The bridge to possible

Deploying Your First SD-Access Project

Based on LISP\VXLAN stack

Sergey Nasonov, Solutions Engineer
CCIE R&S 62572

BRKENS-2824

cisco Live!

Cisco Webex App

https://ciscolive.ciscoevents.com/ciscolivebot/#BRKENS-2824

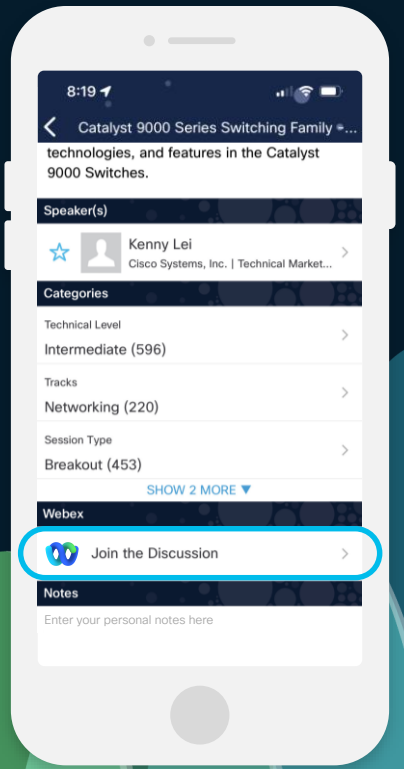
Questions?

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

How

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

Webex spaces will be moderated by the speaker until June 7, 2024.



Introduction

- The session assumes fundamental knowledge of SD-Access solution:
  - BRKENS-2810 – Cisco SD-Access Solution Fundamentals
  - BRKENS-2811 – Connecting Cisco SD-Access to the External World
  - BRKENS-2814 – Role of ISE in SD-Access
  - BRKENS-2827 – Cisco SD-Access Migration Tools and Strategies

- Practical session, no textbook examples.
- My opinion, different people will have different opinions based on their own experiences.
- “Ok, I’ve watched all these videos and read the CVD, where do I start?”

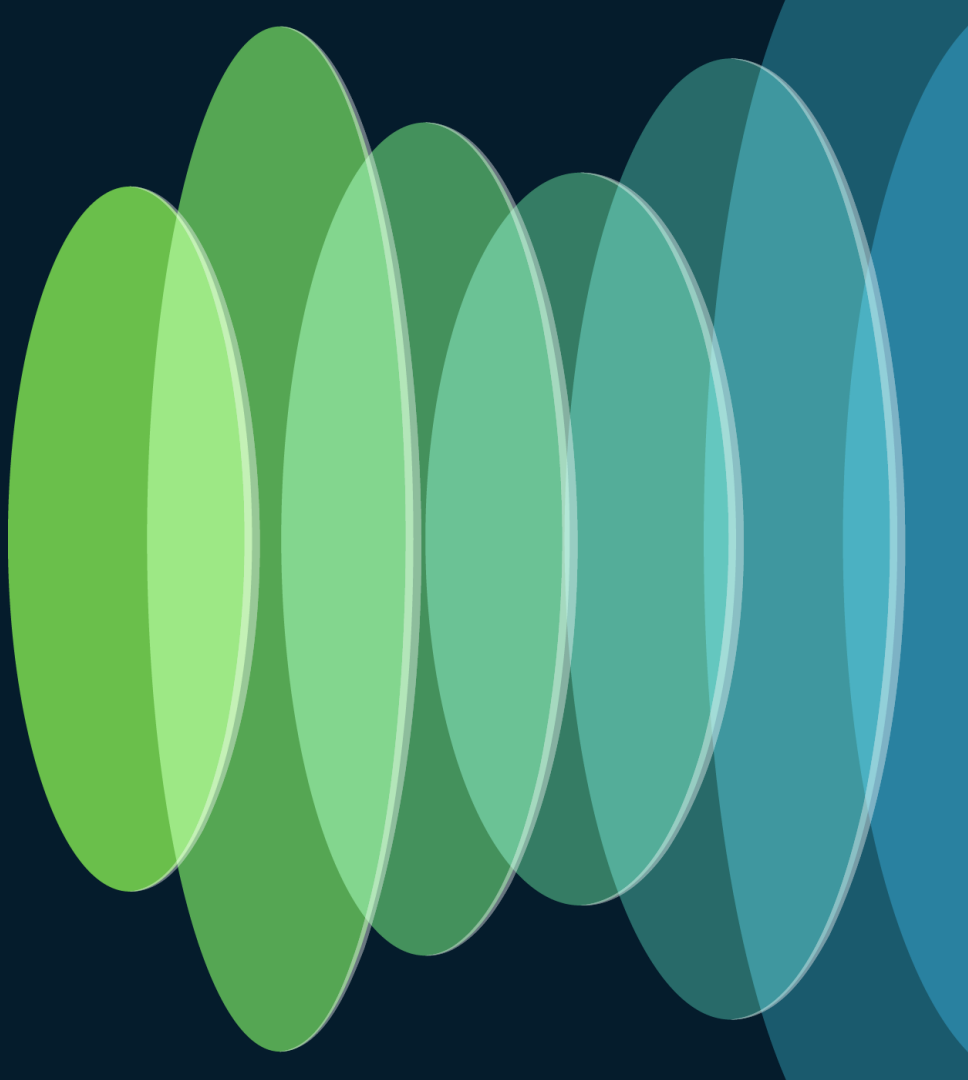
Agenda

- Planning SD-Access Deployment
- Designing SD-Access Deployment
- Implementing or Migrating to SD-Access
- Take-aways or What's Next?

Nomenclature

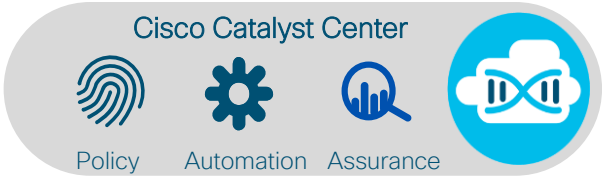
- Catalyst Center. Cisco network management solution, formerly known as DNA Center.
- Endpoint (EP). Connected device that is not routing the traffic. Can be laptop, workstation, server, printer, BMS system and so on.
- SD-Access Fabric. For the purposes of this presentation, an overlay—based connectivity solution implemented by a SD-Access Border Nodes, Control Plane Nodes, Edge Nodes and optionally Fabric-Enabled wireless controllers and Access Points using LISP\VXLAN stack.

Planning SD-Access Deployment



cisco Live!

Collect the Requirements



Cisco SD-Access has a few considerations that network designer needs to be aware of:

- Deployment wide - Catalyst Center:
  - Number of endpoints (EPs – concurrent/transient), number of network devices, number of interfaces, IP pools and L2 overlays.

- Site level: Border and/or Control Plane nodes and Catalyst Center:
  - Logical: Number of concurrent EPs (v4/v6, wired/wireless), RTT to controllers, IP pools, L2 handoffs.
  - Physical: Number of fabric devices per site.

All scalability limits are well documented in Cisco Catalyst Center Data Sheet, but it's hard to apply those to the design when doing it for the first time.

Meet ACME Corporation

Large manufacturing organization – legacy network refresh.

Main site – 3 sub-areas interconnected via dark fibre in ring topology:

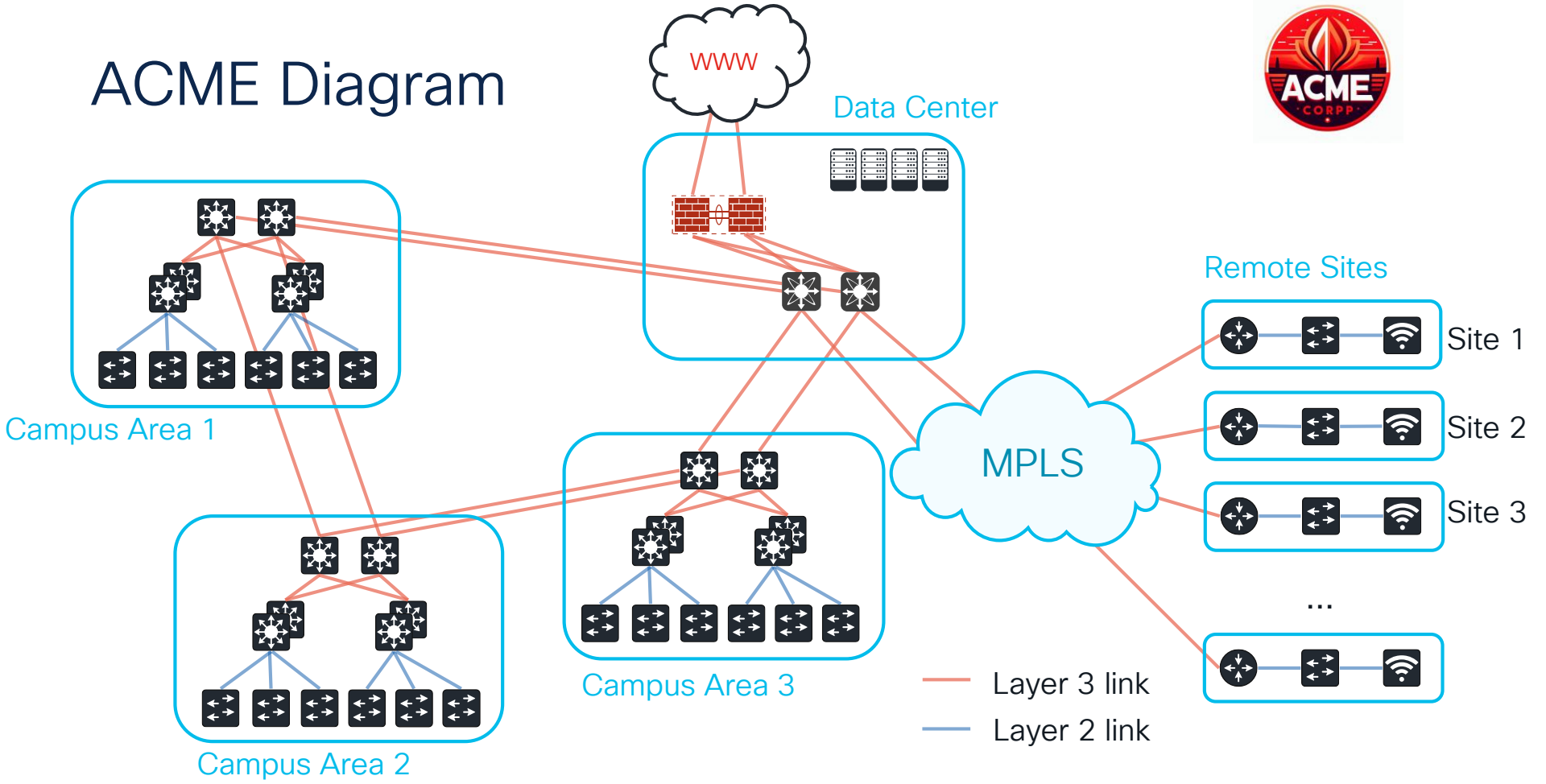
- 25,000 users with 45,000 concurrent devices.
- 2100 x WS-C2960X access switches in 1300 access switch cabinets.
- 5200 x AIR-CAP3702I wireless access points.
- 700 VLANs for users and device segmentation.
- L3 boundary at distribution layer, MPLS for segmentation, DC firewall as enforcement point.
- Multiple business units are sharing the same network.

Two onsite active/active data centers with applications, Internet access and public cloud peering.

Remote sites – 70 small sites, currently connected via MPLS network:

- 1 switch per site.
- 1-2 APs per site.

ACME Diagram



Cisco SD-Access Design Tool

- Cisco SD-Access Design tool is used once high-level requirements (number of sites, number of EPs, Catalyst Center, ISE, wired/wireless, etc) are collected.

- Input the requirements in the tool and it generates HLD.

- Available for everyone with Cisco.com account at http://cs.co/sda-design-tool.

ACME Business Drivers for Cisco Campus Fabric

Network Resiliency

Eliminate Layer 2 protocols from the network, all traffic is routed.

Unified Wired/Wireless

User or device security level, visibility, QoS policy and traffic path is independent from access medium.

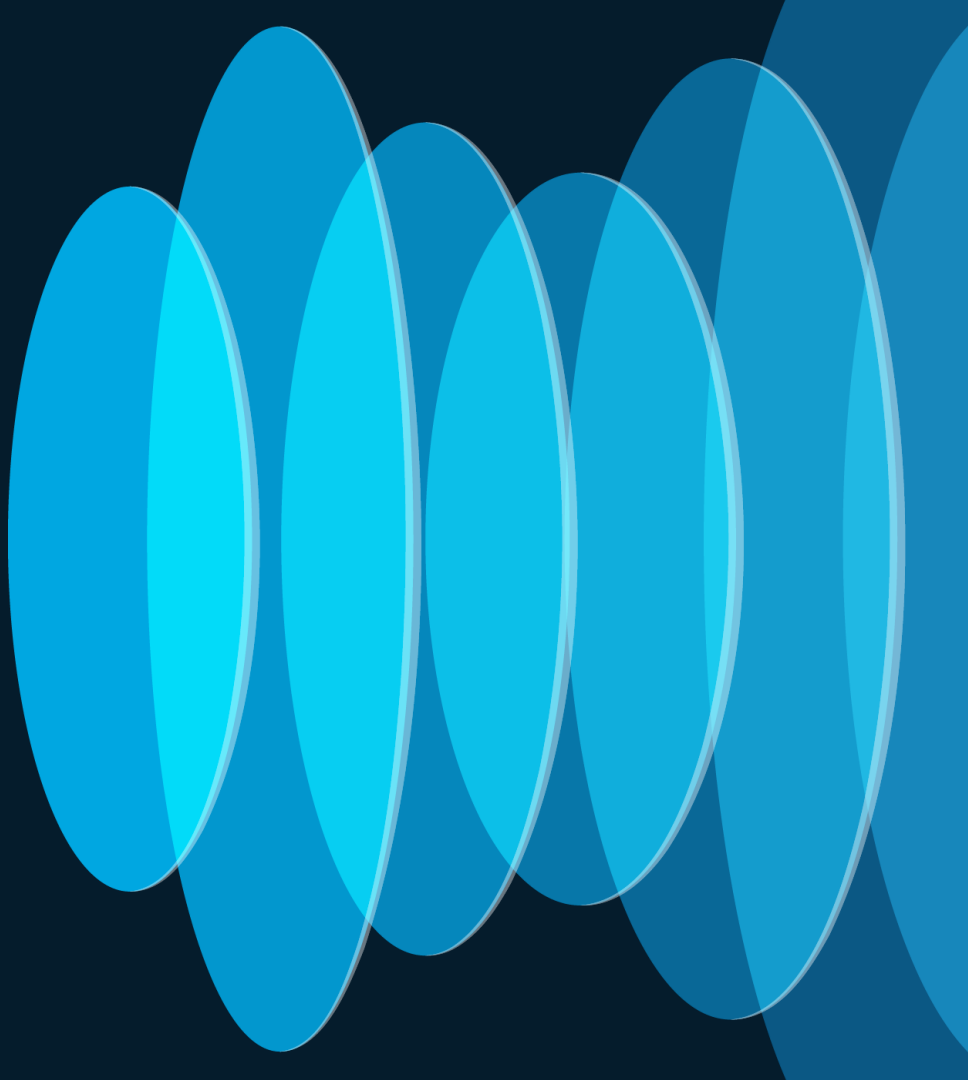
Segmentation

Centrally defined security policy that is enforced line rate at access switches, without tunnelling traffic to the centralised firewalls.

Network Automation

All network operations are performed from central management controller, CLI is no more.

Designing SD-Access Deployment

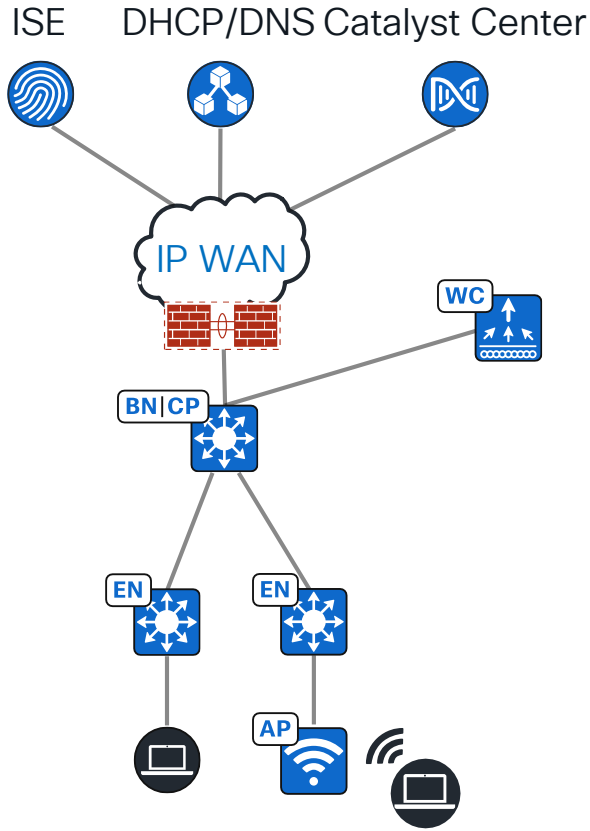


cisco Live!

External Dependencies

Before you spin up your first SD-Access fabric site, you will need:

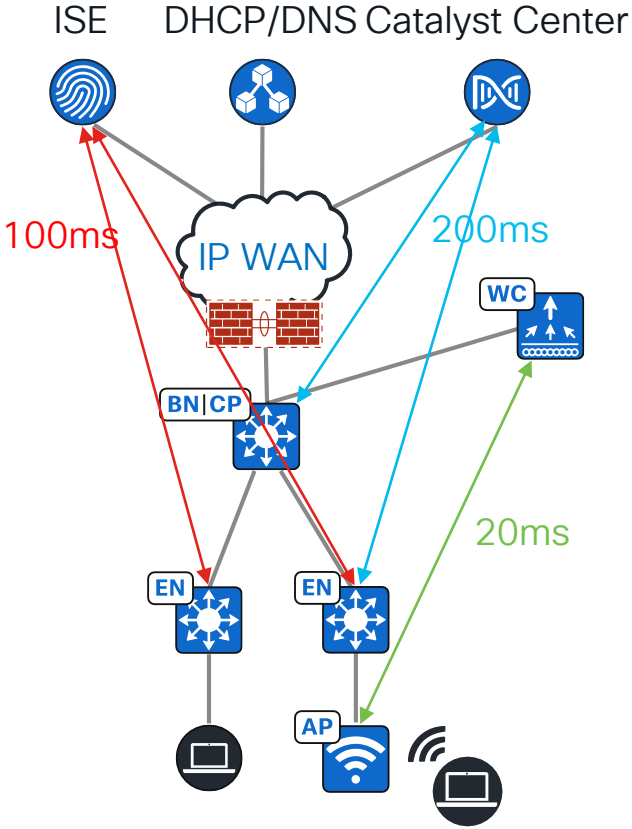
- Catalyst Center – automation engine for SD-Access.
- DHCP / DNS – if you intend to provide these services to users connecting to SD-Access network.
- Cisco ISE – if you want to authenticate and authorize users or devices.
- Cisco WLC – if you want to provide wireless access. WLC can enable fabric-enabled wireless for a single site only.
- Fusion device (typically a firewall) to implement VRF route-leaking and enforce security policy at the leaking point.



External Dependencies

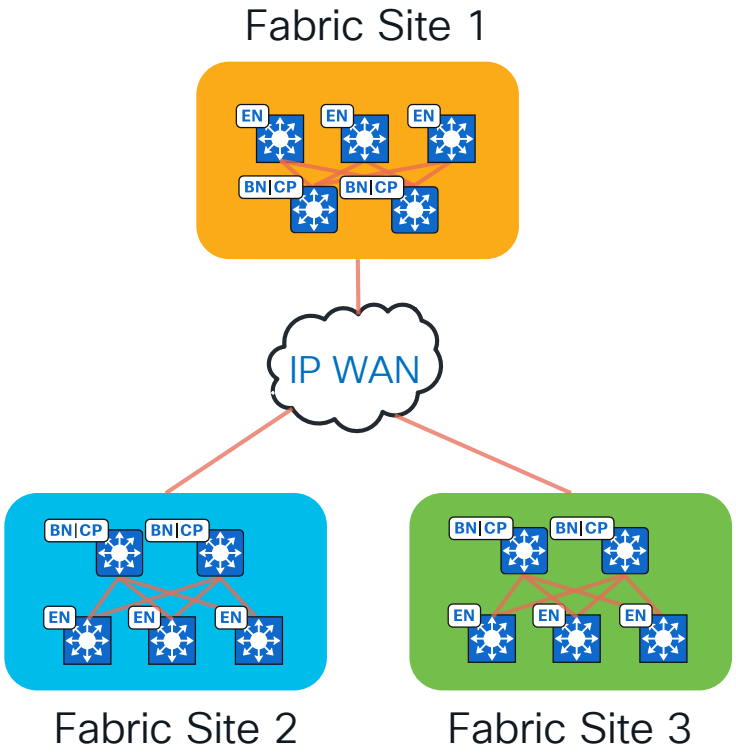
All external dependencies reside outside the fabric site and just need IP (Layer 3) connectivity to fabric devices. Latency requirements:

- Catalyst Center to fabric devices – 200ms RTT.
- ISE to fabric devices – 100ms RTT.
- Fabric WLC to fabric APs – 20ms RTT (put it onsite).



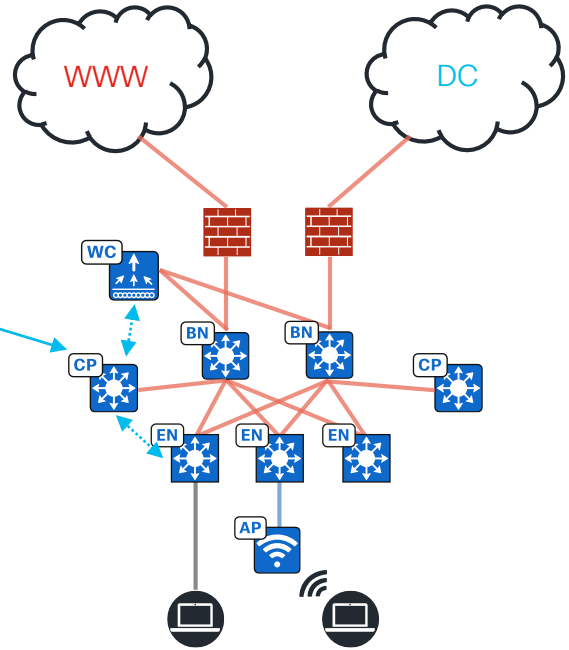
How Would Your Carve Your Fabric Sites?

- Fabric site is an instance of an SD-Access Fabric.
- A collection of Edge Node switches using the same set of CP/BN switches.
- Typically defined by disparate geographical locations, but not always.
- Can also be defined by:
  - Endpoint scale.
  - Failure domain scoping.
  - Underlay connectivity attributes (MTU, multicast).
- Typically interconnected by a “Transit”.



Site Limits – Endpoint Scale

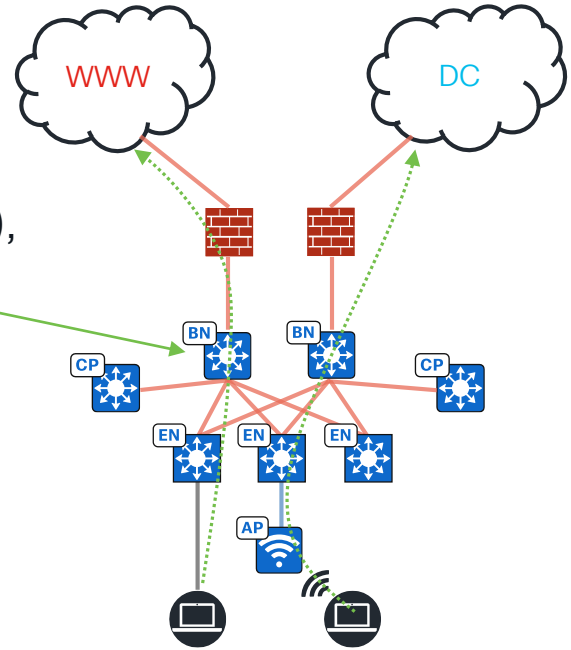
Control Plane Node keeps information about all site endpoints in RAM and uses CPU to process it (including wireless roaming events).



- C9300\L switches can support up to 16,000 EPs as CP node.
- C9500-32C / C9500-48Y4C / C9500-24Y4C switches can support up to 80,000 EPs as CP node.
- Other C9K switches are possible in CP role, sizing values are documented in Catalyst Center Data Sheet.

Site Limits – Endpoint Scale

Border Node keeps all EP information in TCAM as host routes. If EP has multiple IP addresses (v4 + multiple v6), each address is counted as individual entry.

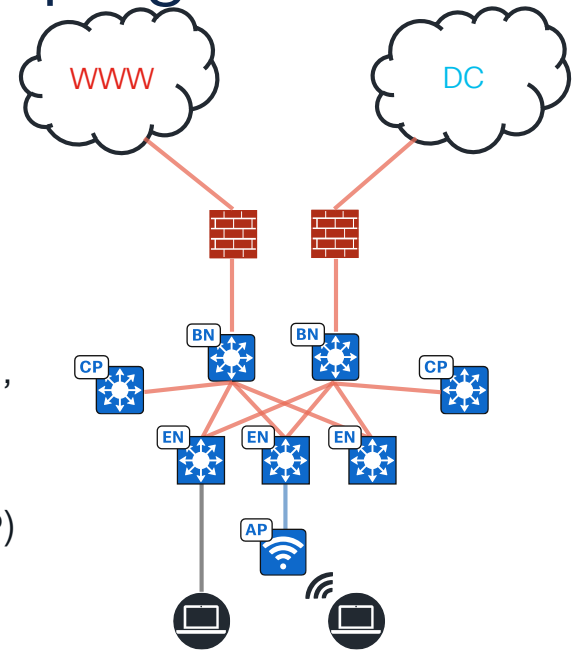


- C9300\L switches can support up to 16,000 IP host routes (/32 or /128) as Border Node.
- C9500-32C / C9500-48Y4C / C9500-24Y4C switches can support up to 150,000 IP host routes as Border Node.

Full border node sizing values for all SD-Access platforms are documented in Catalyst Center Data Sheet.

Site Limits – Failure Domain Scoping

- All Edge Nodes in the site are sharing the same set of Control Plane and Border Nodes. If all CP or BN nodes fail, the site is failed*. SD-Access site with fabric wireless can have 2 CP nodes max.
- A lot of configuration elements (VRF, VLAN, multicast, wireless, default switchport policy) are applied at the site level, to all** fabric site switches at the same time.
- Fabric site is underpinned by a single instance of underlay routing protocol (IGP) as well as overlay routing protocol (LISP) and is visible as single BGP AS from the outside world.

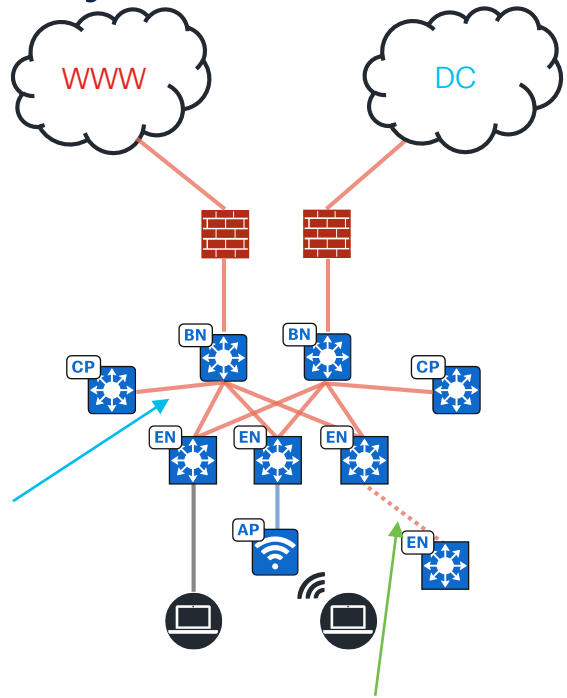


*During a total CP failure, no new endpoints can be onboarded into the fabric and roaming events won't work. Existing traffic flows will be cached for 24 hours.

**Some changes can be scoped to a limited subset of switches via Fabric Zones, see BRKENS-3833 for details.

Site Limits – Underlay Connectivity Attributes

- Avoid mixing different underlay connectivity attributes, such as MTU or multicast support because you will end up dropping to the lowest common denominator within a fabric site.



Multiple Fabric Sites vs Single Fabric Site?

Make large single fabric site within single geographical area until:

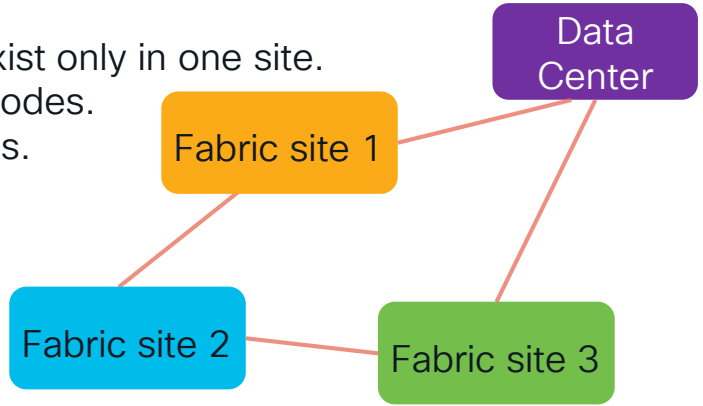
- You hit fabric device (1200 logical switches for -XL Catalyst Center) or endpoint limit (~100,000 EPs).
- Links between parts of your fabric site can support increased MTU (from 1550 to 9000 bytes) and can be multicast-enabled.
- Part of your fabric site needs to be online even if the rest of your site is offline.
- Part of your fabric site needs to provide Direct Internet Access for users in the overlay.

Multiple Fabric Sites vs Single Fabric Site for ACME?

Requirement:
- 2100 x WS-C2960X access switches in **1300 switch cabinets.**

Solution:
- Three fabric sites in main campus because of 1300 switch cabinets (max fabric site is 1200 fabric devices).

Caveats:
- No seamless wireless roaming as IP subnet can exist only in one site.
- Each site needs its own set of WLCs and BN/CP nodes.
- Extra switching hardware for SDA Transit CP nodes.



SD-Access Transit

Allows SD-Access fabric sites to communicate to each other using VXLAN tunnels between Border Nodes leveraging plain IP network between each other.

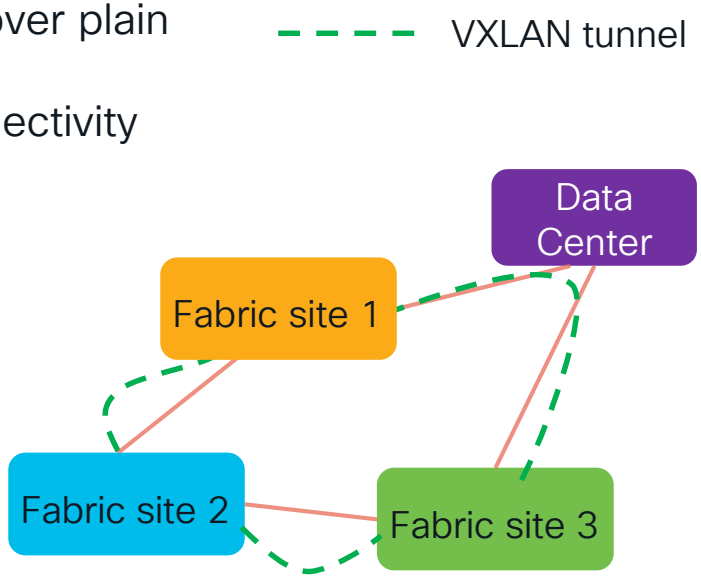
Why VXLAN?

- VXLAN carries VRF and SGT in the header over plain IP network.
- Transit network just need to provide IP connectivity between BN Looback0 interfaces.

Requirements:

- MTU > 1550 bytes.
- Dedicated Transit Control Plane(s).
- Multicast in the transit network*

*If overlay multicast is required



What About 70 Small Sites?

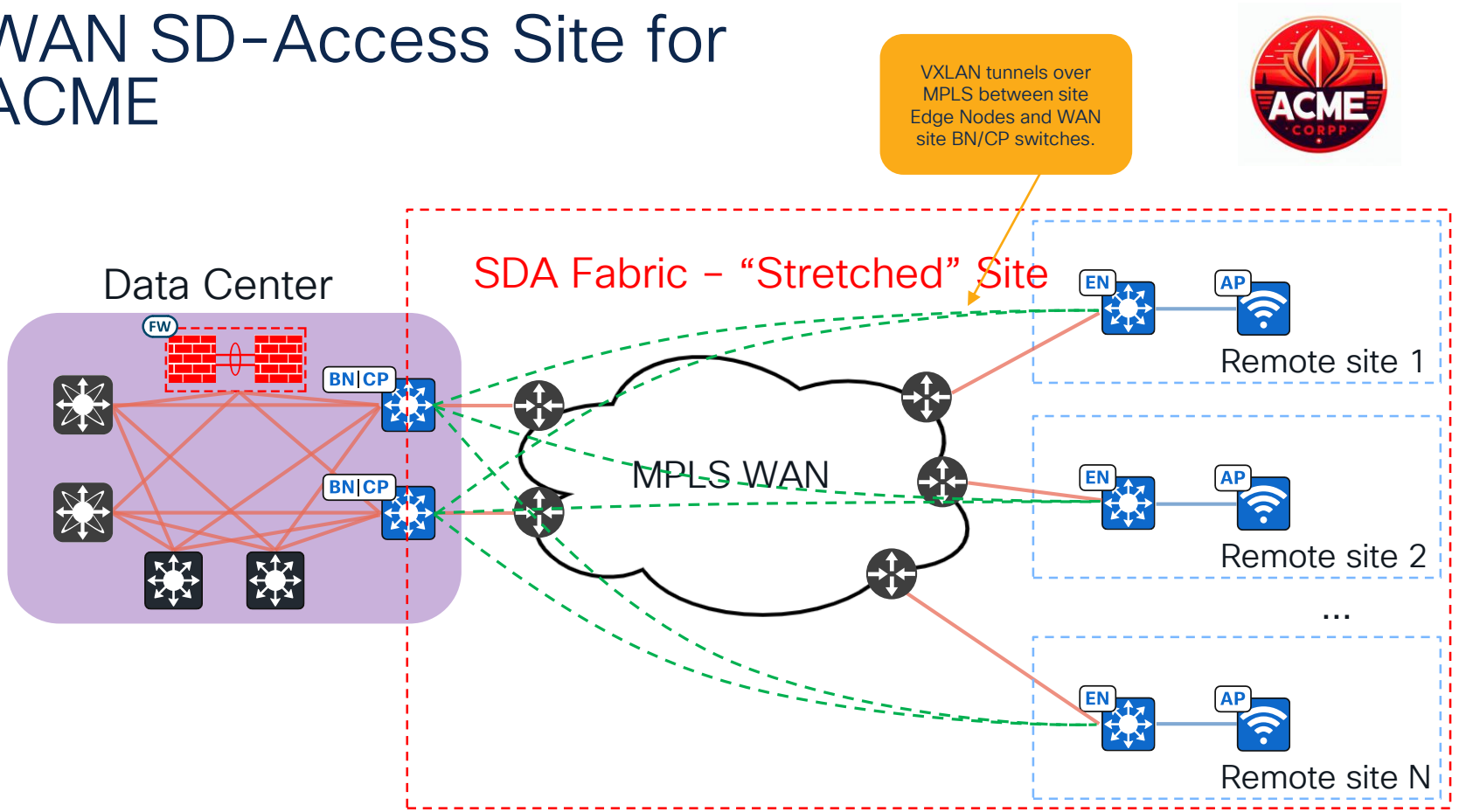
Two main options:

- 70 small individual fabric sites.
- 1 ”Stretched” fabric site.

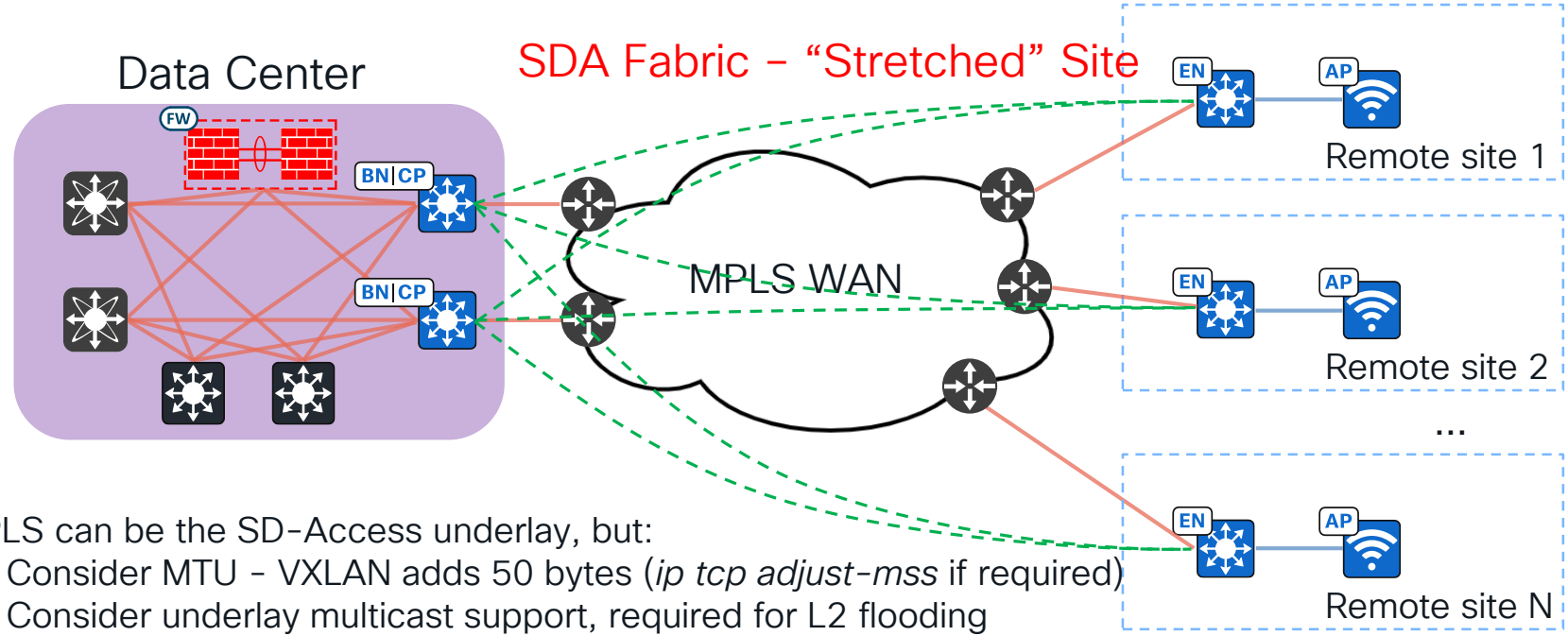
Can always mix and match.

	Individual site	“Stretched” site
Device SD-Access roles	Fabric in a box (FIAB)	Edges Nodes onsite, set of CP+BN nodes in central location
Management overhead	High – need to manage 70 sites individually (VRF, BGP, subnets are defined per site)	Low – all changes are performed on a single site
Survivability	High – each site is running its own set of CP/BN nodes	Low – all sites are running shared set of CP/BN nodes
Flexibility	High – each site can have DIA and unique routing policy	Low – all sites have single egress point – BN/CP at central location

WAN SD-Access Site for ACME



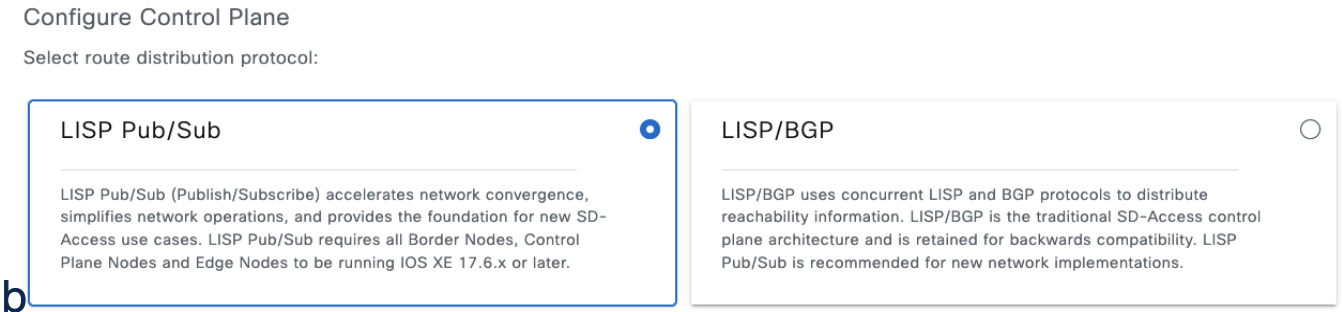
WAN SD-Access Site for ACME



MPLS can be the SD-Access underlay, but:

- Consider MTU - VXLAN adds 50 bytes (*ip tcp adjust-mss* if required)
- Consider underlay multicast support, required for L2 flooding

Control Plane – Pub/Sub or Not?



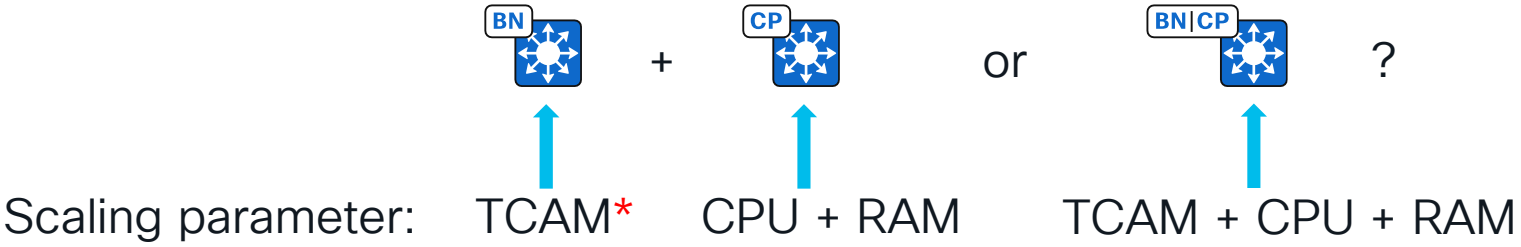
LISP Pub/Sub

- Released in 2022 with Catalyst Center 2.2.3.X and IOS-XE 17.6.X.
- Reliable and stable.
- Less Control Plane load.
- Faster convergence.

Greenfield: deploy LISP Pub/Sub.

- Requires default route (0.0.0.0/0) from upstream to work in External Border capacity.
- No longer need per-VN iBGP peering between Border Nodes.
- All sites connected via SDA Transit need to be on the same CP architecture (Pub/Sub or LISP/BGP).

Control Plane – Colocate with Border or Not?



- Border Node downloads all fabric host routes in switch TCAM.
- Control Plane RAM is non-issue from scale perspective.
- Main CPU stress for CP is handling wireless roaming for Fabric Enabled Wireless endpoints.
- It is safe to colocate until 50,000 EPs**, even in wireless-heavy environment.
- Can split BN and CP for architectural reasons (fault isolation, network modularity), rather than technical (scale).
- Avoid using routing platforms (C8K) as Control Plane and/or Border Nodes if possible.

*Number of host (/32 or /128) routes. **C9500H or above

Underlay Design Options – LAN Automation vs DIY

Underlay build:

- Configure Loopback0 interface (/32) on each SD-Access BN, CP, and Edge node.
- Set increased MTU to accommodate VXLAN header overhead, vtp transparent and enable multicast routing.
- Configure point to point routed links between each switch in the topology.
- Enable routing protocol so that each switch in the topology can reach the Loopback0 of each other in the topology.
- Enable PIM on each point-to-point link, Loopback0 and configure anycast ASM RP on CP/BN nodes.
- Configure SNMP and SSH credentials and that's it!

```
interface Loopback0
 ip address 10.0.0.12 255.255.255.255
 ip pim sparse-mode
 ip ospf network point-to-point
 ip ospf 1 area 0
```

```
system mtu 9100
```

```
vtp mode transparent
```

```
ip multicast-routing
```

```
interface GigabitEthernet1/0/1
 description UNDERLAY ROUTED UPLINK
 no switchport
 ip address 10.0.2.6 255.255.255.252
 no ip redirects
 ip pim sparse-mode
 ip ospf network point-to-point
 ip ospf 1 area 0
 bfd interval 250 min_rx 250 multiplier 3
 no bfd echo
```

```
router ospf 1
 router-id 10.0.0.12
 nsf ietf
 passive-interface Loopback0
 bfd all-interfaces
```

```
ip pim rp-address 10.0.0.1
ip pim register-source Loopback0
ip pim ssm default
```

Underlay Design Options – LAN Automation vs DIY

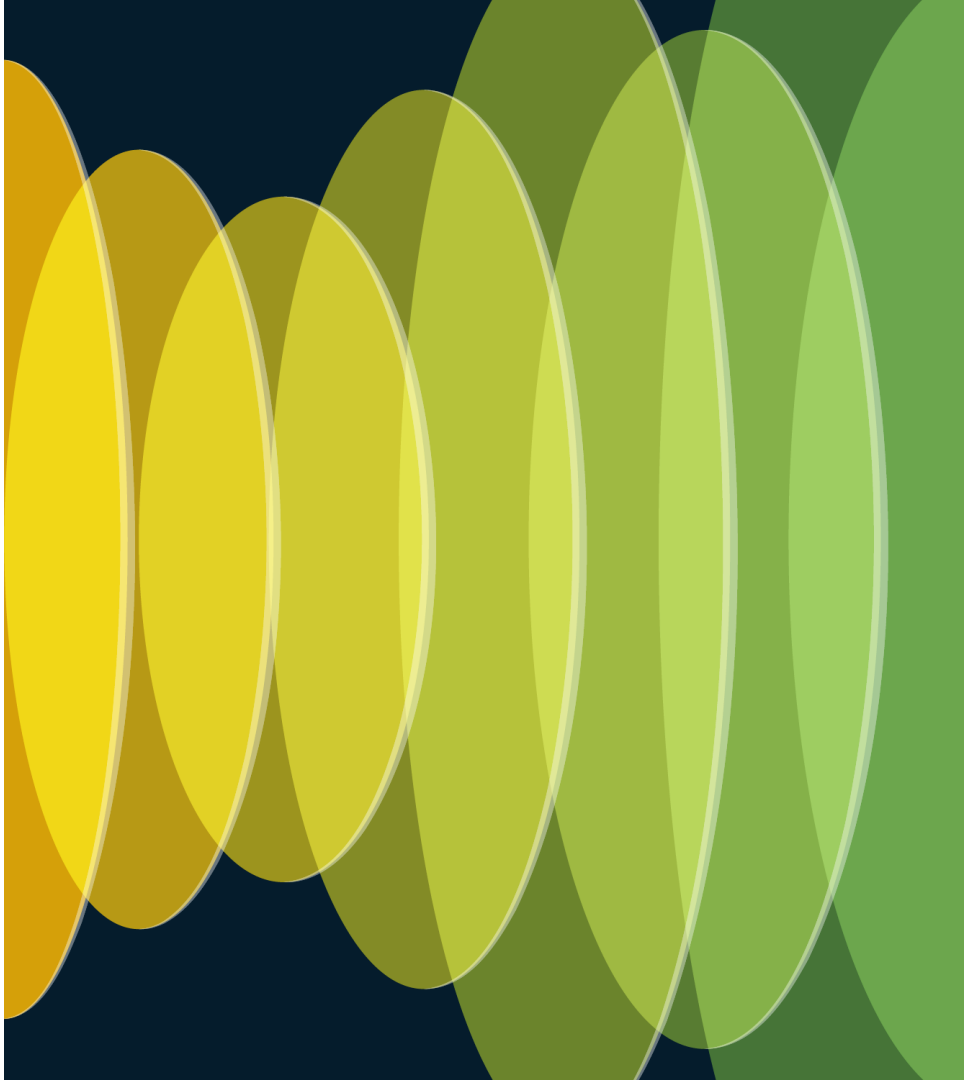
	LAN Automation	DIY
Solution approach	Turnkey automation	CLI template or CLI
Routing Protocol	IS-IS (single Level-2 area)	Any (most organisations deploy OSPFv2)
IPv4 address allocations	Separate pools for loopbacks and P2P interfaces (in CC 2.3.5 and later)	Anything is possible (as long as it's IPv4)
Multicast configuration	Yes	Yes
BFD configuration	Yes	Yes
STP configuration	Yes	Yes
MTU configuration	Yes	Yes

OSPF or IS-IS for SD-Access Underlay?

	LAN Automation	DIY
Routing Protocol	IS-IS (single Level 2 area)	Any (most organisations deploy OSPF)

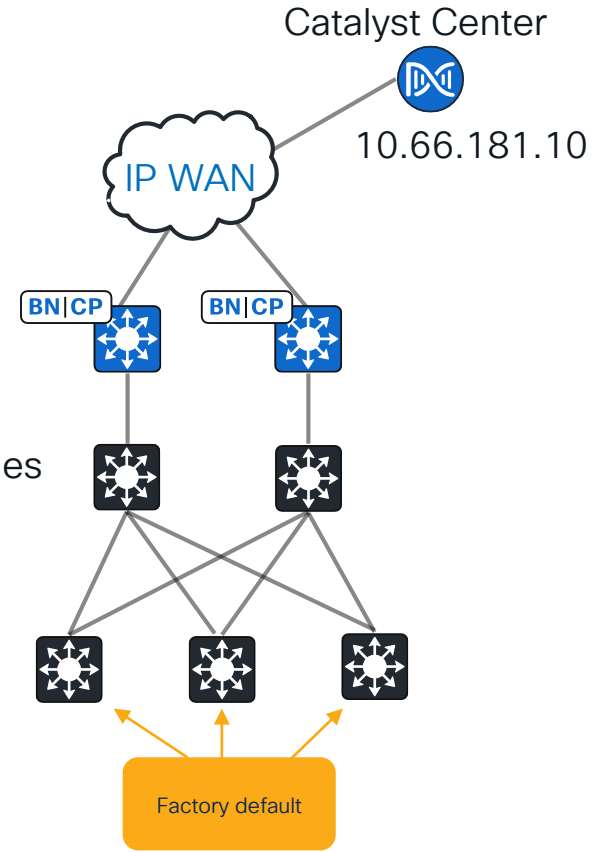
1. LISP needs /32 host route for destination VTEP Loopback0 to be present in forwarding table.
2. Maximum tested/supported L3 switches in link-state protocol area is 250.
3. More than 250 switches in the network will require multi-area deployment.
4. IS-IS Level1 areas filter all inter-area prefixes, including Loopback0 host routes (injects 0/0 route instead). OSPF areas allow inter-area routes by default.
5. Solution?
    a) Implement IS-IS multi-area design and configure Level2->Level1 route leaking (manually).
    b) Implement OSPF multi-area design (manually).

Underlay Automation Demo



Demo Topology

- C9500-1: automation pool 10.250.15.0/24
- C9500-2: automation pool 10.250.16.0/24
- Loopback range – 10.250.255.0/24

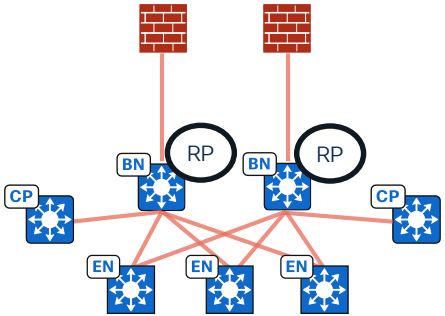


Template sources:
https://github.com/sergeynasonov/sda-underlay-templates

Underlay Multicast

Multicast in underlay is no longer optional. It is required for:

- Layer 2 flooding (broadcasts) in user overlays – most deployments have this.
- Layer 2 border functionality – most deployments have this.
- Multicast support in overlays.



Where should fabric underlay RPs go?

- Configure underlay Anycast RPs for the SDA site on BN/CP nodes.
  - *Use separate Loopback (not Loopback0) interfaces for RP source*
  - *Setup MSDP between two Border Nodes / RPs*
  - *Configure static RPs (no BSR / Auto-RP)*
  - *Enable PIM sparse on all P2P links and Loopback interfaces*

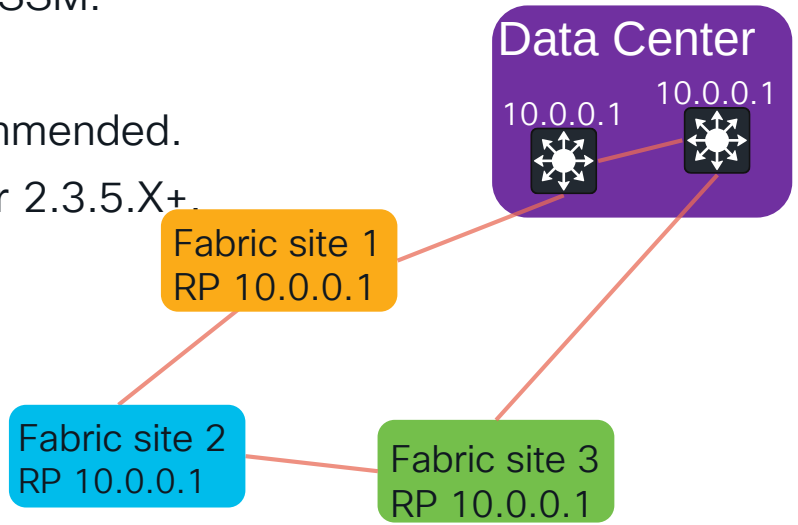
Multicast with SD-Access Transit – Underlay

For your reference

Underlay requirements:

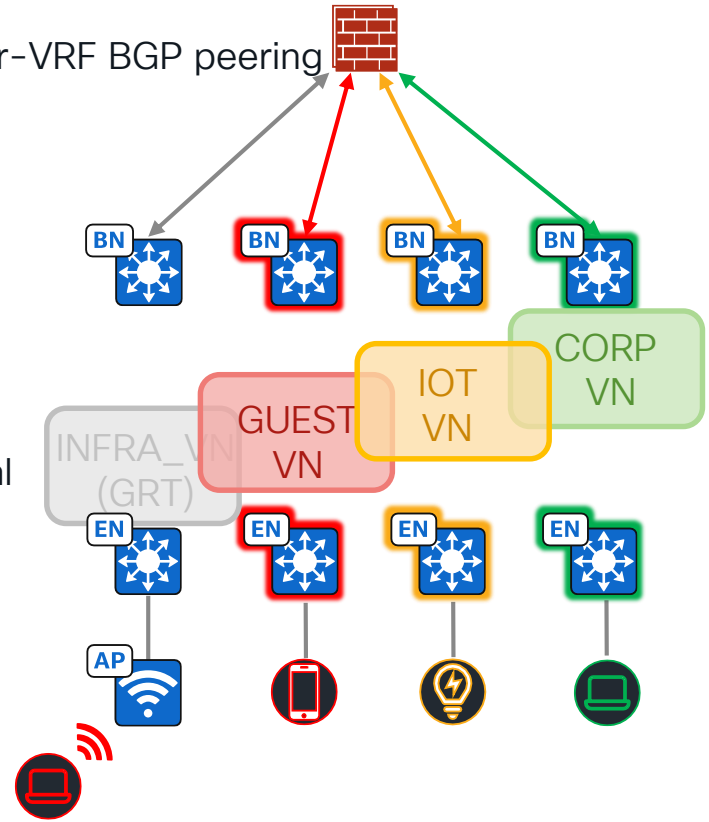
- Underlay links between fabric sites support PIM-SSM.
- All fabric sites use the same set of underlay RPs.
- RPs outside the fabric (external) are highly recommended.
- Minimum SW version is 17.10.1 / Catalyst Center 2.3.5.X+.

```
ip pim rp-address 10.0.0.1
ip pim register-source Loopback0
ip pim ssm default
```



Overlay Unicast

- Broadcasts are supressed by default in SD-Access -> make large subnets for users (10k hosts in IP pool is fine).
- Avoid migrating subnets “as-is” into the fabric.*
- Sum of subnets and pure L2 overlays cannot exceed 1000 per fabric site with Catalyst Center –XL (100 and 300 in smaller appliance versions).
- Catalyst Center has deployment-wide 1.5m physical + logical interface limit. Each IP pool creates 2 interfaces (SVI + LISP tunnel) on each switch in the fabric.
- 700 IP pools will require (2*700*1300+48*2100) 1,920,800 interfaces, which is above 1.5m limit.
- 100 IP pools in 1300 stacked-switch fabric** will contribute (2*100*1300+48*2100) 360,800 ports to that limit.

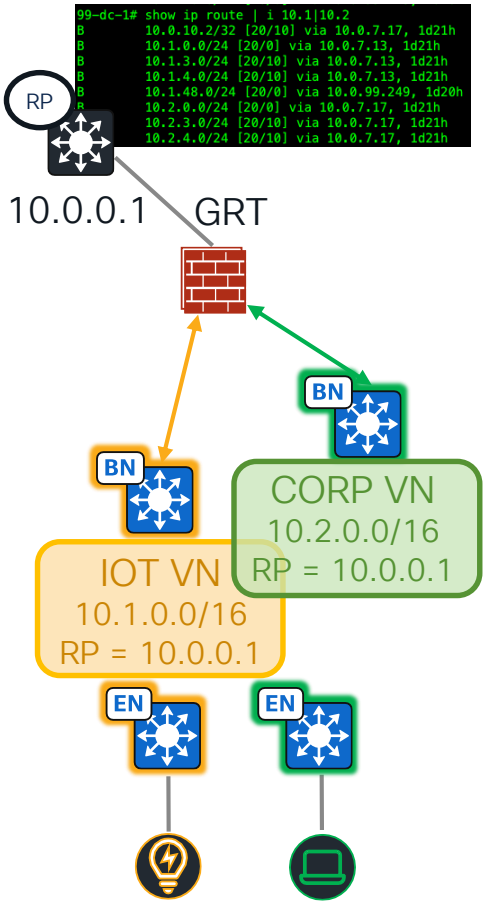


*Requirement reference: 700 VLANs for users and device segmentation.

**Requirement reference: 2100 x WS-C2960X access switches in 1300 switch cabinets.

Overlay Multicast

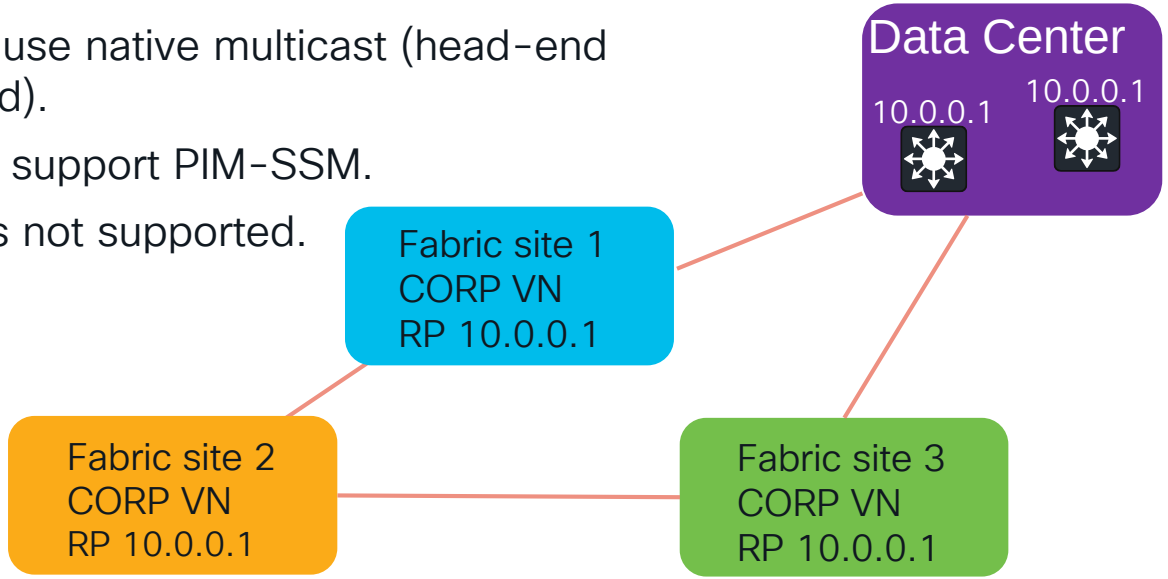
- Overlay multicast requires multicast-enabled underlay (avoid head-end replication).
- Overlay multicast is enabled per Virtual Network (VRF) rather than per IP pool (subnet) and needs an IP pool per multicast-enabled VN.
- Both internal and external RPs are supported (use external RP if possible).
- Multicast route-leaking is not supported on C9K platform.
- If you have sources/receivers in different Virtual Networks, use external RP and perform route-leaking outside of fabric (e.g. on Fusion).
- As of now, SDA fabric supports all multicast flow variations in overlays:
  - ASM and SSM (concurrently)
  - Sources and Receivers inside fabric
  - Sources inside fabric, Receivers outside fabric
  - Sources outside fabric, Receivers inside fabric



Overlay Multicast in SD-Access Transit

Multicast over SDA Transit (in VXLAN) is supported when:

- Multicast-enabled VNs in all sites are configured with the same set of RPs (per VN).
- All sites are configured to use native multicast (head-end replication is not supported).
- Links between fabric sites support PIM-SSM.
- Pub/Sub only, LISP/BGP is not supported.



Upstream Connectivity – Fusion Firewall

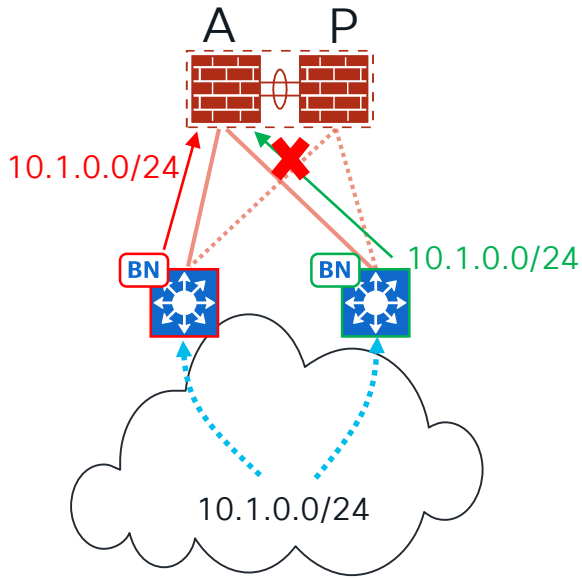
Active/Active Borders with two uplinks to HA firewalls (active/passive pair)

Problem:

- Each BN will register itself as active gateway for fabric.
- Each BN will advertise fabric subnets via BGP with the identical AS-PATH length (and other BGP attributes) to the firewall.
- The firewall will receive two equal routes via two next-hops and will only install one (by default).
- Inevitably half the traffic will arrive to firewall via the other interface (facing BN Green) and will get dropped.

Solution?

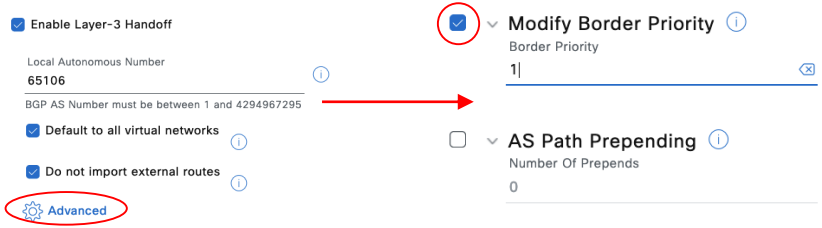
Destination	Interface	Next-hop
10.1.0.0/24	Eth1/1	BN Red



Upstream Connectivity – Fusion Firewall

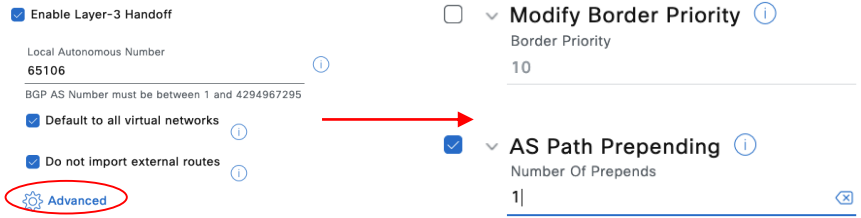
Solution 1 – Make Border Nodes Active/Passive too.

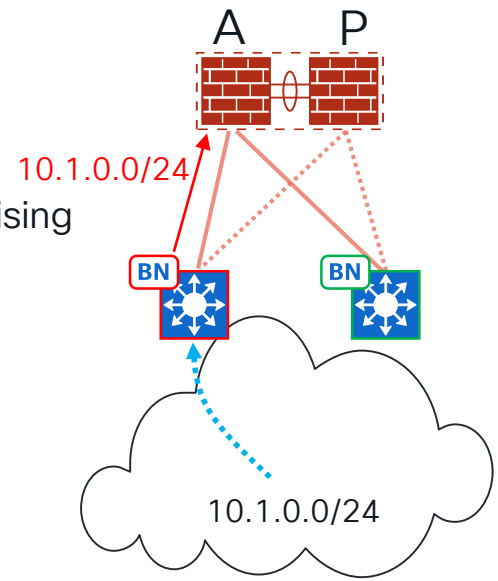
1. Configure Border Red to have better LISP priority as fabric exit (smaller the better, default value is 10).



Destination	Interface	Next-hop
10.1.0.0/24	Eth1/1	BN Red

2. Configure Border Green to add AS-PATH prepend while advertising fabric subnets to the firewall.





cisco Live!

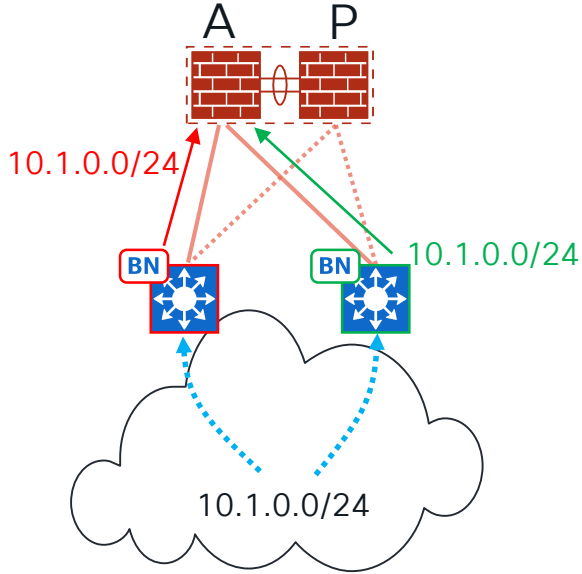
Upstream Connectivity – Fusion Firewall

Solution 2. ECMP on firewall cluster.

Configure Equal-Cost Multipathing (ECMP) on the firewall so that both next-hops are installed in the firewall forwarding table:

- Each mainstream firewall vendor supports this functionality.
- Cisco FTD firewalls support this from FTD 6.5.
- Requires interaction with firewall team (I know!).
- Pay attention to Multicast and ECMP interaction.

Destination	Interface	Next-hop
10.1.0.0/24	Eth1/1	BN Red
10.1.0.0/24	Eth1/2	BN Green



Upstream Connectivity – Fusion Firewall

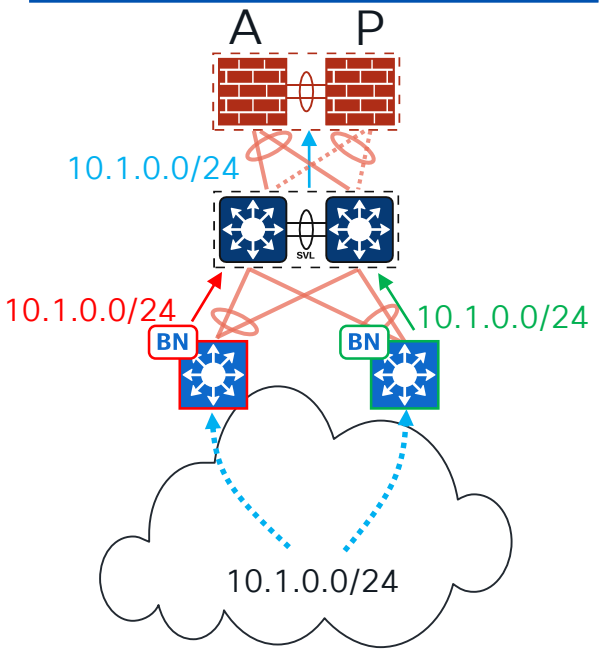
Solution 3. Intermediate hop.

Make only single interface on the firewall by inserting another L3 hop (typically stacked switch) between BNs and the firewall pair. Repeat the configuration per fabric VN (VRF).

This approach:

- Creates single logical point of failure in otherwise highly-available network.
- Requires extra hardware to procure and configure.
- Adds more moving parts, making ongoing operational changes lengthy and more complex, ultimately driving down the network uptime.

Destination	Interface	Next-hop
10.1.0.0/24	Po1	SVL stack



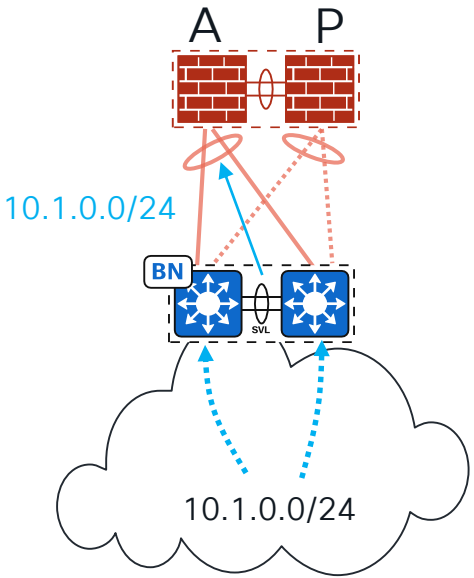
Upstream Connectivity – Fusion Firewall

Solution 4. Stack Border Nodes.

Make only single interface on the firewall by stacking Border Nodes. Please avoid.

- Single point of failure, especially if you collocate CP and BN roles.
- Hardware changes require SVL reboot (=fabric outage).
- No In-Service Software Upgrade (ISSU) for SVL in SD-Access.

Destination	Interface	Next-hop
10.1.0.0/24	Po1	BN Stack

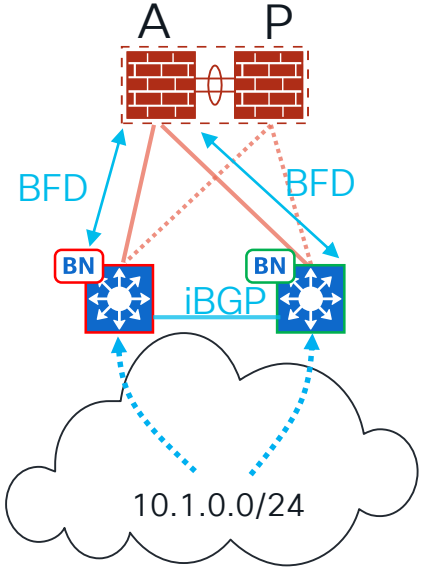


Upstream Connectivity – Fusion Firewall

General Observations

- Network link and per-VN iBGP peering between Border Nodes is no longer required with Pub/Sub fabric control plane.
- Configure BFD to speed up BGP convergence.
- Make sure to research firewall vendor High Availability implementation, to make sure BFD does not trigger BGP adjacency drop during firewall failover.
- Catalyst Center still provisions iBGP peering between BNs in GRT. Configure "*bgp neighbor fall-over*" on that peering to speed up upstream BGP convergence.

Destination	Interface	Next-hop
10.1.0.0/24	Eth1/1	BN Red
10.1.0.0/24	Eth1/2	BN Green

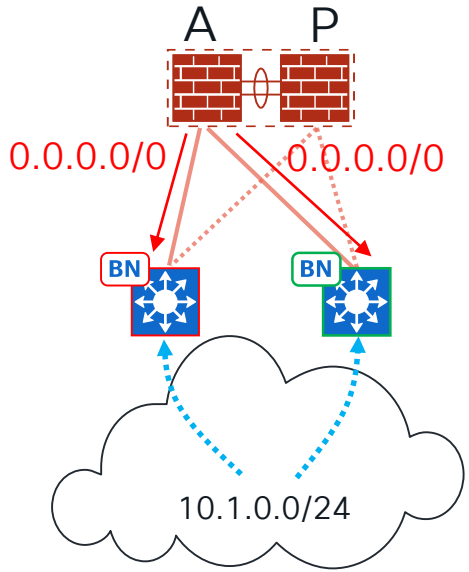


Upstream Connectivity – Fusion Firewall

General Observations

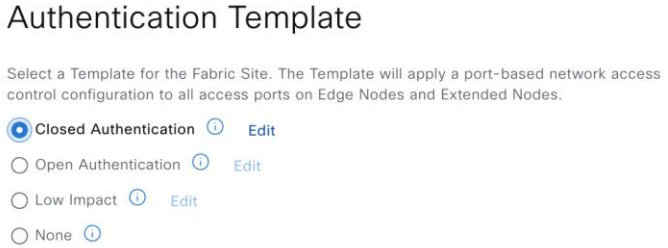
- Network link and per-VN iBGP peering between Border Nodes is no longer required with Pub/Sub fabric control plane.
- Configure BFD to speed up BGP convergence.
- Make sure to research firewall vendor High Availability implementation, to make sure BFD does not trigger BGP adjacency drop during firewall failover.
- Default route (0.0.0.0/0) advertisement is required from the firewall to enable External Border Node functionality.

Destination	Interface	Next-hop
10.1.0.0/24	Eth1/1	BN Red
10.1.0.0/24	Eth1/2	BN Green



Switchport Access Policy

- Closed authentication – 802.1X + MAB (IBNS 2.0 template). No DHCP/ARP before authentication.
- Open Authentication – 802.1X + MAB. Even if you fail authentication, you are still allowed.
- None – no authentication, all ports are statically configured.
- Can always start with None, then change later.
- Migrate the existing switchport policy as part of fabric rollout.



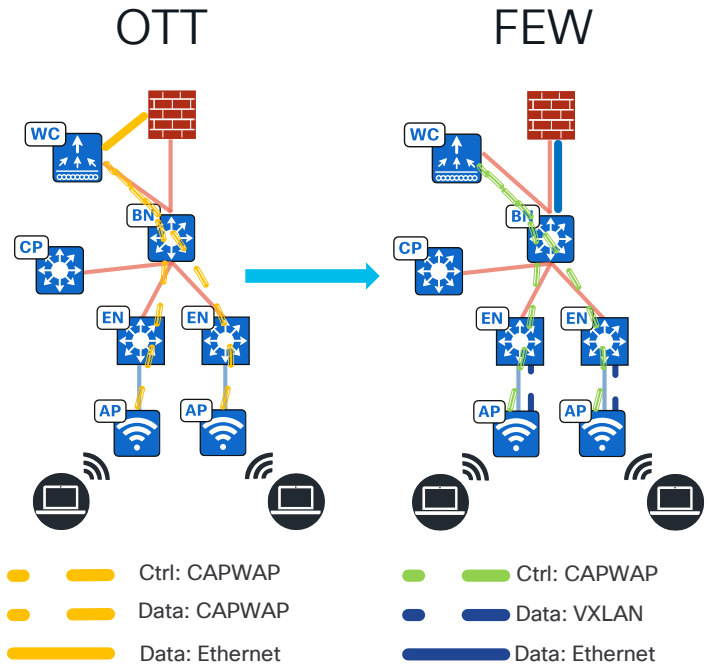

```
interface GigabitEthernet1/0/2
 switchport mode access
 device-tracking attach-policy IPDT_POLICY
 ip flow monitor dnacmonitor input
 ip flow monitor dnacmonitor output
 ipv6 flow monitor dnacmonitor_v6 input
 ipv6 flow monitor dnacmonitor_v6 output
 dot1x timeout tx-period 7
 dot1x max-reauth-req 3
 source template DefaultWiredDot1xClosedAuth
 spanning-tree portfast
 spanning-tree bpduguard enable
 ip nbar protocol-discovery
```

```
template DefaultWiredDot1xClosedAuth
 dot1x pae authenticator
 dot1x timeout supp-timeout 7
 dot1x max-req 3
 switchport mode access
 switchport voice vlan 2046
 mab
 access-session closed
 access-session port-control auto
 authentication periodic
 authentication timer reauthenticate server
 service-policy type control subscriber PMAP_DefaultWiredDot1xClosedAuth_1X_MAB
!
```

Wireless Considerations

For your reference

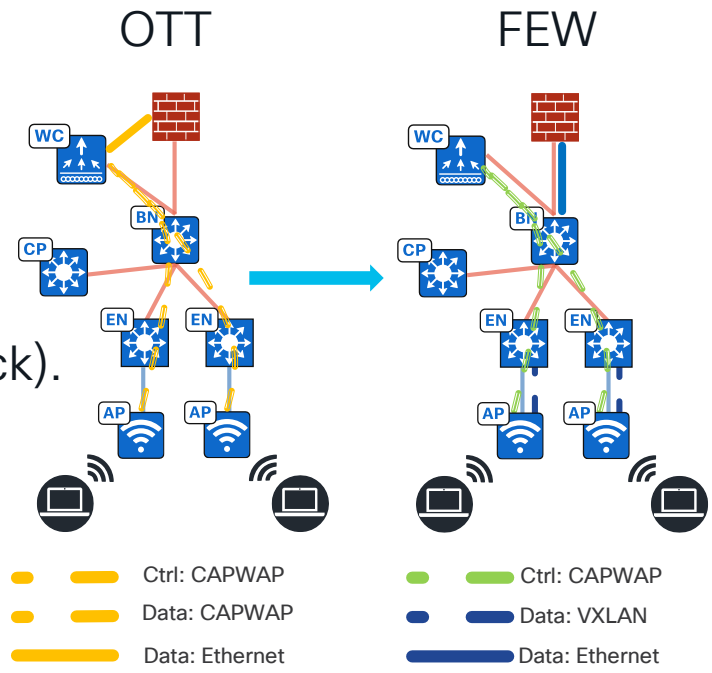
- Wireless configuration needs to be managed by Catalyst Center.
- Can have Fabric-Enabled Wireless (FEW) and Centrally Switched/Flex (OTT in SDA lingo) mode for the same SSID across different sites.
- Can have mix of SSIDs (FEW vs OTT) on the same AP.
- Can have fabric APs and non-fabric APs on the WLC.
- If multicast is required on OTT SSID, AP pool in INFRA_VN needs to be multicast-enabled via CLI template ("*ip pim sparse*" under AP pool SVI).



Wireless Considerations – My Take?

For your reference

1. Keep wireless “as is”.
2. Finalise switching migration first (block by block).
3. Then convert wireless to FEW in one go.



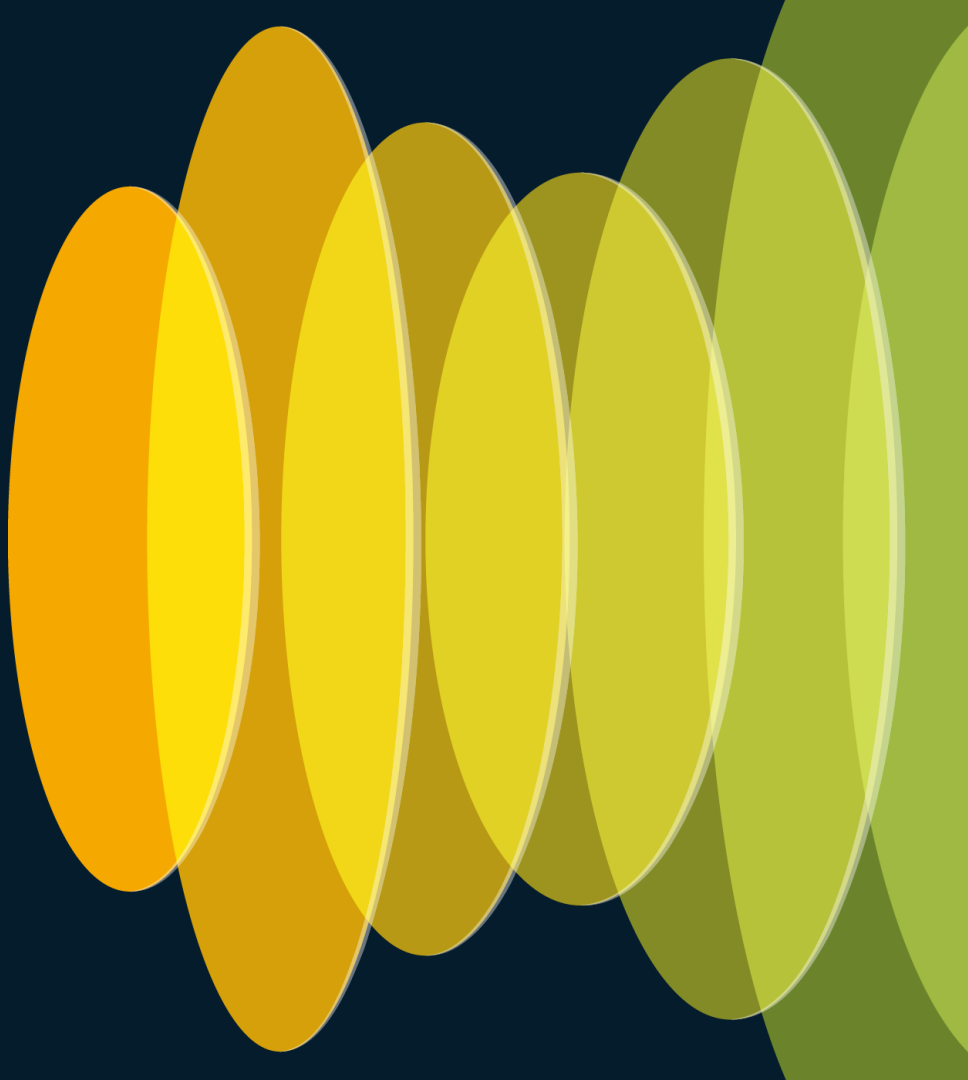
Key Design Decisions

	Design Decision	Rationale
D1	Divide main campus into 3 fabric sites.	Cannot implement single fabric site, number of fabric devices is >1200. Three geographical sub-sites align with proposed fabric site structure.
D2	Implement SDA Transit between 3 fabric sites in the main campus.	Need to maintain unified macro- and micro-segmentation policy across all three fabric sites that make up ACME campus.
D3	Use colocated BN/CP roles.	Each individual site will not exceed more than 50,000 EP. Two BN/CP switches will provide adequate level of resilience of the fabric site.
D4	Implement one “Stretched” fabric site for 70 small branch sites across the WAN.	1. MPLS sites do not have local server resources or DIA and are accessing all resources via the centralised data center. 2. MPLS carrier can support MTU > 1550. 3. Small branch sites do not have overlay multicast and L2F requirements.
D5	Use OSPFv2 as underlay routing protocol for the fabric.	1. LAN Automation (with IS-IS) cannot be used due the scale of the deployment, necessitating multi-area design. 2. ACME IT team has a lot of experience with OSPF and is not comfortable with IS-IS manual deployment.
D6	Use the external set of multicast RPs for overlay VNs.	ACME has multicast sources in IoT VN (AppleTV & Printers) and receivers in Corp VN.

Final BOM

- Catalyst Center: 3 x DN3-HW-APL-XL.
- 4 fabric sites: 8 x C9500-32C core switches for each (running BN+CP roles), 2 per fabric site (include “stretched” site).
- 4 fabric sites: 8 x C9800-40-K9 WLCs, 2 per fabric site (“stretched” site still needs WLC).
- SDA Transit Control Plane: 2 x C9500-24Y4C (per deployment).
- Existing ISE (make sure it has ISE Advantage Licenses for expected concurrent EP quantity).
- Distribution and access switches follow the traditional networking pattern.

Implementing SD-Access



cisco Live!

Project Flow

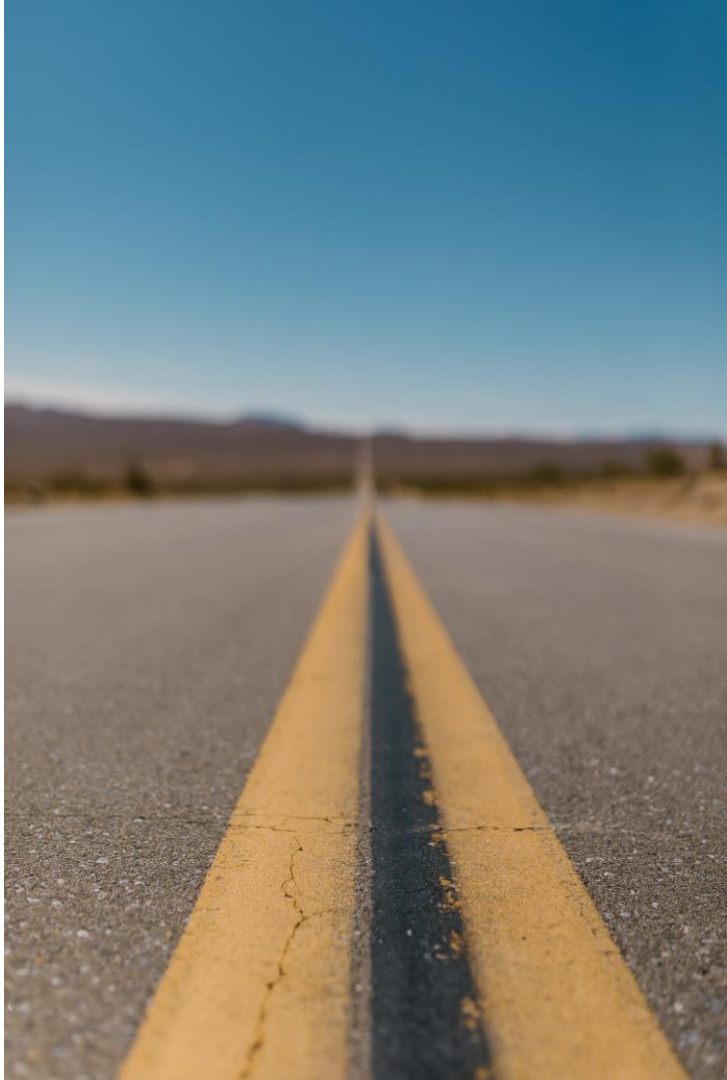
M1. Build management stack (Catalyst Center).

M2. Integrate Catalyst Center with existing ISE.

M3. Deploy new core switches in parallel to the existing (new Border Nodes).

M4. Migrate switching infrastructure – per distribution block (building), keeping existing L2 switchport policy (802.1X, MAB, open).

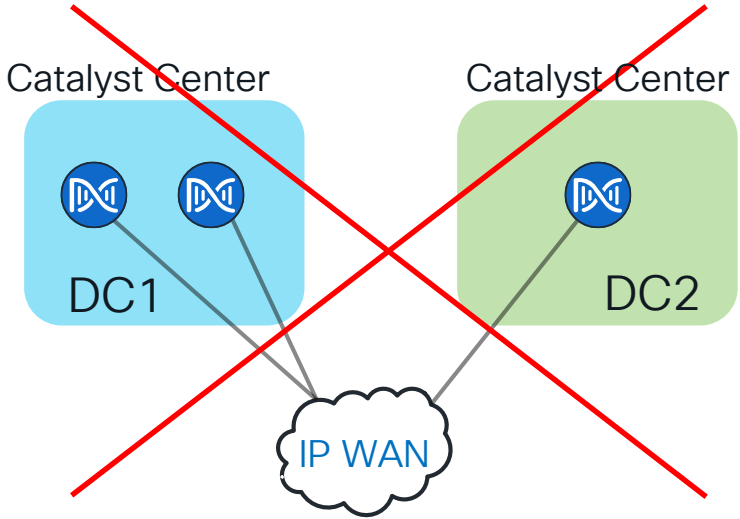
M5. Migrate wireless once wired network is fully converted.



M1. Building Management Stack

SD-Access requires Catalyst Center as automation engine.

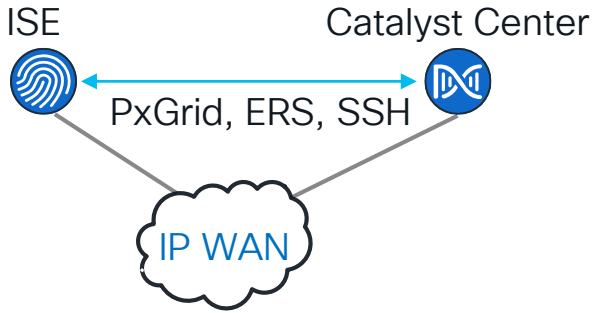
- If high availability (HA) is required – deploy 3 node cluster.
- Avoid splitting 3 cluster nodes across 2 separate locations.
- Deploy Catalyst Center in 1:1 or 3:3 mode if disaster recovery (DR) is required.
- Virtual (AWS or ESXi) Catalyst Center appliance does not have native HA or DR capabilities as of today (June 2024).

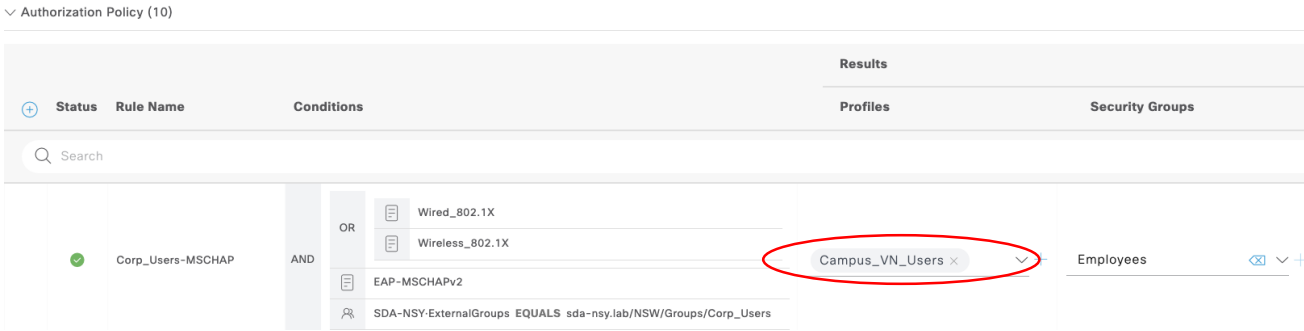


If you lose DC1, single Catalyst Center node in DC2 will shut down automatically.

M2. Integrate with Existing ISE

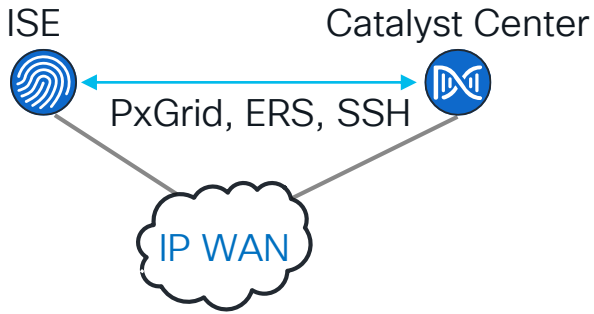
- One Catalyst Center cluster can only be integrated with a single ISE cluster.
- Reuse existing authentication flows and add new SD-Access specific authorization profiles.



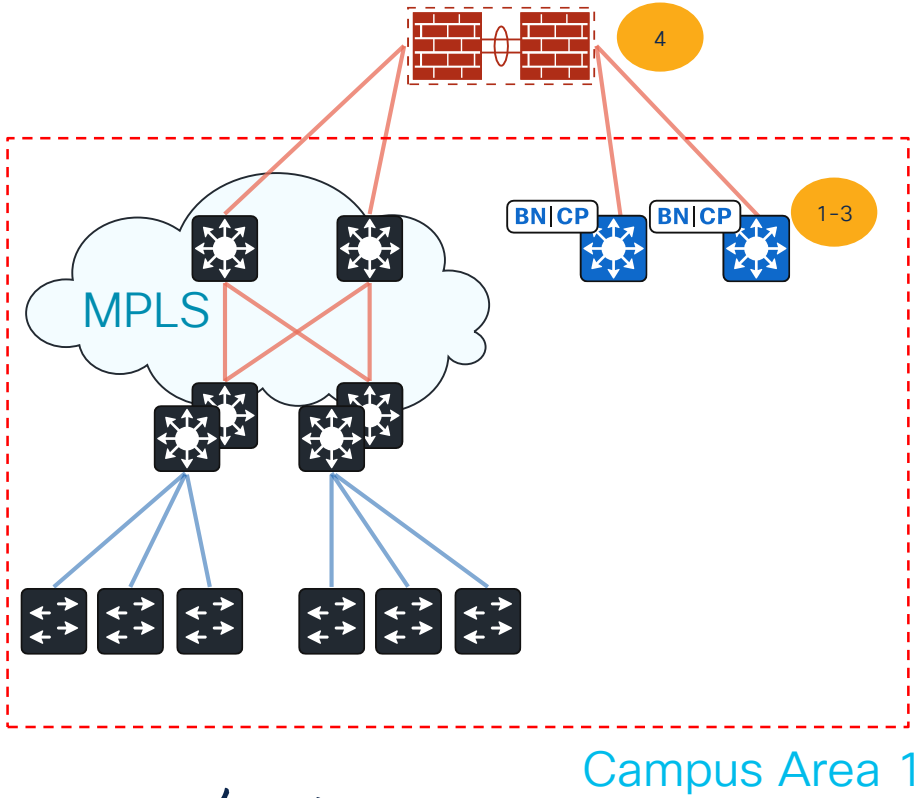


M2. Integrate with Existing ISE

- One Catalyst Center cluster can only be integrated with a single ISE cluster.
- Reuse existing authentication flows and add new SD-Access specific authorization profiles.
- Changing already-integrated ISE cluster requires removal of all SD-Access fabric sites in Catalyst Center.

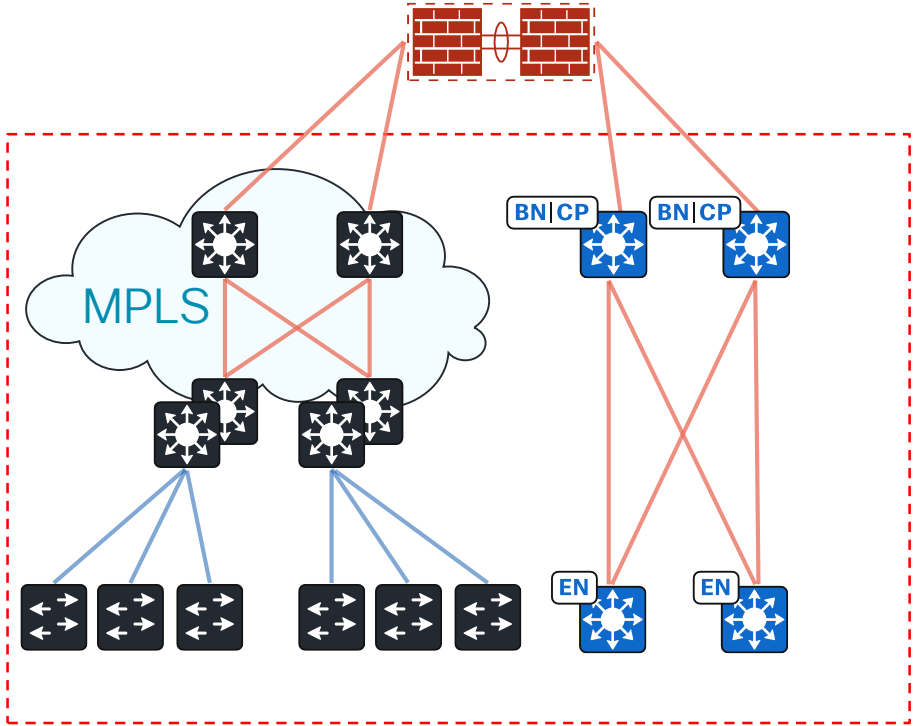


M3. Parallel Core



1. Deploy new core switches in parallel to the old.
2. Add to new switches to the Catalyst Center and enable BN + CP roles for the new fabric site.
3. Configure required VNs (=VRF) in Catalyst Center and assign to the new fabric site.
4. Configure BGP peerings for underlay and new VNs between new Border Nodes and the fusion firewall.

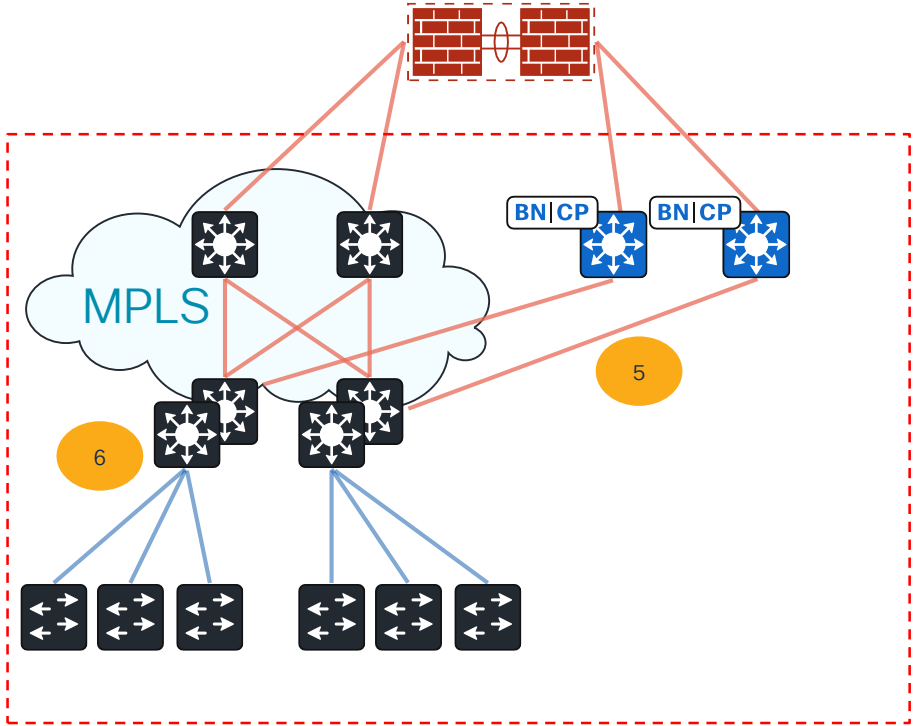
M3. Test Configuration Before Migration



At this point, fabric configuration complexity does not increase with the growth of access switches:

- Create all final-state subnets/Anycast gateways.
- Bring and test all endpoint classes, focusing on authentication, multicast, exotic use-cases (PC imaging, Wake-on-LAN, etc).
- Test fabric failover (shutdown border, unplug links, etc.). Border configuration does not change if fabric has 2 Edge Nodes or 200 Edge Nodes.

M3. Reuse Existing Distribution

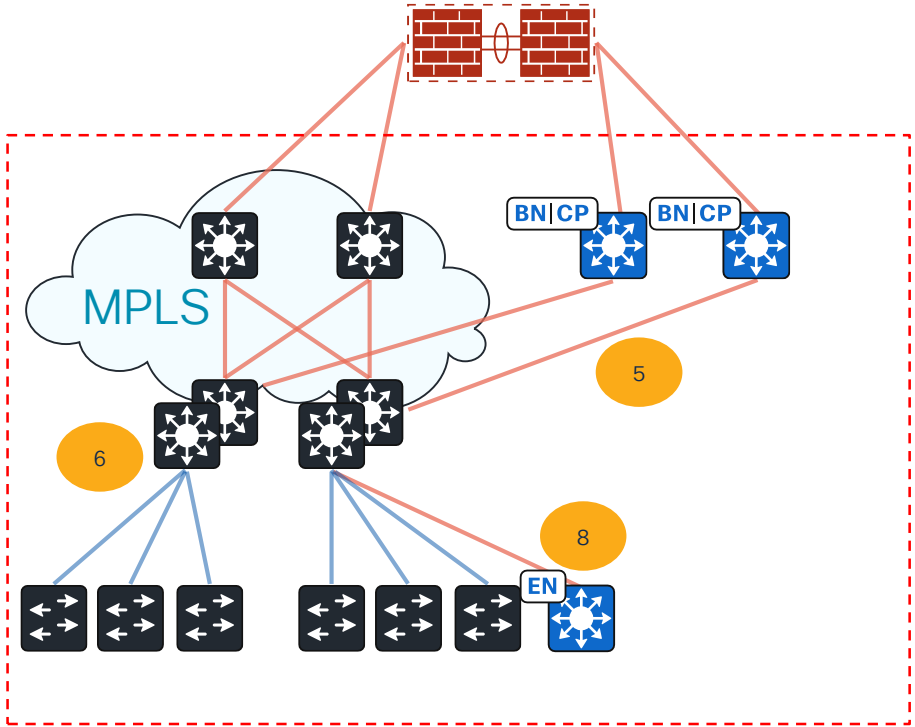


5. Establish routable connection from distribution switches to new core.
6. Adjust MTU and enable multicast if required.
7. Decision point:
   a) Parallel access build
   b) Convert existing access switches

Parallel vs Incremental Access Migration

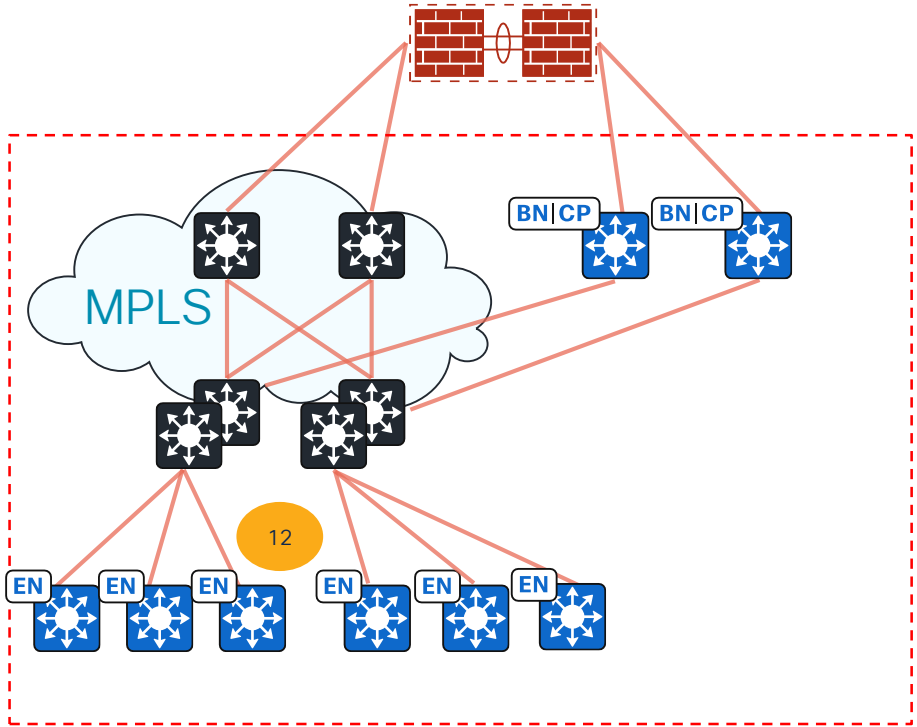
	Parallel Build	Incremental Switch Conversion
Nature of the change	Build SD-Access switch next to traditional and repatch.	Convert existing switch to be SD-Access enabled.
Hardware requirements	Can migrate from previous generation of Cisco switching (e.g., C2K, C3K).	Need C9K switch with DNA Advantage licenses already installed.
Extra Space, Power and Cabling and Requirements	Need additional space and power outlets for at least one switch, as well as additional fibre runs (usually 2 per switch).	None.
Risk	Low – switch build and testing happen outside the maintenance window. EP migration can be incremental. Simple incremental rollback.	Medium – switch build, testing and EP migration happen inside the maintenance window. Rollback requires device wipe and loading old config.

M4. Parallel Build



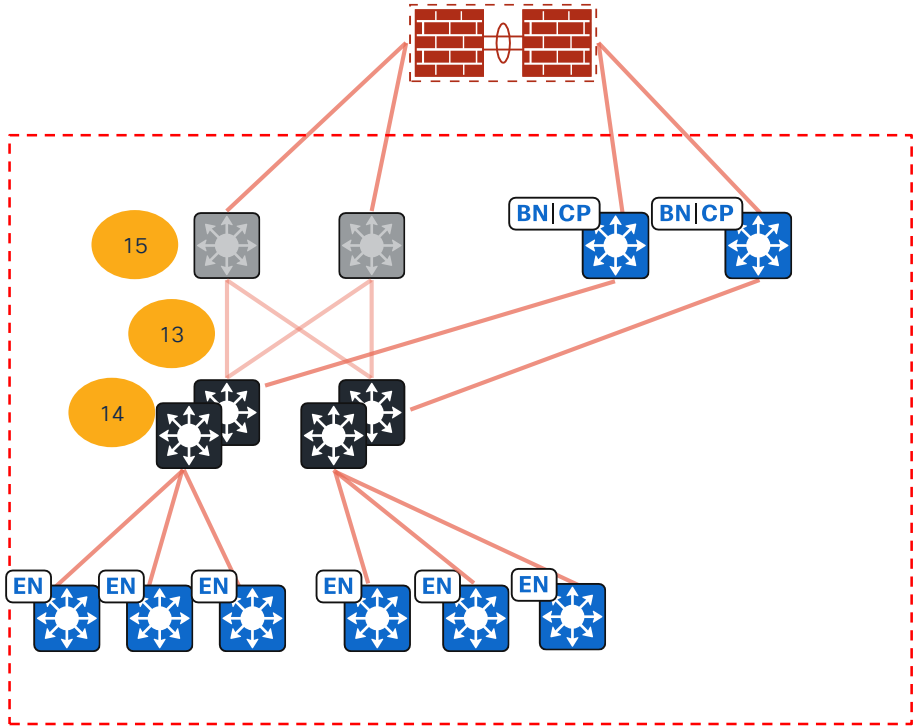
8. Deploy new C9K switch in parallel with the old C2K:
   - Routed P2P links, Loopback0 interface. Advertise in OSPFv2.
   - MTU >1550.
   - PIM sparse on interfaces and multicast RP configuration.
   - SNMP and SSH credentials.
9. Discover new switch in Catalyst Center and assign EN role.
10. Assign switchports to user VLANs if not using dynamic authentication via ISE.
11. Repatch endpoints.

M4. Parallel Build



12. Continue deploying new Edge Node switches and migrating endpoints until all access switches are replaced.

M4. Remove Legacy Configuration



Campus Area 1

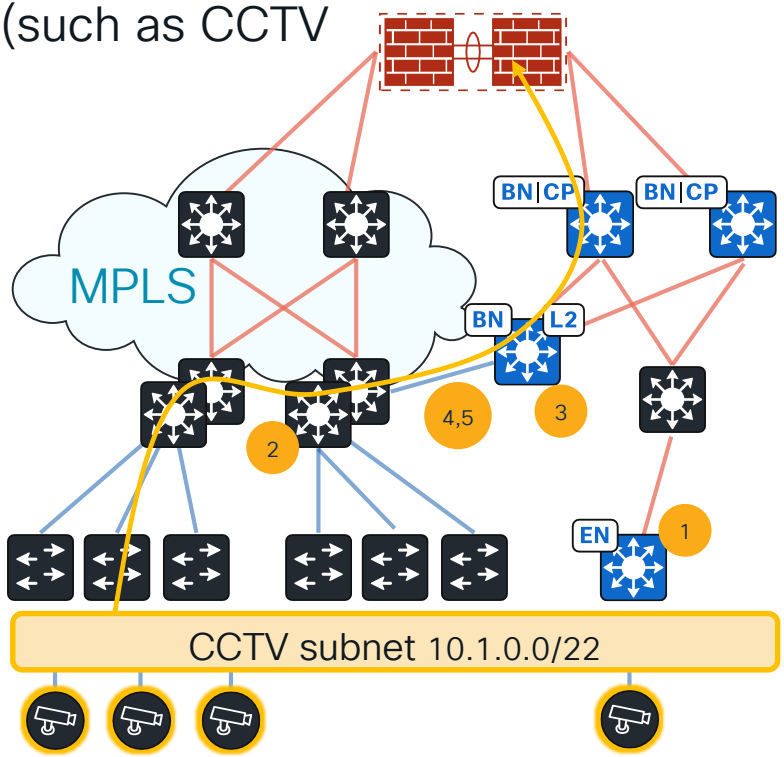
13. Once all access switches are migrated in the distribution block, remove MPLS configuration from distribution switches.
14. Remove stacking configuration from distribution switches.
15. Once all distribution blocks are ”migrated” to the fabric, legacy core switches can be removed.

Layer 2 Border – Gateway Inside the Fabric

Use-case: Stretch VLAN between fabric and traditional network

There are endpoints with static IP addresses (such as CCTV cameras)

1. Create anycast gateway inside the fabric.
2. Shutdown corresponding SVI in traditional network.
3. Configure BN with Layer 2 handoff (gateway inside the fabric) – L2 BN.
4. Optional: configure external VLAN ID if not matching fabric VLAN ID.
5. Allow VLAN on trunk between traditional network and L2 BN.
6. Cameras on old network will use SVI on L2 BN to reach the fabric and egress out.
7. A maximum of 6000 EPs can be connected outside the fabric.

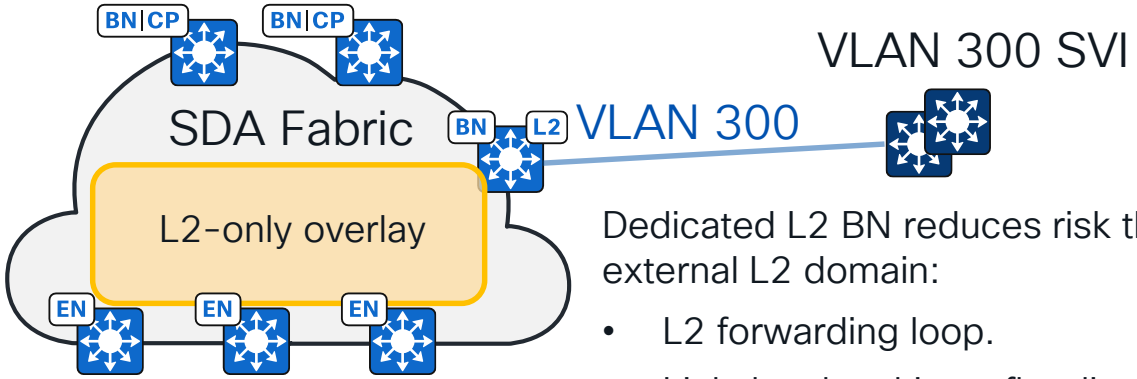


cisco Live!

Layer 2 Border – Gateway Outside the Fabric

Two use-cases:

- Endpoints that are not using IP (Profinet, Bacnet, Modbus and other industrial protocols) and relying on MAC layer / broadcasts for communication.
- Overlapping IP addresses in the overlay (multi-tenancy).



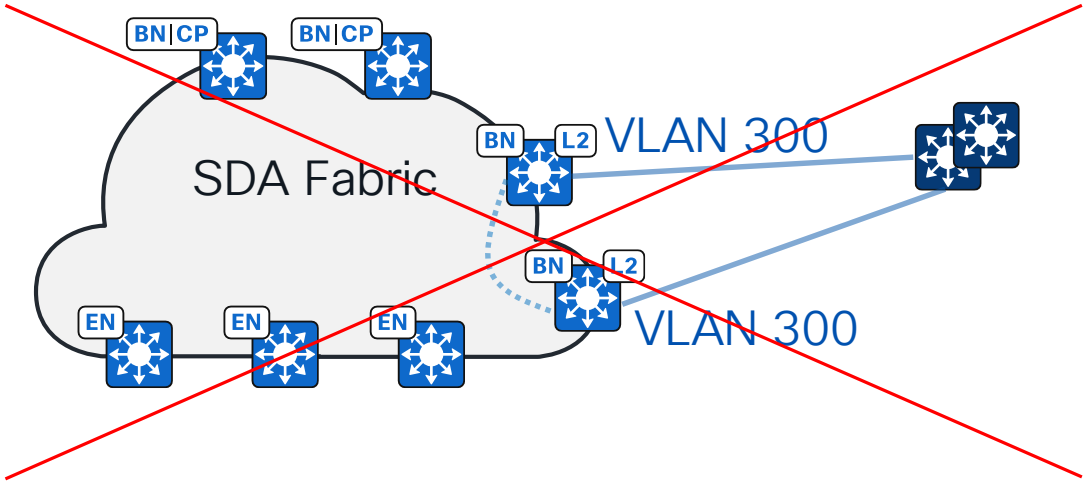
Dedicated L2 BN reduces risk that is created by attaching fabric to external L2 domain:

- L2 forwarding loop.
- Link-local multicast flooding.

L2 BN requires Layer 2 flooding to be enabled for the stretched segment.

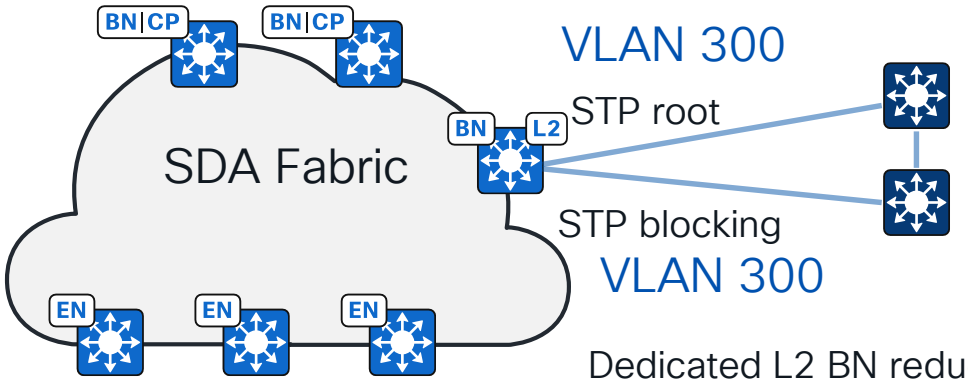
Layer 2 Border – Deployment Model

STP BPDUs are not tunnelled inside the fabric, but broadcasts are -> Same VLAN on two L2 BN handoffs will create a L2 forwarding loop.



Layer 2 Border – Deployment Model

Dual-homing from single L2 BN is supported.



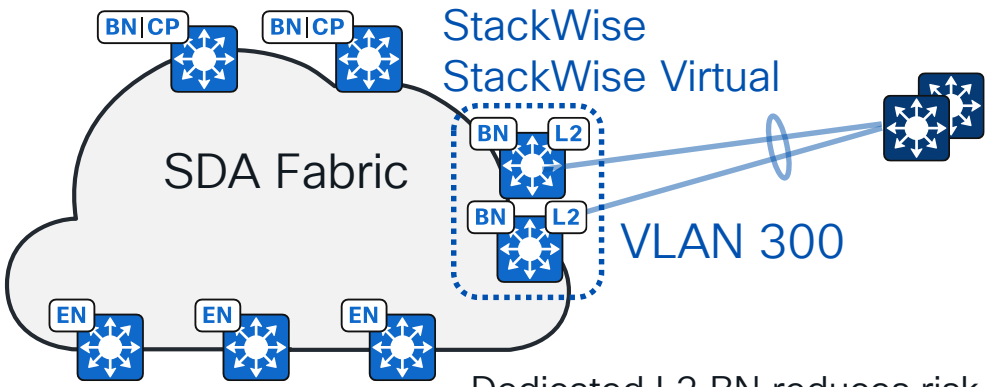
Dedicated L2 BN reduces risk that is created by attaching fabric to external L2 domain:

- L2 forwarding loop.
- Link-local multicast flooding.

L2 BN requires Layer 2 flooding to be enabled for the stretched segment.

Layer 2 Border – Deployment Model

Multi-chassis EtherChannel from stacked L2 BN is supported.



Dedicated L2 BN reduces risk that is created by attaching fabric to external L2 domain:

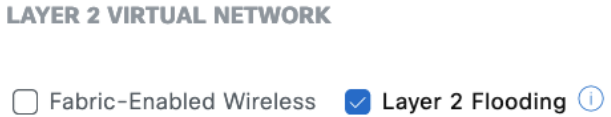
- L2 forwarding loop.
- Link-local multicast flooding.

L2 BN requires Layer 2 flooding to be enabled for the stretched segment.

Broadcast Traffic in Fabric

Also known as Layer 2 Flooding (L2F)

- Disabled by default as not having flooding enables large number of hosts in the same Layer 2 segment.
- Automatically enabled for segments stretched via L2 Border Node with gateways outside the fabric.
- Can enable manually (per subnet) – broadcast rules apply.
- Enabling L2F floods Ethernet broadcast and link-local multicast (TTL=1) in overlay.
- Requires multicast in underlay.
- Every deployment will have hosts that need L2F, so put them to separate VLAN/VNI and enable L2F there. Do not enable L2F on main VLANs with conventional endpoints.



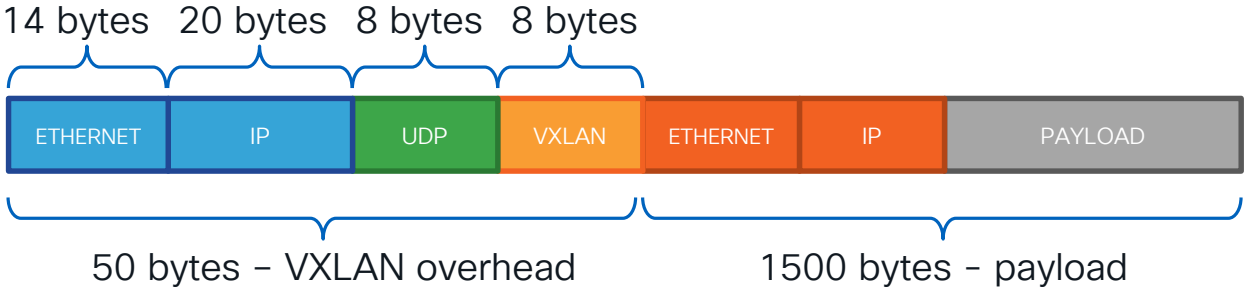
Fragmentation in VXLAN

RFC 7348 “Virtual eXtensible Local Area Network (VXLAN): A Framework for Overlaying Virtualized Layer 2 Networks over Layer 3 Networks” says:

“4.3 VTEPs MUST NOT fragment VXLAN packets. Intermediate routers may fragment encapsulated VXLAN packets due to the larger frame size. The destination VTEP MAY silently discard such VXLAN fragments.”

Solution?

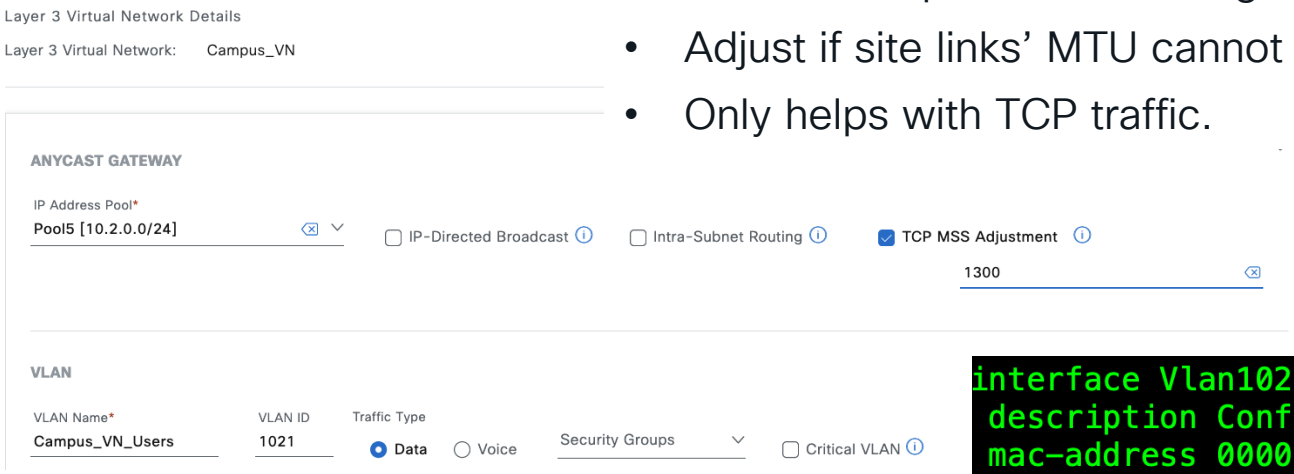
- Increase link MTU – within the campus.
- Adjust TCP MSS – over WAN (1300 is the magic number).



Fragmentation in VXLAN

ip tcp adjust-mss

- Per VLAN – pushed to all Edge Nodes within fabric site.
- Adjust if site links' MTU cannot be set to >1550 bytes.
- Only helps with TCP traffic.

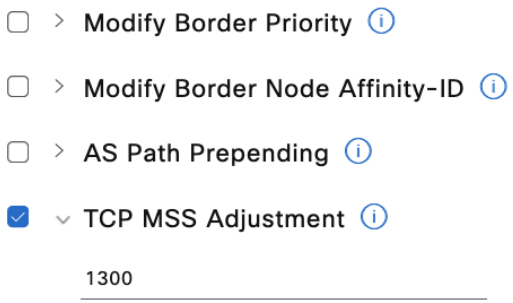



```
interface Vlan1021
 description Configured from Cisco DNA-Center
 mac-address 0000.0c9f.f11f
 vrf forwarding Campus_VN
 ip address 10.2.0.1 255.255.255.0
 ip helper-address 10.66.117.23
 no ip redirects
 ip route-cache same-interface
 ip tcp adjust-mss 1300
```

Fragmentation in VXLAN

ip tcp adjust-mss

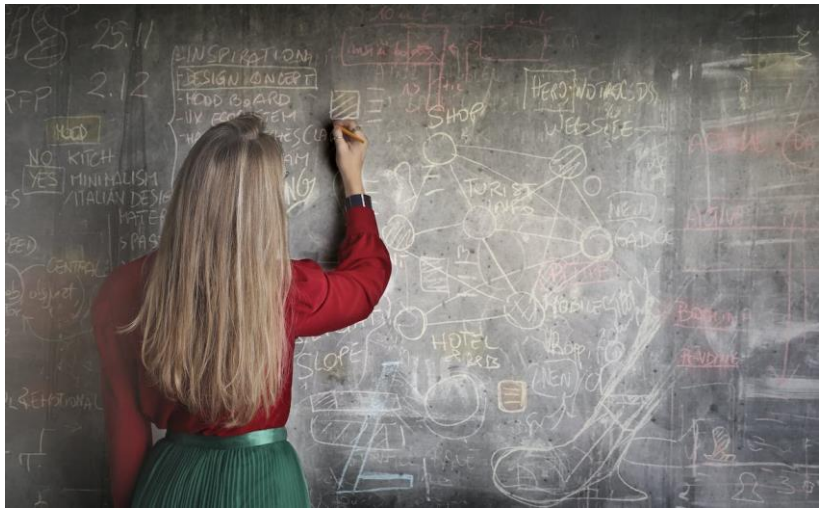
- Per BN – pushed to all L3 VRF handoff interfaces.
- Adjust if WAN MTU cannot be adjusted and you use SDA Transit.
- Only helps with TCP traffic.



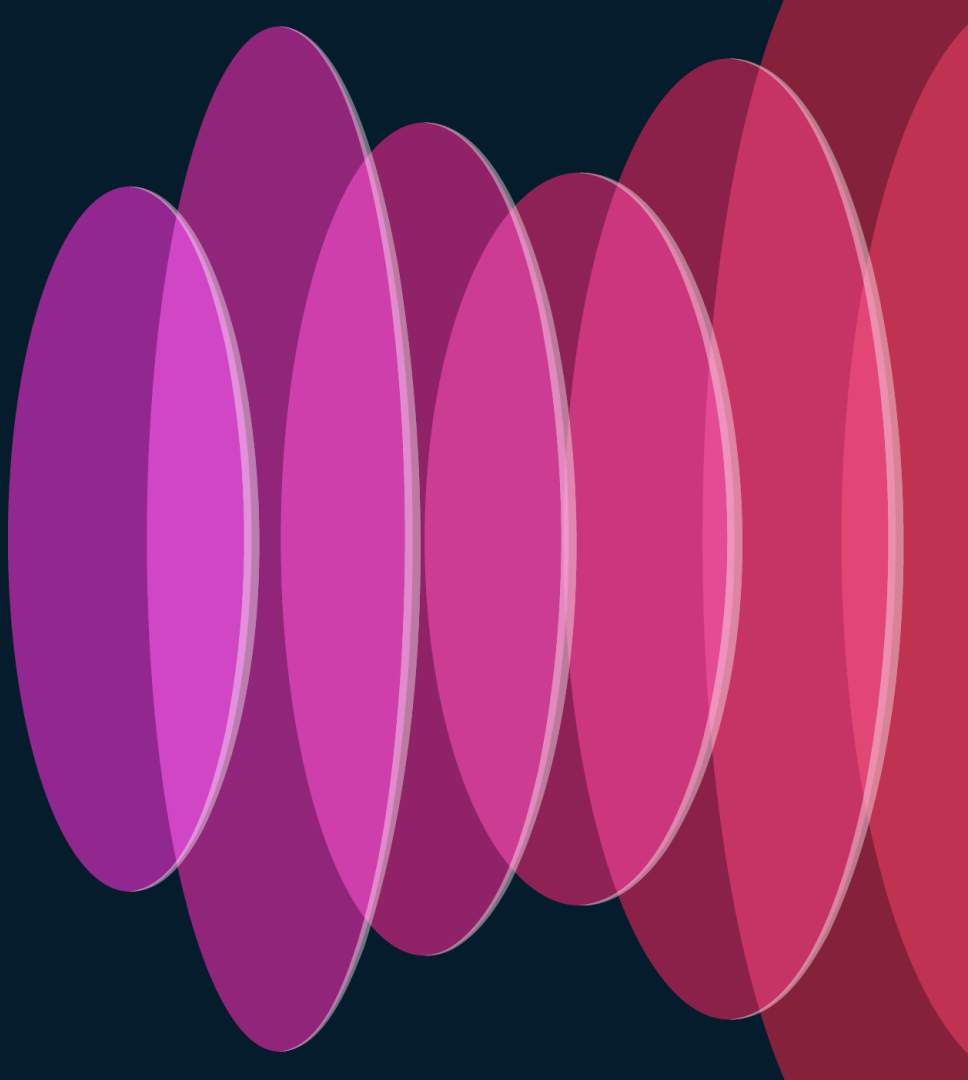

```
interface Vlan3000
 description vrf interface to External router
 vrf forwarding Campus_VN
 ip address 10.0.254.1 255.255.255.252
 no ip redirects
 ip route-cache same-interface
 ip tcp adjust-mss 1300
 ipv6 address 2001:DB8:0:254::1/126
 ipv6 enable
end
```

Lessons Learned From Previous Migrations

- Most struggles during SDA deployments are found with underlay routing design (IGP, BGP) and misbehaving endpoints – iron those out before the deployment.
- Using IS-IS without experience – do you really want to learn new IGP while troubleshooting fabric operations?
- Migrating subnets into the fabric “as is” – quickly reach subnet limit in Catalyst Center.
- Trying to approach the project as “transformational”: HW refresh + Fabric + Fabric Wireless + Transition to 802.1x + Micro-segmentation + changes in shared services (WLC, DHCP, authentication) in a single project. Better split into multiple phases.

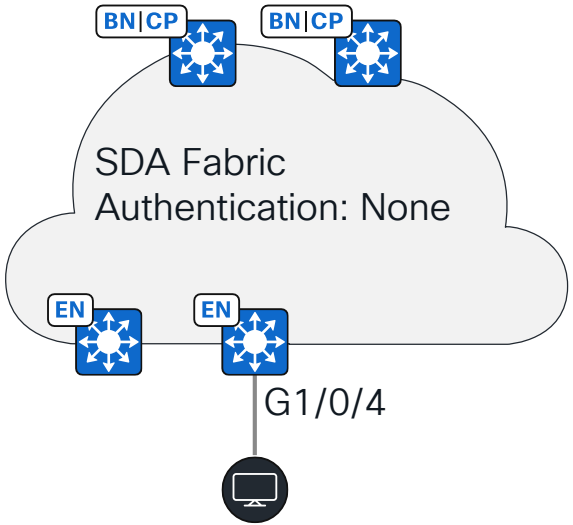


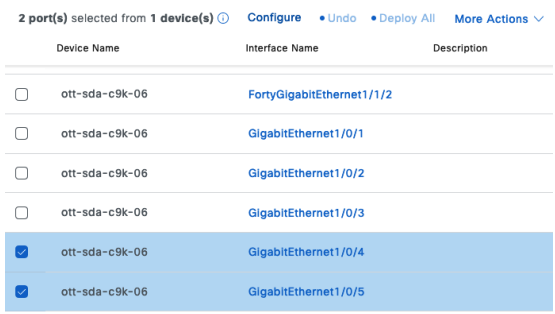
What's Next for ACME?

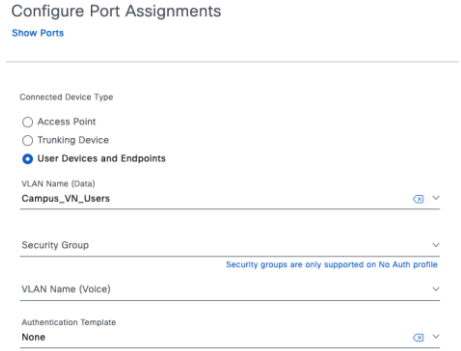


What's Next for ACME?

Switchport Authentication Policy: None



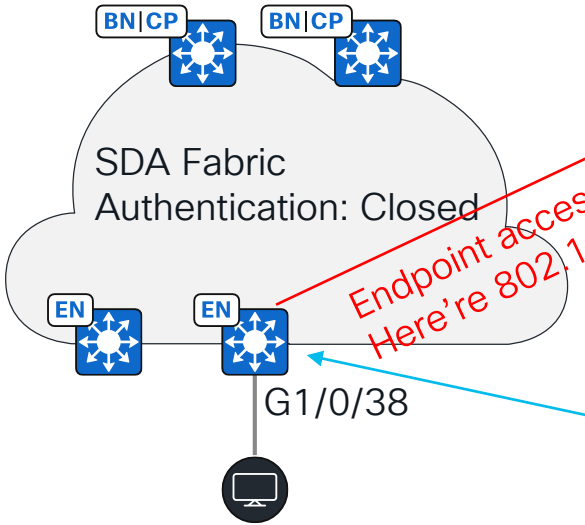




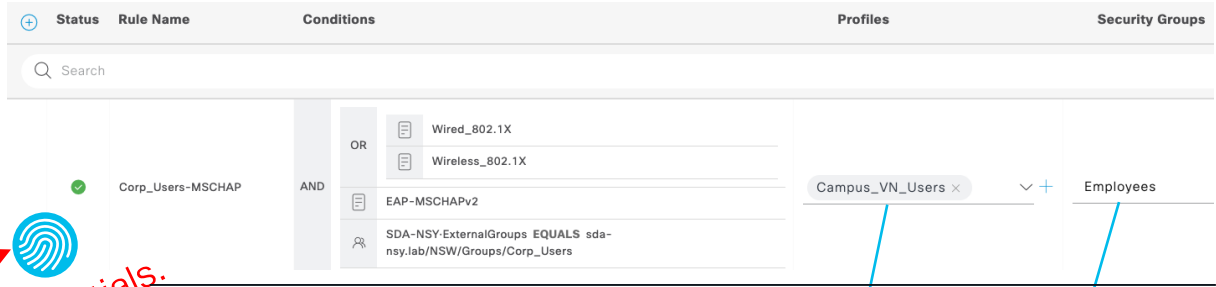

```
interface GigabitEthernet1/0/4
 switchport access vlan 1021
 switchport mode access
 device-tracking attach-policy IPDT_POLICY
 ip flow monitor dnacmonitor input
 ip flow monitor dnacmonitor output
 load-interval 30
 ipv6 flow monitor dnacmonitor_v6 input
 ipv6 flow monitor dnacmonitor_v6 output
 access-session inherit disable interface-template-sticky
 access-session inherit disable autoconf
 no macro auto processing
 spanning-tree portfast
 spanning-tree bpduguard enable
 ip nbar protocol-discovery
```

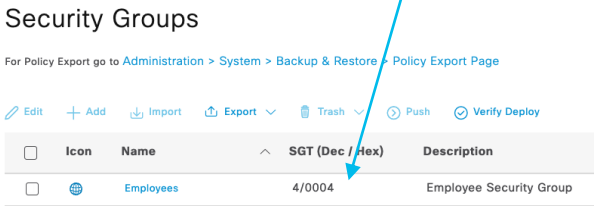
What's Next for ACME?

Upgrade Switchport Policy to 802.1X + MAB



Endpoint access?
Here're 802.1x credentials.

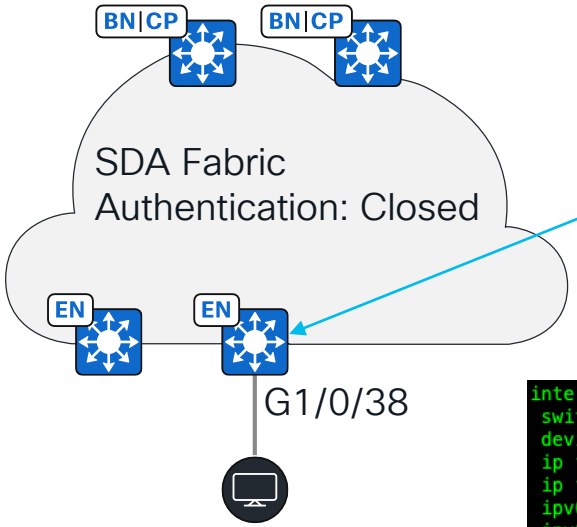




Authorize endpoint into VLAN
"Campus_VN_Users" and assign SGT 4

What's Next for ACME?

Upgrade Switchport Policy to 802.1X+MAB



Authorize endpoint into VLAN "Campus_VN_Users" and assign SGT 4

```
SE-22-07_U22_101_EN1#show vlan brief

VLAN Name                             Status
---- -------------------------------- ---------
1    default                          active

1002 fddi-default                     act/unsup
1003 token-ring-default               act/unsup
1004 fddinet-default                  act/unsup
1005 trnet-default                    act/unsup
1021 Campus_VN_Users                  active
1022 Guest_VN_Users                   active
1023 INFRA_VN_AP                      active
1024 IoT_VN_Devices                   active
1028 MY_L2VN                          active
2000 10_2_0_0-TEST                    active
2046 VOICE_VLAN                       active
```

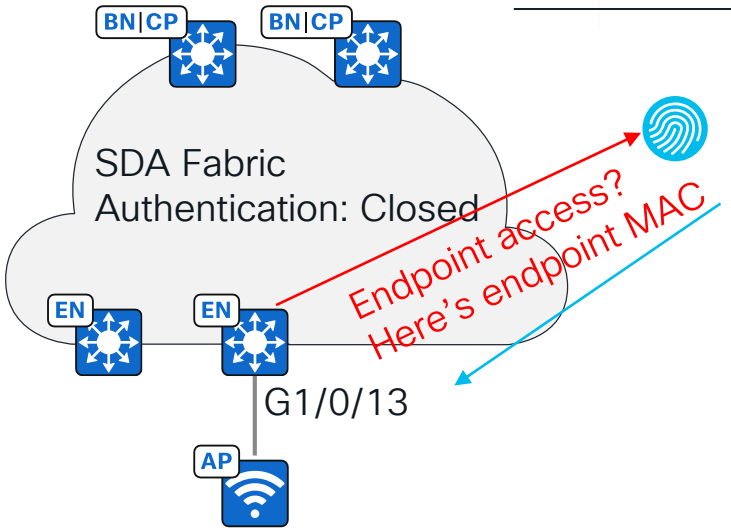
```
interface GigabitEthernet1/0/38
 switchport mode access
 device-tracking attach-policy IPDT_POLICY
 ip flow monitor dnacmonitor input
 ip flow monitor dnacmonitor output
 ipv6 flow monitor dnacmonitor_v6 input
 ipv6 flow monitor dnacmonitor_v6 output
 dot1x timeout tx-period 7
 dot1x max-reauth-req 3
 source template DefaultWiredDot1xClosedAuth
 spanning-tree portfast
 spanning-tree bpduguard enable
 ip nbar protocol-discovery
```

```
SE-22-07_U22_101_EN1#show authentication sessions interface g1/0/38 details
            Interface:  GigabitEthernet1/0/38
              IIF-ID:  0x1B3BFD10
         MAC Address:  a036.9fd2.f1aa
        IPv6 Address:  fe80::1ec6:418b:91b2:faa1
                       2001:db8:2:0:76b5:da0f:a8f:533f
        IPv4 Address:  10.2.0.11
           User-Name:  101wired2
         Device-type:  Microsoft-Workstation
         Device-name:  SDA-VM-1
                 VRF:  Campus_VN
              Status:  Authorized
              Domain:  DATA
     Oper host mode:  multi-auth
   Oper control dir:  both
    Session timeout:  N/A
  Acct update timeout:  172800s (local), Remaining: 172531s
   Common Session ID:  0602000A00000064C62024D6
     Acct Session ID:  0x0000004f
              Handle:  0xd000005a
      Current Policy:  PMAP_DefaultWiredDot1xClosedAuth_1X_MAB

Local Policies:

Server Policies:
         Vlan Group:  Vlan: 1021
           SGT Value:  4
```

cisco Live!

Cisco AI Endpoint Analytics

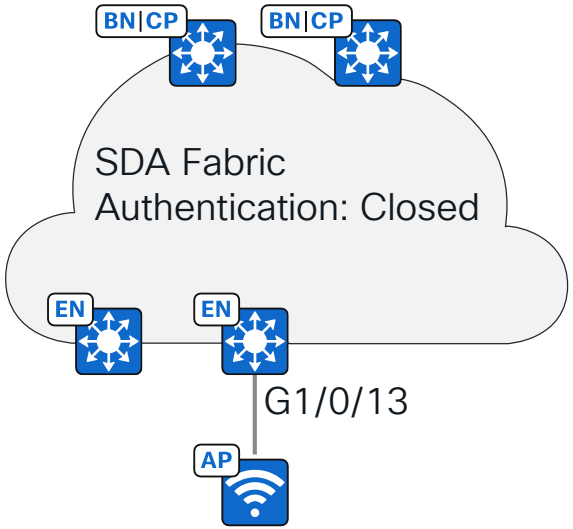
What if endpoint does not support 802.1X?



Authorize endpoint into VLAN
"Guest_VN_Users" and assign SGT 2

Cisco AI Endpoint Analytics

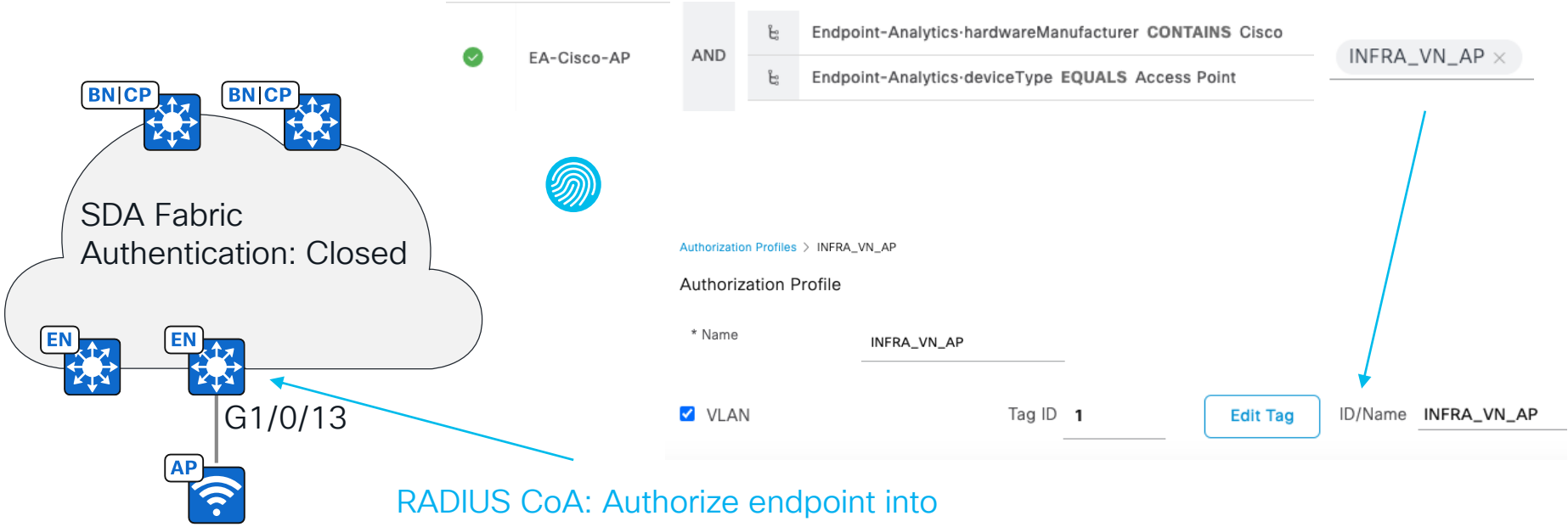
What if endpoint does not support 802.1X?



EA-DeviceType	Access Point, Network Device
EA-HardwareManufacturer	Cisco Systems, Inc.
EA-HardwareModel	Cisco-Access-Point 9120, Cisco-Access-Point, Cisco-Device
EA-OperatingSystem	Cisco IOS 17.11.0.155, Cisco IOS
EA-SpuriousAccessResult	false
EA-TrustScore	6
ElapsedDays	215
EndPointMACAddress	F4-BD-9E-9A-24-20
EndPointPolicy	Cisco-AP-Catalyst-9120AX

Cisco AI Endpoint Analytics

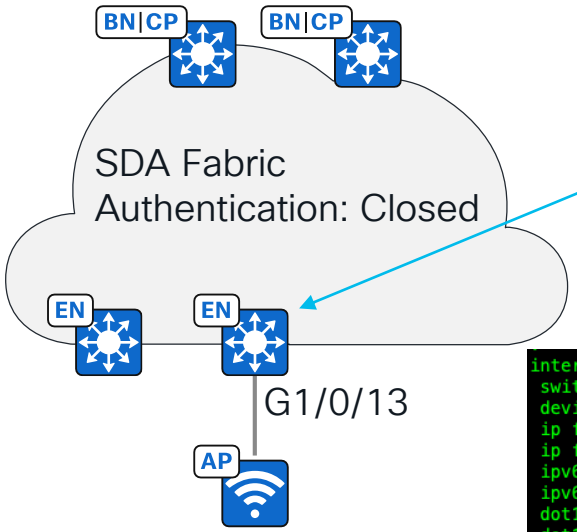
What if endpoint does not support 802.1X?



RADIUS CoA: Authorize endpoint into VLAN “INFRA_VN_AP”

Cisco AI Endpoint Analytics

What if endpoint does not support 802.1X?



RADIUS CoA: Authorize endpoint into VLAN “INFRA_VN_AP”

```
SE-22-07_U22_101_EN1#show vlan brief

VLAN Name                             Status
---- -------------------------------- ---------
1    default                          active

1002 fddi-default                     act/unsup
1003 token-ring-default               act/unsup
1004 fddinet-default                  act/unsup
1005 trnet-default                    act/unsup
1021 Campus_VN_Users                  active
1022 Guest_VN_Users                   active
1023 INFRA_VN_AP                      active
1024 IoT_VN_Devices                   active
1028 MY_L2VN                          active
2000 10_2_0_0-TEST                    active
2046 VOICE_VLAN                       active
```

```
interface GigabitEthernet1/0/13
 switchport mode access
 device-tracking attach-policy IPDT_POLICY
 ip flow monitor dnacmonitor input
 ip flow monitor dnacmonitor output
 ipv6 flow monitor dnacmonitor_v6 input
 ipv6 flow monitor dnacmonitor_v6 output
 dot1x timeout tx-period 7
 dot1x max-reauth-req 3
 et-analytics enable
 source template DefaultWiredDot1xClosedAuth
 spanning-tree portfast
 spanning-tree bpduguard enable
 ip nbar protocol-discovery
```

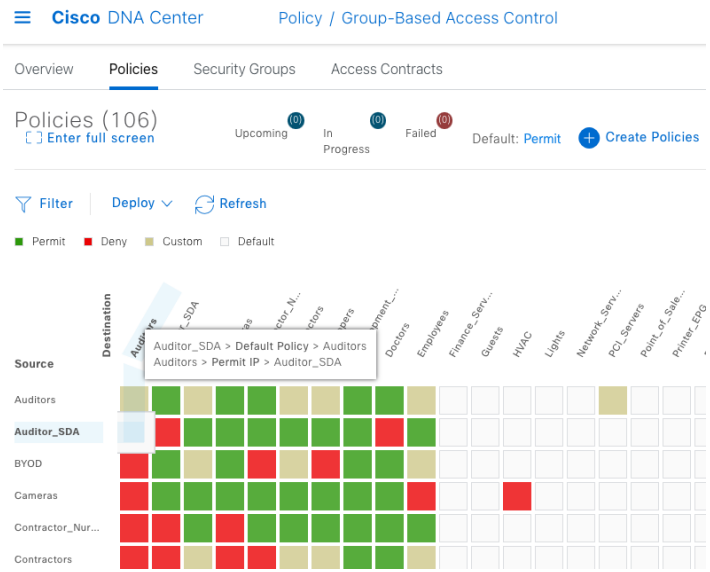
```
SE-22-07_U22_101_EN1#show authentication sessions interface g1/0/13 details
            Interface:  GigabitEthernet1/0/13
              IIF-ID:  0x19FD31AD
         MAC Address:  f4bd.9e9a.2420
        IPv6 Address:  fe80::47df:9e4c:b69c:ddf2
        IPv4 Address:  10.1.0.12
           User-Name:  F4-BD-9E-9A-24-20
         Device-type:  Cisco-CAT-LAP
        Device-name:  cisco C9120AXI-Z
              Status:  Authorized
              Domain:  DATA
     Oper host mode:  multi-auth
   Oper control dir:  both
    Session timeout:  N/A
  Acct update timeout:  172800s (local), Remaining: 85979s
   Common Session ID:  0602000A00000047952F8333
     Acct Session ID:  0x0000003e
              Handle:  0xe700003d
     Current Policy:  PMAP_DefaultWiredDot1xClosedAuth_1X_MAB

Local Policies:

Server Policies:
        Vlan Group:  Vlan: 1023
         SGT Value:  2
```

What's Next for ACME?

Implement micro-segmentation

Micro-segmentation "gotchas":

- Default deny will deny everything including broadcasts/ARP/DHCP traffic.
- Micro-segmentation policy is applied to unicast traffic only:
  - Broadcast (including DHCP) traffic is not filtered.
  - Multicast traffic is not filtered.
- Statically assigned SGTs (in switch CLI) are not shown in Policy Analytics visualisation (classified as "Unknown").
- Avoid "default deny" unless you have very specific reasons.



Summary

- Thank you! Without you SD-Access will remain in the CVD ☺
- Keep sharing feedback. We are listening.
- Go deep, check DGTL-BRKENS-3822 for brownfield migration details.
- Engage with Cisco Sales – we will always help you.
- Get virtual Catalyst Center and try SD-Access next week!

Complete Your Session Evaluations

Complete a minimum of 4 session surveys and the Overall Event Survey to be entered in a drawing to **win 1 of 5 full conference passes** to Cisco Live 2025.

Earn 100 points per survey completed and compete on the Cisco Live Challenge leaderboard.

Level up and earn **exclusive prizes!**

Complete your surveys in the **Cisco Live mobile app.**

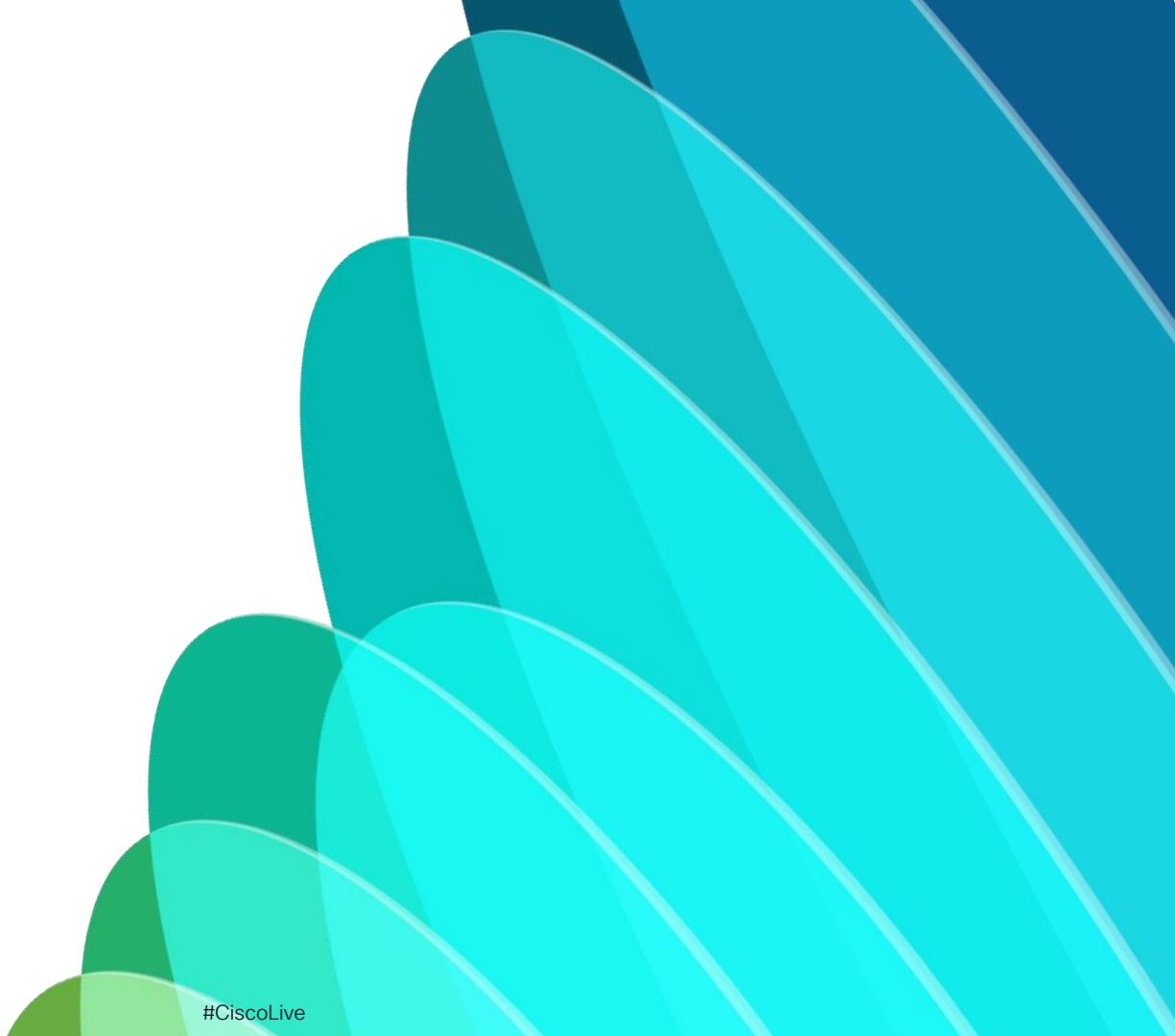
Continue your education

- Visit the Cisco Showcase for related demos
- Book your one-on-one Meet the Engineer meeting
- Attend the interactive education with DevNet, Capture the Flag, and Walk-in Labs
- Visit the On-Demand Library for more sessions at www.CiscoLive.com/on-demand

Contact me at: snasonov@cisco.com via email

The bridge to possible

Thank you



cisco Live!