



EVPN

One Control Plane to rule them all

RONOG 11

Bucharest

Patrick Prangl - Technical Lead, Systems Engineering <pprangl@arista.com>

Traditional VPN Service Landscape

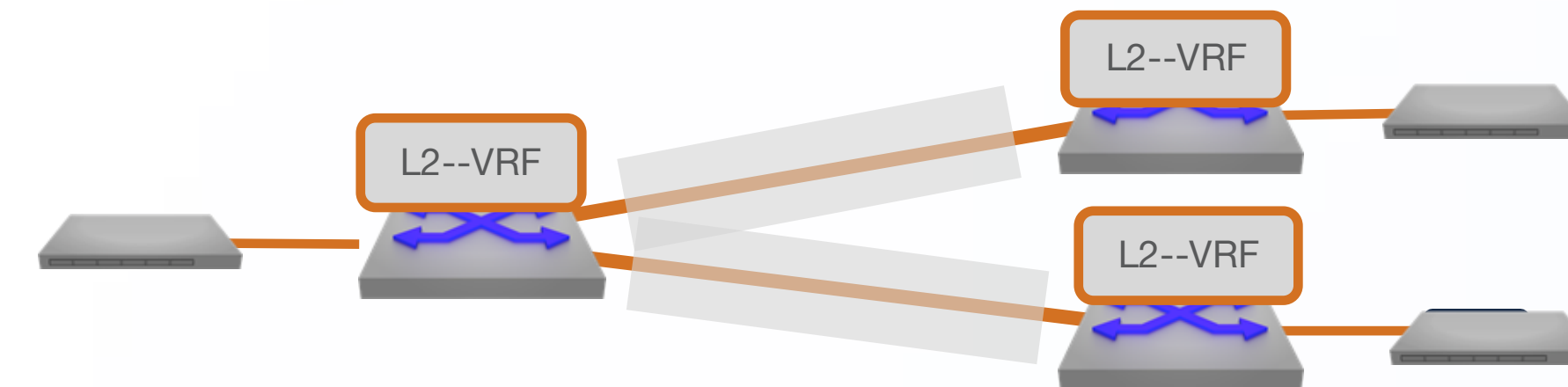


Simplification with EVPN

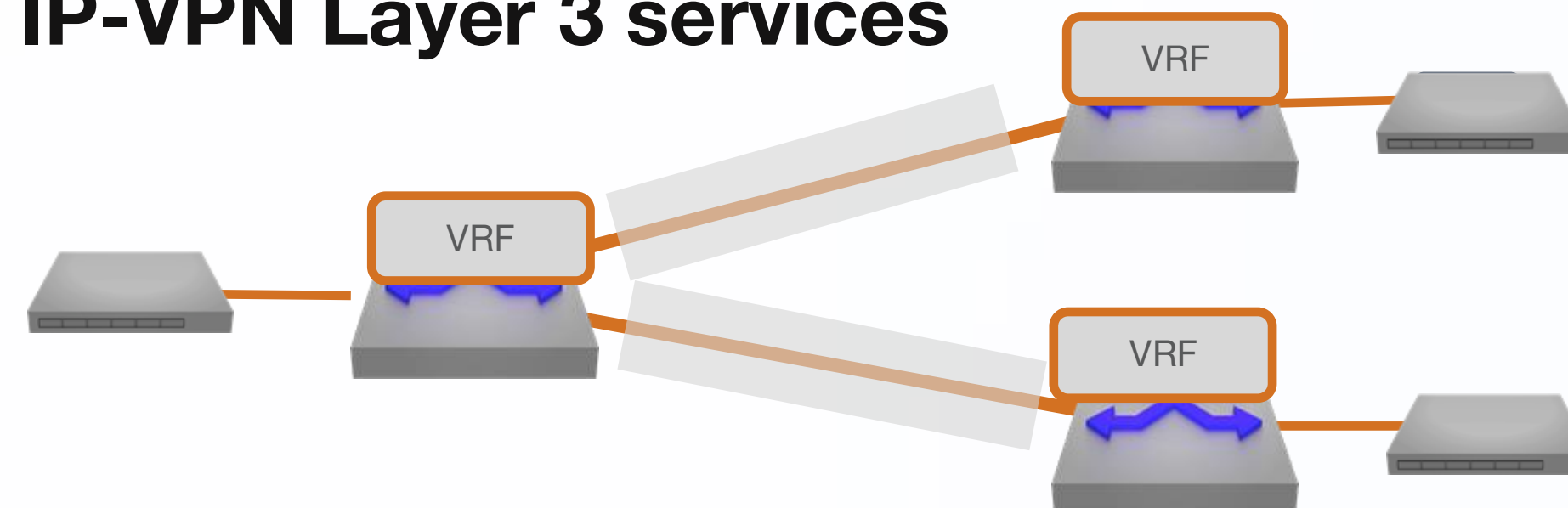
E-LINE (Point-to-Point) services



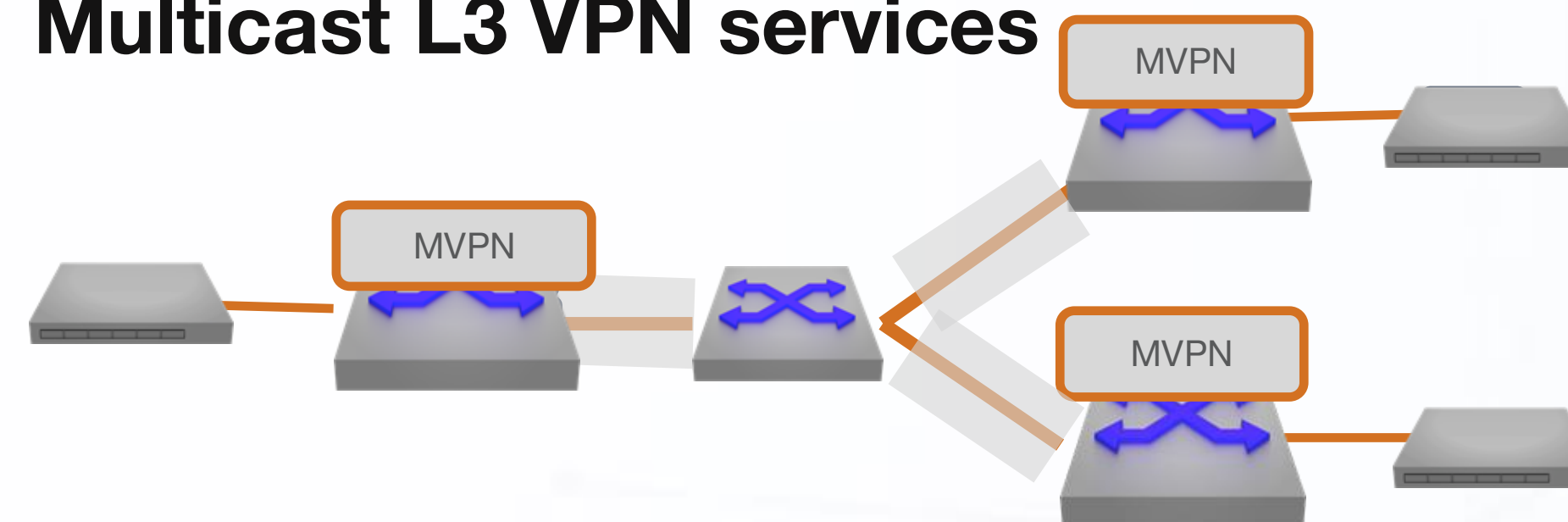
E-LAN (Point-to-multipoint) services



IP-VPN Layer 3 services



Multicast L3 VPN services



Traditional for Brownfield

LDP/BGP signaled PW

- LDP control-plane for signaling
- Type-4, Type-5 modes, Standby PW support
- PW coloring for traffic engineering

VPLS (RFC 7432)

- LDP signaled VPLS support
- LDP signaled with BGP A-D
- Hierarchical VPLS, with standby PWs

IP-VPN (RFC 4364)

- BGP signaling and auto-discovery
- Control-plane learning via BGP
- Layer 3 only solution no L2 VPN support

NG MVPN (RFC 6513)

- BGP signaling and auto-discovery RFC 6514
- P2MP transport with mLDP
- Profile 12, 14 and 22 Supported!

EVPN for Greenfield

EVPN VPWS (RFC 8214)

- Consistent BGP EVPN Control-plane
- Auto-discovery via BGP EVPN (Type-1)
- Support for standards-based A-A/A-S models

EVPN L2 VPNs (RFC 7432)

- Consistent BGP EVPN Control-plane
- Control-plane learning with Type-2 routes
- Support for standards based A-A/A-S models

EVPN L3 VPNs (RFC 9136)

- Consistent BGP EVPN Control-plane
- Control-plane learning with Type-5 routes
- Seamless integration with L2VPN models

EVPN-Multicast (RFC 9625)

- EVPN control-plane with MPLS, replace for NG-MVPN
- Support for Layer 2 and 3 multicast with A-A
- Only mLDP Signaling (today - more to come)

IP Unicast and Ethernet VPNs

EVPN Multihoming
All-Active/Active-Standby

Multicast VPNs

Why EVPN...

... for service delivery ...

Simplification

Protocol Reduction – Single BGP AF for all Ethernet and IP VPN services removing the need for dedicated protocols per services, PW, VPLS, MVPN, IP-VPN

Consistency

Repeatable model – Consistent implementation and operational model for deploying any service type E-LINE, E-LAN, IP unicast/multicast enabling road to automation

Resiliency

Flexible multi-homing – Standard based consistent multi-homing procedures across all Ethernet and IP VPN services

Flexible

Any encapsulation – Support for multiple encapsulation models VXLAN/MPLS/SR and SRv6, allowing a consistent operation model from the WAN to the DC and Campus

Skillset

Converged Teams - One operational and implementation team for managing VPN services across both the WAN and DC infrastructure



CapEx

The simplicity of EVPN removes the need for dedicated HW platforms to for VPN services, service can now be achieved on a standard DC ToR



OpEx

A consistent operational models for MPLS/SR and VXLAN means OpEX saving trough automation and the merging of teams

EVPN

— One Control Plane to rule them all —

MPLS

SR

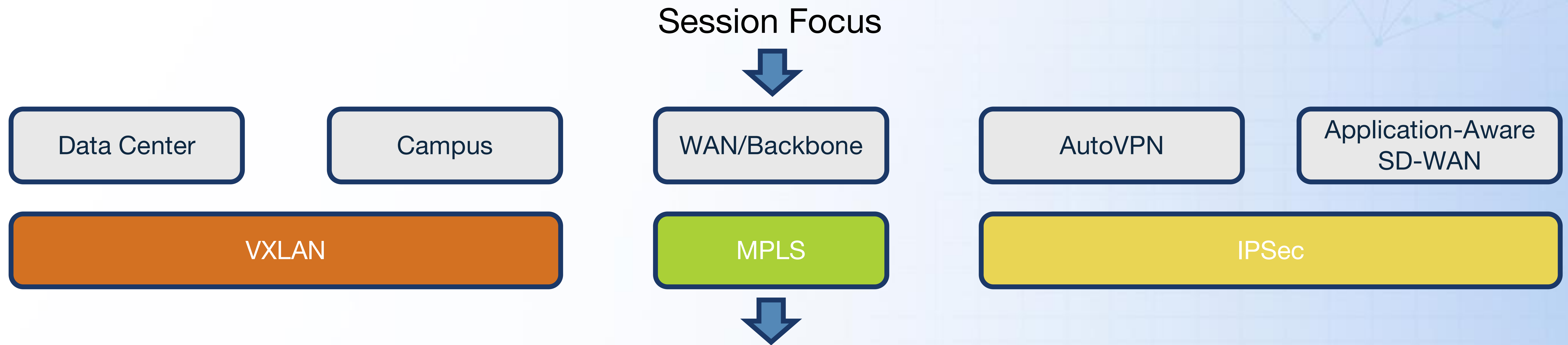
SRv6

VXLAN



Recap: transport models ...

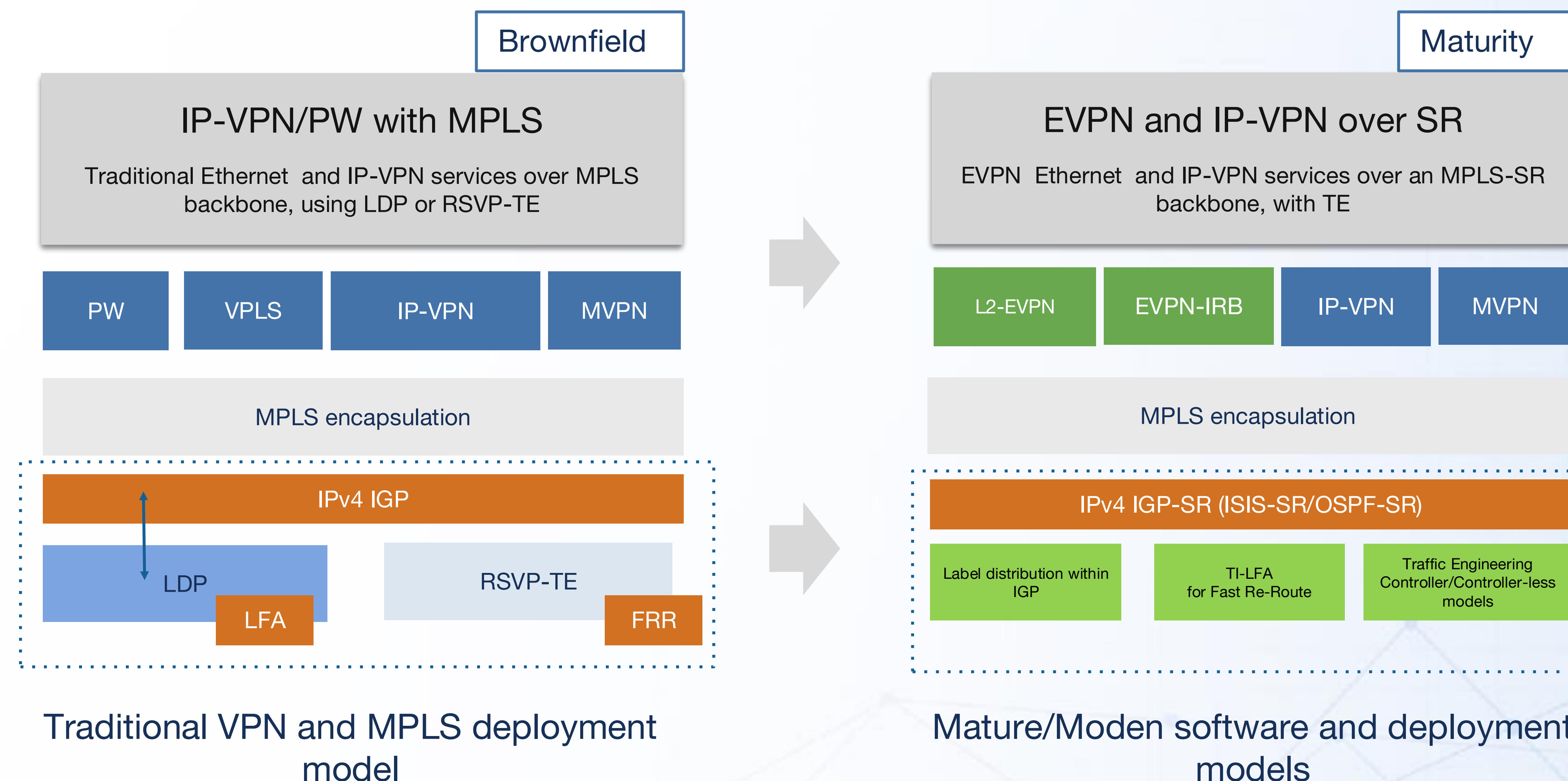
Picking the correct transport for the use cases is critical...



- Market adoption : Segment Routing in combination with EVPN/IP-VPN/L2VPN
- Overlay requirement : L2/L3
- Traffic Engineering: Yes as BW is not freely available
- Service stitching: Yes to the DC
- Cost: feature rich and high BW platforms

SR strategic direction for transport simplification

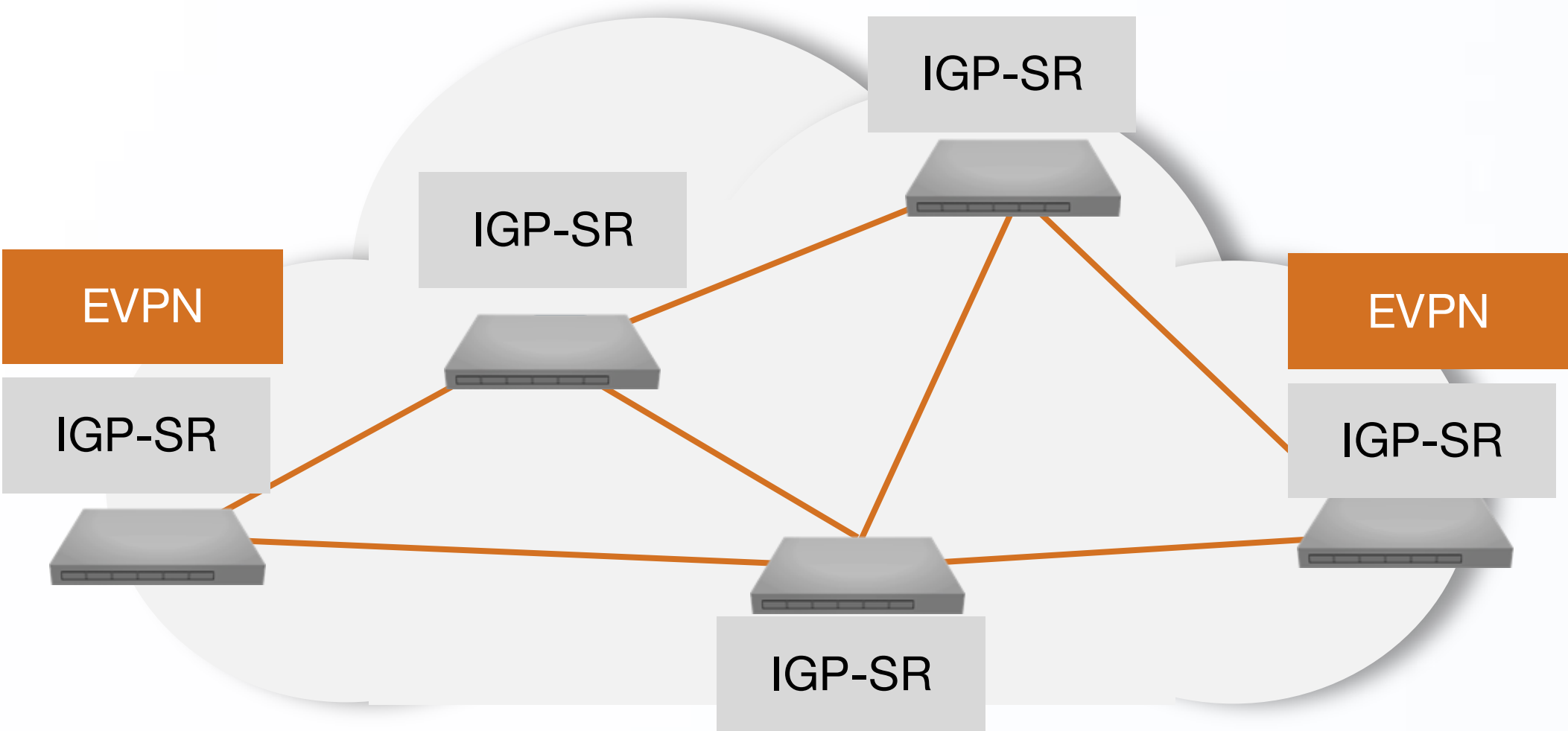
General trend across the industry (Operators and Enterprise) is to simplify Protocol and state reduction to simplify operations and scale



New problems to solve ?

MPLS-SR

Segment Routing Non-Traffic Engineering Applications



- Re-use the IGP for SID distribution
- Protocol reduction no need for LDP and IGP-sync
- Fast reroute capability (50ms) with TI-LFA
- Native ECMP and Anycast capabilities

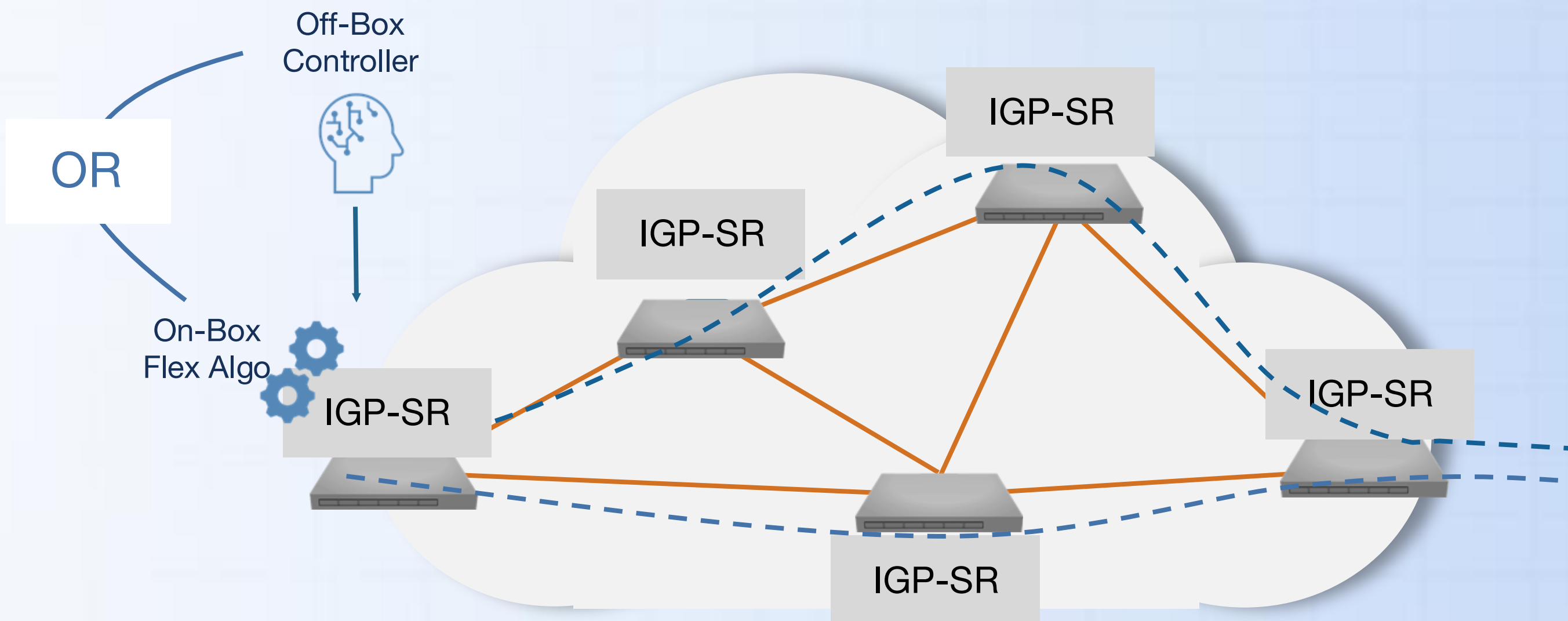


ISIS-SR
TI-LFA



OSPF-SR
TI-LFA

Segment Routing Traffic Engineering Applications



- Replacement of stateful RSVP-TE solutions
- Macro Traffic Engineering with FlexAlgo (Controller-less)
- Micro Traffic Engineering with Controller-based solutions
- Simplification with improved scale due to reduction of state
- Full coverage of SP and Enterprise use cases



Static
SR-TE



FlexAlgo



Dynamic
SR-TE



BGP SR-TE policies
(incl. UCMP for LBW)

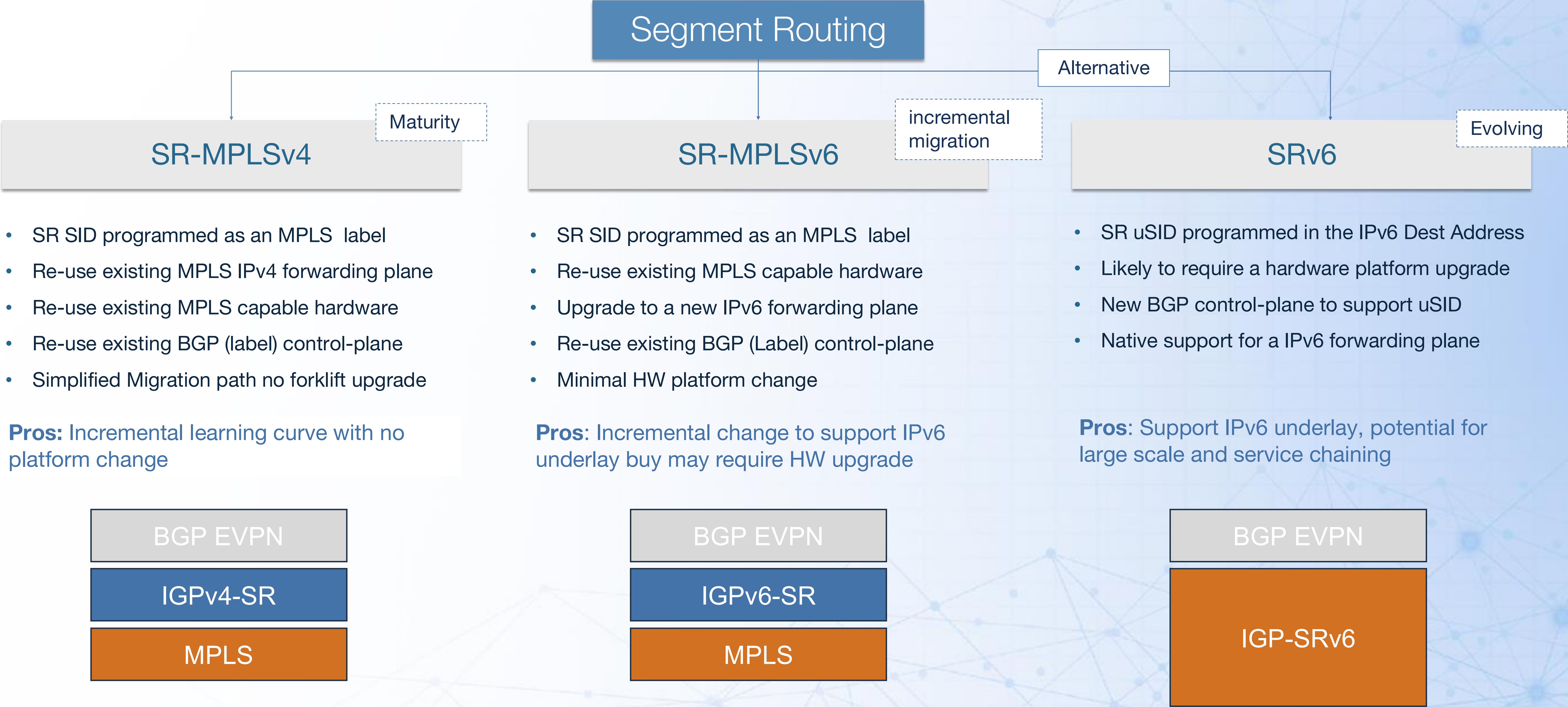


FlexAlgo + CBF



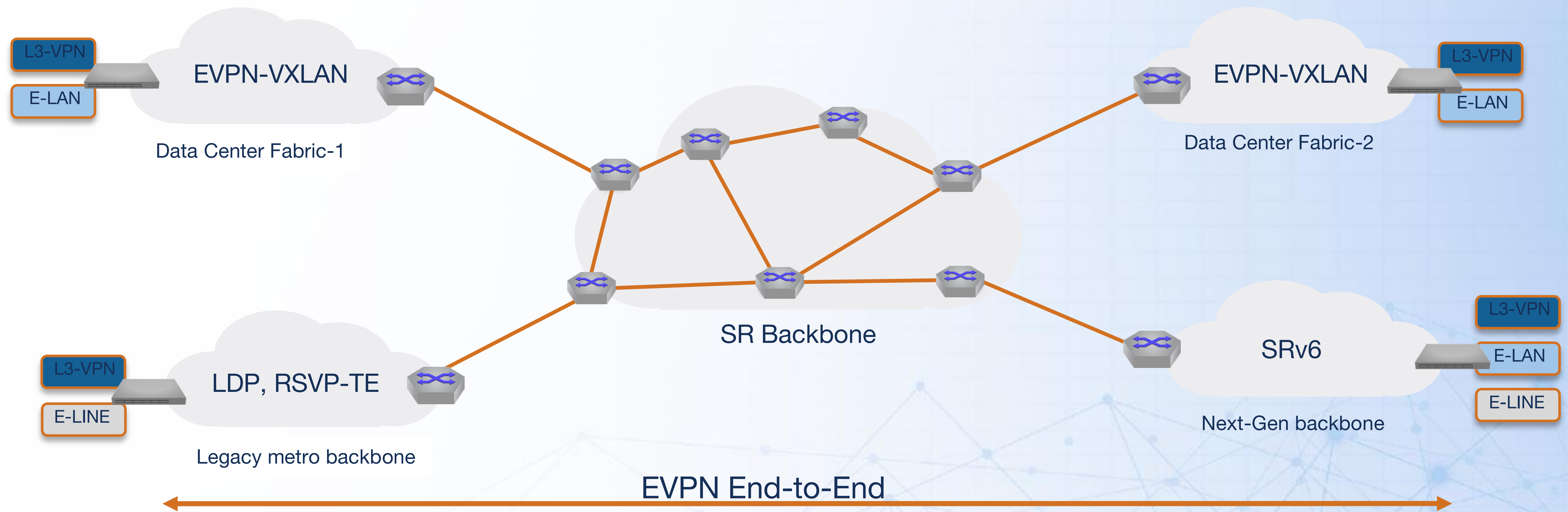
Prefix/Anycast
Segment for FlexAlgo

Simplification with SR



Why Do We Need EVPN Gateways?

- Scaling limits in large EVPN fabrics lead to excessive control-plane state
- Fault containment is harder in flat, single-domain EVPN topologies
- Service continuity demands seamless L2/L3 VPN across multiple domains or sites
- Hybrid interconnects (VXLAN ↔ MPLS) require translation and encapsulation at domain boundaries



Standards That Define EVPN Gateways

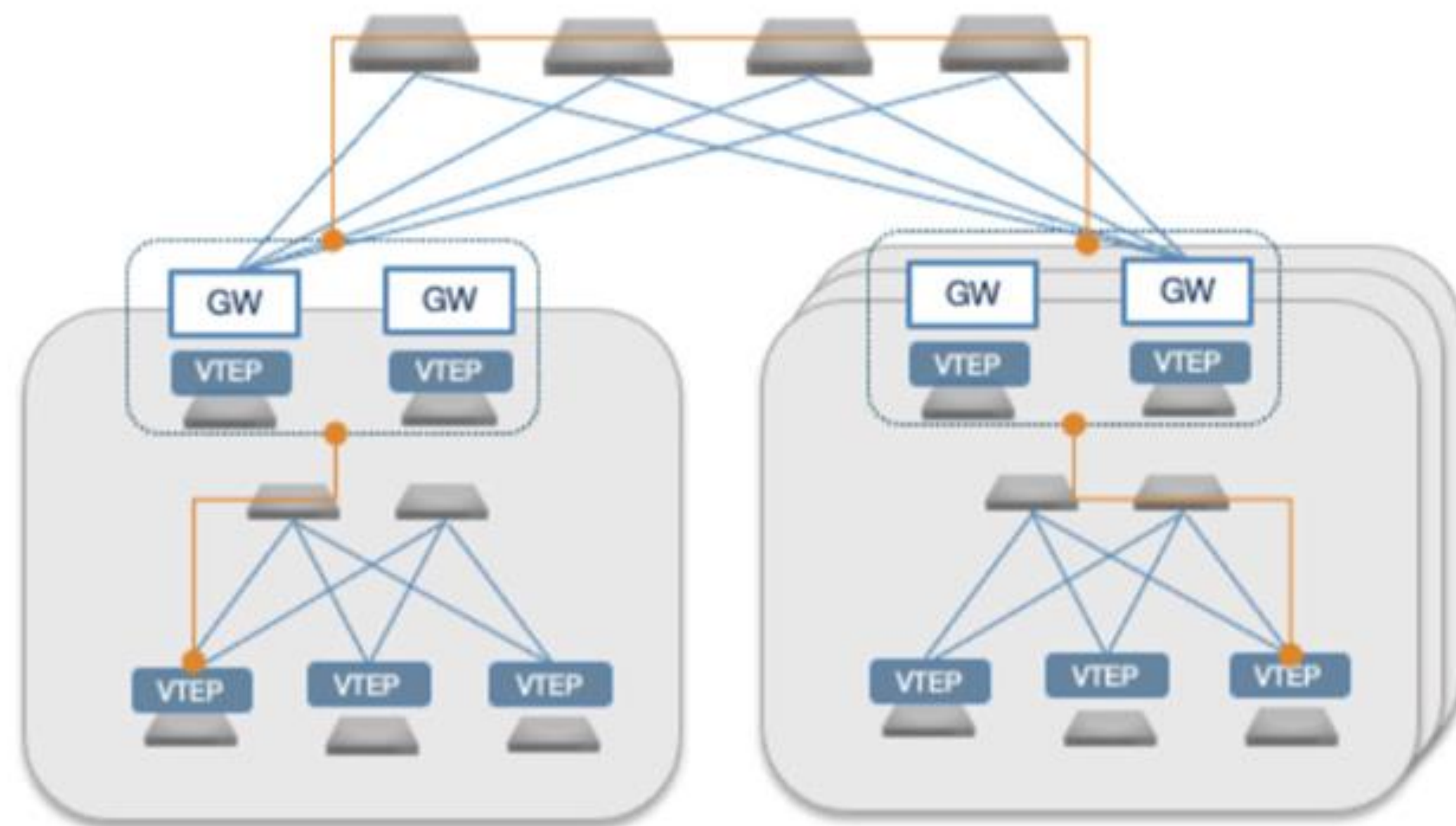
- The IETF BESS Working Group has defined a number of RFCs and drafts that govern EVPN gateway behaviour. These standards enable:
 - Support for both Layer 2 (Type-2 and Type-3) and Layer 3 (Type-5) DCI solutions
 - Interoperability across different BGP address families and control planes
 - Data-plane encapsulation interworking between VXLAN and MPLS

Draft	Overview	
A Network Virtualization Overlay Solution using EVPN RFC 8365	EVPN control plane for L2 VPNs with an NVO environment with VXLAN, NVGRE and GENEVE encap– DCI using GWs and DCI using ASBRs	
EVPN and IP-VPN Integrated Solution draft-ietf-bess-evpn-ipvpn-interworking	Layer 3 DCI interop between EVPN-VXLAN/MPLS and IP-VPN WAN for layer 3 DCI	L3 GW solution
Multi-site EVPN based VXLAN using Border Gateways draft-sharma-bess-multi-site-evpn	GW DCI solution focused only on EVPN-VXLAN, support for a single control planes (EVPN) and single data-plane (VXLAN)	
Interconnect Solution for EVPN Overlay networks RFC 9014	EVPN GW solution for L2 interconnecting of multiple control planes (VPLS/EVPN) and data-planes (MPLS, VXLAN, PBB)	Industry adopted L2 solution
EVPN multicast forwarding for EVPN to EVPN GWs draft-rabnic-bess-evpn-mcast-eeg	EVPN GW solution for providing seamless multicast interconnect between EVPN domains, across VXLAN and MPLS data-planes	
Domain Path (D-PATH) for Ethernet VPN (EVPN) Interconnect Networks draft-sr-bess-evpn-dpath	D-path community for EVPN routes to provide loop-free route advertisement between EVPN domains for layer 2.	

Real-World EVPN Gateway Use Cases

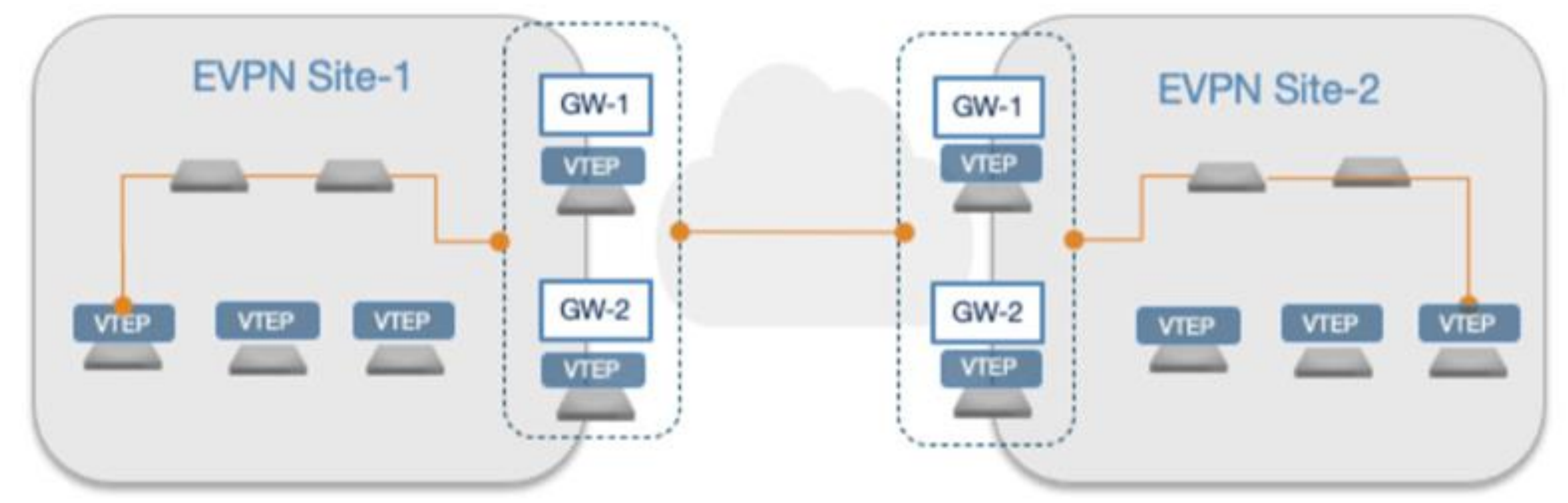
Intra-DC POD Interconnect

- Interconnects multiple EVPN PODs within a single data centre
- Improves scalability by limiting EVPN state to local PODs
- Enhances fault isolation and simplifies control plane convergence



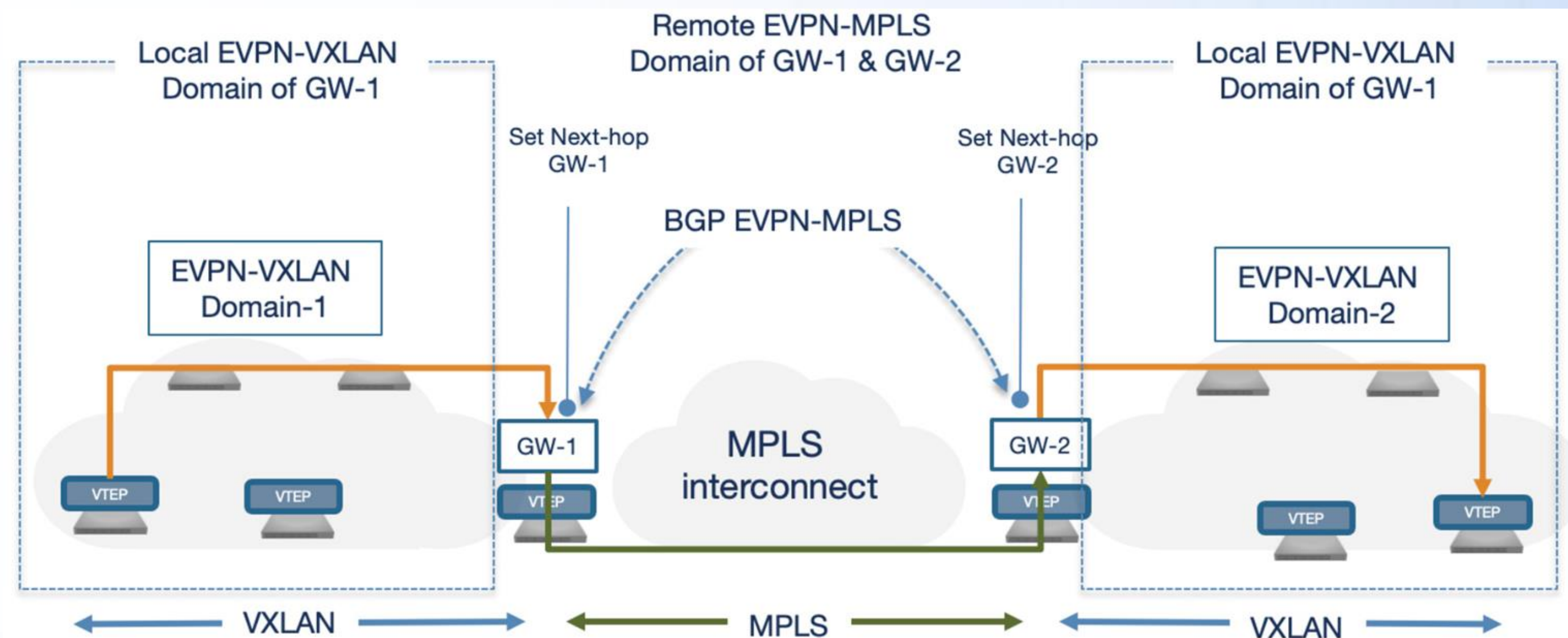
Data Center Interconnect (DCI)

- Seamless L2/L3 VPN connectivity between geographically dispersed sites
- Works over both IP and MPLS backbones
- Enables active-active multi-site designs with local gateway functionality



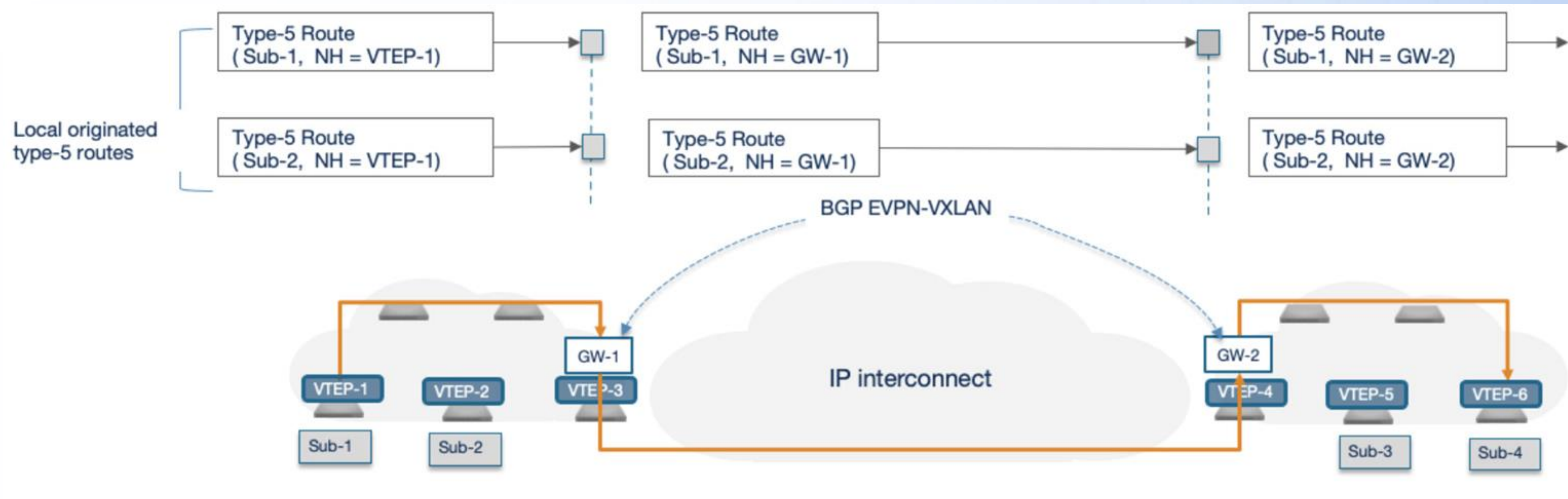
EVPN Gateway: VXLAN ↔ MPLS Interworking

- Used when interconnecting EVPN-VXLAN domains across an MPLS backbone
- Gateway performs control and data-plane translation between VXLAN and MPLS
- EVPN-MPLS routes use VPN and LSP labels; VXLAN routes use VNIs
- Type-2 and Type-5 routes are re-advertised across domains with appropriate encapsulation and next-hop
- Type-1, 3, and 4 routes remain local to each domain
- Supports both Layer 2 bridging and Layer 3 routing between VXLAN and MPLS environments



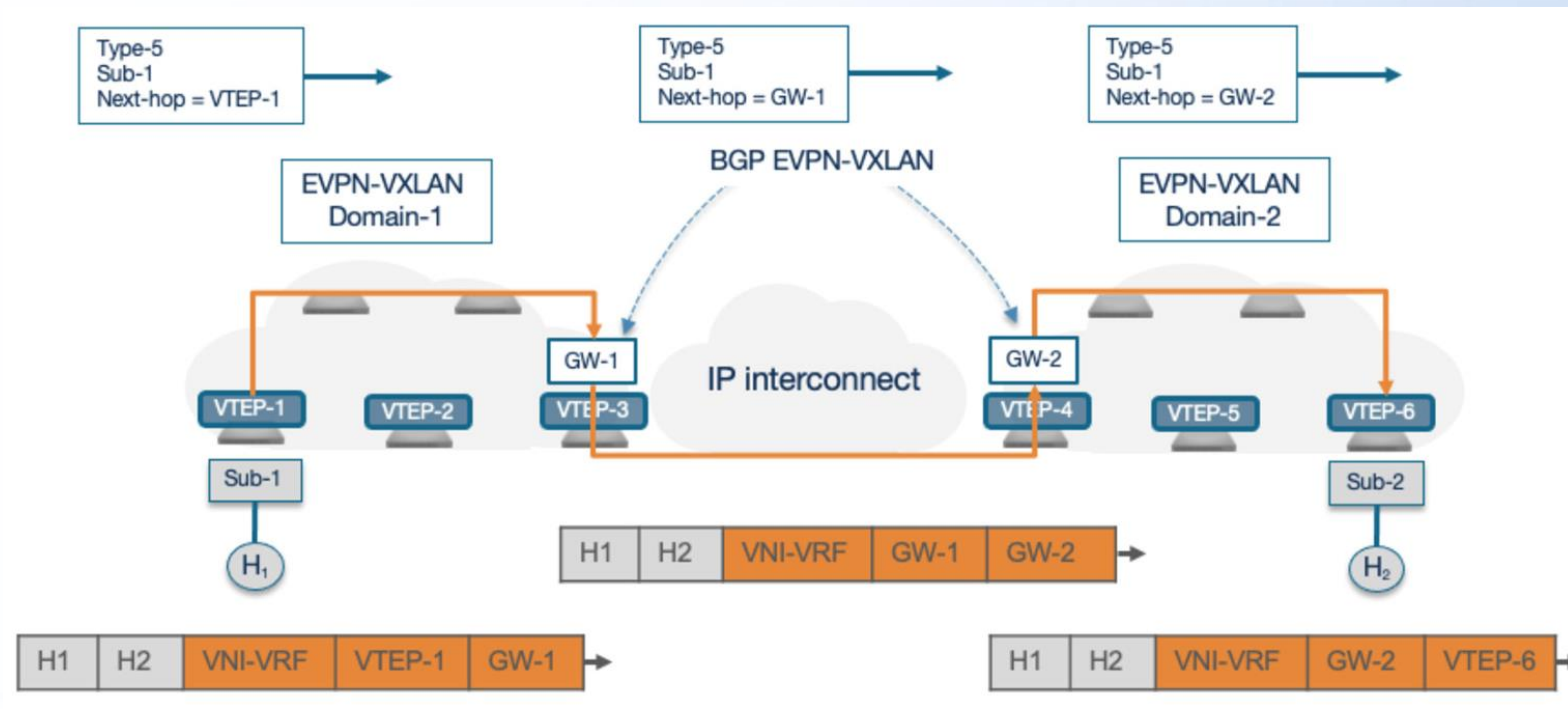
GW Control Plane Operations (Layer 3 Example)

- VTEP advertises a Type-5 (IP prefix) route with its own IP as next-hop
- Local gateway imports the route and re-advertises it to the remote gateway with its own IP as next-hop
- Remote gateway re-advertises the route into its local domain using its own IP as next-hop
- Only gateways require end-to-end IP reachability between domains
- Leaf nodes learn next-hops for local nodes and their local gateway only



GW Data Plane Operations (VXLAN-VXLAN Example)

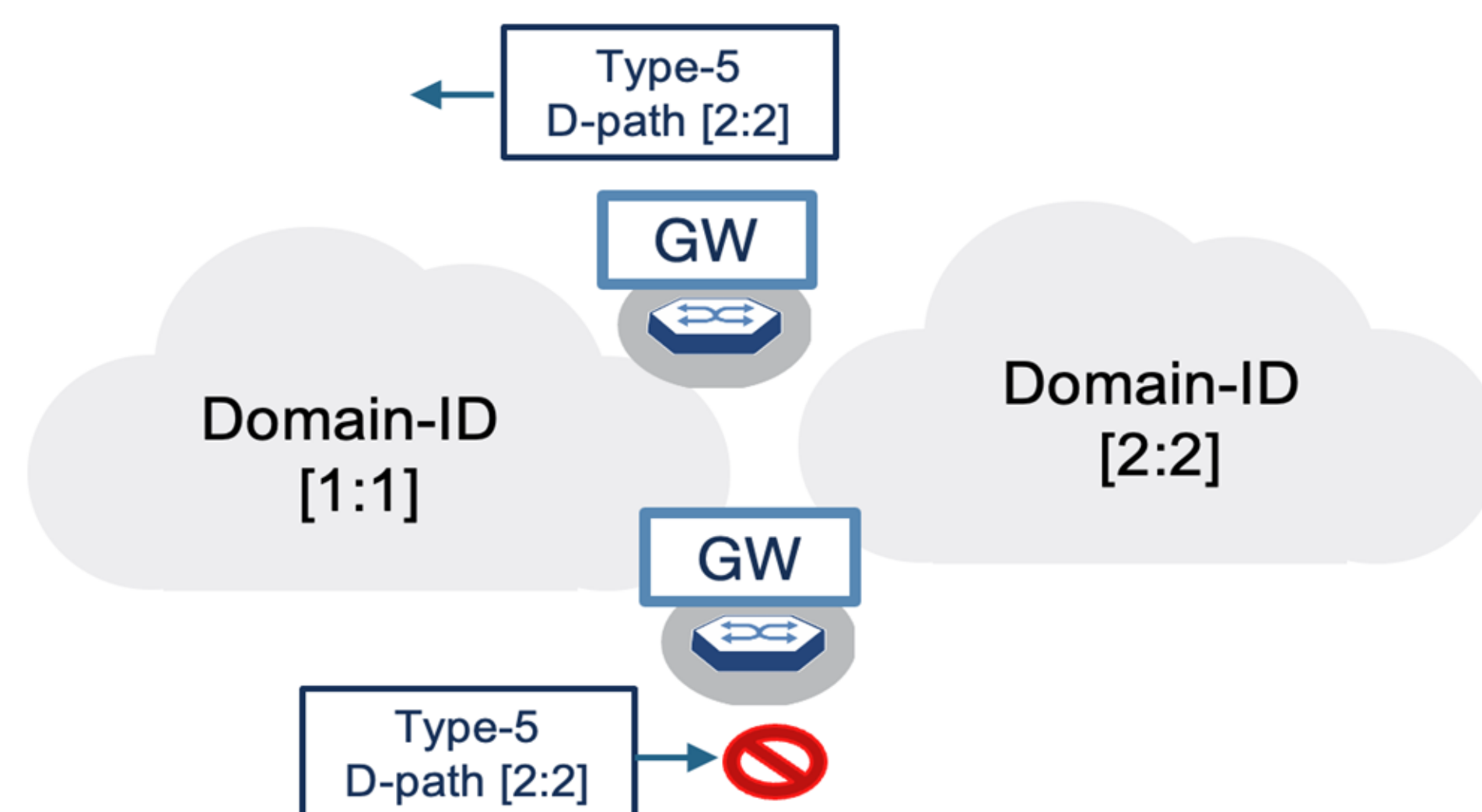
- Traffic from Host-1 is routed by its local VTEP to the local gateway
- Gateway performs a VXLAN decapsulation and L3 route lookup
- Packet is re-encapsulated in VXLAN and forwarded to the remote gateway
- Remote gateway decapsulates, routes the packet, and re-encapsulates for delivery to destination VTEP
- Destination VTEP removes VXLAN and forwards to Host-2



Gateway Enhancements

Domain Path (D-PATH)

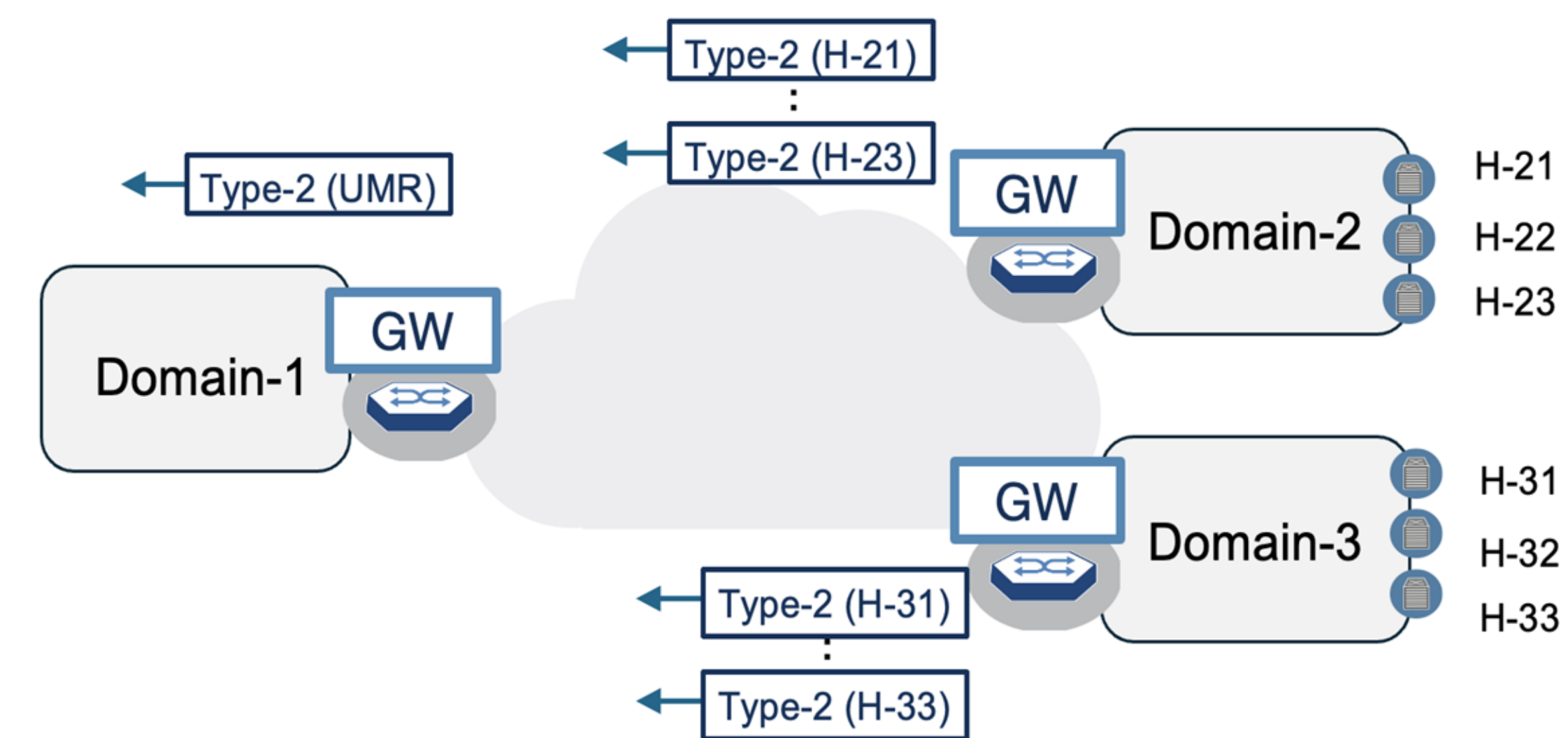
Loop avoidance in EVPN gateway deployments



- Domain Path (D-PATH) BGP attribute is optional and transitive
- Prevents EVPN/IPVPN route loops by tagging routes with source domain ID
- Gateways do not re-advertise routes containing their own domain ID in D-PATH
- Enables control-plane loop prevention across inter-domain mesh

Unknown MAC Route (UMR)

Enables MAC-scale efficiency while preserving correct forwarding

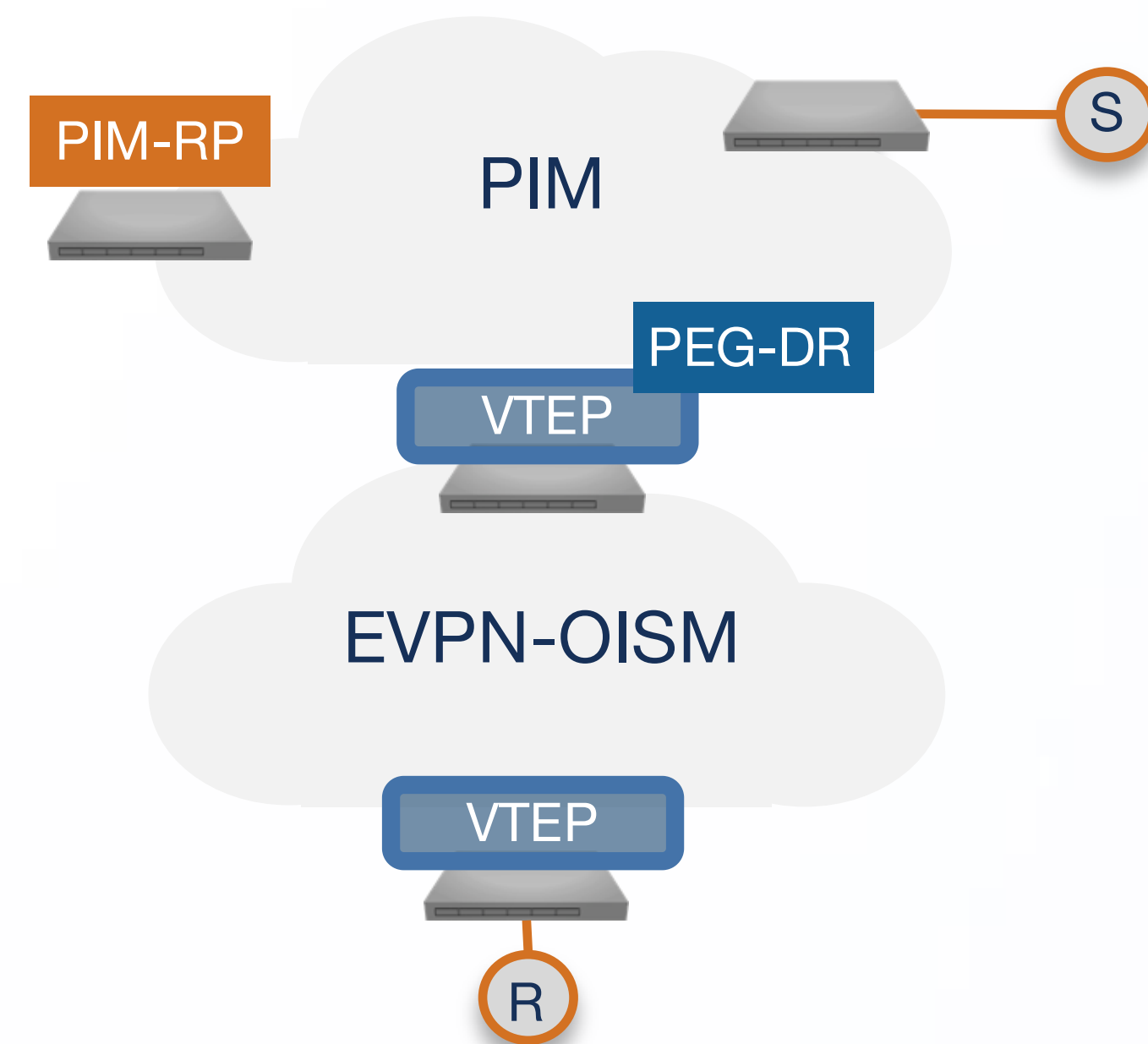


- Unknown MAC Route (UMR) is a special EVPN MAC route with MAC=0 and ESI set to gateway
- Used to reduce MAC scale in multi-domain/DC EVPN, avoiding flooding full MAC tables
- Leaf VTEPs forward unknown traffic to gateway instead of flooding
- Gateway may selectively inject real MAC routes to handle MAC mobility

Completing the Picture: Multicast Joins the Party

PIM EVPN Gateway (PEG)

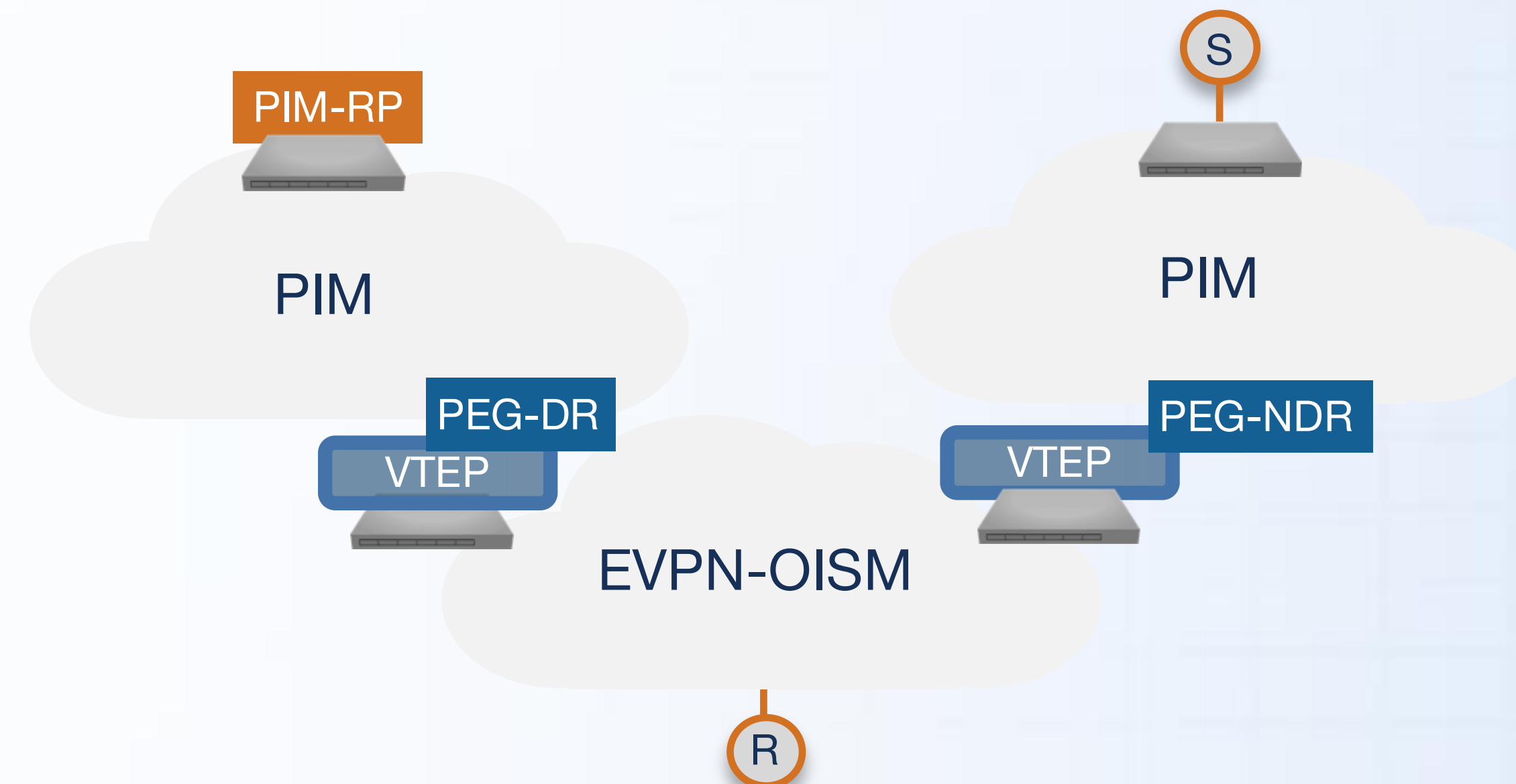
Provide connectivity to external sources and receivers outside of the EVPN OISM domain



- Provide multicast connectivity to external PIM sources and receivers
- PEG node acts as a GW between the EVPN and PIM domains
- Sending PIM joins/registers upstream to the RP or the source

PEG GW for Transit

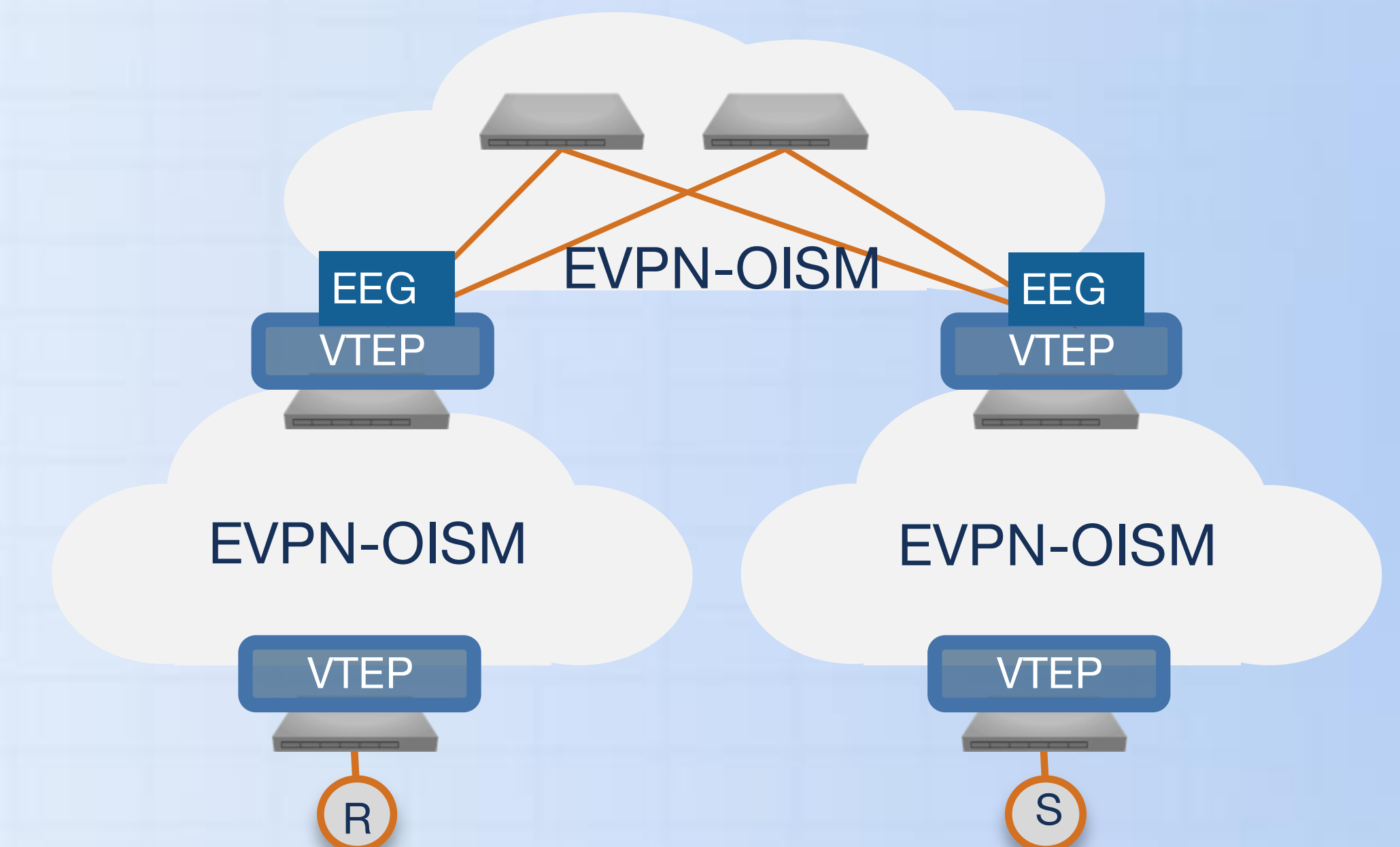
Provide connectivity to external sources and receivers when the connectivity to the PIM source or RP needs to transit the EVPN domain



- PIM domain is split between the EVPN domain
- Provide transit connectivity to both the Source and RP via the EVPN domain
- All PEG nodes are able to process and send PIM joins and registers

EVPN to EVPN GW (EEG)

Provide connectivity to sources and receivers in a different EVPN OISM domain



- Need OISM connectivity between EVPN domain via the EVPN GW
- EEG GW responsible for proxying SMET routes and OISM forwarding across EVPN domains
- GW provides seamless EVPN-VXLAN between domains

Routing Strategy

Single Service Plane

EVPN end-to-end from the Campus to the Data Center, across the WAN or securely internet via SD-WAN

Single consistent end-to-end operational model for Service delivery – EVPN

- Open, standards, No vendor lock-in to prevent end-to-end connectivity
- Convergence of skill-set across` domains (WAN, Campus, DC) – Reduction, in OpEx cost
- Single operational model to troubleshoot, model and automate

Any Transport

Support for the appropriate, cost-effective transport. Campus, DC, WAN and SD-WAN

Connectivity agnostic, allowing appropriate transport layer for the use case

- Cost sensitive high bandwidth - Campus and DC – VXLAN
- Security sensitive WAN – VXLAN over IPSec
- Bandwidth, connectivity restrictive WAN – SR, SR-TE or SRv6 (future)

Future proof

Zero technology lock-in provides incremental upgrades to next-Gen solutions i,e, SRv6

Enables, incremental network upgrade without a lock-in

- Ability to change transport and service layer at the customer's pace
- No need for a big bang approach
- No vendor lock-in to prevent end-to-end connectivity

The background of the slide features a complex network of blue lines and dots, resembling a data or communication network, set against a light blue gradient. The Arista logo is positioned in the top left corner.

ARISTA

Thank You

www.arista.com