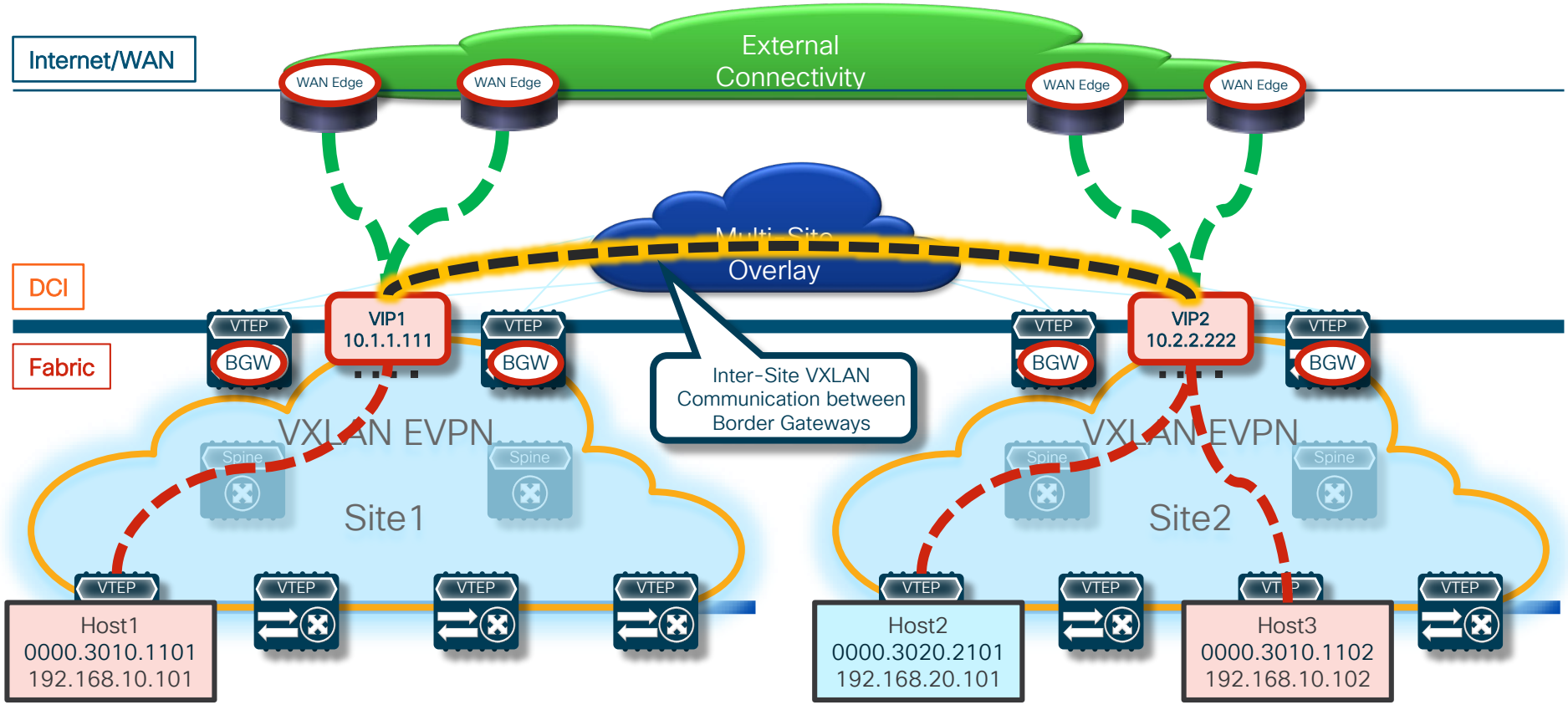


- Single MP-BGP EVPN peering established with the external routers to exchange routes for all the VRFs
- VXLAN Data-Plane between the BGWs and the external routers
- Same spine uplinks used for all VXLAN encapsulated traffic (North-South and East-West)
  - Required because of the use of DCI link tracking
- Various northbound hand-off options depending on specific HW support: VRF-Lite, MPLS-VPN, LISP

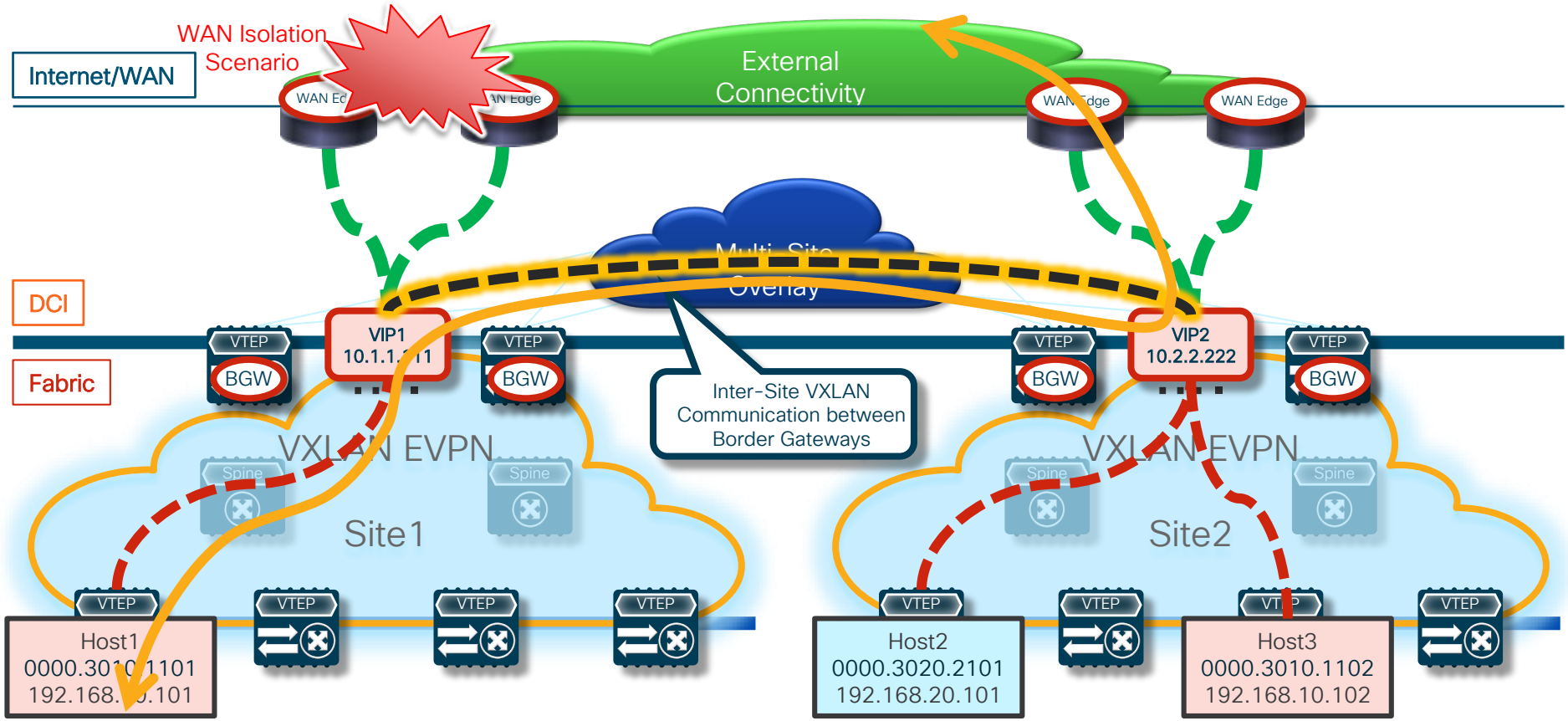
VXLAN Multi-Site
Internet/WAN Gateways Shared between Sites



VXLAN Multi-Site
Per Site Internet/WAN Gateways



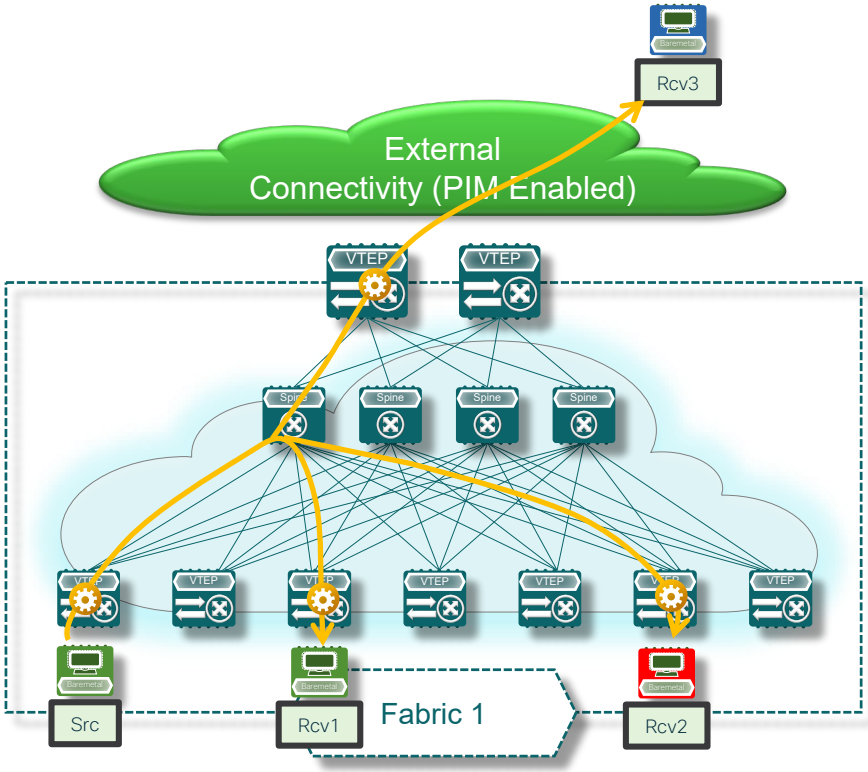
VXLAN Multi-Site
Per Site Internet/WAN Gateways



Tenant Routed Multicast (TRM) and Multi-Site Integration

Tenant Routed Multicast

Single Fabric Deployment

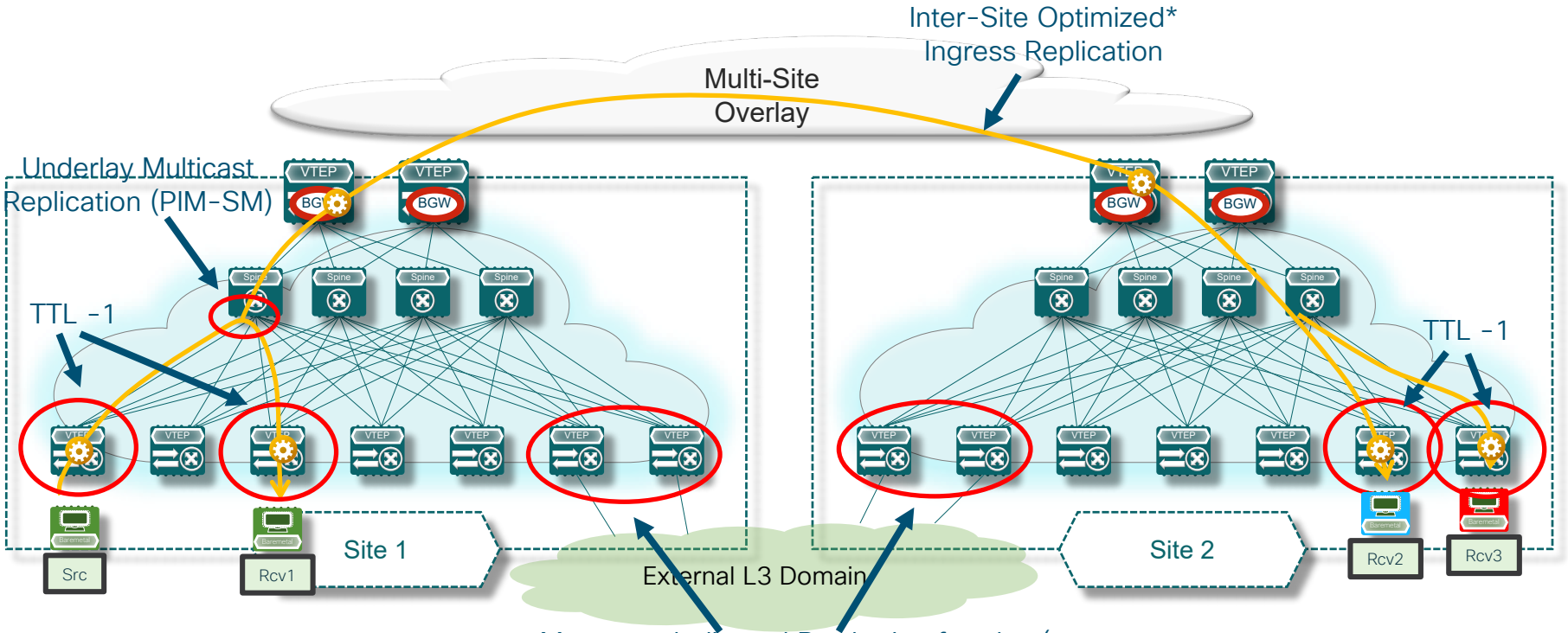


- Built as **Routing-First Approach**
  - Intra-subnet IP Multicast is always routed
- Underlay: VXLAN encapsulated traffic destined to a dedicated VRF Multicast group
  - Mandates the use of underlay multicast (PIM-SM only)
- Overlay: PIM-SM and PIM-SSM supported for TRM
  - For PIM-SM, three RP deployment models are supported
    1. RP-less: Anycast-RP on the fabric leaf nodes)
    2. External RP
    3. RP Anywhere: coexistence of RP-less and External RP models (Anycast RP or MSDP for syncing sources information)

Tenant Routed Multicast

East-West Forwarding via VXLAN Data-Plane

NX-OS 9.3(1)



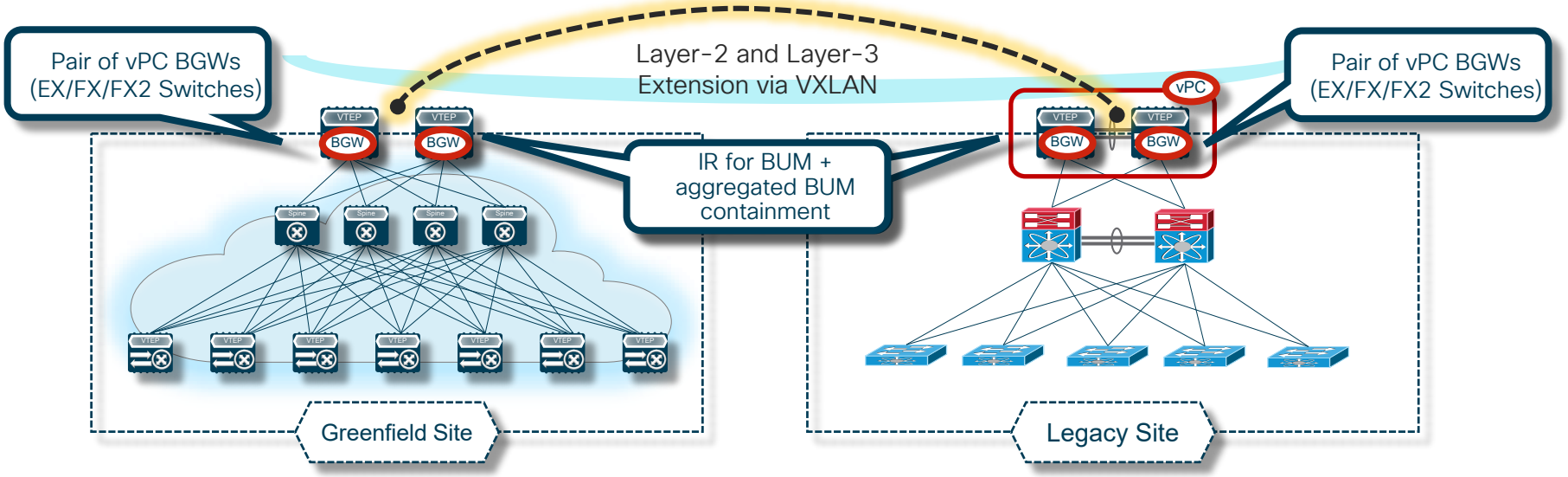
* Multicast streams are only replicated to Sites with interested receivers

Legacy Site Integration

Legacy Site Integration

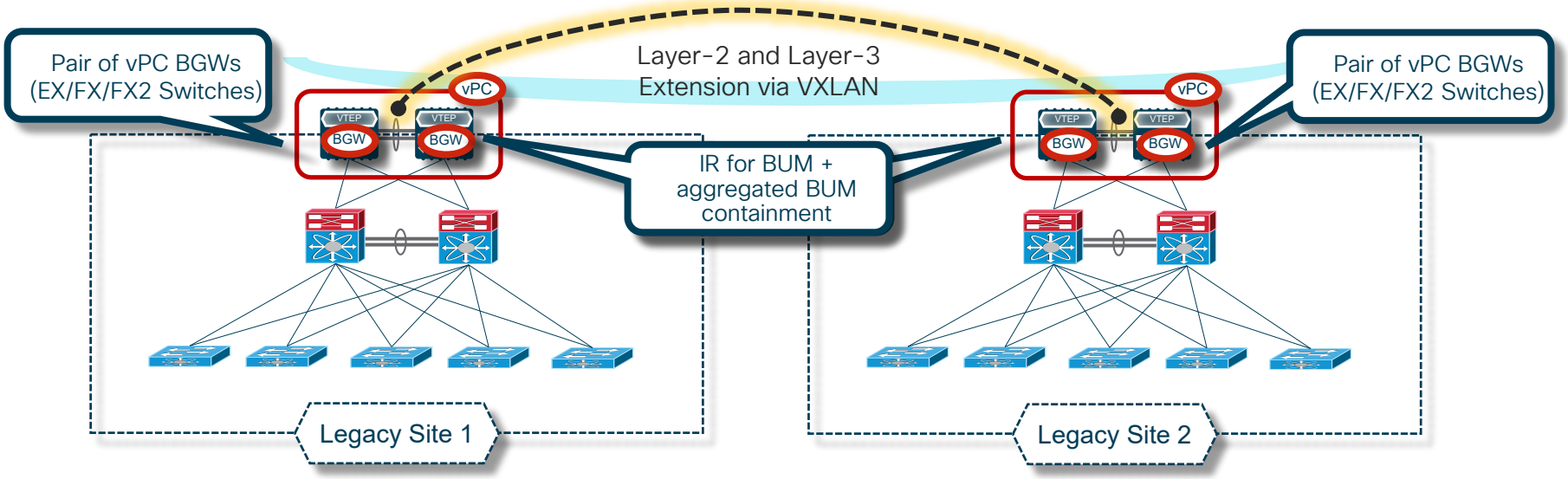
Main Use Cases

VXLAN Multi-Site with vPC BGWs
Migration/Coexistence Use Case



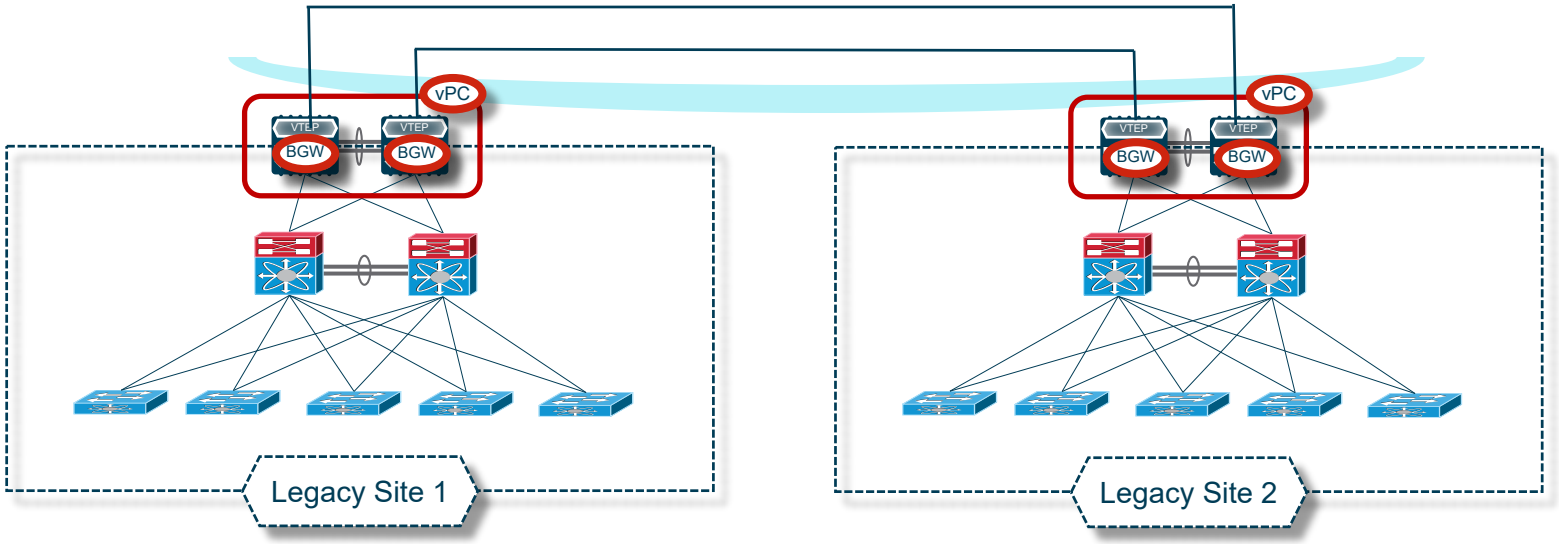
- Coexistence and/or migration use cases
  - Need to extend Layer-2 and Layer-3 multi-tenant connectivity across sites
- Deploy a pair of vPC BGWs in the legacy site
  - Seamless connectivity extension via VXLAN
  - Leveraging native Multi-Site functions (Ingress Replication for BUM, BUM containment, etc.)

VXLAN Multi-Site with vPC BGWs
Next-Gen DCI to Interconnect Legacy Networks



- A pair of vPC BGWs inserted in each legacy site to extend Layer-2 and Layer-3 connectivity between sites
  - Replacement of traditional DCI technologies (EoMPLS, VPLS, OTV, ...)
- Provides the option of slowing phasing out the legacy networks and replace them with modern VXLAN EVPN fabrics

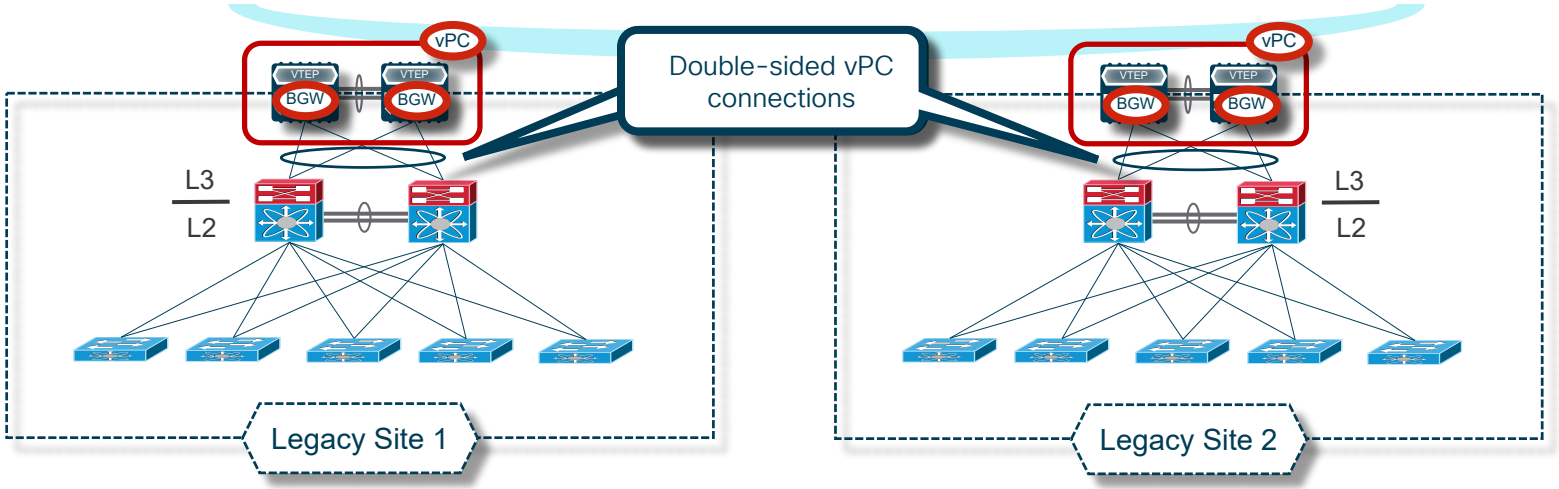
VXLAN Multi-Site with vPC BGWs
Next-Gen DCI Use Case with Back-to-Back BGWs



- Typical topology leveraging dedicated dark fiber links or DWDM circuits
- ‘Squared’ and ‘full mesh’ topologies are both fully supported
- Recommended to limit the back-to-back deployment to two sites
  - 2 sites topology can be fully automated using DCNM
  - Recommended to insert Layer 3 core network with 3+ sites

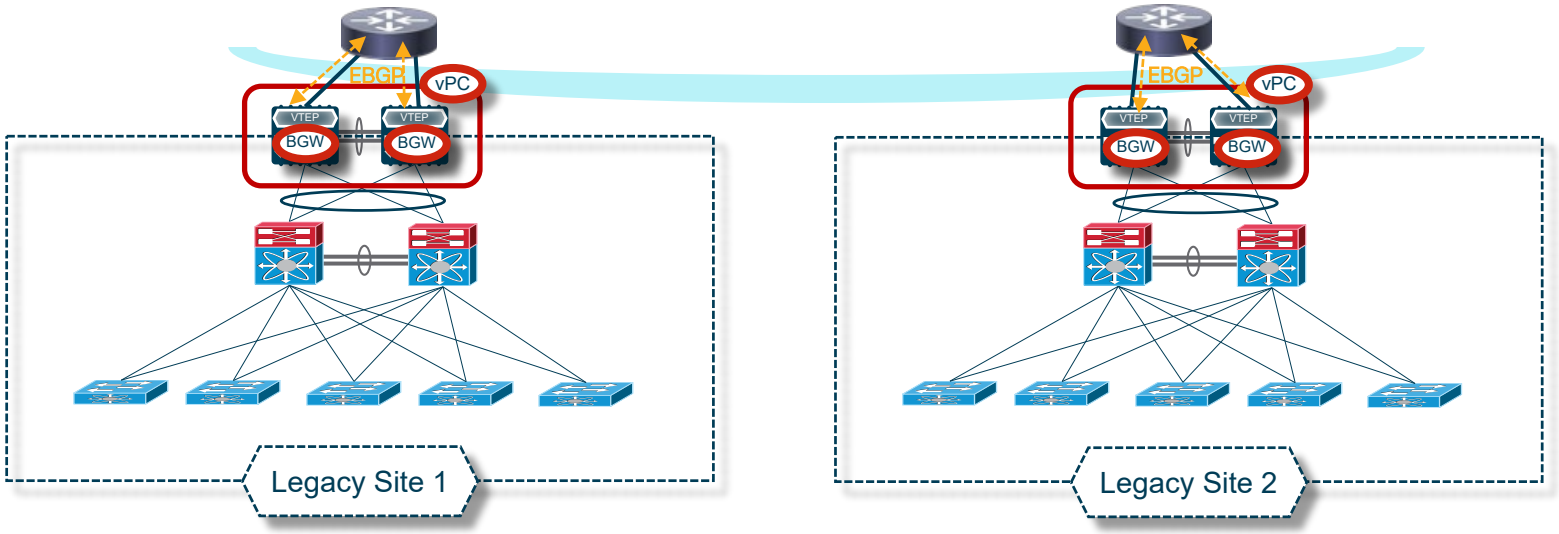
Migrating Legacy DCs to VXLAN EVPN Fabrics

Migrating Legacy DCs to VXLAN EVPN Fabrics
Step 1 – Insert a Pair of vPC BGWs in Each Legacy Site



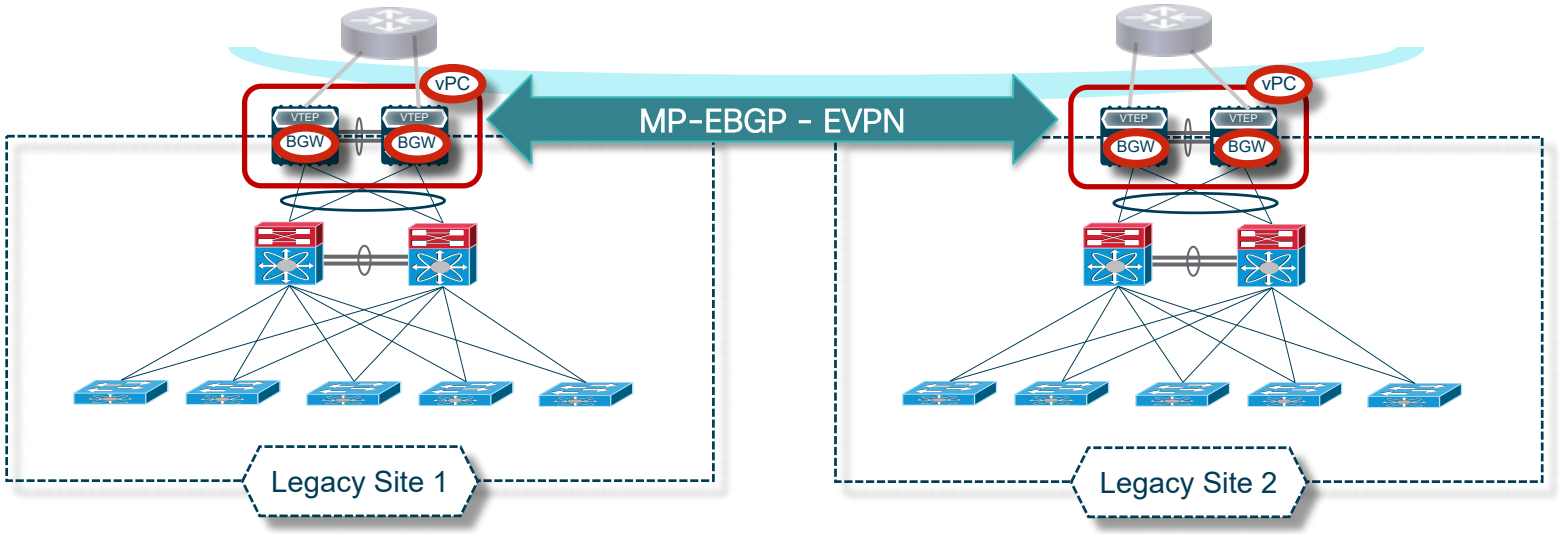
- Recommended to deploy double-sided vPC connections between legacy aggregation devices and vPC BGWs
  - Allows to create a single L2 logical connection with all links actively forwarding traffic
  - Can apply BPDU filtering between aggregation devices and vPC BGWs to mitigate impact of TCNs
- Default gateway functions still offered on the legacy aggregation devices (Active/Standby across sites)

Migrating Legacy DCs to VXLAN EVPN Fabrics
Step 2 - Configure vPC BGWs DCI Underlay Peerings



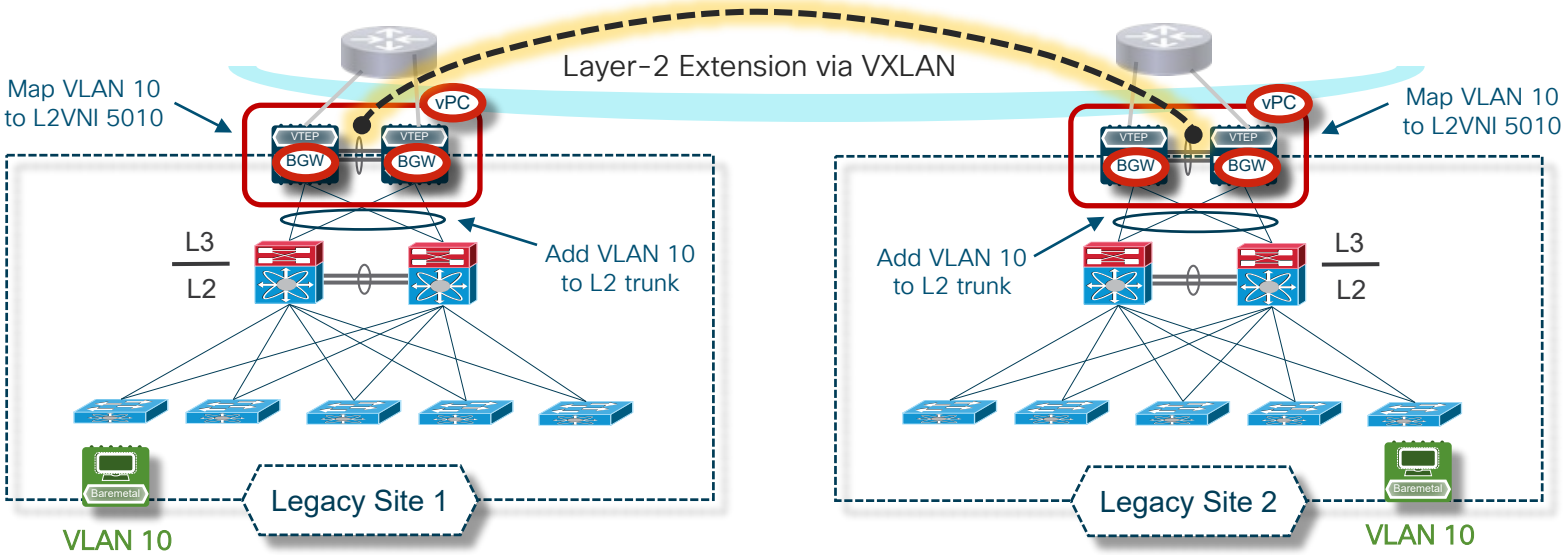
- Establish underlay routing adjacencies with the first-hop L3 devices in the core network
  - EBGP is the recommended protocol of choice
  - Establish EBGP point-to-point peerings using the physical interfaces IP addresses
- Underlay connectivity across the core network required to exchange BGW loopback addresses with the remote vPC BGWs

Migrating Legacy DCs to VXLAN EVPN Fabrics
Step 3 - Configure vPC BGWs DCI Overlay Peerings



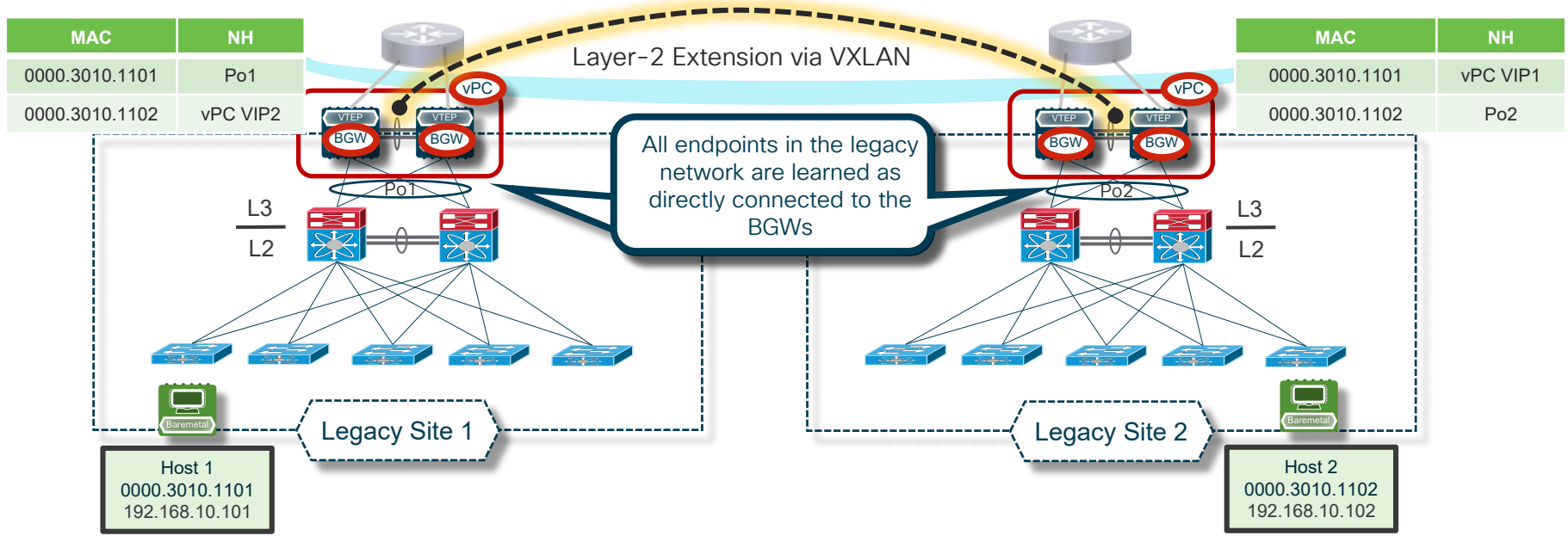
- Establish overlay routing adjacencies between vPC BGWs deployed in separate sites
  - Mandatory establishment of EBGP session across sites
  - Full-mesh EBGP peering is required
  - Alternatively, can use route-server services in the core network

Migrating Legacy DCs to VXLAN EVPN Fabrics
Step 4 - Configure vPC BGWs for DCI Layer 2 Extension across Sites



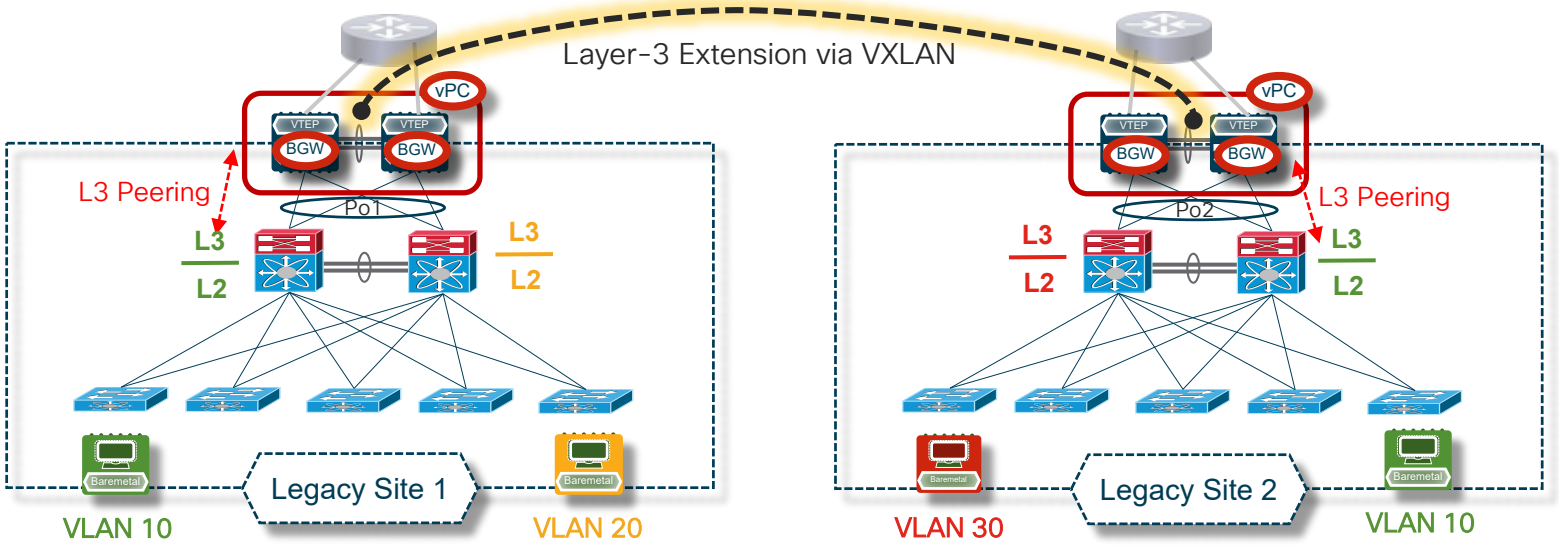
- Layer-2 extension can now start being performed between vPC BGWs pairs
  - Add the VLANs that need to be extended on the L2 trunk between legacy network and vPC BGWs
  - Map the VLANs to L2VNI segments on the vPC BGW devices
  - MAC information would start being advertised across sites for endpoints connected to those VLANs

Migrating Legacy DCs to VXLAN EVPN Fabrics
Step 4 - Configure vPC BGWs for DCI Layer 2 Extension across Sites



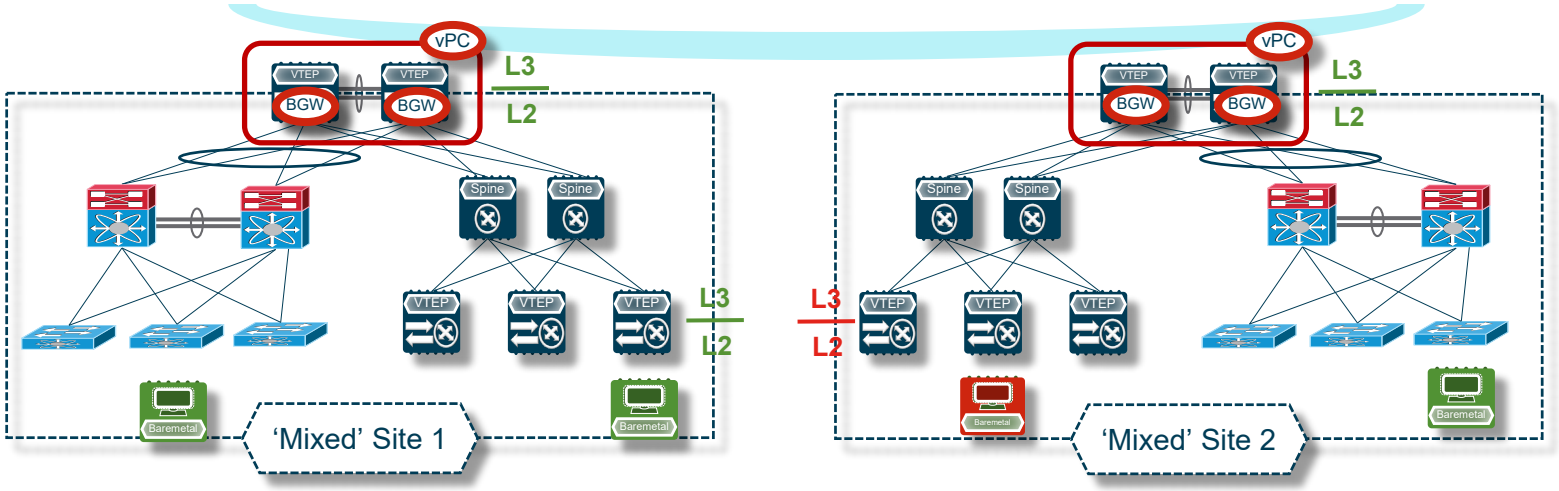
- Endpoints connected to the legacy network are discovered as directly connected to the local vPC BGW pair
- VXLAN tunnels for intersite Layer-2 connectivity are established between the vPC VIP addresses

Migrating Legacy DCs to VXLAN EVPN Fabrics
Step 5 - Migrate Default Gateway to the vPC BGWs



- The migration of the default gateway on the vPC BGW can be performed on a subnet by subnet basis
- Allows to provide an all-active default gateway in both sites
- Until the gateway for all the IP subnets is migrated, it is required to create a L3 peering between the legacy network and the vPC BGW

Migrating Legacy DCs to VXLAN EVPN Fabrics
Step 6 – Start Deploying a New Local VXLAN Fabric



- Introduce VXLAN EVPN spines and additional VTEPs in each site
- Migrate endpoints between the legacy network and the new VXLAN EVPN fabric

Conclusions

Multi-Site Advantages – "The Multiple"

- **Multiple** Overlay Domains – Interconnected & Controlled
  - Scaling and Segregating VXLAN EVPN Networks
- **Multiple** Overlay Control-Plane Domains – Interconnected & Controlled
  - Limited Overlay Control-Plane Update Propagation
- **Multiple** Underlay Domains – Isolated
  - Isolated Underlay Domains – No need for Extension
- **Multiple** Replication Domains for BUM – Interconnected & Controlled
  - Individual BUM flooding domain with Traffic control

Resources

- VXLAN EVPN Multi-Site Design and Deployment White Paper
  https://www.cisco.com/c/en/us/products/collateral/switches/nexus-9000-series-switches/white-paper-c11-739942.html
- NextGen DCI with VXLAN EVPN Multi-Site Using vPC Border Gateways White Paper
  https://www.cisco.com/c/en/us/products/collateral/switches/nexus-9000-series-switches/whitepaper-c11-742114.html
- Cisco Live Online - VXLAN BGP EVPN based Multi-POD, Multi-Fabric and Multi-Site - BRKDCN-2035
  https://www.ciscolive.com/global/on-demand-library/?search=BRKDCN-2035&showMyInterest=false#/
- Cisco DCNM 11.3(1) - Multi-Site Domain for VXLAN BGP EVPN Fabrics

  https://www.cisco.com/c/en/us/td/docs/switches/datacenter/sw/11_3_1/config_guide/lanfabric/b_dcnm_fabric_lan/border-provisioning-multisite.html

Complete your online session survey

- Please complete your session survey after each session. Your feedback is very important.
- Complete a minimum of 4 session surveys and the Overall Conference survey (starting on Thursday) to receive your Cisco Live t-shirt.
- All surveys can be taken in the Cisco Events Mobile App or by logging in to the Content Catalog on ciscolive.com/emea.

Cisco Live sessions will be available for viewing on demand after the event at ciscolive.com.

cisco Live!

Continue your education

Demos in the Cisco campus

Walk-in labs

Meet the engineer 1:1 meetings

Related sessions

Thank you

You make **possible**