



Massimo Magnani  
Arista Networks  
Systems Engineer  
massimo@arista.com

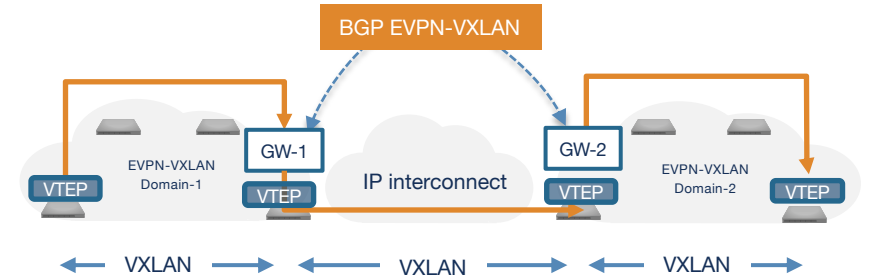
# Hierarchical EVPN for scaling and DCI

RFC 9014 - Interconnect Solution for Ethernet VPN (EVPN) Overlay Networks

# EVPN GW – Hierarchical EVPN for scaling and DCI

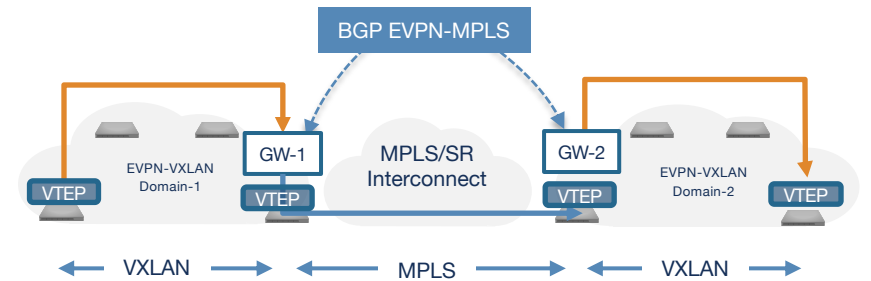
Interconnect EVPN domains same campus or geographically separate locations

- Single administration domain with IP connectivity between domains.
- Providing layer 2 and 3 connectivity across the EVPN domains?



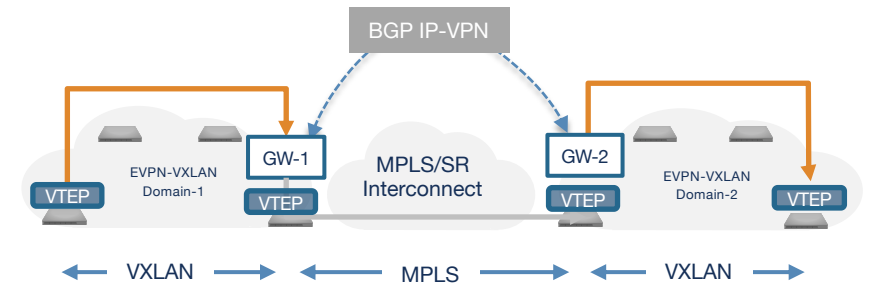
Interconnect EVPN domains geographically separate locations

- Single administration domain with EVPN-MPLS connectivity between domains.
- Providing layer 2 and 3 connectivity across the EVPN domains?



Interconnect EVPN domains geographically separate locations

- Single administration domain with MPLS IP-VPN connectivity between domains.
- Providing layer 3 only connectivity across the EVPN domains?



# EVPN GW – Hierarchical EVPN for scaling and DCI

- IETF BESS working group, number of drafts for DCI and MPLS interop.
  - Support for both Layer 2 and 3 DCI solutions
  - Interop across BGP Address Families and data-plane encapsulations (VXLAN, PBB, MPLS)

Draft	Overview	
A Network Virtualization Overlay Solution using EVPN <b>RFC 8365</b>	EVPN control plane for L2 VPNs with an NVO environment with VXLAN, NVGRE and GENEVE encap– DCI using GWs and DCI using ASBRs	
E-VPN and IP-VPN Integrated Solution <b>draft-ietf-bess-evpn-ipvpn-interworking-07</b>	Layer 3 DCI interop between EVPN-VXLAN/MPLS and IP-VPN WAN for layer 3 DCI	L3 IPVPN solution
Interconnect Solution for EVPN Overlay networks <b>RFC 9014</b>	GW DCI solution with multiple control planes (VPLS/EVPN) and data-planes (MPLS, VXLAN, PBB)	
Multi-site EVPN based VXLAN using Border Gateways <b>draft-sharma-bess-multi-site-evpn-01</b>	GW DCI solution focused only on EVPN-VXLAN, support for a single control planes (EVPN) and single data-plane (VXLAN)	Competing for Solutions

# EVPN GW – EVPN VXLAN/MPLS GW

## Standard based solution

- EVPN-DCI, RFC 9014
- Multiple encap support (VXLAN/MPLS)
- Standards based Multi-homing support

## EVPN GW for Hierarchical scaling

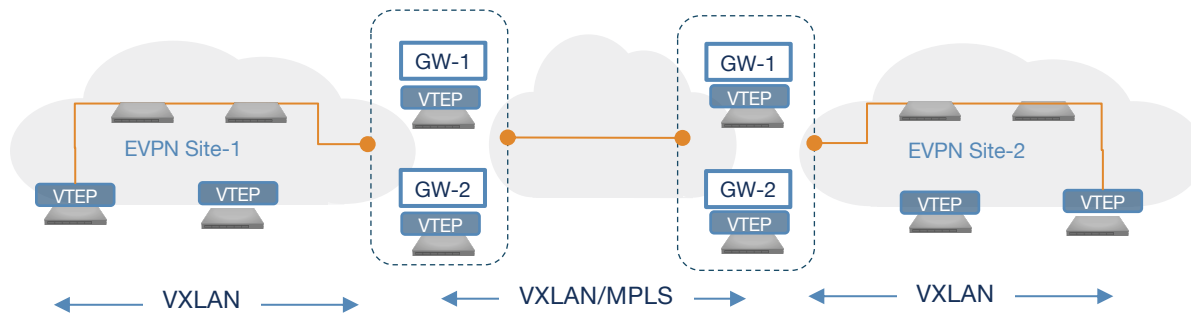
EVPN GW for scaling EVPN-VXLAN deployments inter-POD and intra-site (DCI) by introducing hierarchy

## Scalable L2 interconnect

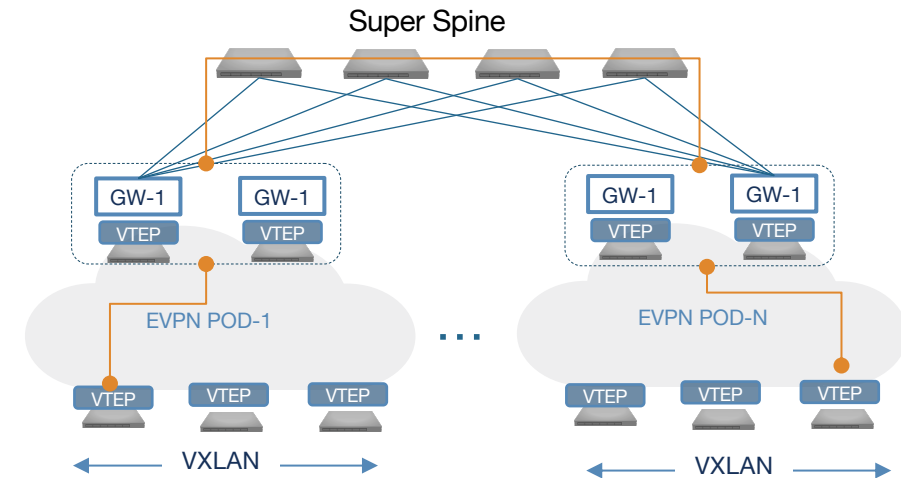
- GW scoping of Type 1,4 and 3 routes
- Flood-list scale with split-horizon forwarding of BUM traffic on GW
- Type-2 re-originated with GW next-hop

## Layer 3 interconnect

- Layer 3 (type-5) interconnect between domains
- Type-5 routes re-advertised with GW next-hop



DCI for interconnecting Datacenters

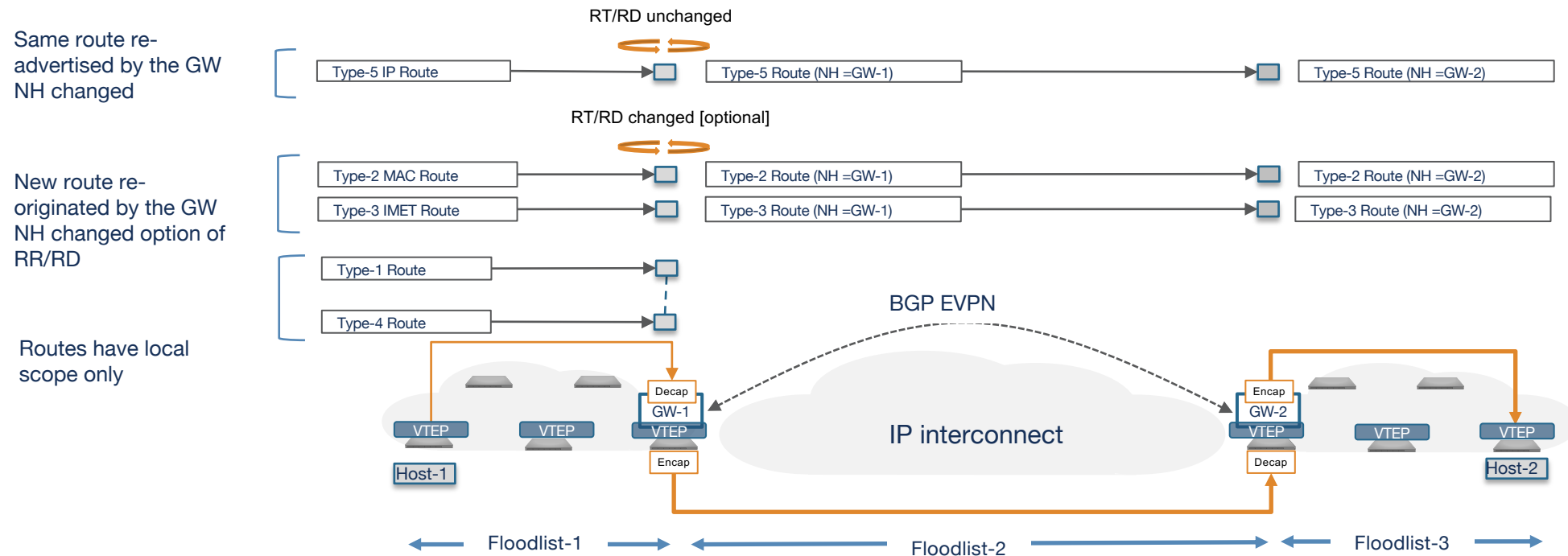


Interconnecting PODs for hierarchical scaling

# EVPN GW – Route forwarding behavior

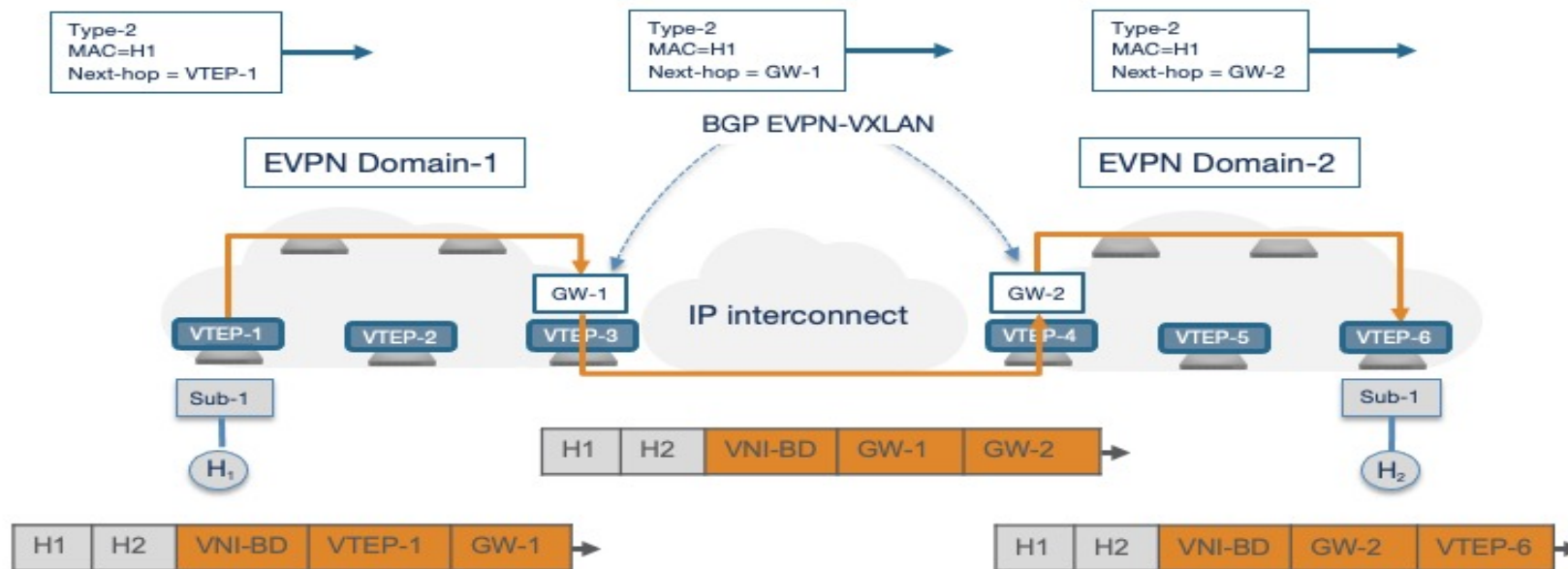
- EVPN GW model

- Does NOT re-advertise local Type-1, 4 and Type-3 routes
- Re-originates local Type-2 and 3 routes, NH changed and RT/RD changed [optional]
- Re-advertises local Type-5 routes, NH changed, RT/RD are unchanged



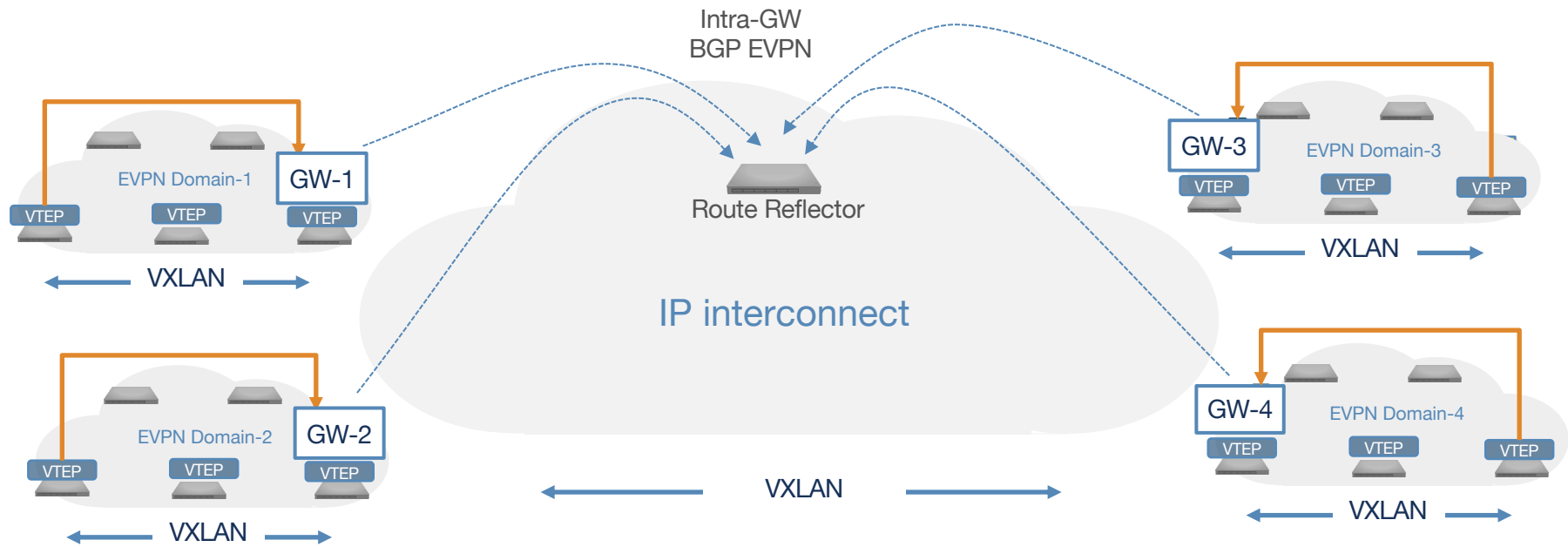
# EVPN GW – EVPN GW layer 2 with EVPN interconnect

- The GW node is in the VXLAN forwarding plane
  - Traffic forwarded to the GW, remote hosts learnt with a next-hop of the EVPN GW
  - Performs VXLAN decap/encap for L2 and L3 VNIs stretched across the DCI

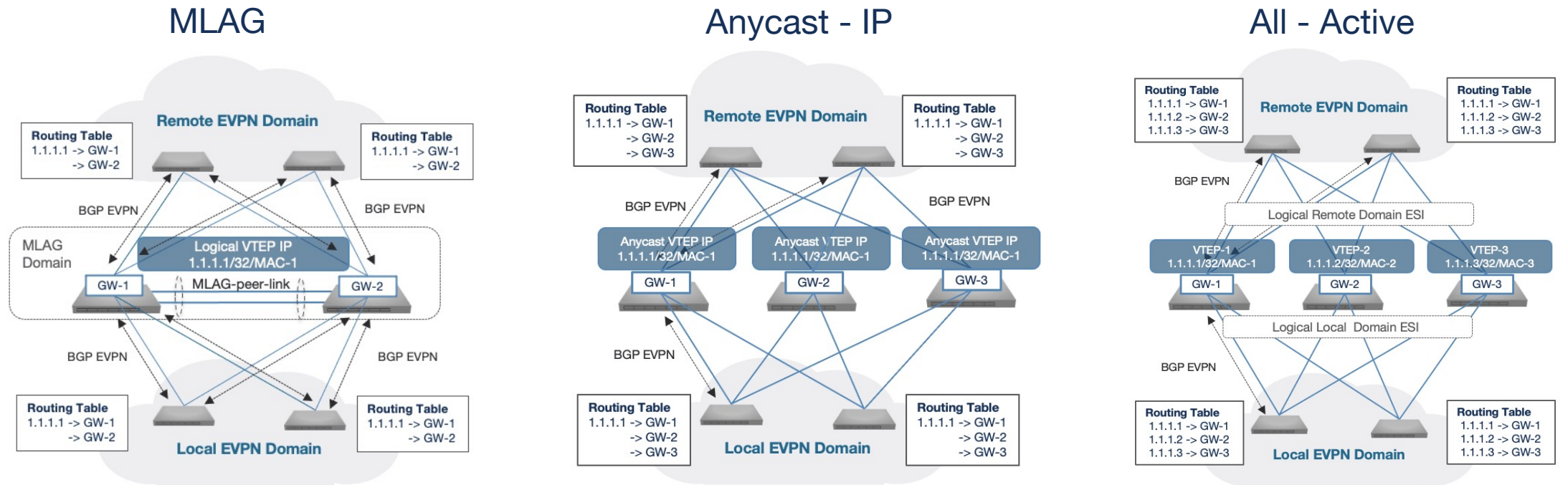


# EVPN GW – EVPN multi-domain topologies

- EVPN GW for multi-domain architectures
  - Hierarchical EVPN topology with EVPN route scoping between domains
  - Seamless layer 2 and 3 connectivity between domains
  - Resilient A-A GW at each domain for active-active forwarding



# EVPN GW – Resiliency Models at a glance

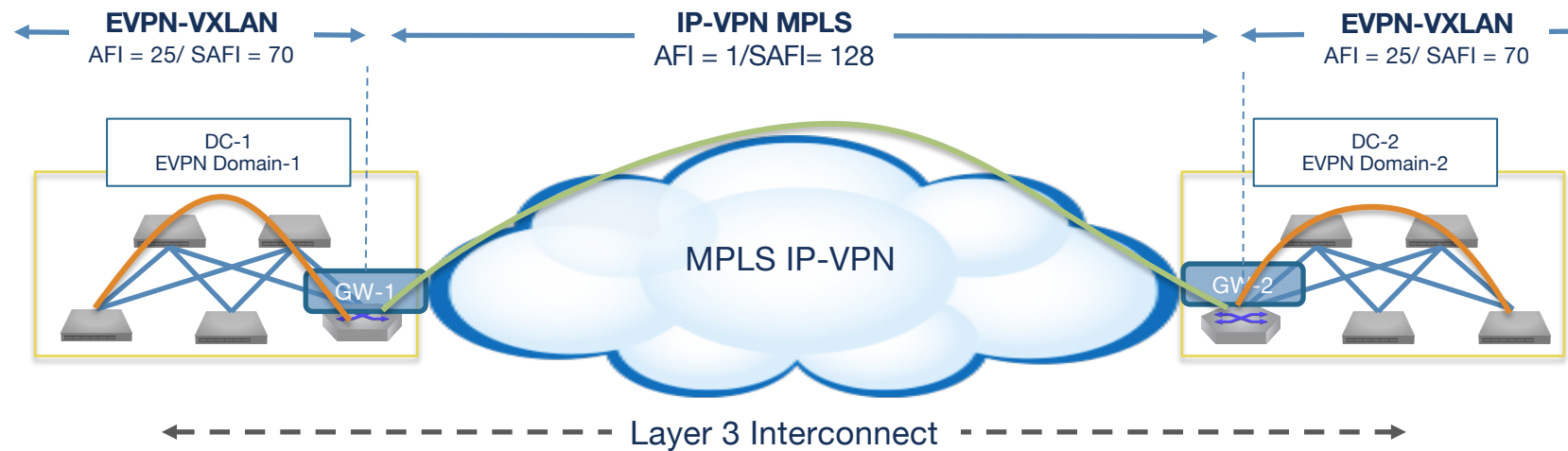


<b>Max # of Nodes</b>	Limited to 2 nodes - Single MLAG domain x site	No Restriction on # of GW per site	Up to 16 GW nodes
<b>VTEP IP</b>	Single Logical VTEP IP	Single Logical VTEP IP	VTEP IP per Node
<b>EVPN Routes</b>	No additional routes – MLAG Sync	No additional routes for state sync	Type-1 (AD per ES/EVI) – Type-4 (ES)
<b>Type 2/3/5 Routes</b>	Next-hop Logical VTEP IP	Next-hop Anycast VTEP IP	VTEP IP of node, with ES on Type-2 routes
<b>BUM</b>	Shared Logical IP – GW receive and forward a single copy	Shared Logical IP – GW receive and forward a single copy	Each GW receives a copy only DF forwards BUM out
<b>ECMP</b>	Underlay Load Balancing	Underlay Load Balancing	Overlay Load Balancing
<b>Attached hosts</b>	Supported	Not Supported	Not Supported



# EVPN GW - EVPN to IPVPN

- Integrated EVPN-VXLAN and IP-VPN GW
  - Provides layer 3 interconnect between different EVPN domains
  - Based on the the BESS *evpn-ipvpn-interworking* draft
  - BGP Address Family translation and forwarding plane translation (VXLAN to MPLS)
  - Support introduced on Jericho in 4.23.2F supported on J/J+/J2 platforms
  - **Use case** : Integration with an existing IP-VPN MPLS WAN topology



# EVPN – Vendor Interop testing

- European Advanced Networking Test Center (EANTC)
  - Independent/vendor neutral testing facility based in Berlin
  - Currently only event globally - hence wide multi-vendor attendance and importance
  - Verify vendor interoperability of new and involving IETF standards,
  - Vendor proposed test cases to mimic EVPN DC and MPLS core deployments
  - EANTC responsible for providing independent validation of test



# Questions?

