

draft-malhotra-bess-evpn-centralized-anycast- gw-00

N. Malhotra(Cisco), K. Ananthamurthy (Cisco),
A. Sajassi (Cisco), L. Krattiger (Cisco)

IETF 118, Nov 2023

Prague

Agenda

- Introduction
- Control Plane Operations
- Data Plane Operations
- GW/MAC IP Mobility
- Deployment Models
- Feedback

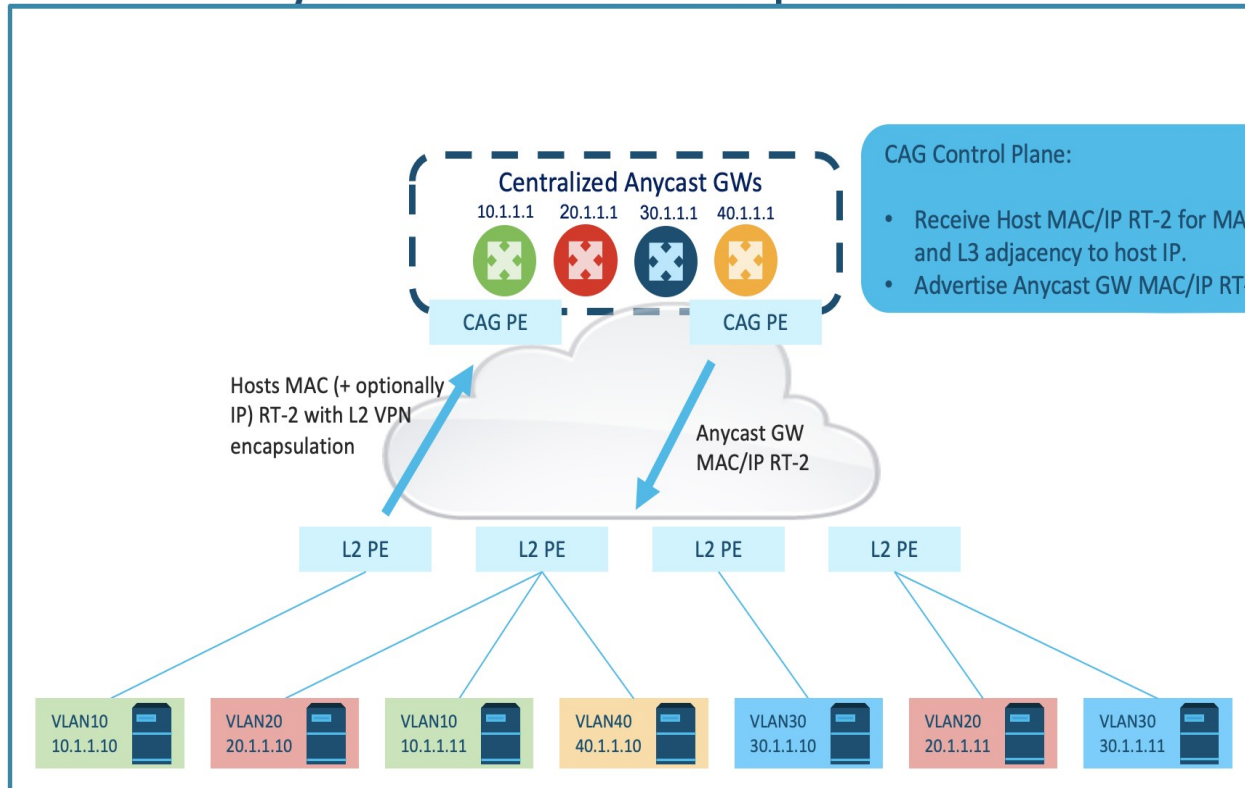
Introduction

- EVPN IRB with Distributed Anycast Gateway (DAG) provides First Hop Routing (FHR) and bridging at PEs where Tenant Systems (TSs) are attached. This draft provides the architecture and operation model for the Centralized Anycast Gateway (CAG) where FHR is centralized, and bridging located at PEs where TSs are attached
- CAG architecture essentially uses a Layer-2 EVPN overlay and FHR operation on CAG is identical to asymmetric routing as defined in rfc9135

Control Plane Operations (1)

- CAG PE Procedures

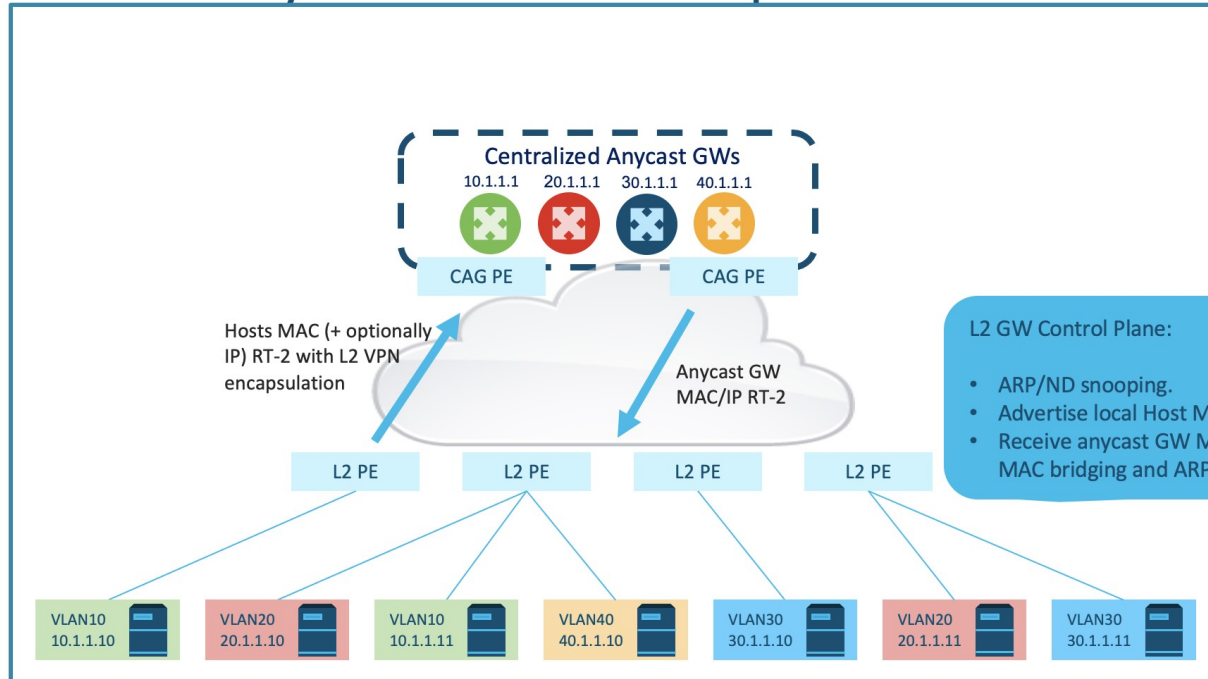
Centralized Anycast GW – control plane



- A set of redundant CAG PEs that act as the FHR for the same subnet MUST be provisioned with the same anycast GW IP and MAC
- Set of redundant CAG PEs may advertise anycast VTEP IP as NH to enable MAC multi-pathing
- CAG PEs MUST import MACs received via RT-2 from Layer-2 PE into MAC-VRFs, to establish Host MAC reachability via Layer-2 EVPN encapsulation and tunnel to Egress L2 PE. CAG will perform the Asymmetric IRB procedures defined in RFC 9135 with IP-to-MAC binding resolved via respective adjacency/next-hop table entry.
- Remote host layer-3 adjacency is still resolved by host MAC reachability via a Layer-2 VPN tunnel to the Egress L2-PE.

Control Plane Operations (2)

Centralized Anycast GW – control plane

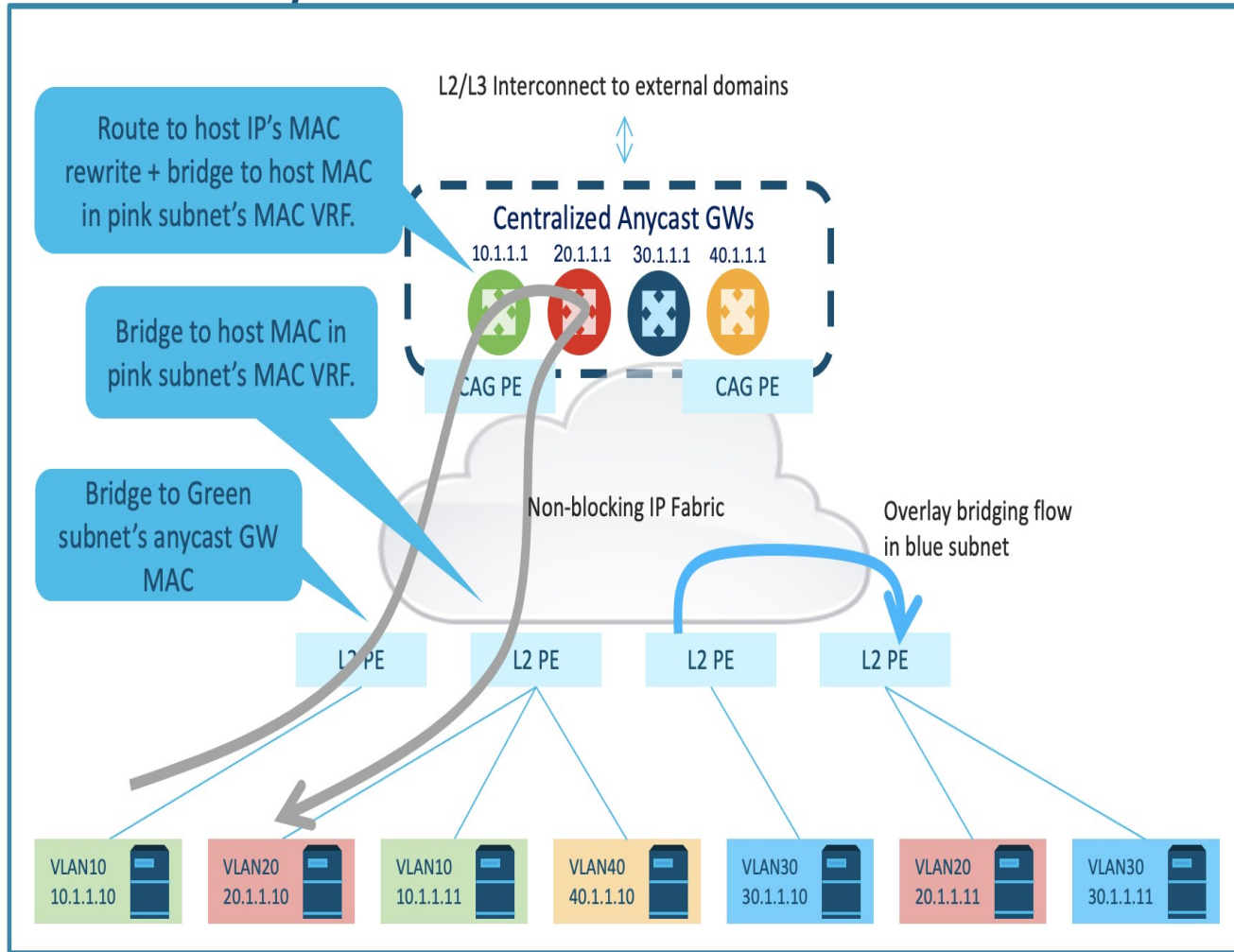


- Layer-2 GW PE Procedures

- ARP/ND snooping MUST be enabled on L2-PEs such that TSs host IP and MAC is advertised using EVPN RT-2 with a single label that represents the EVI. This simplifies CAG operation to avoid the need for data plane learning and syncing of ARP/ND tables across redundant CAGs.
- L2-PEs MUST handle ARP refreshes when MAC/ARP ages out. If the host responds to host MAC/IP, then ARP MUST be refreshed. EVPN RT-2 SHOULD NOT be re-advertised
- L2-PEs MAY install a GW MAC/IP entry in the ARP/ND suppression cache to act as a proxy for CAG by responding to ARP requests from hosts for GW MAC/IP, this avoids flooding of ARP/ND requests across the fabric and avoids duplicate ARP responses from redundant CAGs.

Data Plane Operations

Centralized Anycast GW - Data Path



- Bridging

- Intra-subnet host to host flow is identical to that in symmetric or asymmetric distributed anycast GW based IRB deployments as defined in rfc9135

Routing

- Inter-subnet host to host flow destined to default GW MAC is bridged to CAG PE with the L2 VPN encapsulation learnt via default GW RT-2 from the CAG PE.
- CAG does a destination MAC lookup in the local MAC VRF that results in my MAC and local IRB VRF interface. IP lookup yields a
- If the destination subnet is local/attached or within the local CAG domain, IP lookup yields a local host adjacency on the destination subnet IRB interface. This results in host MAC rewrite, followed by host MAC lookup that results in the packet being bridged to the egress L2 PE with L2 VPN.
- If the destination subnet is not local to the CAG node, IP lookup yields the destination subnet prefix that resolves via L3 VPN encapsulation and tunnel to the next-hop CAG PE that is the FHR for the destination subnet.

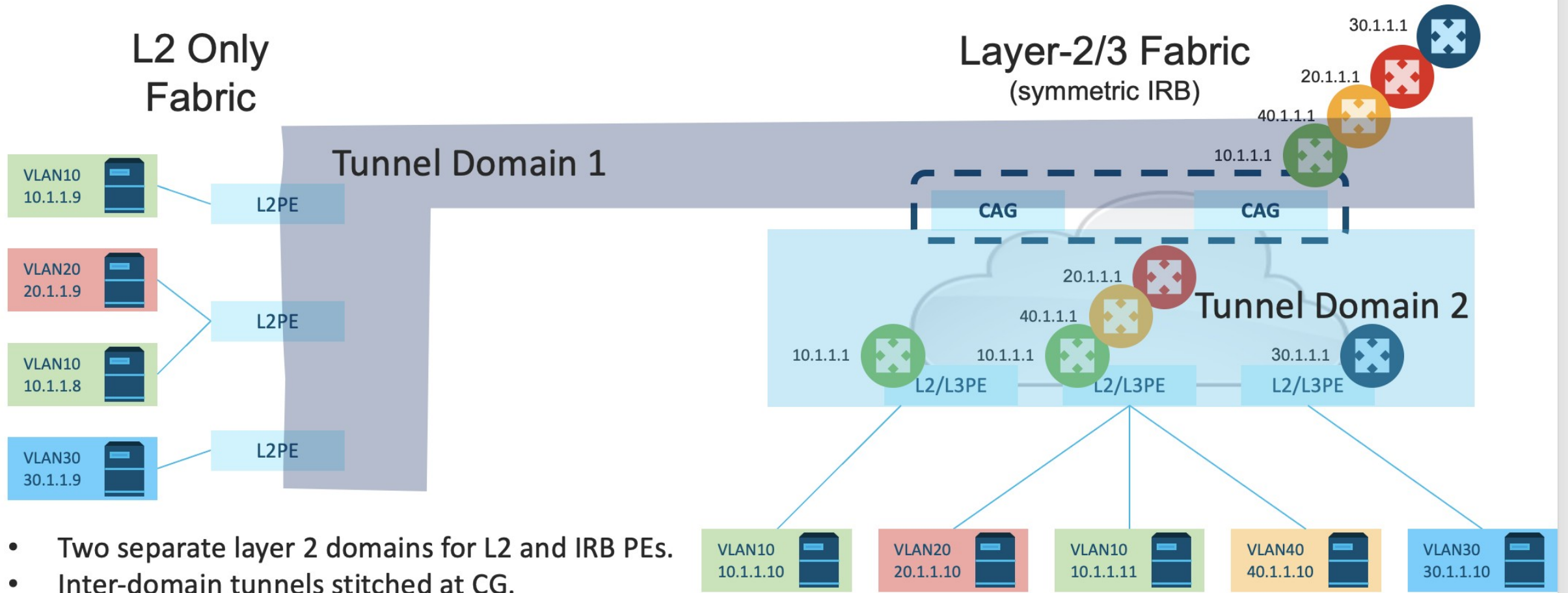
GW MAC/IP Mobility

- L2 PE MUST treat remote GW MAC/IP route learnt from the CAG PE as static and MUST NOT apply EVPN mobility procedures to GW MAC on a local learning event.
- If there is a prior remote MAC learnt via BGP on CAG that is same as the GW MAC, GW MAC/IP MUST be advertised with a higher sequence number.
- L2-PEs MUST prefer MAC with BGP produced "Default GW" attribute over a locally learnt MAC route, independent of the sequence number. This is to ensure that the GW MAC/IP RT-2 will be preferred at L2-PE and installed in MAC-VRF and ARP suppression tables, when there is a prior conflicting non-GW MAC/IP.
- The L2 PE MUST alert the operator when duplicate non-GW MAC/IP in the presence of a GW MAC-IP

CAG Deployment Models

- CAG Placement as an Interconnect

Inter-op Model 1 – CAG Placement as Interconnect

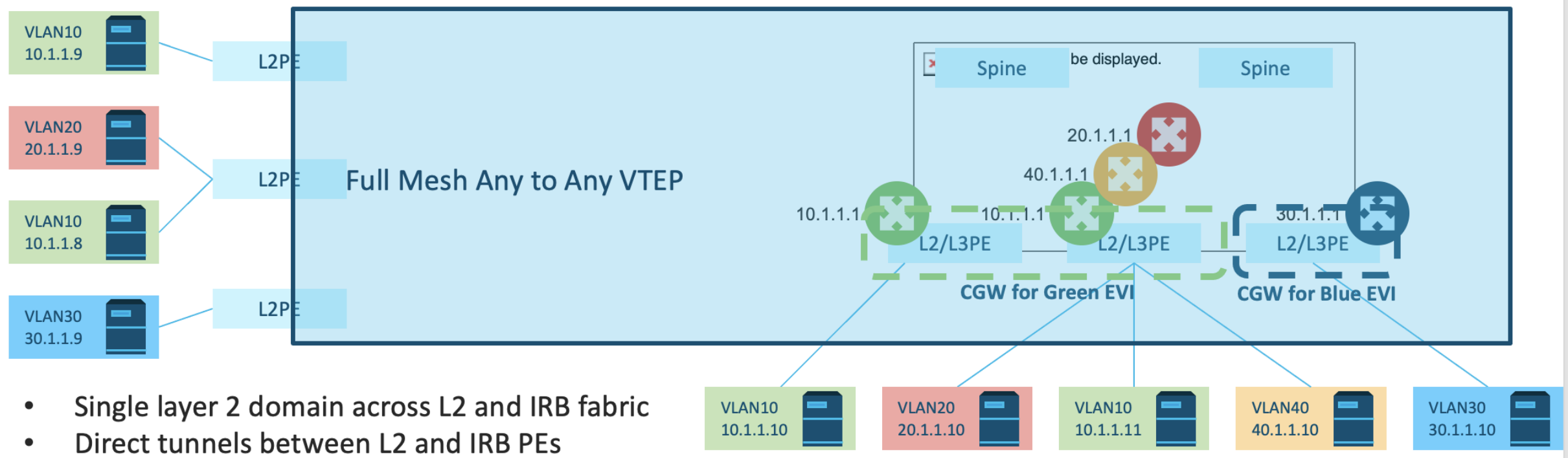


- Two separate layer 2 domains for L2 and IRB PEs.
- Inter-domain tunnels stitched at CG.

Inter-op Model 2 – CAG Placement on DAG L2/L3 PEs

L2 Fabric

Layer-2/3 Fabric (symmetric IRB)



- Single layer 2 domain across L2 and IRB fabric
- Direct tunnels between L2 and IRB PEs

Next Steps

- Request feedback/comments from WG members