ETHERNET VPN (EVPN)
OVERLAY NETWORKS FOR ETHERNET SERVICES

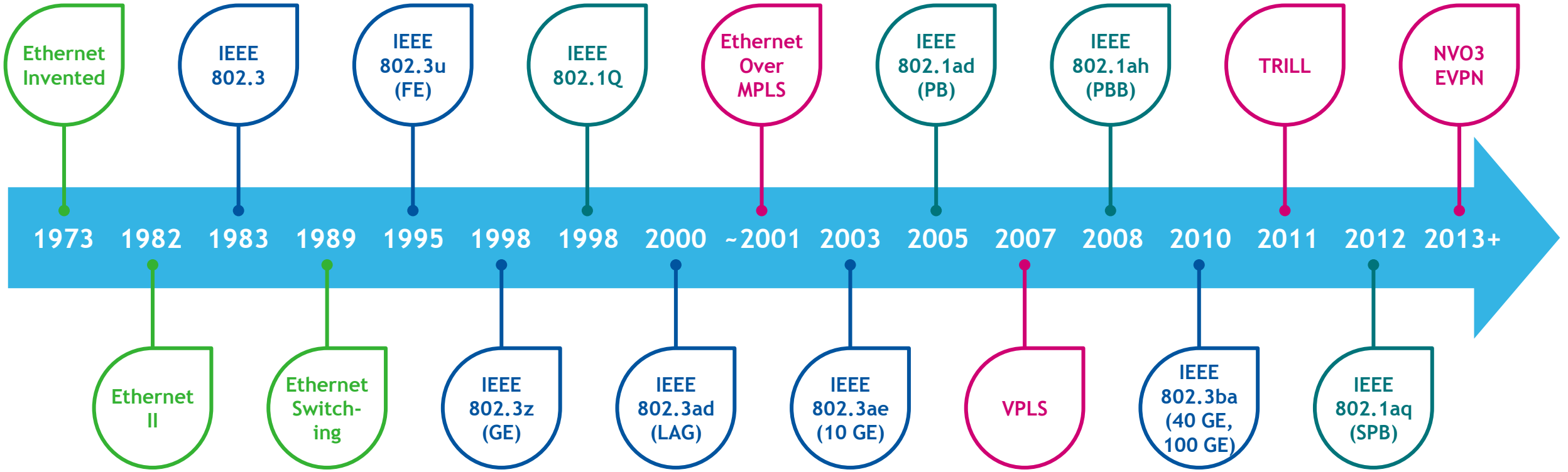
Greg Hankins <greg.hankins@alcatel-lucent.com>
NANOG 61

Alcatel·Lucent

AGENDA

1. EVPN Background and Motivation
2. EVPN Operations
3. EVPN Use Cases

ETHERNET SERVICES TECHNOLOGY CONTINUES TO EVOLVE
HIGHER SPEEDS AND ADVANCED CARRIER-GRADE SERVICES



"The **widespread adoption of Ethernet L2VPN** services and the advent of **new applications** for the technology (e.g., data center interconnect) have culminated in a **new set of requirements** that are **not readily addressable** by the current Virtual Private LAN Service (VPLS) solution." — draft-ietf-l2vpn-evpn-req

WHY ANOTHER VPN TECHNOLOGY?

- MPLS/VPLS and PBB are both proven technologies for Ethernet services, but
  - The control plane approach hasn't changed
  - Still relies on flooding and learning to build the Layer 2 forwarding database (FDB)
- EVPN introduces a new model for delivery of Ethernet services
  - Inherits a decade of VPLS operational experience in production networks
  - Incorporates flexibility for service delivery over Layer 3 networks
  - Abstracts and separates the control and data planes: MP-BGP carries MAC/IP routing information, choice of data plane encapsulation
- Enables network operators to meet emerging needs in their networks
  - Data center interconnect (DCI)
  - Cloud and virtualization services
  - Integrated Layer 2 and Layer 3 VPN services
  - Overlay technologies that simplify topologies, and remove protocols from the network

EVPN KEY OPERATIONAL BENEFITS

Integrated Services

- Delivering Layer 2 and Layer 3 services over the same interface, VLAN and VPN
- L3VPN-like operation for scalability and control

Network Efficiency

- Multihoming with all-active forwarding, load balancing between PEs
- Optimized multidestination frame (BUM) delivery
- More efficient hybrid service delivery over a single interface or VLAN

Design Flexibility

- MPLS or IP data plane encapsulation choices
- VXLAN encapsulation enables EVPN over a simple IP network
- Simpler provisioning and management with a single VPN technology

Greater Control

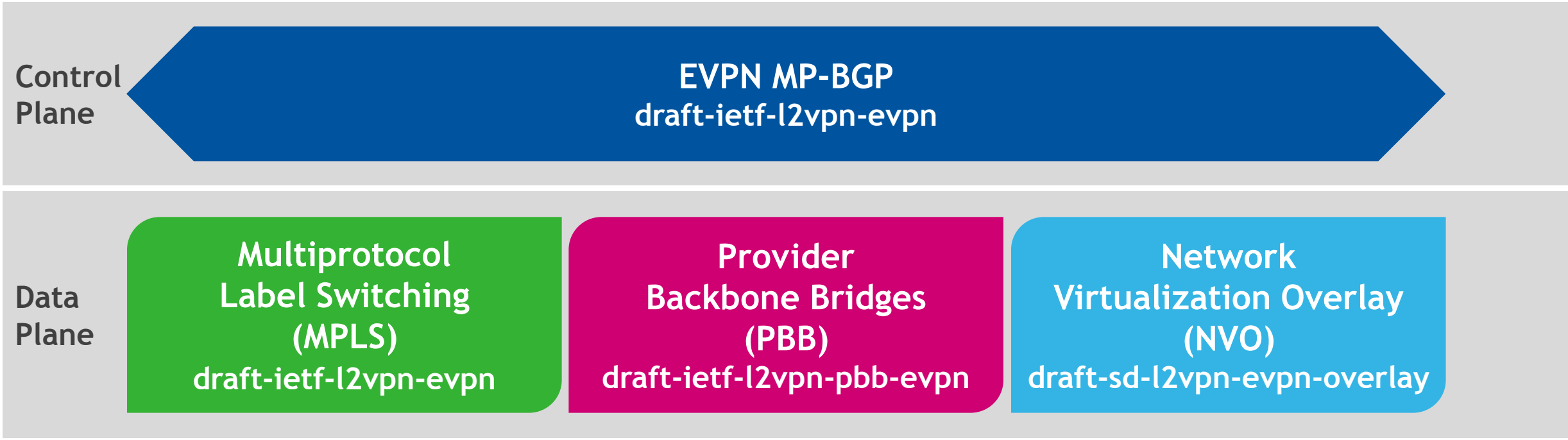
- MAC/IP provisioning enables programmatic network control
- Consistent signaled FDB in control plane vs. flood-and-learn FDB in data plane
- Proxy ARP/ND functionality allows PEs to respond to ARP/ND requests

EVPN STATUS

- Hot new technology in the IETF L2VPN WG
- Many mature base I-Ds becoming RFCs, many new I-Ds
  - RFC 7209: Requirements for Ethernet VPN (EVPN)
  - draft-ietf-l2vpn-evpn base specification: WG last call for -07 on May 9, 2014
  - draft-ietf-l2vpn-pbb-evpn: no more changes expected
- Diverse authors on requirements and base specification
  - Vendors: Alcatel-Lucent, Cisco, Juniper
  - Network operators: Arktan, AT&T, Bloomberg, Verizon
- Shipping implementations
  - Alcatel-Lucent
  - Cisco
  - Juniper

draft-allan-l2vpn-mldp-evpn
draft-boutros-l2vpn-evpn-vpws
draft-boutros-l2vpn-vxlan-evpn
draft-ietf-l2vpn-evpn
draft-ietf-l2vpn-pbb-evpn
draft-ietf-l2vpn-spbm-evpn
draft-ietf-l2vpn-trill-evpn
draft-jain-l2vpn-evpn-lsp-ping
draft-li-l2vpn-evpn-mcast-state-ad
draft-li-l2vpn-evpn-pe-ce
draft-li-l2vpn-segment-evpn
draft-rabadan-l2vpn-dci-evpn-overlay
draft-rabadan-l2vpn-evpn-prefix-advertisement
draft-rp-l2vpn-evpn-usage
draft-sajassi-l2vpn-evpn-etree
draft-sajassi-l2vpn-evpn-inter-subnet-forwarding
draft-sajassi-l2vpn-evpn-ipvpn-interop
draft-sajassi-l2vpn-evpn-vpls-integration
draft-salam-l2vpn-evpn-oam-req-frmwk
draft-sd-l2vpn-evpn-overlay
draft-vgovindan-l2vpn-evpn-bfd
draft-zhang-l2vpn-evpn-selective-mcast
draft-zheng-l2vpn-evpn-pm-framework
RFC 7209: Requirements for Ethernet VPN (EVPN)

EVPN DATA PLANES
ONE EVPN CONTROL PLANE WITH MULTIPLE DATA PLANE OPTIONS



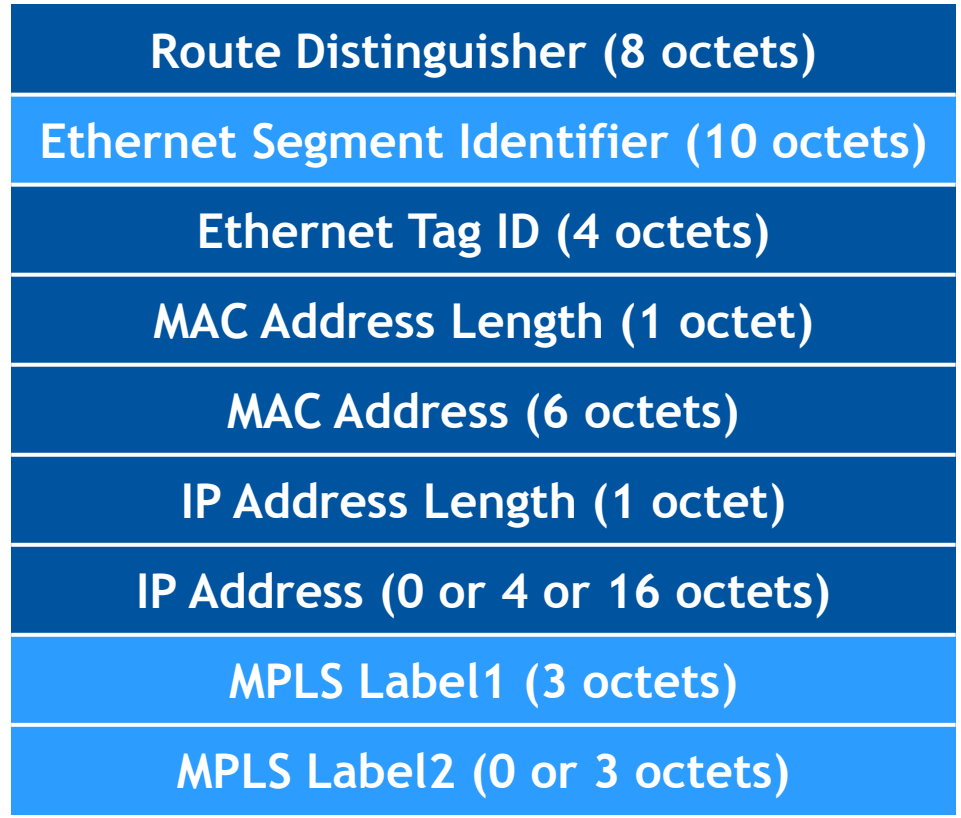
- EVPN over MPLS for E-LAN services
- All-active multihoming for VPWS
- RSVP-TE or LDP MPLS protocols

- EVPN with PBB PE functionality for scaling very large networks over MPLS
- All-active multihoming for PBB-VPLS

- EVPN over NVO tunnels (VXLAN, NVGRE, MPLSoGRE) for data center fabric encapsulations
- Provides Layer 2 and Layer 3 DCI and overlays over simple IP networks

EVPN CONTROL PLANE LEARNING WITH MP-BGP

- Brings proven and inherent BGP control plane scalability to MAC routes
  - Consistent signaled FDB in any size network instead of flooding
  - Even more scalability and hierarchy with route reflectors
- BGP advertises MACs and IPs for next hop resolution with EVPN NLRI
  - AFI = 25 (L2VPN) and SAFI = 70 (EVPN)
  - Fully supports IPv4 and IPv6 in the control and data plane
- Offers greater control over MAC learning
  - What is signaled, from where and to whom
  - Ability to apply MAC learning policies
- Maintains virtualization and isolation of EVPN instances
- Enables traffic load balancing for multihomed CEs with ECMP MAC routes



MAC Advertisement Route

(Light Blue Fields are Not Used in all Data Planes)

EVPN CONCEPTS OVERVIEW

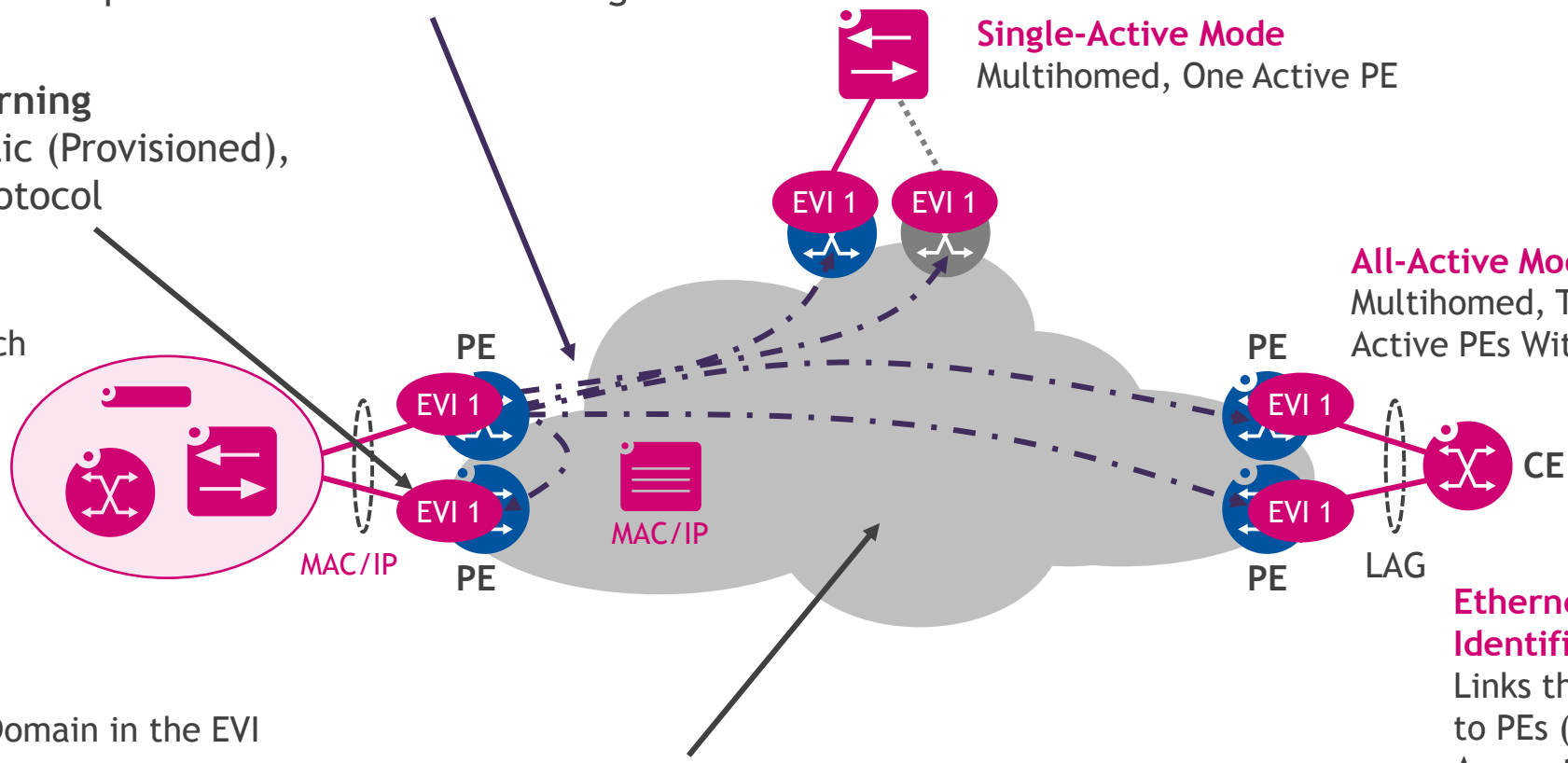
Control Plane Learning
PEs Advertise MAC Addresses and Next Hops From Connected CEs Using MP-BGP

Data Plane Learning
Dynamic or Static (Provisioned), Management Protocol

Customer Edge (CE)
Host, Router or Switch

EVPN Instance (EVI)
Identifies a VPN

Ethernet Tag
Broadcast or Bridge Domain in the EVI



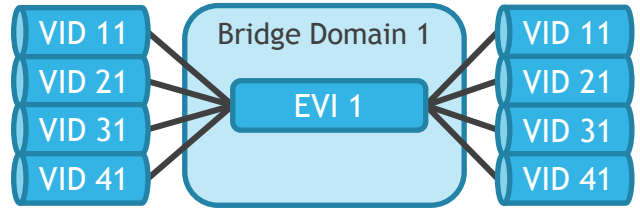
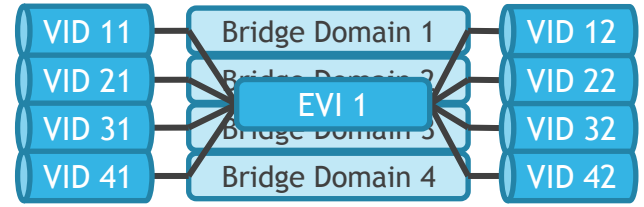
Single-Active Mode
Multihomed, One Active PE

All-Active Mode
Multihomed, Two or More Active PEs Without MC-LAG

Ethernet Segment Identifier (ESI)
Links that Connect the CE to PEs (ESIs are Unique Across the Network)

Data Plane Encapsulation
MPLS or IP

EVPN SERVICE INTERFACES OVERVIEW

	VLAN Based Service Interface	VLAN Bundle Service Interface	VLAN Aware Bundle Service Interface
Service Interface Diagram			
Mapping Between VLAN ID and EVI	1:1	N:1	N:1
Customer VID per EVI	Single	Multiple	Multiple
Bridge Domains per EVI	Single	Single	Multiple
Overlapping MACs Across VLANs	✓	✗	✓
VLAN Translation	✓	✗	✓

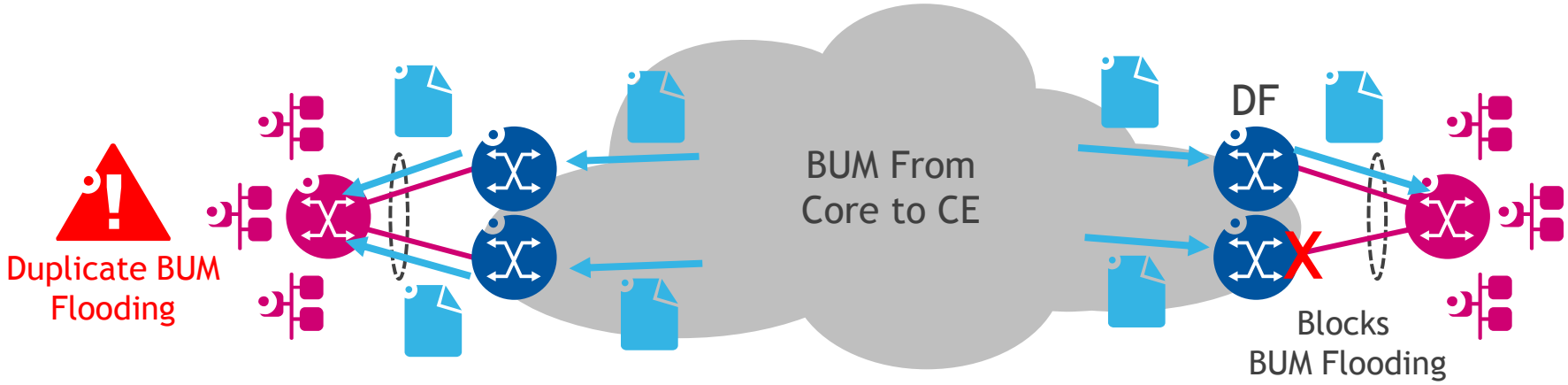
AGENDA

1. EVPN Background and Motivation
2. **EVPN Operations**
3. EVPN Use Cases

EVPN OPERATION

- Key features control plane features
  - All-Active Multihoming and Designated Forwarder Election
  - All-Active Multihoming and Split Horizon
  - Proxy ARP/ND and Unknown Unicast Flooding Suppression
  - Aliasing
  - MAC Mobility
  - MAC Duplication
  - MAC Mass-Withdraw
  - Default Gateway Inter-Subnet Forwarding
- Data planes
  - MPLS: EVPN-MPLS
  - PBB: PBB-EVPN
  - VXLAN: EVPN-VXLAN

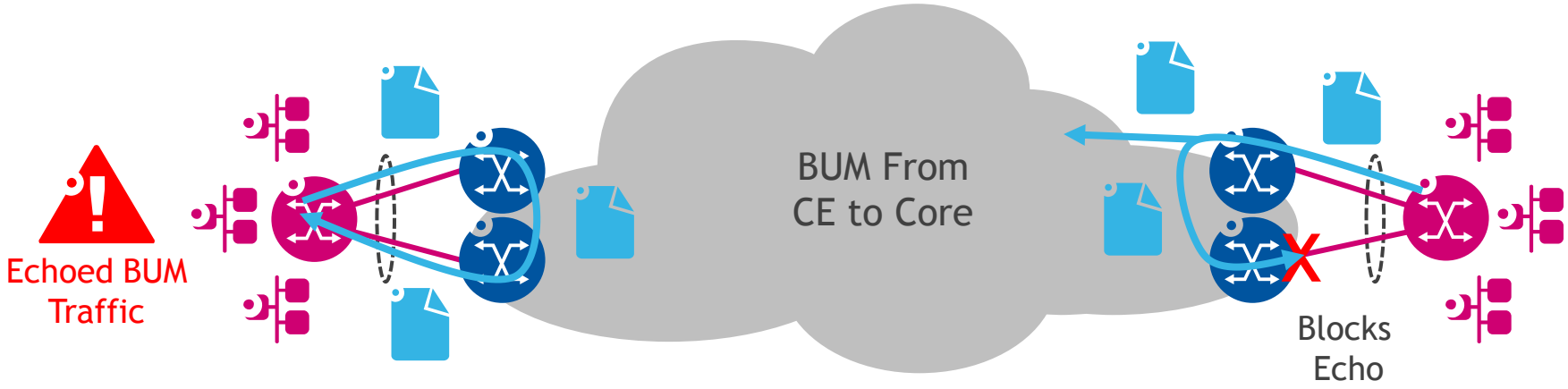
EVPN OPERATION
ALL-ACTIVE MULTIHOMING AND DESIGNATED FORWARDER ELECTION



- Avoids duplicate BUM flooding to all-active CEs
- PEs connected to multihomed CEs know about each other through ESI routes
- Elects a designated forwarder (DF) responsible for BUM flooding to the Ethernet segment
- Non-DF PEs block BUM flooding to the CE

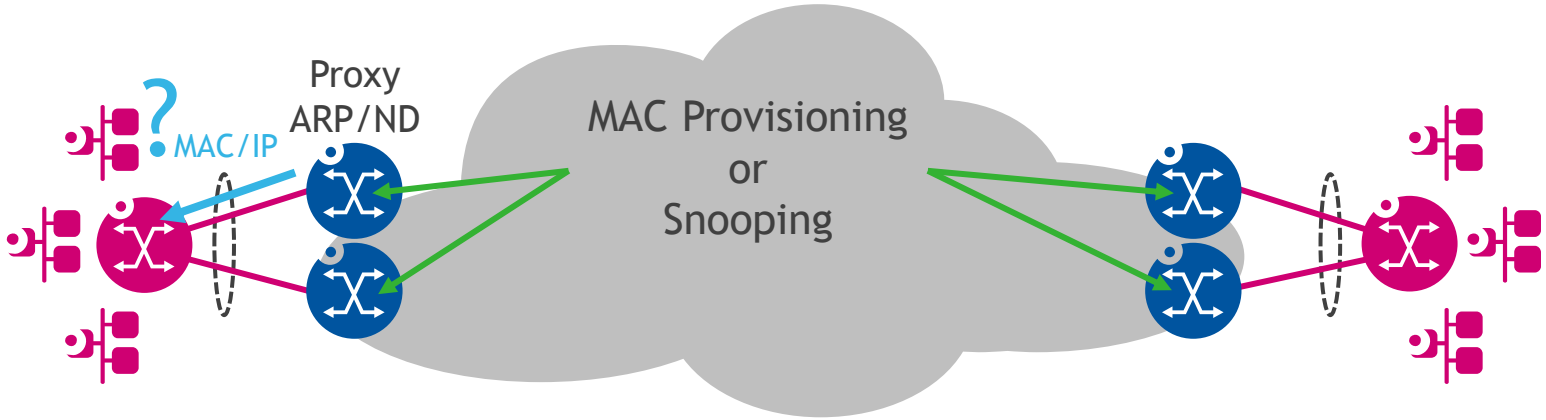
- Flexible DF election and functionality
  - Same DF for all ESIs
  - Different DF per ESI
- Unicast still follows all-active paths

EVPN OPERATION
ALL-ACTIVE MULTIHOMING AND SPLIT HORIZON



- Ensures that BUM traffic from an ESI is not replicated back to the same ESI to an all-active CE
- PE advertises a split horizon label for each all-active Ethernet segment
- When an ingress PE floods BUM traffic, it pushes the split horizon label to identify the source Ethernet segment
- Egress PEs use this label for split horizon filtering and drop packets with the label destined to the Ethernet segment
- Implicit split horizon for core, since PEs won't flood received BUM traffic back into core

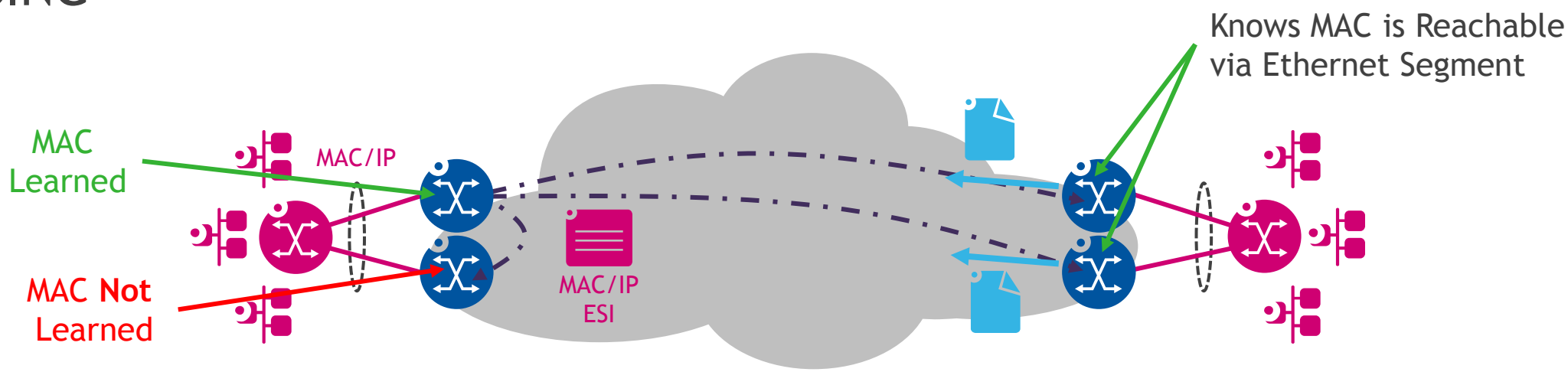
EVPN OPERATION
PROXY ARP/ND AND UNKNOWN UNICAST FLOODING SUPPRESSION



- ARP/ND is a security issue and a scalability issue in large networks
  - Unknown unicast traffic levels, especially in large data center and IXP networks
- We really don't need it anymore in orchestrated or provisioned networks where all MACs/IPs are known

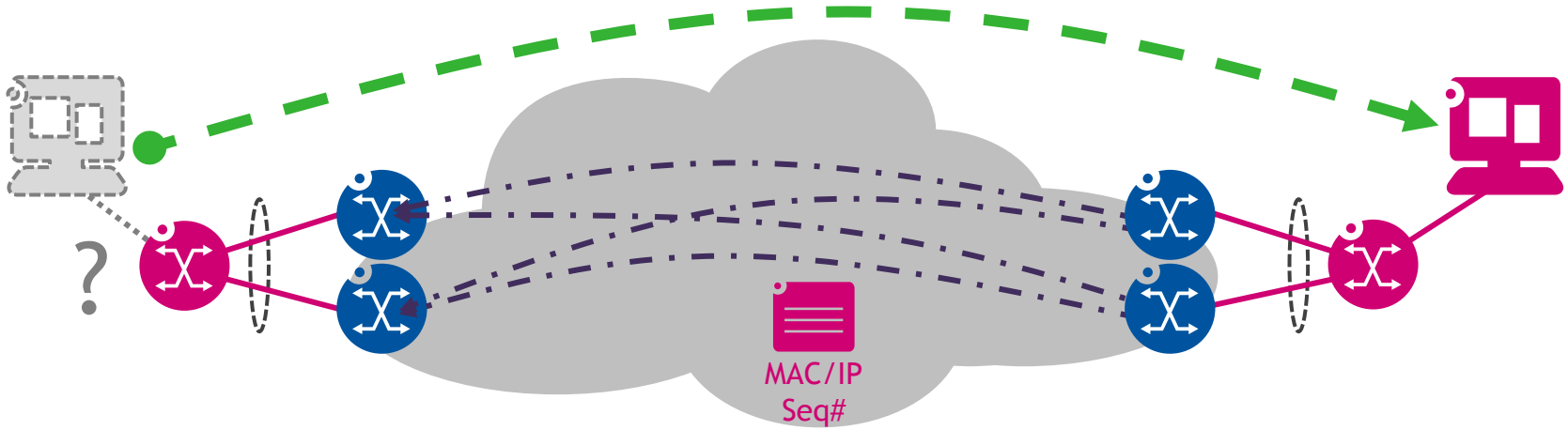
- EVPN can reduce or suppress unknown unicast flooding since all active MACs and IPs are advertised by PEs
  - PEs proxy ARP/ND based on MAC route table to CEs
  - ARP/ND/DHCP snooping optimizes and reduces unknown unicast flooding, useful in dynamic data center networks
  - Provisioning MAC addresses can reduce or eliminate unknown unicast flooding entirely
  - Can disable learning and snooping for programmatic network control

EVPN OPERATION
ALIASING



- Provides load balancing to all-active CE when the MAC address is only learned by one PE
  - First MAC learning by PE is usually from a Layer 2 broadcast (ARP/ND/DHCP)
  - Broadcasts are sent on the primary link in a LAG
  - Can have periods of time when the MAC is only learned by the PE connected to the primary link
- PEs advertise the ESI in MAC routes with all-active mode
- Remote PEs can load balance traffic across all PEs advertising the same ESI
  - Multipathing to CE always works, does not depend on random learning situations or hashing at CE
- Can also be used for a backup path in single-active mode with a standby link

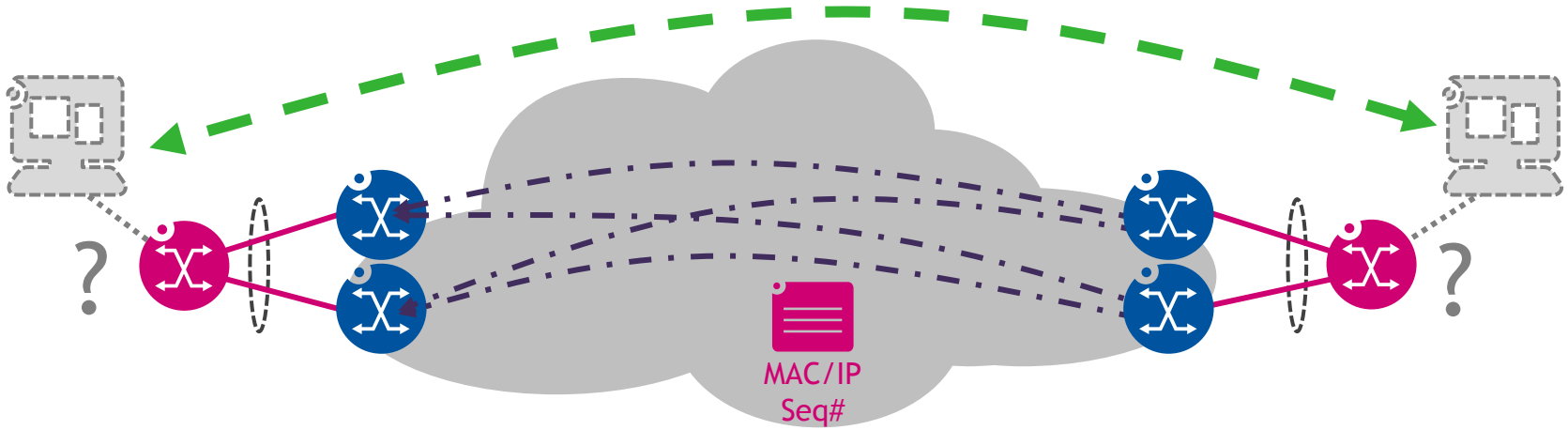
EVPN OPERATION
MAC MOBILITY



- MAC addresses may move between ESIs
- If local learning is used, the PE may not detect that a MAC address has moved and won't send a withdraw for it
- New PE sends a new MAC route
- Now there are two routes for the MAC address: an old wrong one and a new correct one

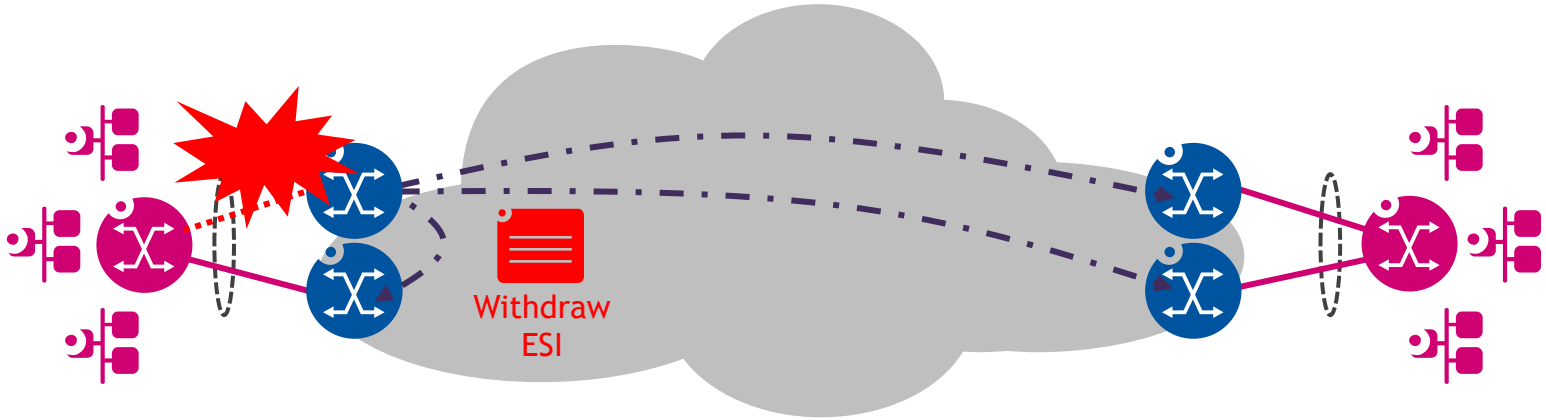
- Each MAC is advertised with a MAC mobility sequence number in an extended community with the MAC route
  - PE selects the MAC route with the highest sequence number
  - Triggers withdraw from PE advertising MAC route with the lower sequence number
  - Lowest PE IP address is used as the tie breaker if the sequence number is the same

EVPN OPERATION
MAC DUPLICATION



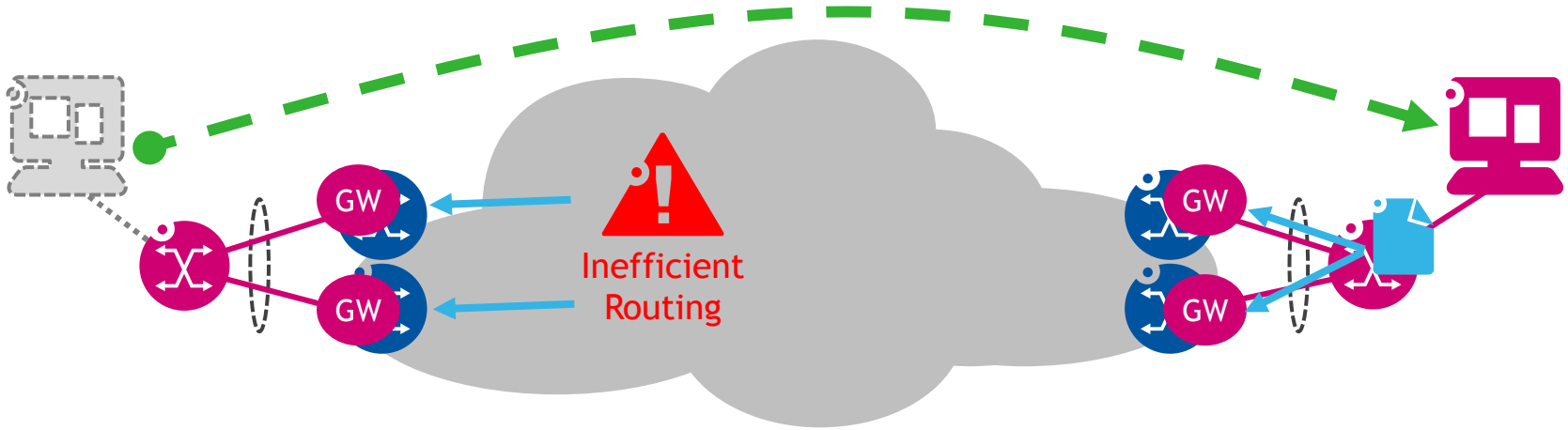
- In certain bad situations, the same MAC could be learned by two PEs
  - MAC duplication
  - Rapid movement
  - Loops
- MAC duplication detection mechanism uses a configurable timer and move counter
  - Provides per-MAC duplication control vs. per-port control in Layer 2 bridging
- If five (*N*) moves (*M*) are detected in 180 s, then the MAC is considered duplicated (default timers)
- PEs stop advertising its route, PEs will use the route with the highest sequence number for forwarding
- Condition can be cleared manually or by implementing a retry timer to clear it automagically

EVPN OPERATION
MAC MASS-WITHDRAW



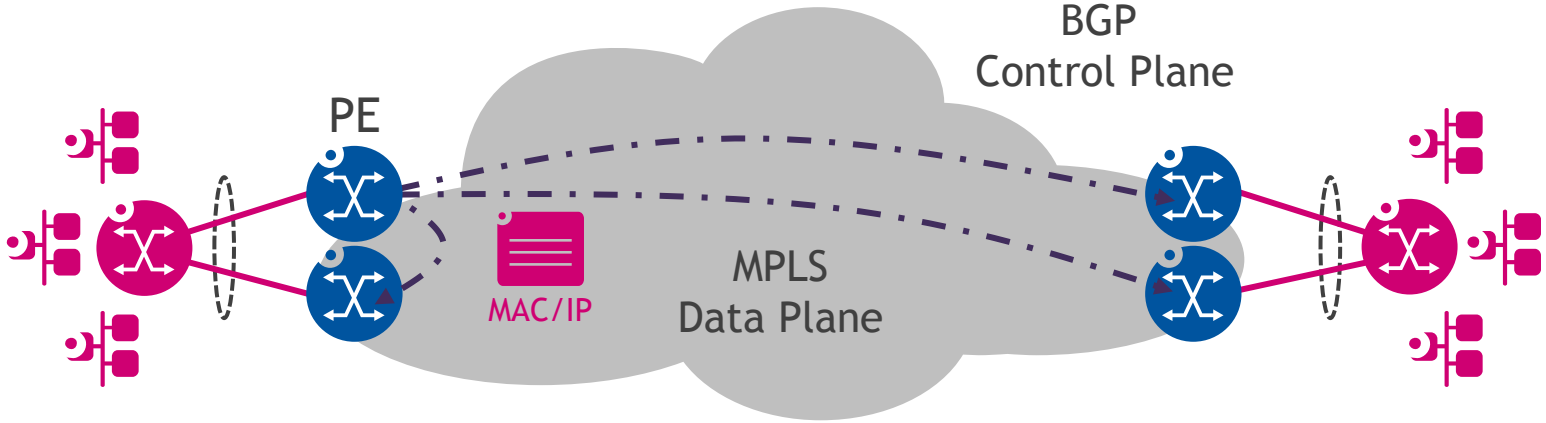
- Provides rapid convergence when a link failure affects many MAC addresses
- PEs advertise two routes
  - MAC/IP address and its ESI
  - Connectivity to ESIs
- If a failure affects an ESI, the PE simply withdraws the route for the ESI
- Remote PEs remove failed PE from the path for all MAC addresses associated with an ESI
- Functions as a MAC mass-withdraw and speeds convergence during link failures
- No need to wait for individual MAC addresses to be withdrawn

EVPN OPERATION
DEFAULT GATEWAY INTER-SUBNET FORWARDING



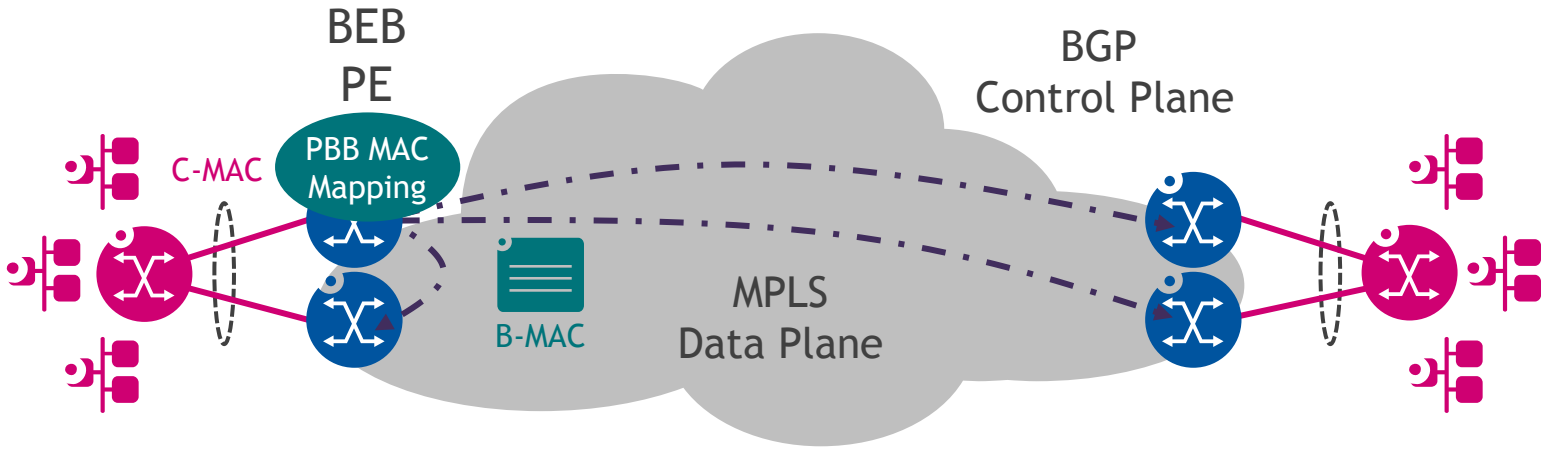
- EVPN supports inter-subnet forwarding when IP routing is required
- No additional separate L3VPN functionality is needed, uses EVPN default gateway
- One or more PEs is configured as the default gateway, 0.0.0.0 or :: MAC route is advertised with default gateway extended community
- Local PEs respond to ARP/ND requests for default gateway
- Enables efficient routing at local PE
- Avoids tromboning traffic across remote PEs to be routed after a MAC moves, if all default gateways use the same MAC address

EVPN MULTIPROTOCOL LABEL SWITCHING (MPLS) DATA PLANE
DRAFT-IETF-L2VPN-EVPN (EVPN-MPLS)



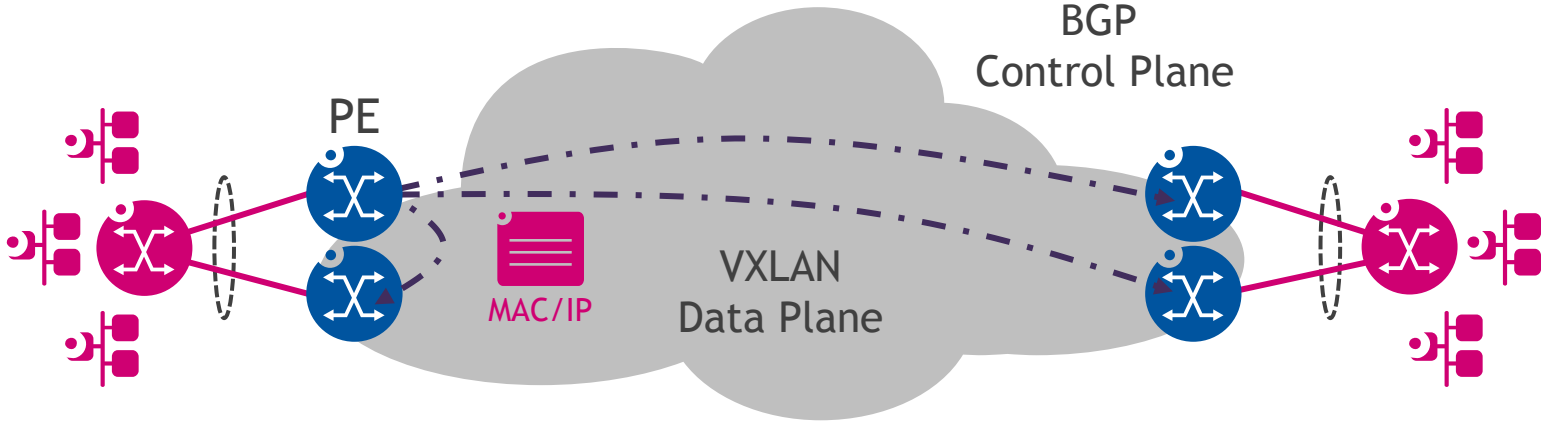
- EVPN over an MPLS data plane is the original EVPN solution in the base specification
- Requires IGP, RSVP-TE or LDP, BGP
- No pseudowires
- MPLS runs in the core network's control plane and data plane
- Core network supports all the MPLS features we know and love, since EVPN uses MPLS as the data plane (TE, FRR, …)

PROVIDER BACKBONE BRIDGES (PBB) EVPN DATA PLANE
DRAFT-IETF-L2VPN-PBB-EVPN (PBB-EVPN)



- PBB-EVPN combines IEEE 802.1ah PBB with EVPN
- PEs are PBB Backbone Edge Bridges (BEB)
- Reduces number of MACs in EVPN by aggregating customer MACs with backbone MACs
  - Same concept as route aggregation in IP
- Scales EVPN networks to a very large number of MACs
  - PEs only advertise backbone MACs with BGP
  - Customer MAC and backbone MAC mapping is learned in the data plane
  - Useful for providing services to networks where the MACs are not under your control
- MPLS runs in the control plane and data plane

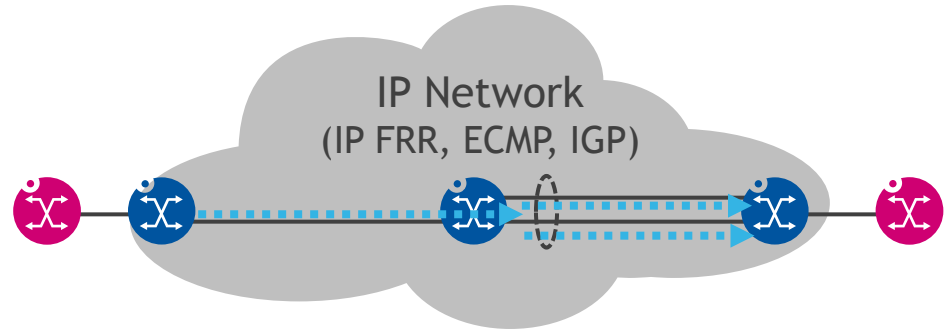
EVPN VIRTUAL EXTENSIBLE LAN (VXLAN) DATA PLANE
DRAFT-SD-L2VPN-EVPN-OVERLAY (EVPN-VXLAN)

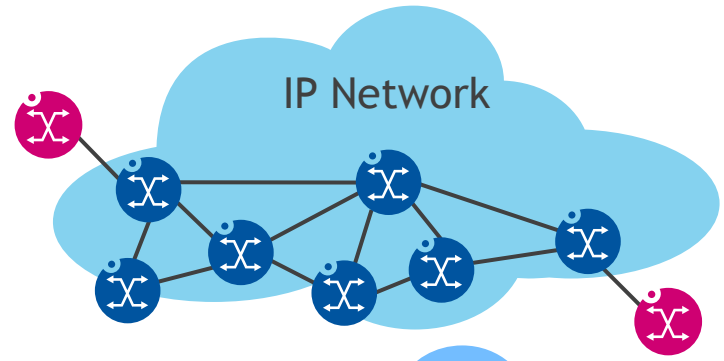


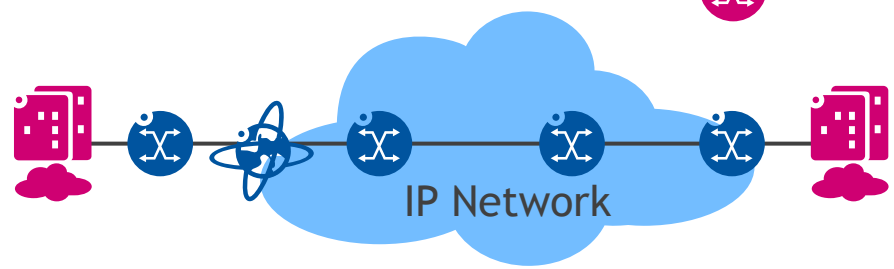
- EVPN-VXLAN uses EVPN over a VXLAN data plane
  - VXLAN is typically used for data center extension over WAN
  - Can also be used as an overlay in any IP network for IP/Ethernet services
  - Useful when MPLS is unavailable or unwanted
  - Alternative to NVGRE or MPLSoGRE (NVO3)
  - PIM is not needed with ingress BUM replication
- VXLAN provides the Layer 2 overlay over IP
  - IP reachability is required between PEs
  - EVPN uses BGP control plane for MAC route advertisements
  - VXLAN data plane uses UDP to encapsulate the VXLAN header and Layer 2 frame
- Provides all the benefits of EVPN for DCI and virtualized networks

VXLAN DATA PLANE FLEXIBILITY

- VXLAN encapsulates Ethernet in IP
  - Runs over IPv4 or IPv6 and uses UDP
  - Source port in ranges 49152 – 65535 is a hash of fields from the encapsulated frame to provide load balancing entropy
  - Destination port is 4789
  - 8 byte VXLAN header provides 24 bit VXLAN Network Identifier (VNI) and flags
- VXLAN is routable with IP, so the underlay network may be any network that uses existing resiliency and load balancing mechanisms
  - ECMP
  - IGPs/BGP
  - IP FRR
- VXLAN tunnel endpoints can be on network equipment or computing infrastructure
  - Deliver a VPN to a hypervisor attached to a VM



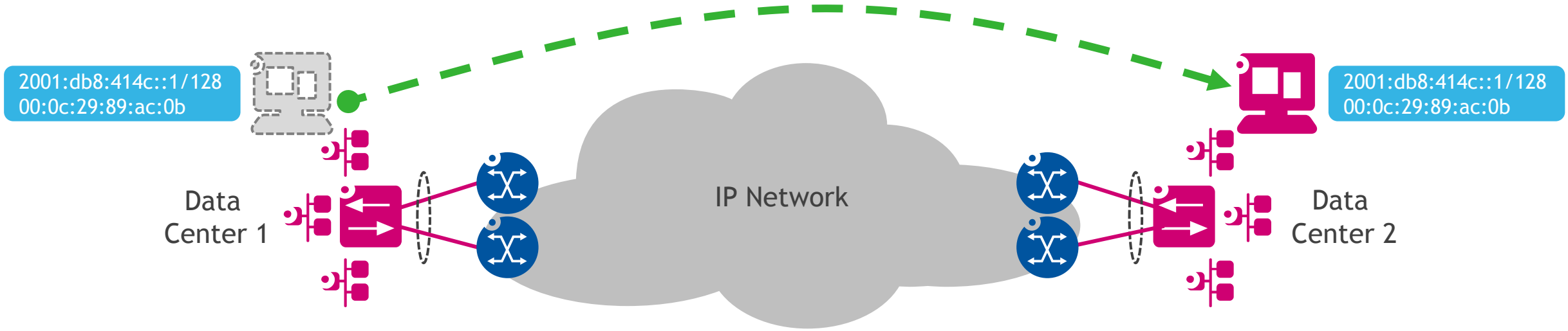




AGENDA

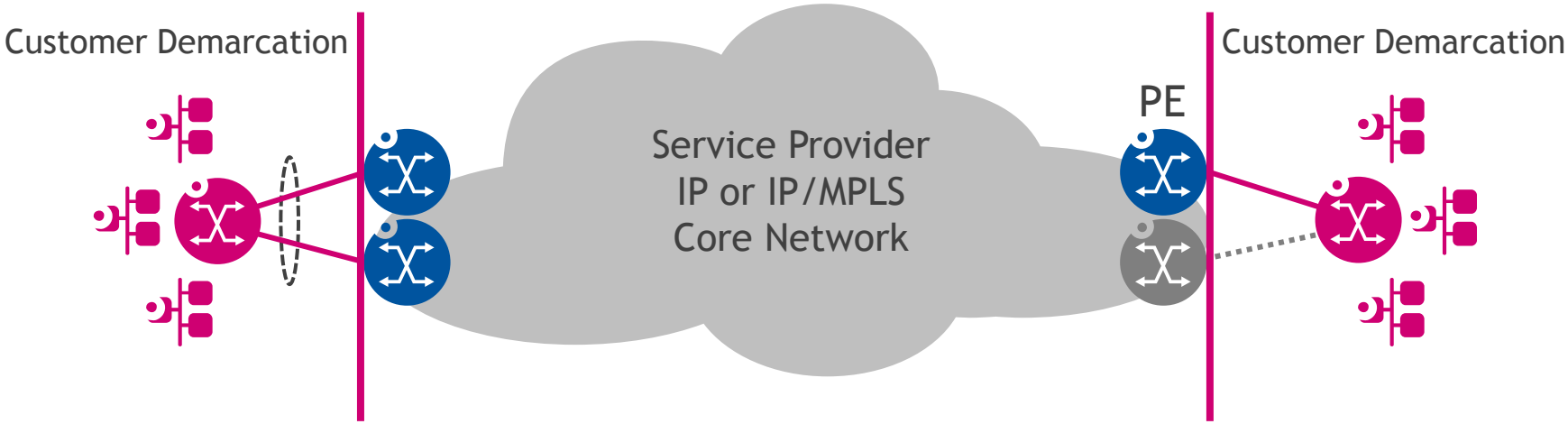
1. EVPN Background and Motivation
2. EVPN Operations
3. **EVPN Use Cases**

LAYER 2 OR LAYER 3 DATA CENTER INTERCONNECT



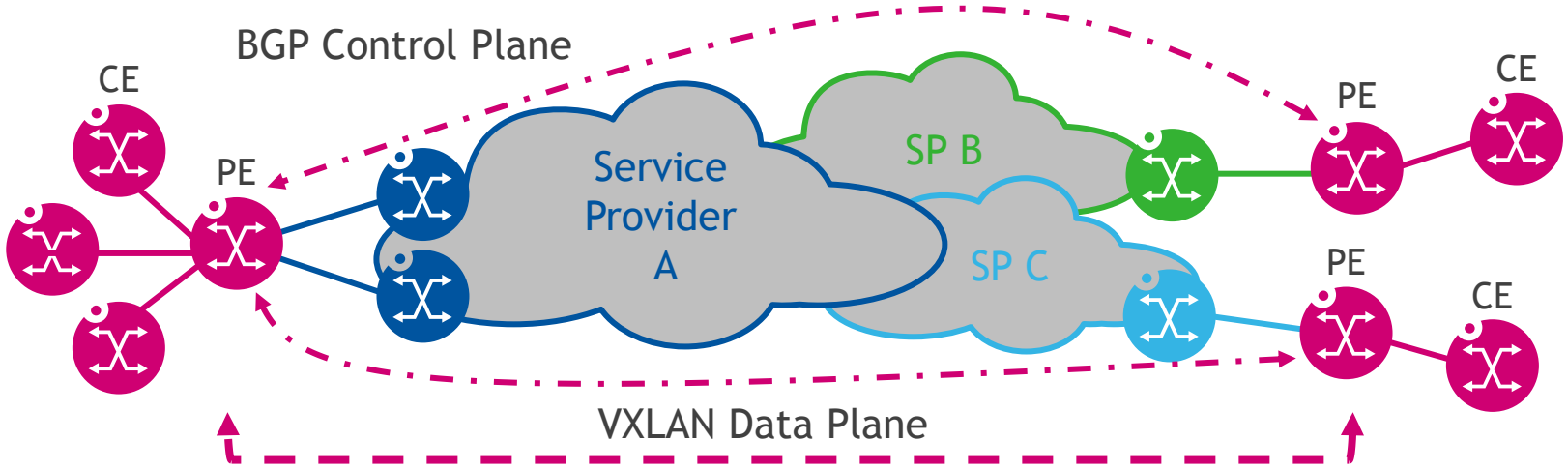
- Enables scalable Layer 2 or Layer 3 DCI services for virtualized data centers
- IP/MAC mobility for VMs that move between data centers
  - Faster moves while maintaining correct FDB on all routers
- Local IP gateway at each PE optimizes routing
- Provides all the benefits of EVPN for DCI and virtualized networks
  - All-active multihoming
  - Eliminates ARP/ND flooding for MAC learning
  - Integrated Layer 2 switching and Layer 3 routing over the same interface or VLAN

BUSINESS SERVICES AND INFRASTRUCTURE NETWORKS
LAYER 2 AND LAYER 3 SERVICES



- EVPN enables service providers to offer integrated Layer 2 and Layer 3 services
  - Single interface, single VLAN to customer
  - One technology for both services, no need for multiple VPN protocols
  - All-active or single-active PE to CE connection
- EVPN service can be provided over any core network
  - MPLS core can use EVPN-MPLS
  - IP core can use EVPN-VXLAN

SITE TO SITE NETWORKS OVER IP
FLEXIBLE LAYER 2 AND LAYER 3 NETWORKS



- EVPN-VXLAN works over any IP service to provide a flexible site to site network
- Just requires IP connectivity between sites, no MPLS or any special configuration by IP service provider
  - Service provider network is transparent to EVPN
  - EVPN overlay is transparent to service providers
- VPN routing between endpoints can be controlled with BGP and routing policies to service providers
- Routing and MAC/IP advertisement within EVPN controlled via IBGP between PEs

SUMMARY

- EVPN provides next-generation VPN solutions for Layer 2 and Layer 3 services over Ethernet
  - Consistent signaled FDB in control plane using MP-BGP vs. flood-and-learn FDB in data plane
  - L3VPN-like operation for scalability and control
  - Flow-based load balancing and all-active multipathing
  - Delivering Layer 2 and Layer 3 services over the same interface, VLAN and VPN
  - Simpler provisioning and management with a single VPN technology
  - ARP/ND security and MAC provisioning
  - MPLS or IP data plane encapsulation choices

- More information
  - IETF Layer 2 Virtual Private Networks (l2vpn) Working Group
    http://datatracker.ietf.org/wg/l2vpn/
  - RFC 7209: Requirements for Ethernet VPN (EVPN)
    http://tools.ietf.org/html/rfc7209
  - Base specification: draft-ietf-l2vpn-evpn
    http://tools.ietf.org/html/draft-ietf-l2vpn-evpn
  - Use case examples: draft-rp-l2vpn-evpn-usage
    http://tools.ietf.org/html/draft-rp-l2vpn-evpn-usage

www.alcatel-lucent.com

QUESTIONS?

EVPN REQUIREMENTS AND BENEFITS

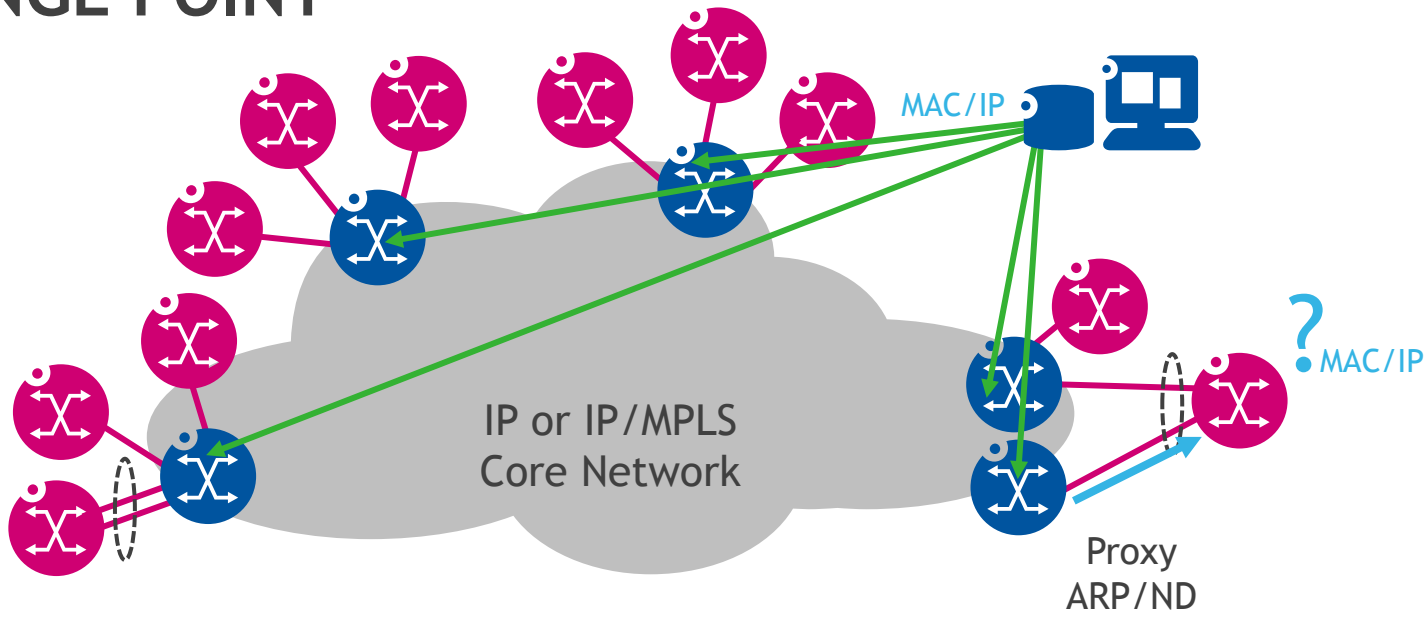
	VPN Requirements	VPLS	EVPN	What does it do for me?
Address Learning	Control Plane Address Learning in the Core	✘	✓	Greater Scalability and Control
Provisioning	L3VPN-Like Operation	✘	✓	Simpler Provisioning and Automation
	Auto Discovery and Configuration	PEs Only	✓	Simpler Provisioning and Automation
Resiliency	Active-Standby Multihoming (Service-Based Load Balancing)	✓	✓	Standby Redundancy
	All-Active Multihoming (Flow-Based Load Balancing)	✘	✓	Active Redundancy and Link Utilization
Services	VLAN Based Service Interfaces	✓	✓	Virtualization and Advanced Services
	VLAN Aware Bundling Service Interfaces	✘	✓	Virtualization and Advanced Services
	Inter-Subnet Forwarding	✘	✓	Layer 2 and Layer 3 Over the Same Interface
Flow Optimization	Proxy ARP/ND	✘	✓	Security and MAC Provisioning
	MAC Mobility	✘	✓	Virtualization and Advanced Services

EVPN NLRI ROUTE TYPES AND EXTENDED COMMUNITIES

Route Type	Route Description	Route Usage	Reference
1	Ethernet Auto-Discovery (A-D) Route	Endpoint Discovery, Aliasing, Mass-Withdraw	draft-ietf-l2vpn-evpn
2	MAC Advertisement Route	MAC/IP Advertisement	draft-ietf-l2vpn-evpn
3	Inclusive Multicast Route	BUM Flooding Tree	draft-ietf-l2vpn-evpn
4	Ethernet Segment Route	Ethernet Segment Discovery, DF Election	draft-ietf-l2vpn-evpn
5	IP Prefix Route	IP Route Advertisement	draft-rabadan-l2vpn-evpn-prefix-advertisement

Extended Community Type	Extended Community Description	Extended Community Usage	Reference
0x06/0x01	ESI Label Extended Community	Split Horizon Label	draft-ietf-l2vpn-evpn
0x06/0x02	ES-Import Route Target	Redundancy Group Discovery	draft-ietf-l2vpn-evpn
0x06/0x00	MAC Mobility Extended Community	MAC Mobility	draft-ietf-l2vpn-evpn
0x03/0x030d	Default Gateway Extended Community	Default Gateway	draft-ietf-l2vpn-evpn, bgp-extended-communities

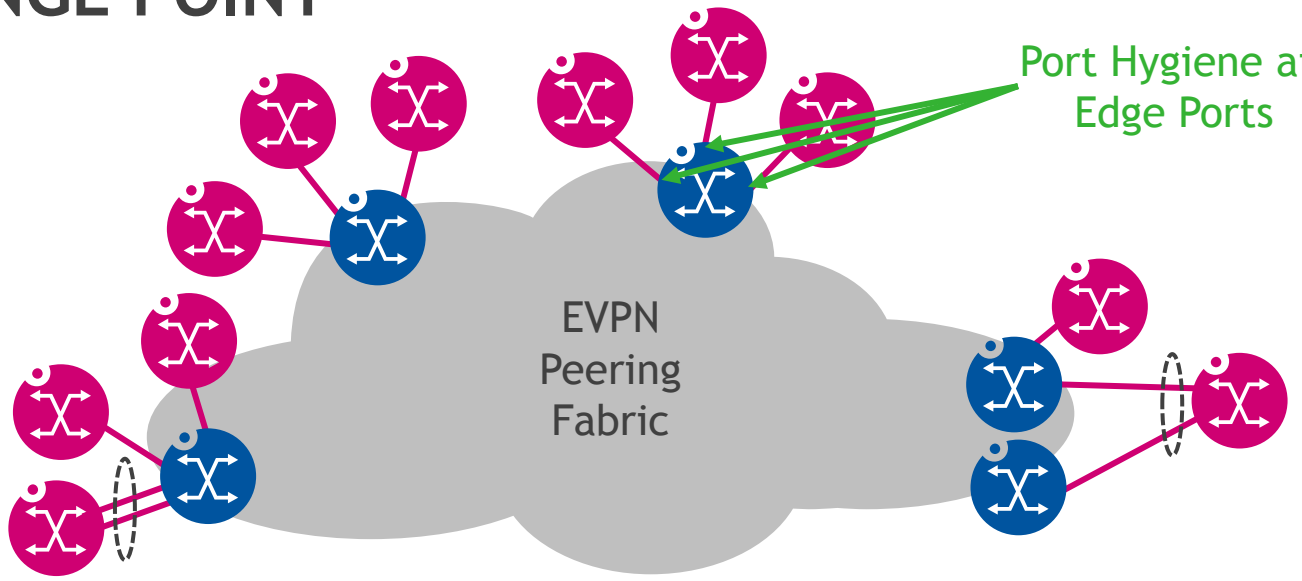
INTERNET EXCHANGE POINT
PEERING FABRIC



- Provides Layer 2 interconnection over an EVPN peering fabric
  - IP/MPLS core network with MPLS data plane
  - IP core network with VXLAN data plane
- Supports single or all-active multihoming to the peering fabric VLAN
- Supports PNIs and/or other overlay VLANs

- Enables precise fine-grained control over MAC addresses
  - Static MAC provisioning and proxy ARP/ND from PEs can reduce or eliminate unknown unicast
  - Per-MAC loop control vs per-port or per-VLAN isolates potential loops
  - Works together with edge port hygiene features to provide a clean and secure peering fabric

INTERNET EXCHANGE POINT
PEERING FABRIC



- EVPN provides the technology for the peering fabric and MAC/IP management over the core
- Still need to use existing port security mechanisms and follow BCPs for port hygiene and allowed traffic
  - Typically allow IPv4, IPv6, ARP and block unwanted traffic types
  - MAC address locking
  - BUM control