CISCO

You make **possible**

Application Centric Infrastructure (ACI)

The Intent Driven Data Center

Max Ardica – Principal Engineer, Cisco
Mike Herbert – Principal Engineer, Cisco
Takuya Kishida – Technical Marketing Engineer, Cisco
Carlos Pereira – Distinguished Systems Engineer, Cisco
Dang Ngo– Technical Solutions Architect, Cisco

TECACI-2009

Barcelona | January 27-31, 2020

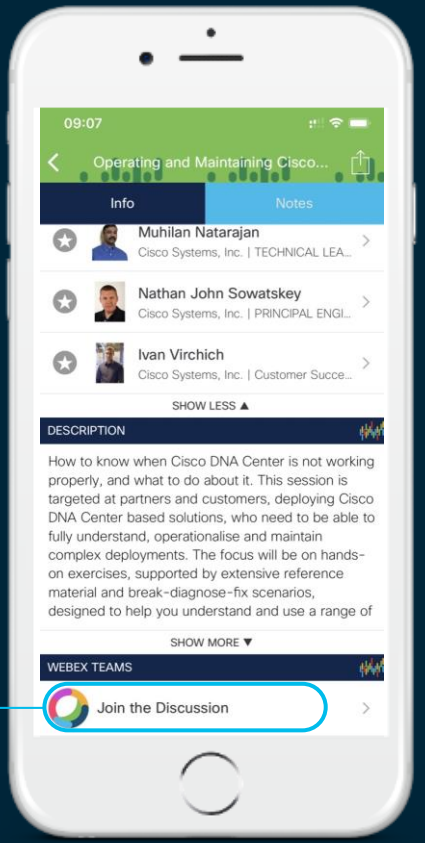
Cisco Webex Teams

Questions?

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

How

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

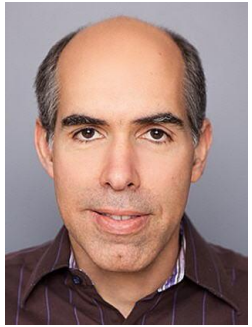


Who Are We??

Max Ardica
Principal Engineer – IBNG

Mike Herbert
Principal Engineer – IBNG

Takuya Kishida
TME - IBNG



Carlos Pereira
Distinguished Systems Engineer

Dang Ngo
Technical Solutions Architect

Agenda for the Day
(Yes your brain will hurt by the end)

- 8:30 – 9:00: Introduction to the Intent Based Data Center (ACI)
- 9:00 – 10:30: ACI Fabric Design Best Practices Part I
- 10:45 – 11:15: ACI Fabric Design Best Practices Part II
- 11:15 – 12:45: ACI Anywhere Design and Principles Part I
- 14:30 – 15:00: ACI Anywhere Design and Principles Part II
- 15:00 – 16:00: ACI and Cisco's Multi-Domain Architecture
- 16:00 – 16:30: How to Operate ACI Fabric(s) Part I
- 16:45 – 18:45: How to Operate ACI Fabric(s) Part II

A key question for the Day
When do I get caffeine?

- 8:30 - 10:30: Technical Seminar
- 10:30 – 10:45: Coffee Break
- 10:45– 12:45: Technical Seminar
- 12:45 – 14:30: Lunch
- 14:30 – 16:30: Technical Seminar
- 16:30 – 16:45: Coffee Break
- 16:45 – 18:45: Technical Seminar

Introduction to the Intent/Application Based Data Center

Why do we build Networks?

Experience versus Risk

Expect Great Experience

Business & Operations must balance Risk and Experience

End-Users (consumers)

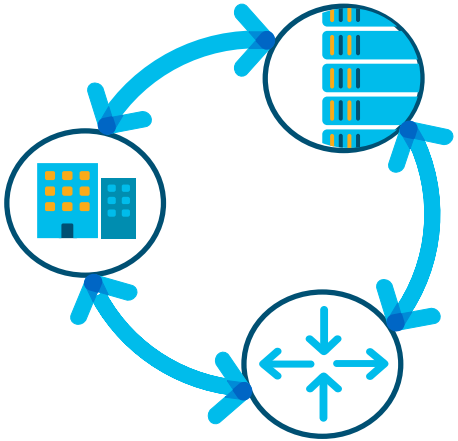
Applications (providers)

Experience versus Risk

Between users and applications, there is a network

Expect Great Experience

End-Users

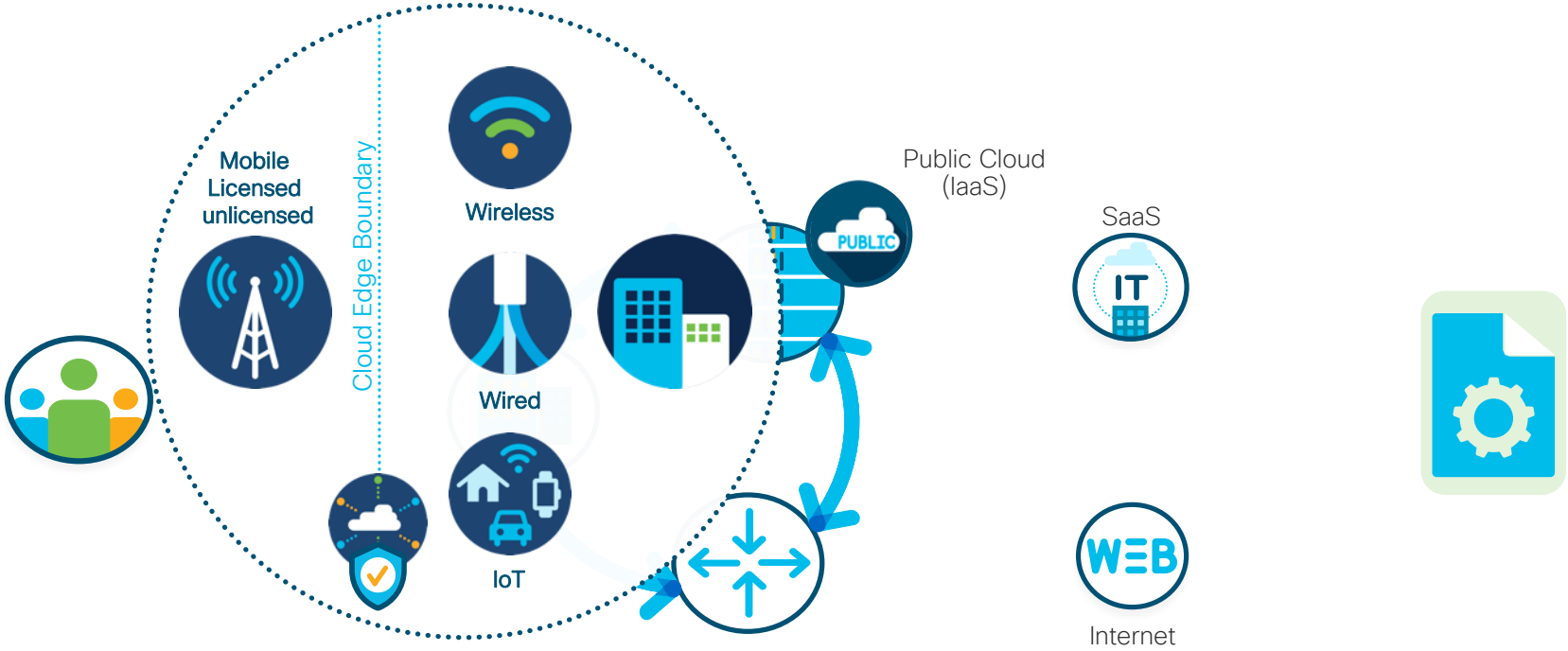


Business & Operations must balance Risk and Experience

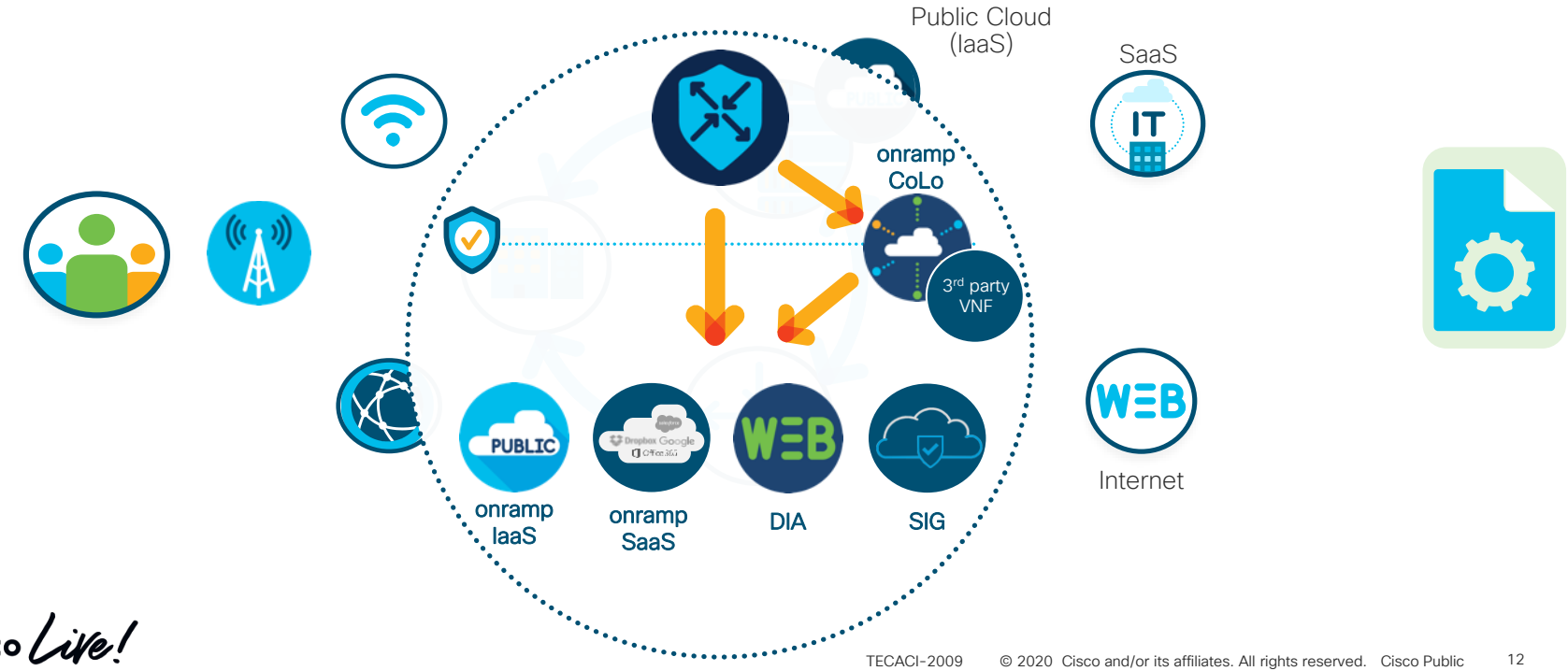
Applications

It's a multi-access world

With trust boundary's and a cloud edge function

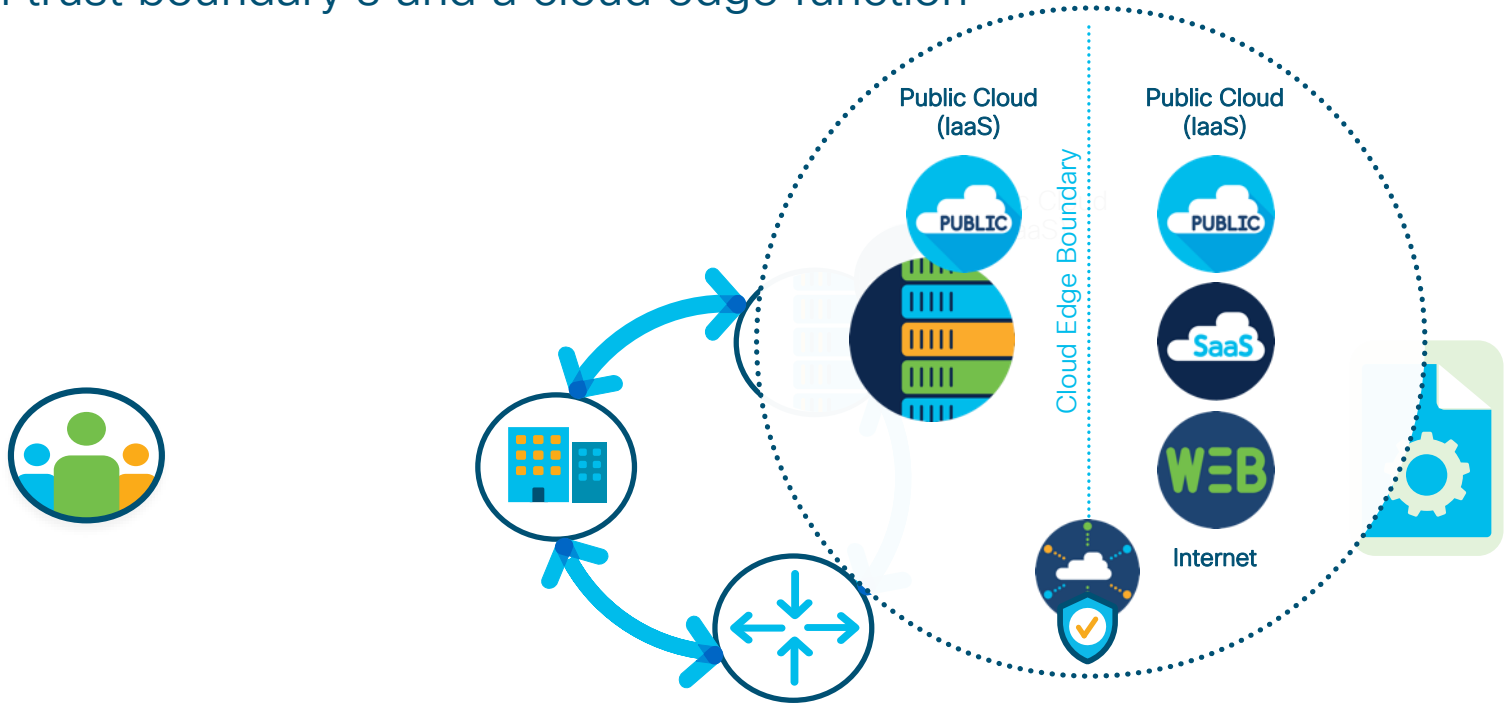


With Hybrid Cloud Connectivity World

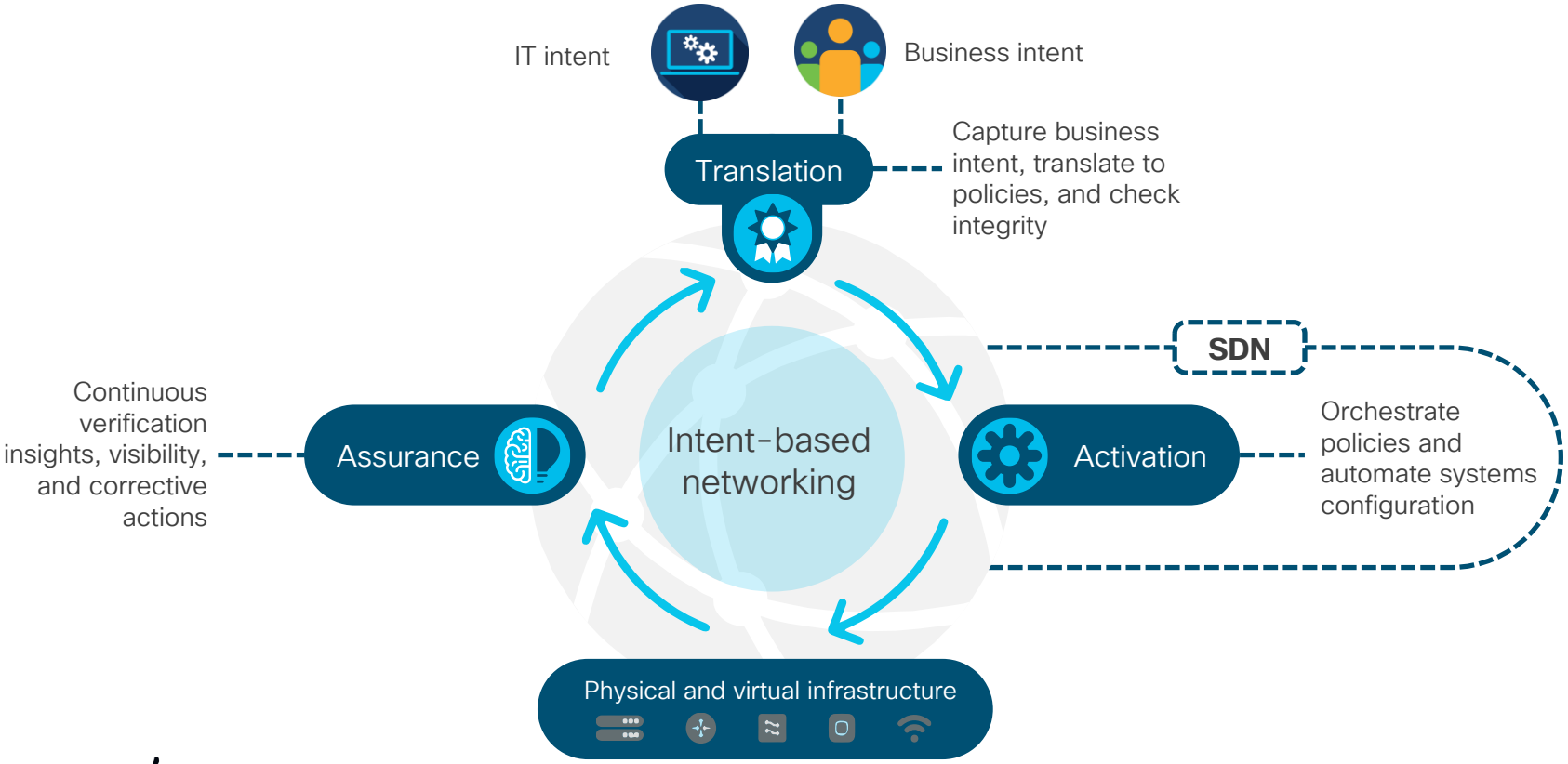


It’s a multi-cloud world

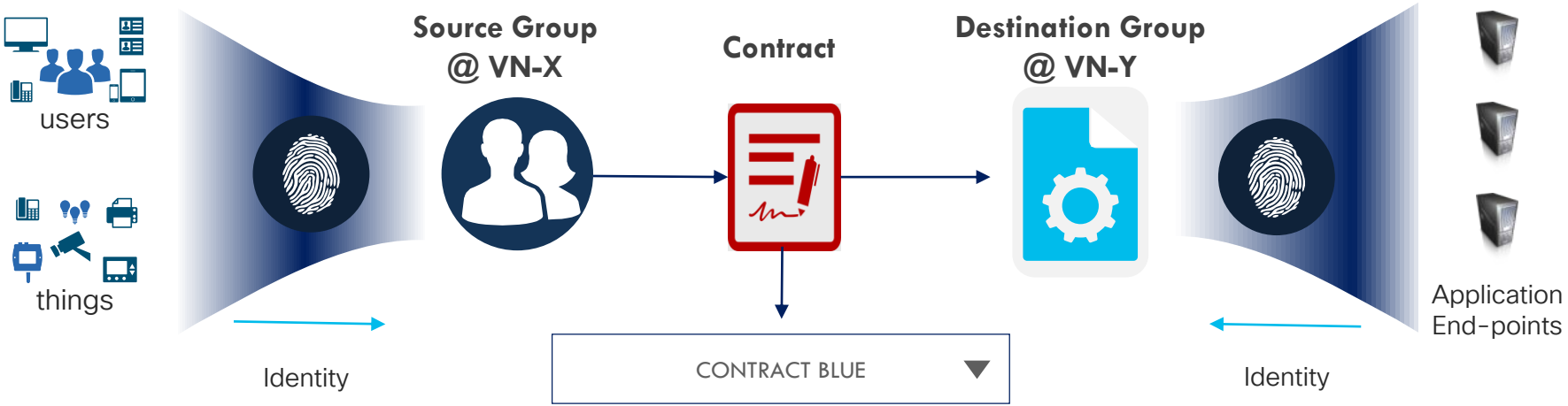
With trust boundary’s and a cloud edge function



Introducing Intent Based Networking



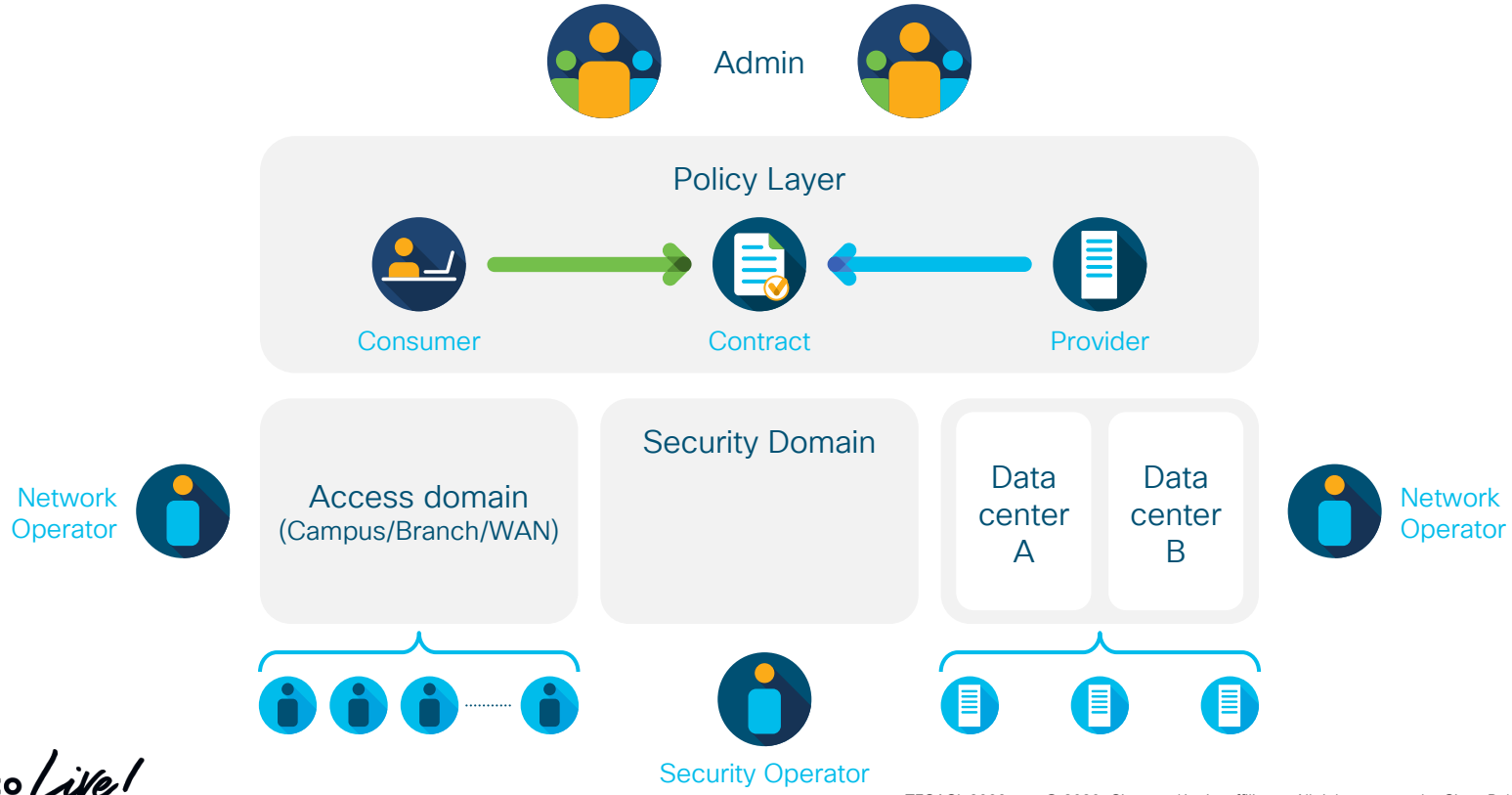
This is the Intent of the Infrastructure
User to Application Policies



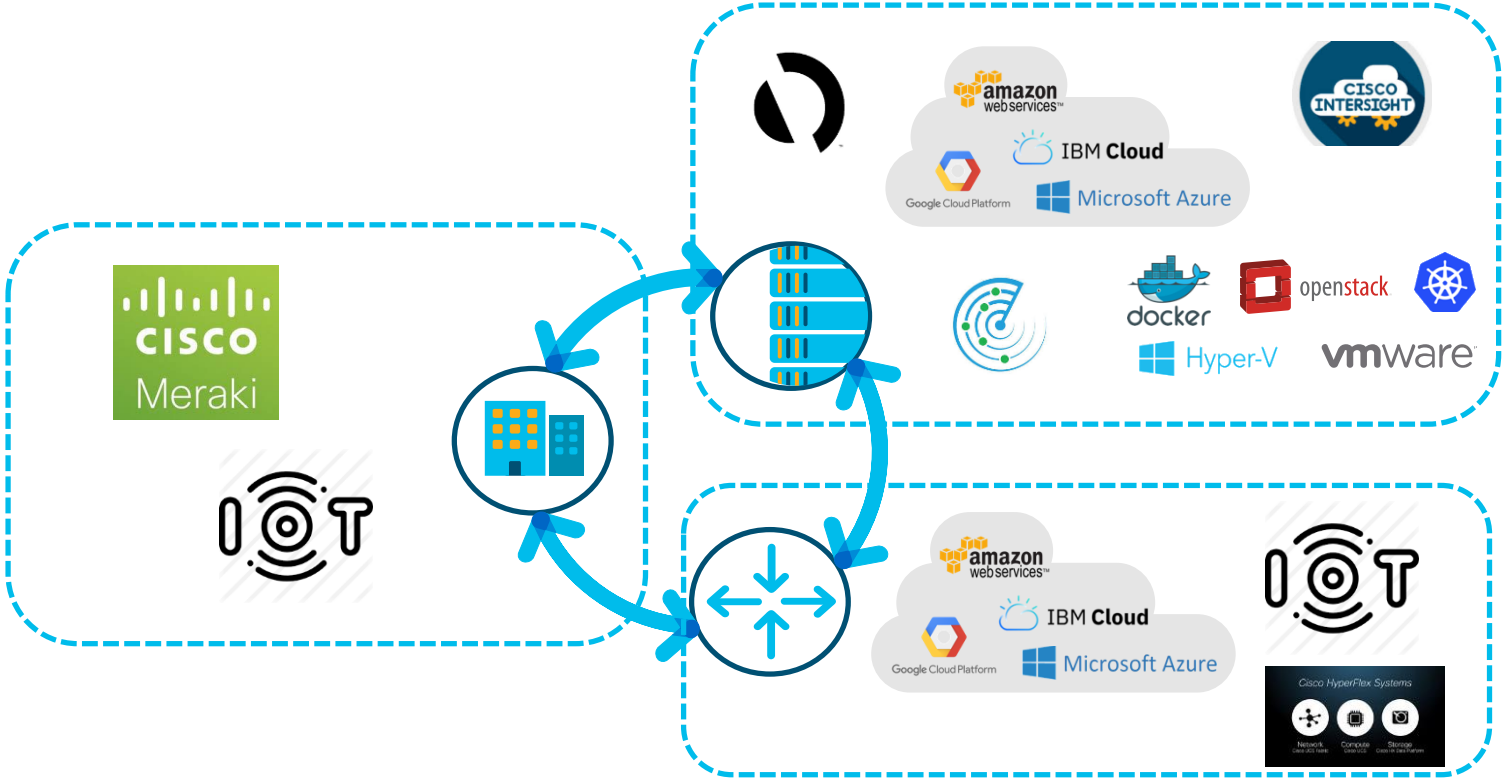
Classifier	Action
Port Number IP Address Application Type	Permit Deny Copy Gold Service

X-Domain Intent Based Networking

Unified policy language across domains



X-Domain Intent - Architectural View
Horizontal and Vertical 'Multi-Domain'



X-Domain IBN is part of a Zero Trust Policy Framework

Zero Trust Security Approach to Network Access

SDA, ACI, SDWAN

Zero Trust Security Approach to App Access

Duo Security is now part of Cisco.

Unified Access across Hybrid IT Enterprises

IoT Access Problem

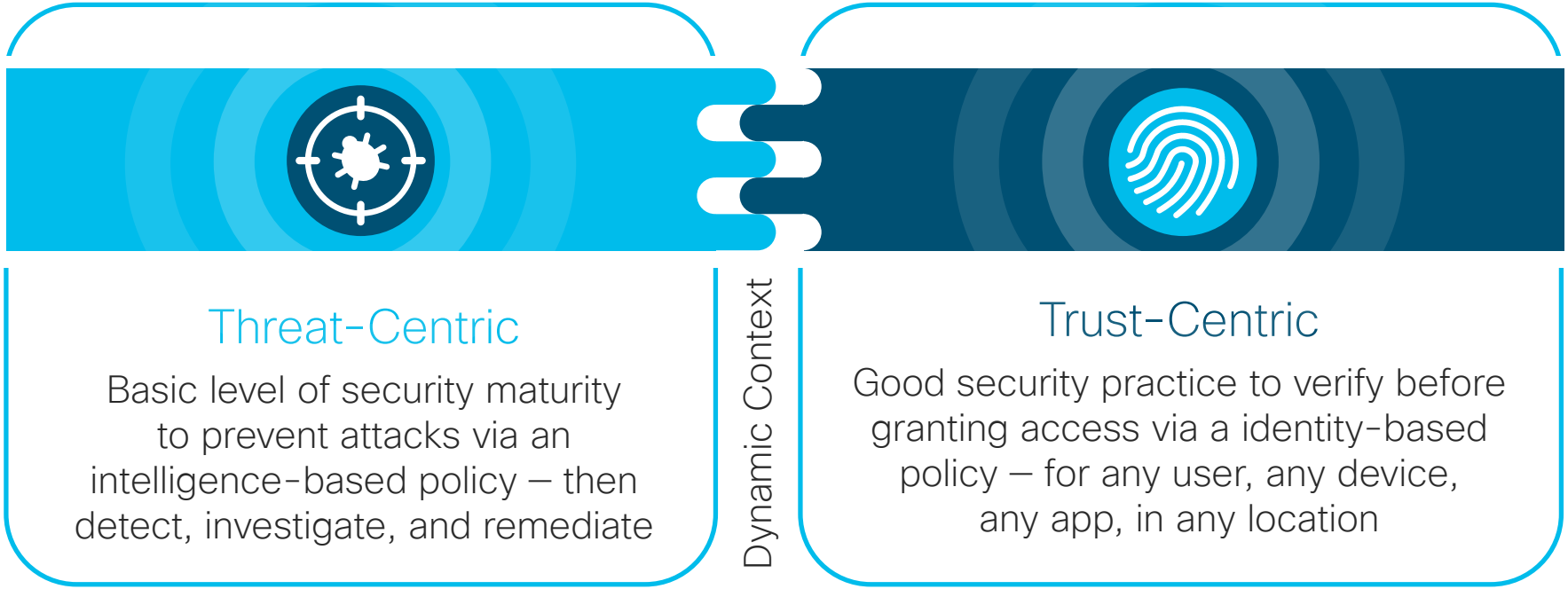
		App / Services	
		On-Prem ⟷	Cloud
Head-less Device	On-Prem	ISE	ISE

Mobile & BYOD Access Problem

		App / Services		
		On-Prem ⟷	Cloud	
User + Device	On-Prem	ISE	ISE	Duo MFA
	Off-Prem	ISE* or Duo†	Duo	

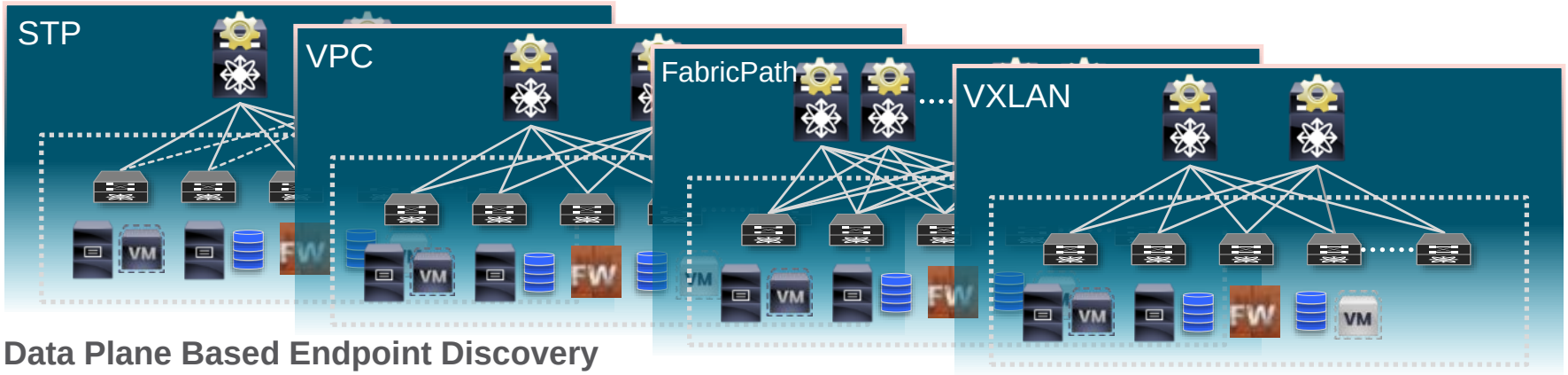
* Integrated with AnyConnect
† Duo Network Gateway (i.e. reverse proxy)

IBN X-Domain is a component of security
Complementary security elements



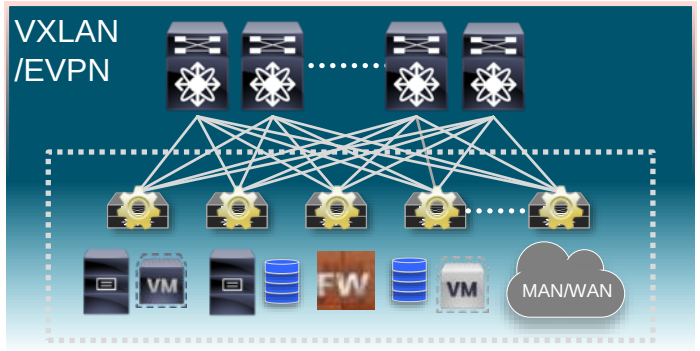
What is ACI?

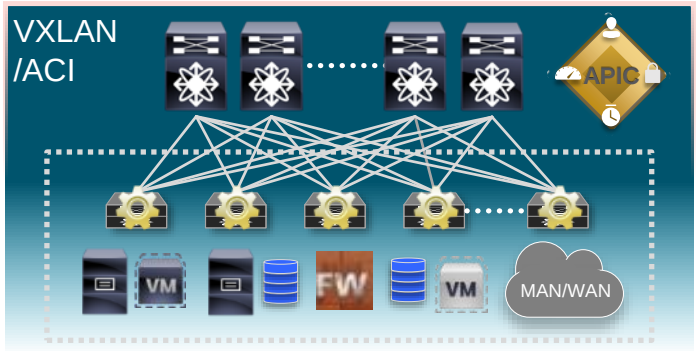
Next Gen Forwarding & Networking



Data Plane Based Endpoint Discovery

Control Plane Based Endpoint Location Tracking

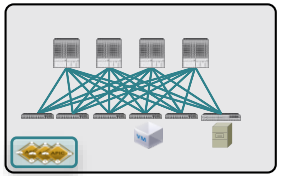




ACI Anywhere

Extending the Reach of the New Network

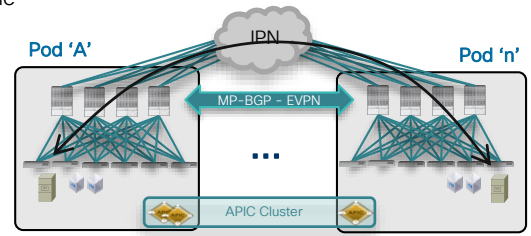
ACI Single Pod Fabric



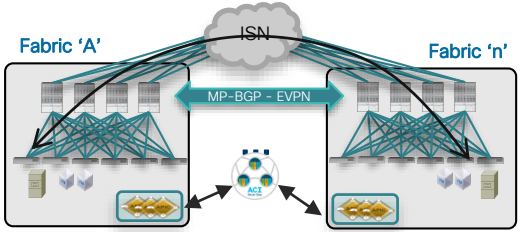
ACI 1.0 – Leaf/Spine Single Pod Fabric

ACI 2.0 – Multiple Networks (Pods) in a single Availability Zone (Fabric)

ACI Multi-Pod Fabric

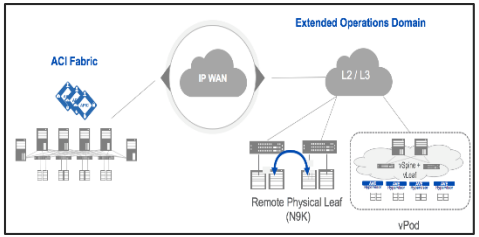


ACI Multi-Site



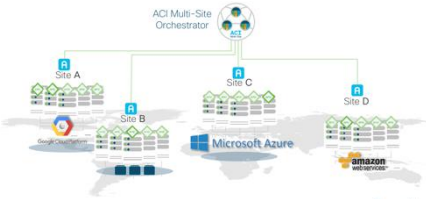
ACI 3.0 – Multiple Availability Zones (Fabrics) in a Single Region 'and' Multi-Region Policy Management

ACI Remote Leaf



ACI 3.1/4.0 – Remote Leaf and vPod extends an Availability Zone (Fabric) to remote locations

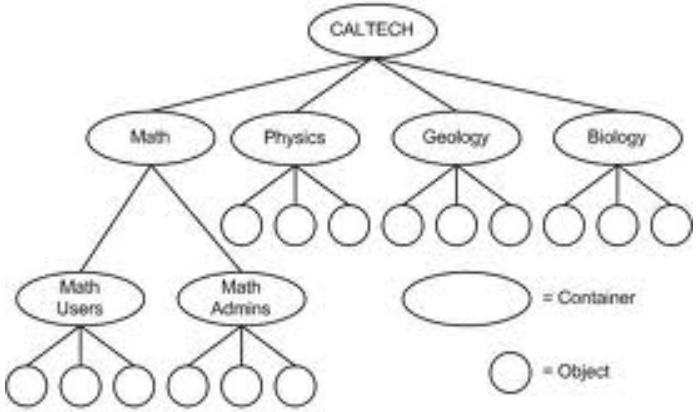
ACI Multi-Cloud

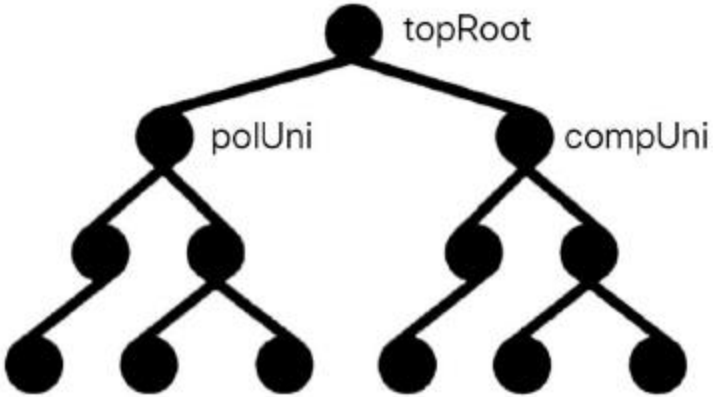


ACI 4.1 – ACI Extensions to Multi-Cloud

Directory Enabled Networking

Data Base Defined Networking



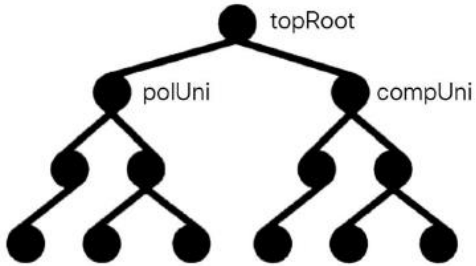


Common Operational Properties – AD, LDAP, …

System Management, Change Management, System Integrity, Correlation

APIC as an Operations Platform

More than just the Controller DataBase



Network Availability

Network Insights Advisor

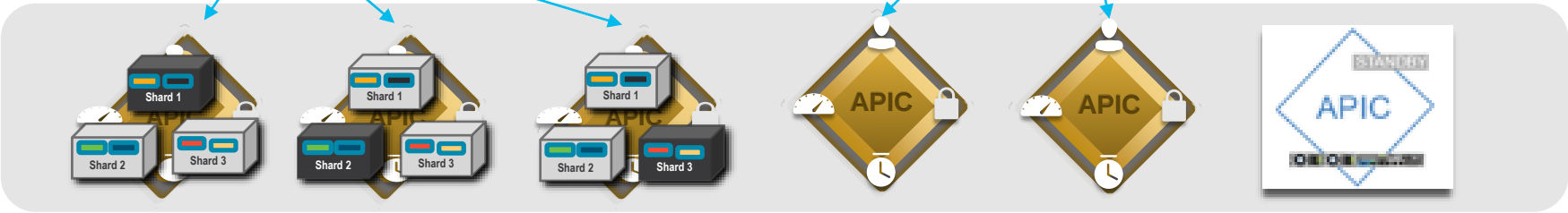
Proactive Software Recommendations/Notifications
Issue Vulnerability Detection & Remediation

Network Health

Network Insights Resources

Physical/Logical Network Capacity & Utilization
Data & Control Plane & Environmental Health

Framework for Cisco + 3rd Party Apps Running Across all APIC cluster nodes



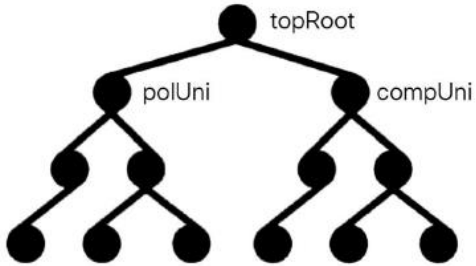
The DME Data Base runs on APIC servers (physical or virtual) and is replicated across APIC nodes (3 copies)

Dedicated “Service Engine” Nodes can be added to an APIC Cluster

Backup Servers for the APIC Cluster

APIC as an Operations Platform

More than just the Controller DataBase



Network Availability

Network Insights Advisor

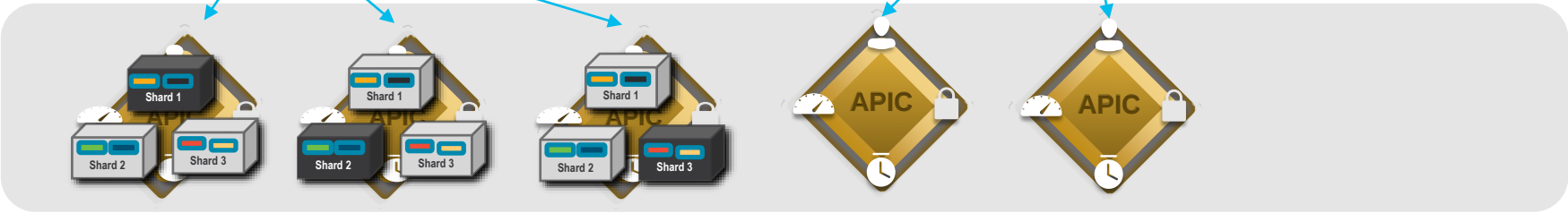
Proactive Software Recommendations/Notifications
Issue Vulnerability Detection & Remediation

Network Health

Network Insights Resources

Physical/Logical Network Capacity & Utilization
Data & Control Plane & Environmental Health

Framework for Cisco + 3rd Party Apps Running Across all APIC cluster nodes



The DME Data Base runs on APIC servers (physical or virtual) and is replicated across APIC nodes (3 copies)

Dedicated “Service Engine” Nodes can be added to an APIC Cluster

Backup Servers for the APIC Cluster

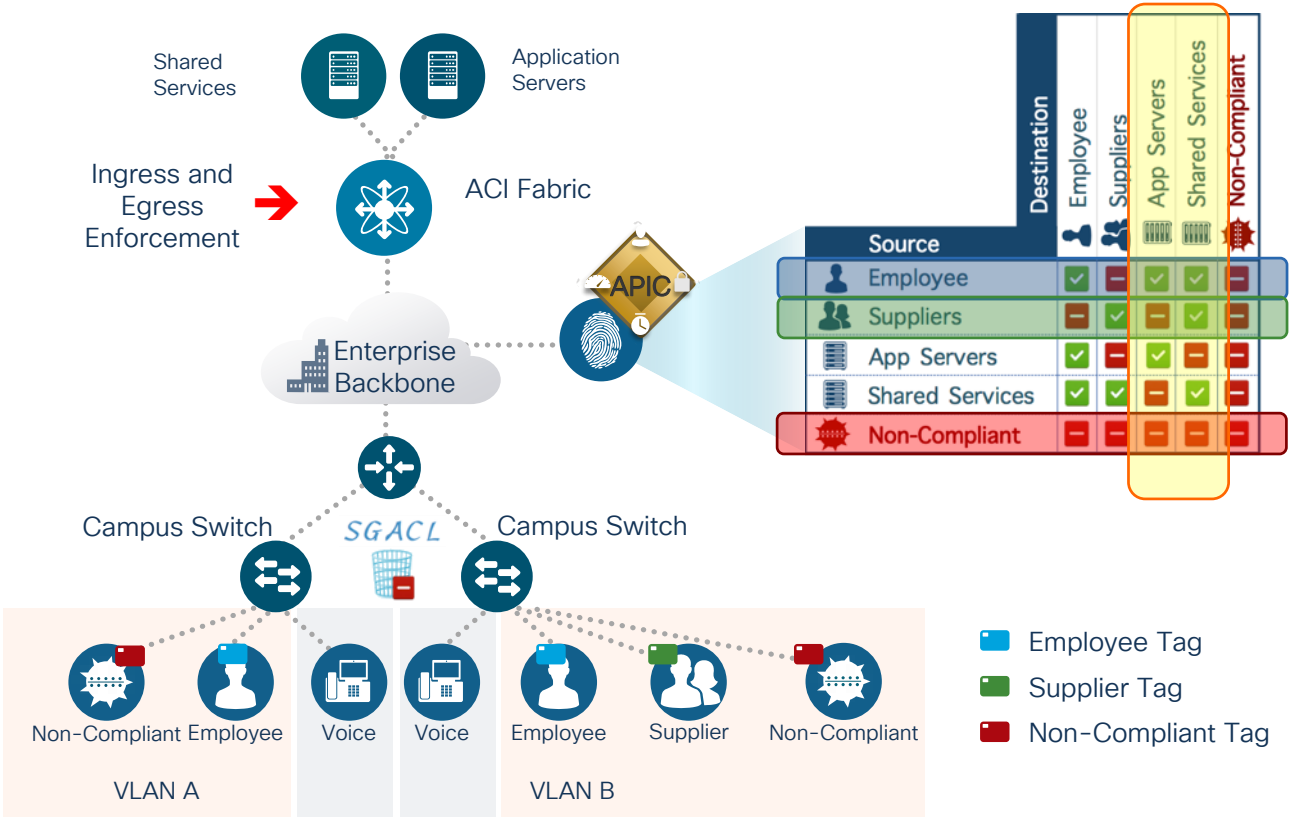
Identity Based Networking

VTEP IP	Group Policy	VNID	Tenant Packet

Devices and users are authenticated and authorized into end-point groups (aka EPG's or SGT's)

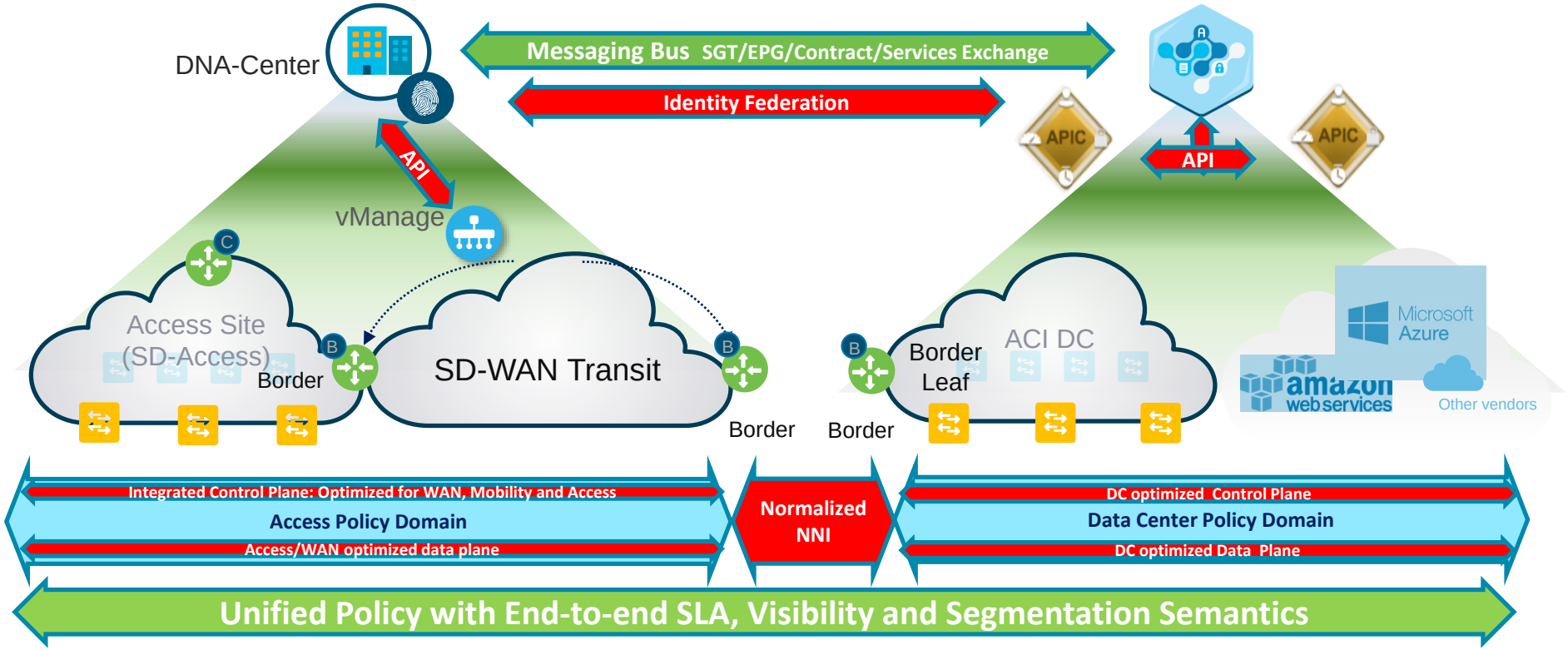
End Point Group Tags (EPG's, SGT's) are encoded in a VXLAN header

Policies between scalable groups are established following the provider/consumer model



X-Domain Intent Based Networking

Enhanced Data Plane and Messaging BUS Services Plane



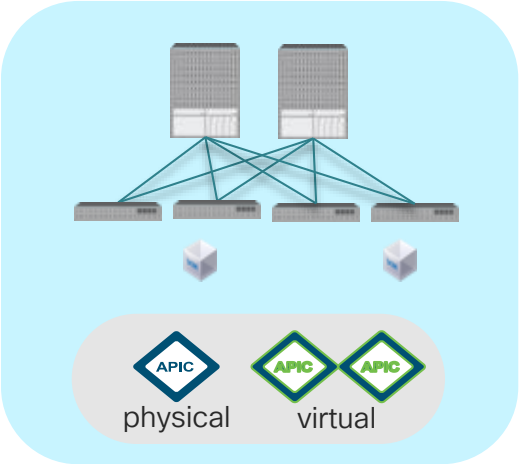
ACI Fabric Design

Agenda

- Single Fabric/Pod Topology Options
- ACI Design Considerations
  - ACI Fabric Bring Up
  - L2 Connectivity to Existing Networks
  - Loop prevention in ACI
  - Moving the default gateway to ACI
  - Connecting servers (Physical, VMM Integration, UCSM Integration)
  - Teaming Options
  - Allowing Traffic through ACI
  - Configuring L3 Connectivity to the Outside
- Application Centric Design
  - Network Centric & Application Centric
  - Contract Priority and Optimization

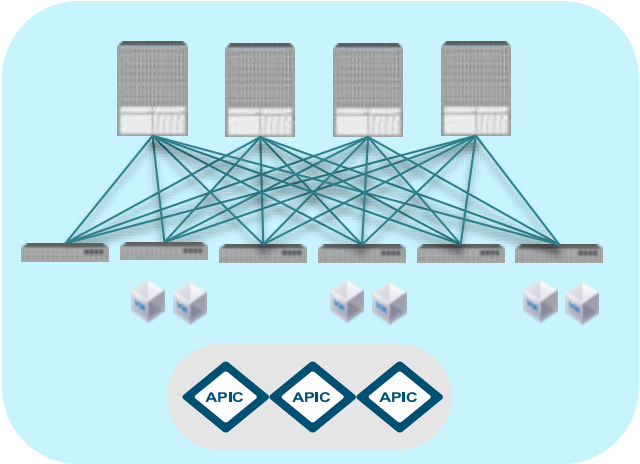
Single Fabric/Pod Topology Options

ACI mini



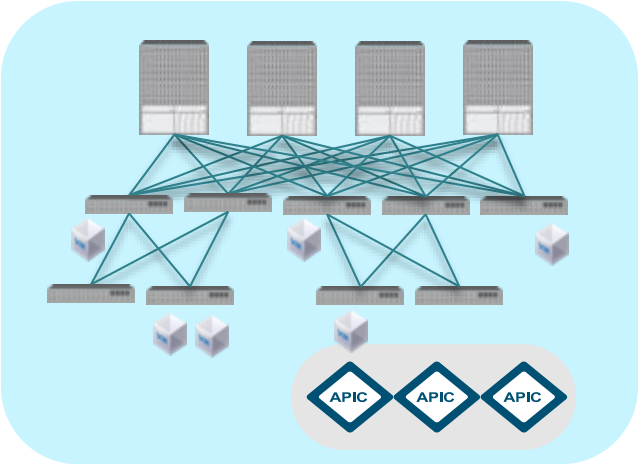
- From release 4.0
- For small deployment
- 1 APIC + 2 virtual APICs
- Up to 2 spine + 4 leaf switches
- Replacing virtual APICs with physical APICs
  - ➢ Expand to a full ACI fabric

Standard (2-Tier)



- From the first release (1.0)
- Most popular and standard ACI topology

Multi-Tier (3-Tier)



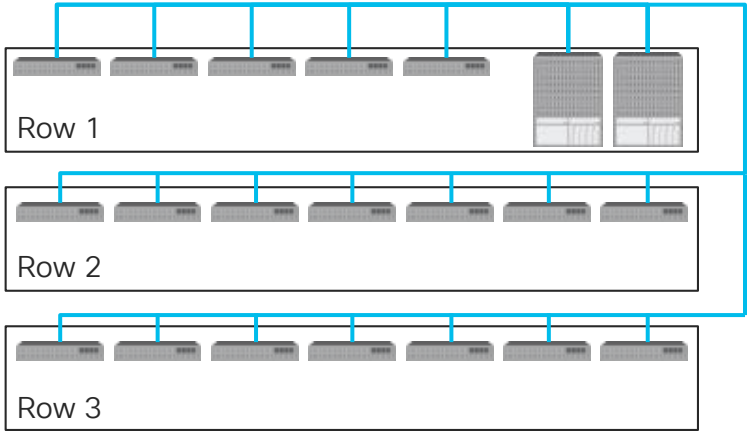
- From release 4.1
- For DCs with cabling restrictions

Not sure what to pick? Go with 2-Tier!

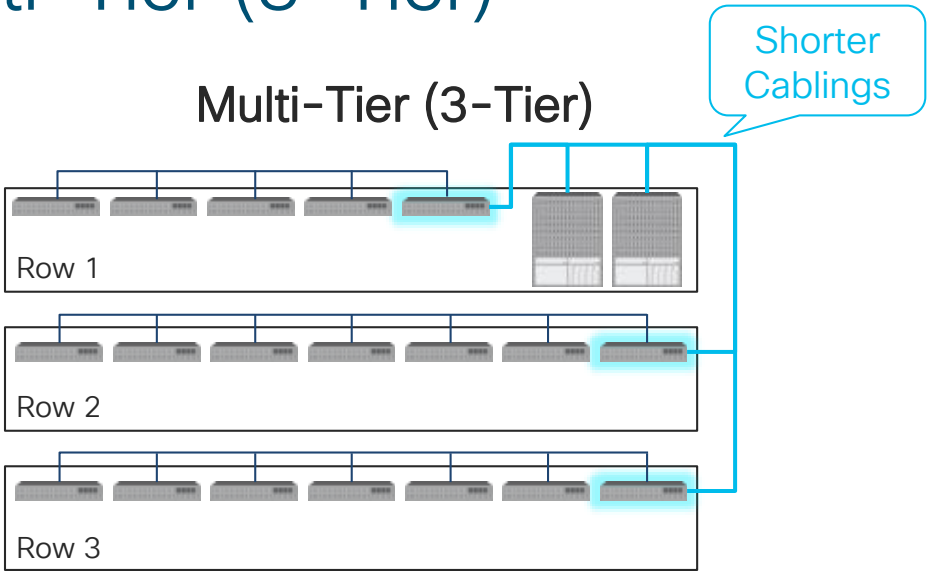
More about ACI Mini:
https://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/sw/kb/Cisco-Mini-ACI-Fabric-and-Virtual-APICs.html

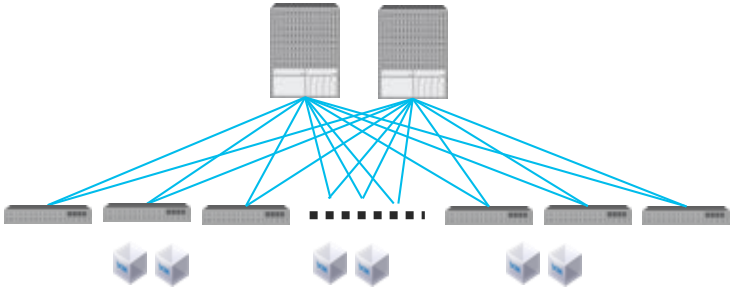
Standard (2-Tier) vs. Multi-Tier (3-Tier)

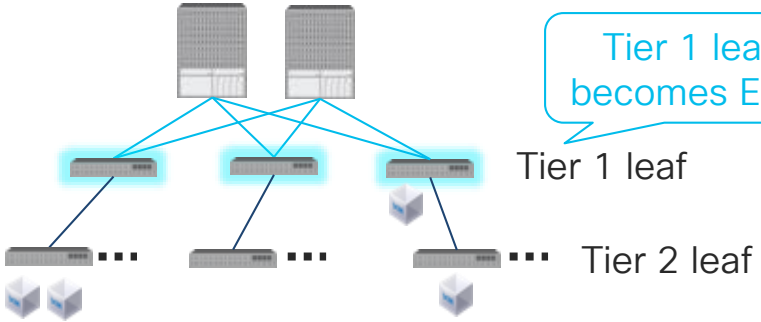
Standard (2-Tier)



Multi-Tier (3-Tier)







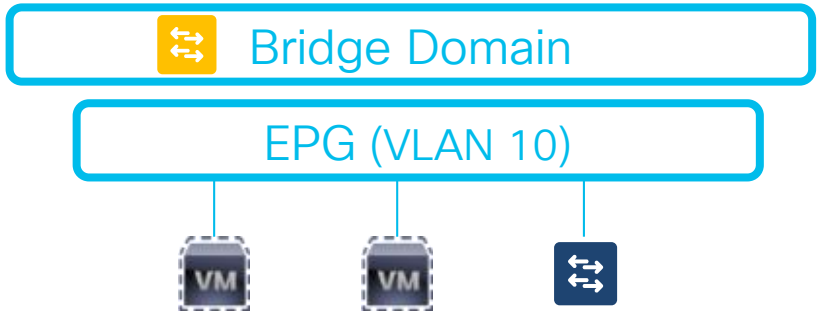
Agenda

- Single Fabric/Pod Topology Options
- ACI Design Considerations
  - ACI Fabric Bring Up
  - L2 Connectivity to Existing Networks
  - Loop prevention in ACI
  - Moving the default gateway to ACI
  - Connecting servers (Physical, VMM Integration, UCSM Integration)
  - Teaming Options
  - Allowing Traffic through ACI
  - Configuring L3 Connectivity to the Outside
- Application Centric Design
  - Network Centric & Application Centric
  - Contract Priority and Optimization

Network Centric & Application Centric Basic

Network Centric

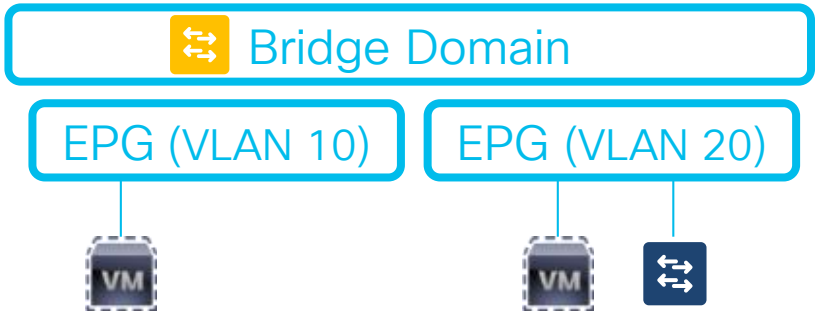
- 1 VLAN = 1 EPG = 1 BD



- Similar to traditional network
  - VLAN as a broadcast domain
  - Easy to connect to legacy networks
- Typically simple or no contracts (no ACLs)

Application Centric

- Multiple EPGs per BD



- Multiple security domains (EPGs) in one broadcast domain
- Flexible network and security design

Details are later

Agenda

- Single Fabric/Pod Topology Options
- ACI Design Considerations
  - ACI Fabric Bring Up
  - L2 Connectivity to Existing Networks
  - Loop prevention in ACI
  - Moving the default gateway to ACI
  - Connecting servers (Physical, VMM Integration, UCSM Integration)
  - Teaming Options
  - Allowing Traffic through ACI
  - Configuring L3 Connectivity to the Outside
- Application Centric Design
  - Network Centric & Application Centric
  - Contract Priority and Optimization

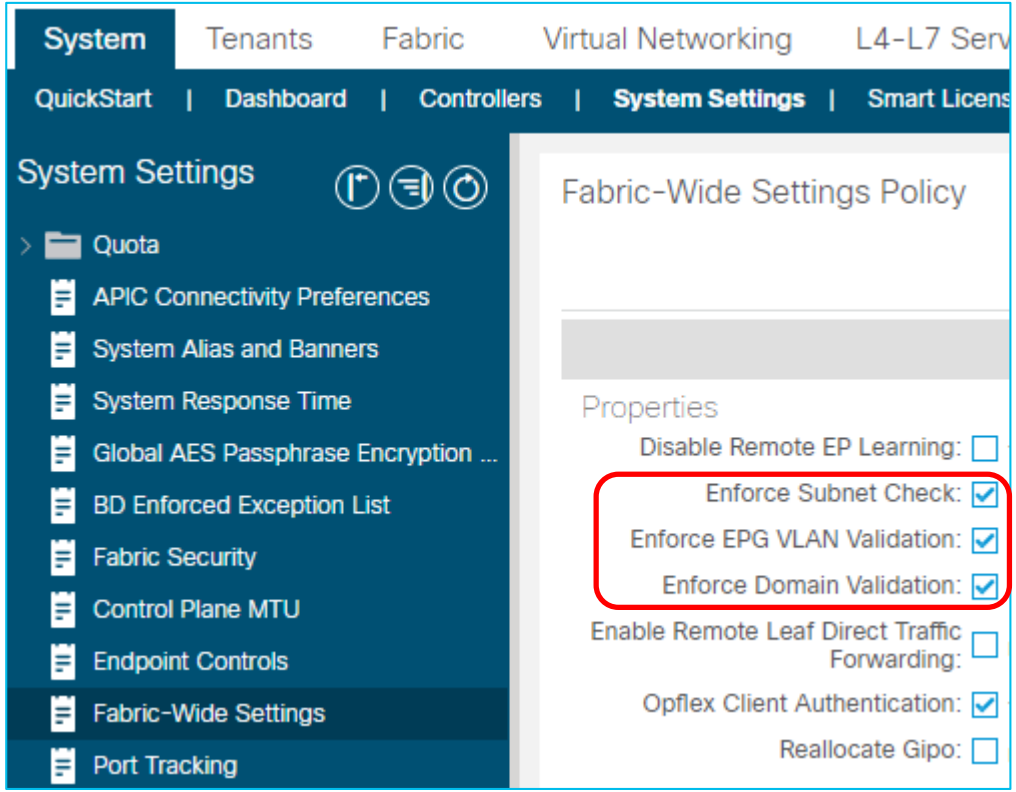
ACI Fabric Bring Up

- CCO Document:
- Setting Up an ACI Fabric: Initial Setup Configuration Example
- https://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/white_papers/Cisco-ACI-Initial-Deployment-Cookbook.html

- Breakouts:
- How to setup an ACI fabric from scratch - BRKACI-2004

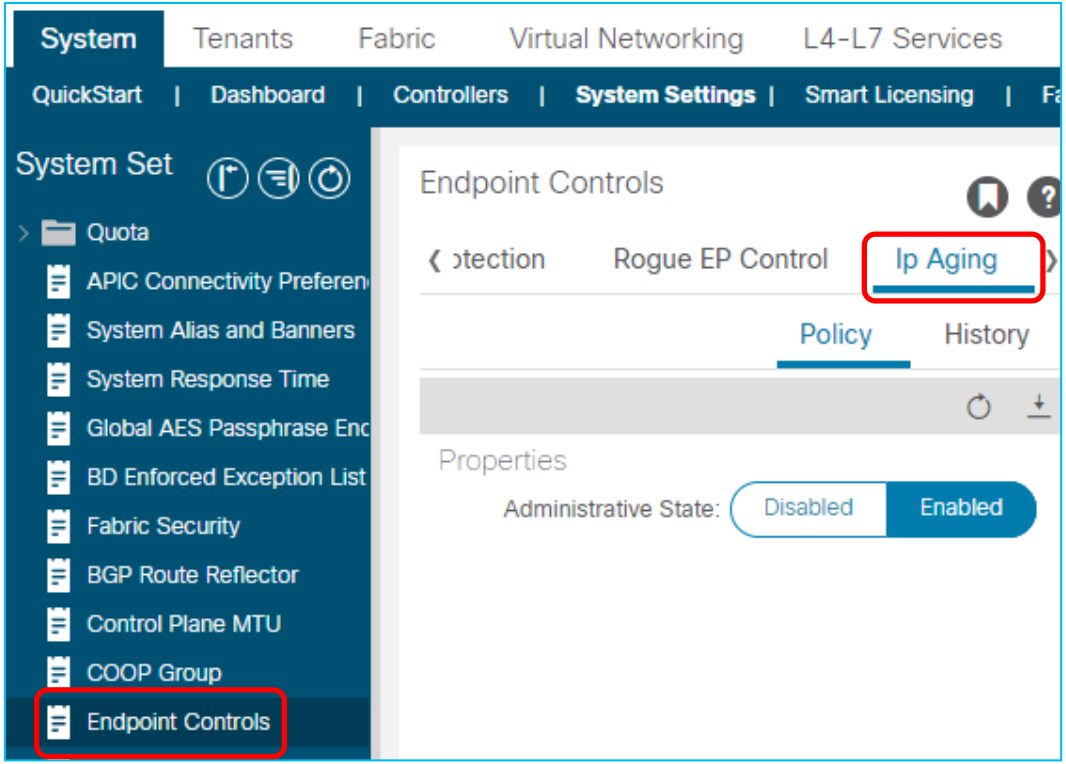
Best Practice Global settings

- Enforce Subnet Check Globally
  - Prevent endpoint learning outside of BD subnets
- Enforce EPG VLAN Validation
  - Prevent overlapping VLAN pool in one EPG
- Enable Domain Validation
  - Enforce domain association to an EPG
  - This cannot be disabled.
    You may have configurations that were working even if they were incorrect. Before enabling it make sure you verify the domain assignment to the EPGs and the associated AEPs.
- IP Aging Enabled
- Rogue EP Protection (more details later)
- MCP per VLAN enabled



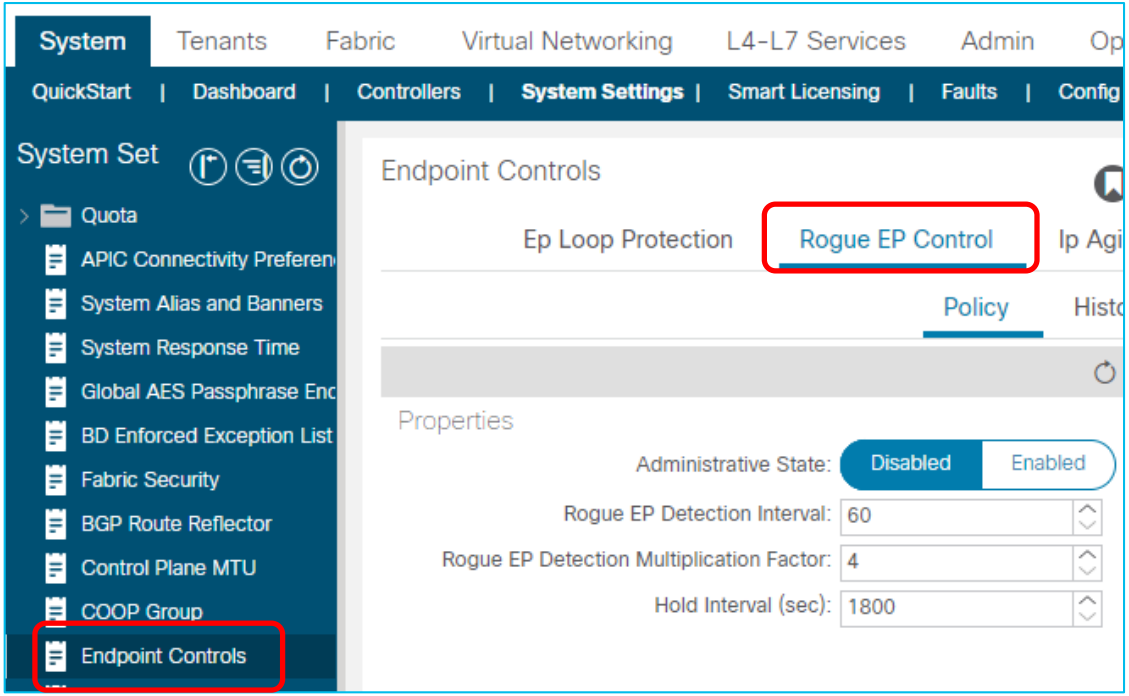
Best Practice Global settings

- Enforce Subnet Check Globally
  - Prevent endpoint learning outside of BD subnets
- Enforce EPG VLAN Validation
  - Prevent overlapping VLAN pool in one EPG
- Enable Domain Validation
  - Enforce domain association to an EPG
- IP Aging Enabled
  - Endpoint IPs tied to a MAC can age out individually
- Rogue EP Protection (more details later)
- MCP per VLAN enabled



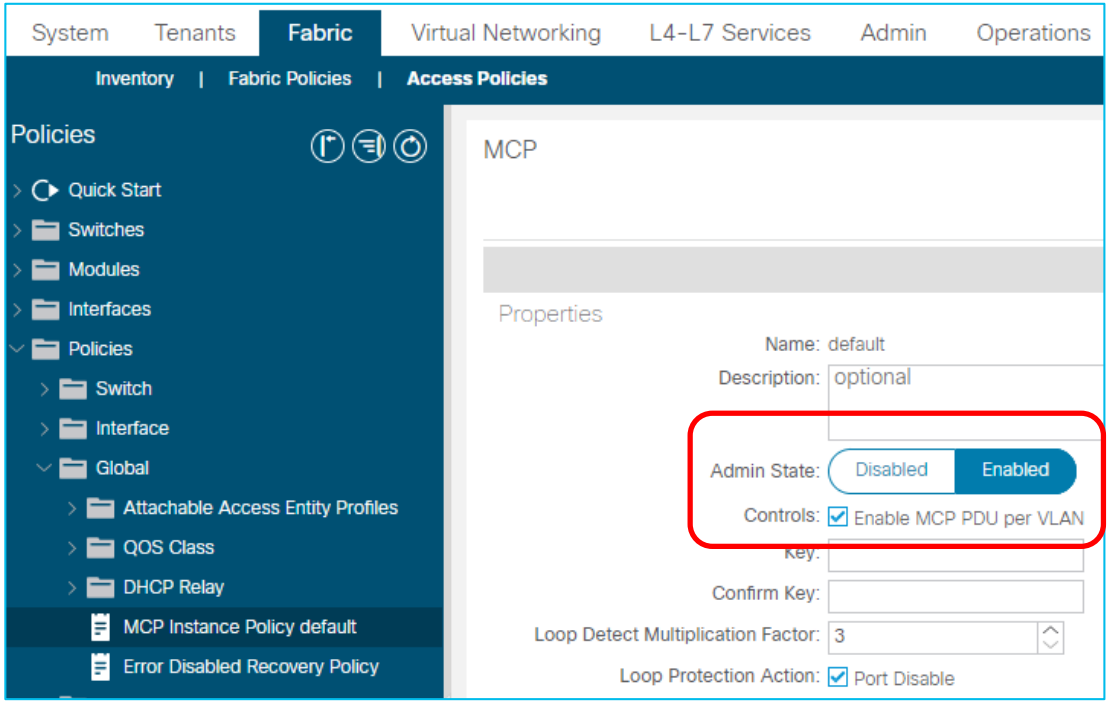
Best Practice Global settings

- Enforce Subnet Check Globally
  - Prevent endpoint learning outside of BD subnets
- Enforce EPG VLAN Validation
  - Prevent overlapping VLAN pool in one EPG
- Enable Domain Validation
  - Enforce domain association to an EPG
- IP Aging Enabled
  - Endpoint IPs tied to a MAC can age out individually
- Rogue EP Protection (more details later)
  - Mitigate impacts of endpoint flap issues
- MCP per VLAN enabled



Best Practice Global settings

- Enforce Subnet Check Globally
  - Prevent endpoint learning outside of BD subnets
- Enforce EPG VLAN Validation
  - Prevent overlapping VLAN pool in one EPG
- Enable Domain Validation
  - Enforce domain association to an EPG
- IP Aging Enabled
  - Endpoint IPs tied to a MAC can age out individually
- Rogue EP Protection (more details later)
  - Mitigate impacts of endpoint flap issues
- MCP per VLAN enabled
  - Granular Mis Cabling Protection

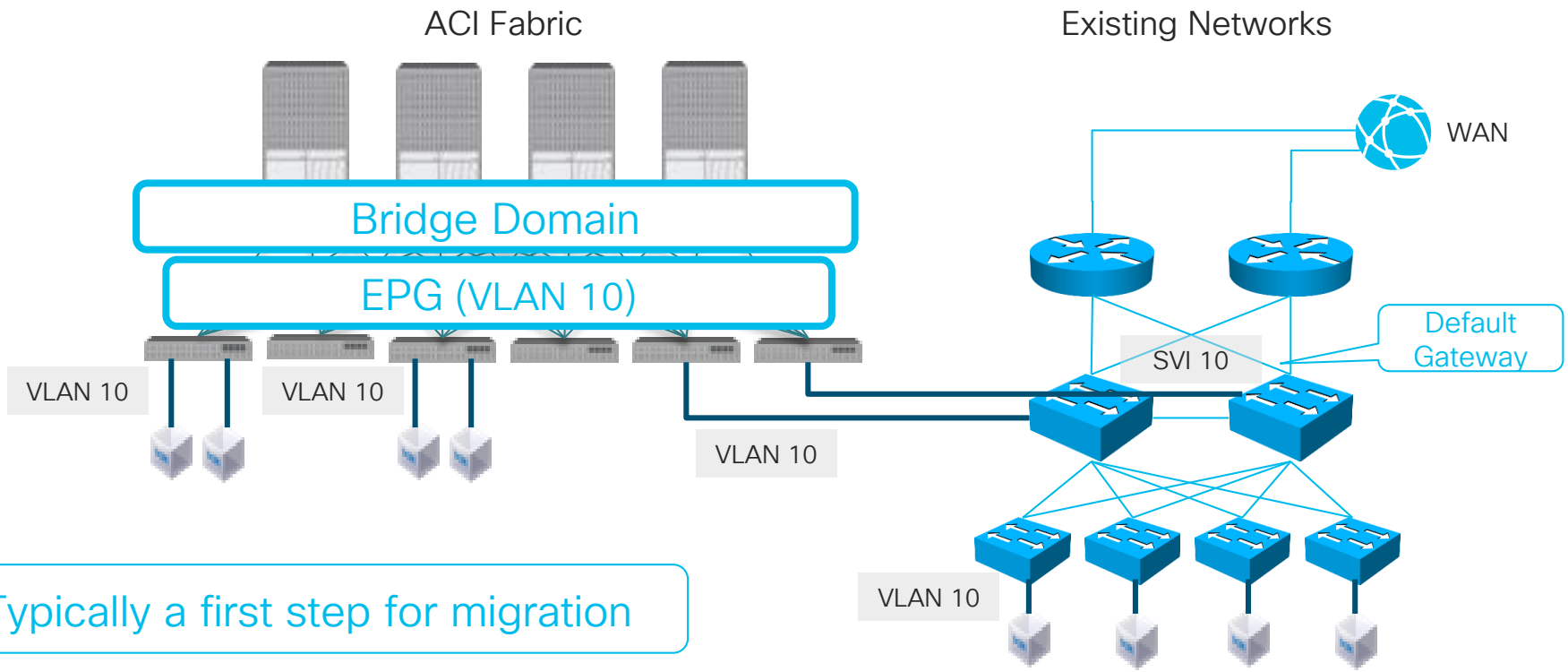


Agenda

- Single Fabric/Pod Topology Options
- ACI Design Considerations
  - ACI Fabric Bring Up
  - L2 Connectivity to Existing Networks
  - Loop prevention in ACI
  - Moving the default gateway to ACI
  - Connecting servers (Physical, VMM Integration, UCSM Integration)
  - Teaming Options
  - Allowing Traffic through ACI
  - Configuring L3 Connectivity to the Outside
- Application Centric Design
  - Network Centric & Application Centric
  - Contract Priority and Optimization

Layer 2 Connectivity to Existing Networks

with Network Centric Design (1 BD = 1 VLAN)



Typically a first step for migration

Layer 2 BD mode recommendations

Rule of thumb

Flood everything in L2 BD

The default gateway is on external device

Disable Unicast Routing (= L2 BD)

Unicast Routing: ☐

<u>best practice when connecting with external networks</u>

Make ACI behave in the same way as the external networks
=> Flood everything

L2 Unknown Unicast:	Flood	Hardware Proxy	
L3 Unknown Multicast Flooding:	Flood	Optimized Flood	
IPv6 L3 Unknown Multicast:	Flood	Optimized Flood	
Multi Destination Flooding:	Flood in BD	Drop	Flood in Encapsulation
PIM:	☐		
IGMP Policy:	select an option		
ARP Flooding:	☑		

Optimize L2 Unknown Unicast (L2UU) in ACI BD

Flood mode

L2 Unknown Unicast: **Flood** | Hardware Proxy

- Traffic with unknown destination MAC is flooded
  - Same as traditional switch

== Note ==

If there are any silent hosts or sensitive applications, use Flood just to be on the safe side.

Hardware Proxy mode

L2 Unknown Unicast: Flood | **Hardware Proxy**

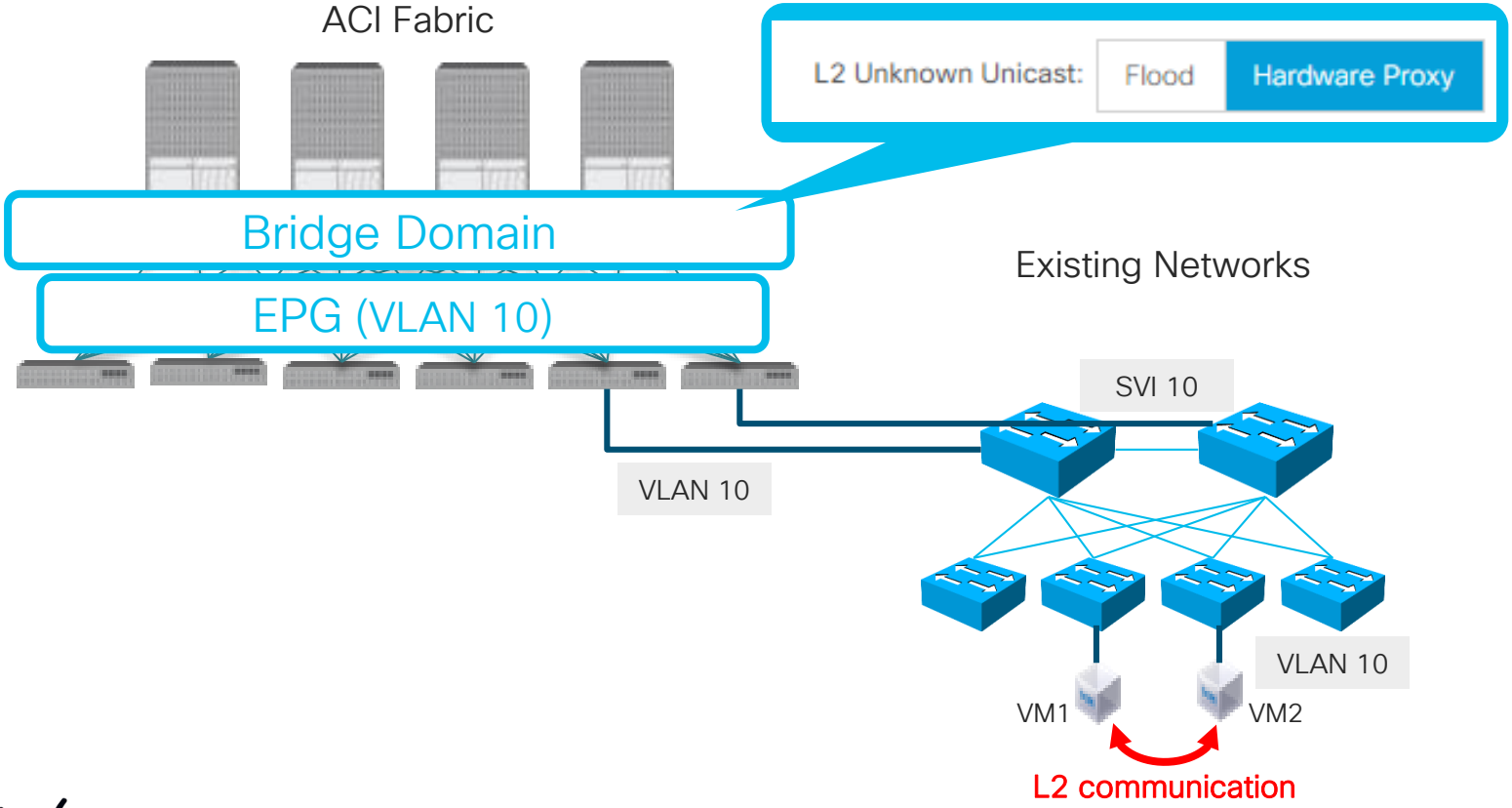
- Traffic with unknown destination MAC is sent to Spine Proxy.
  - Saves bandwidth
  - No unnecessary remote endpoint learning everywhere

== Caution ==

If the spine also does not know the MAC address, the packet is dropped. No silent host detection.

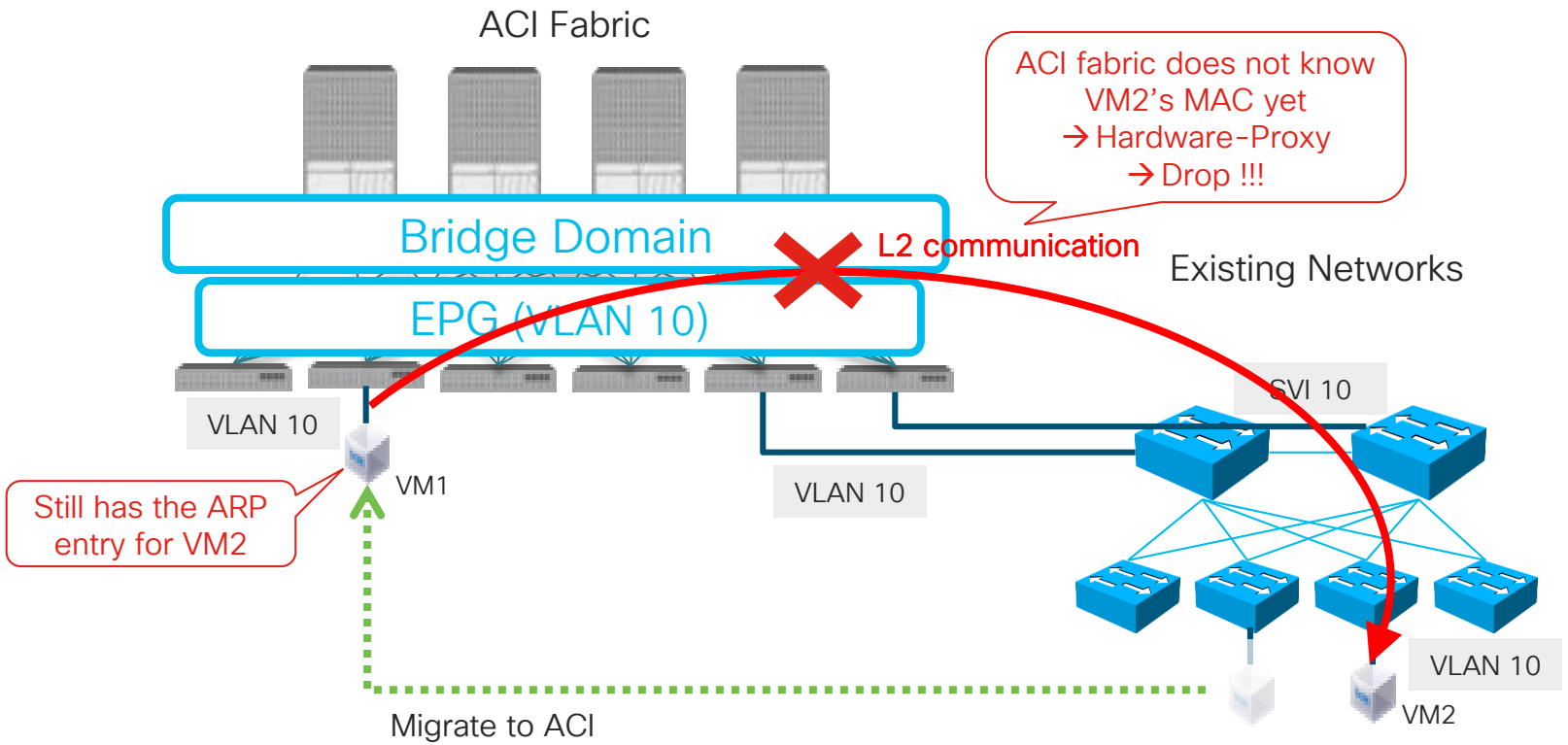
L2UU Hardware-Proxy drop example

Migration Scenario



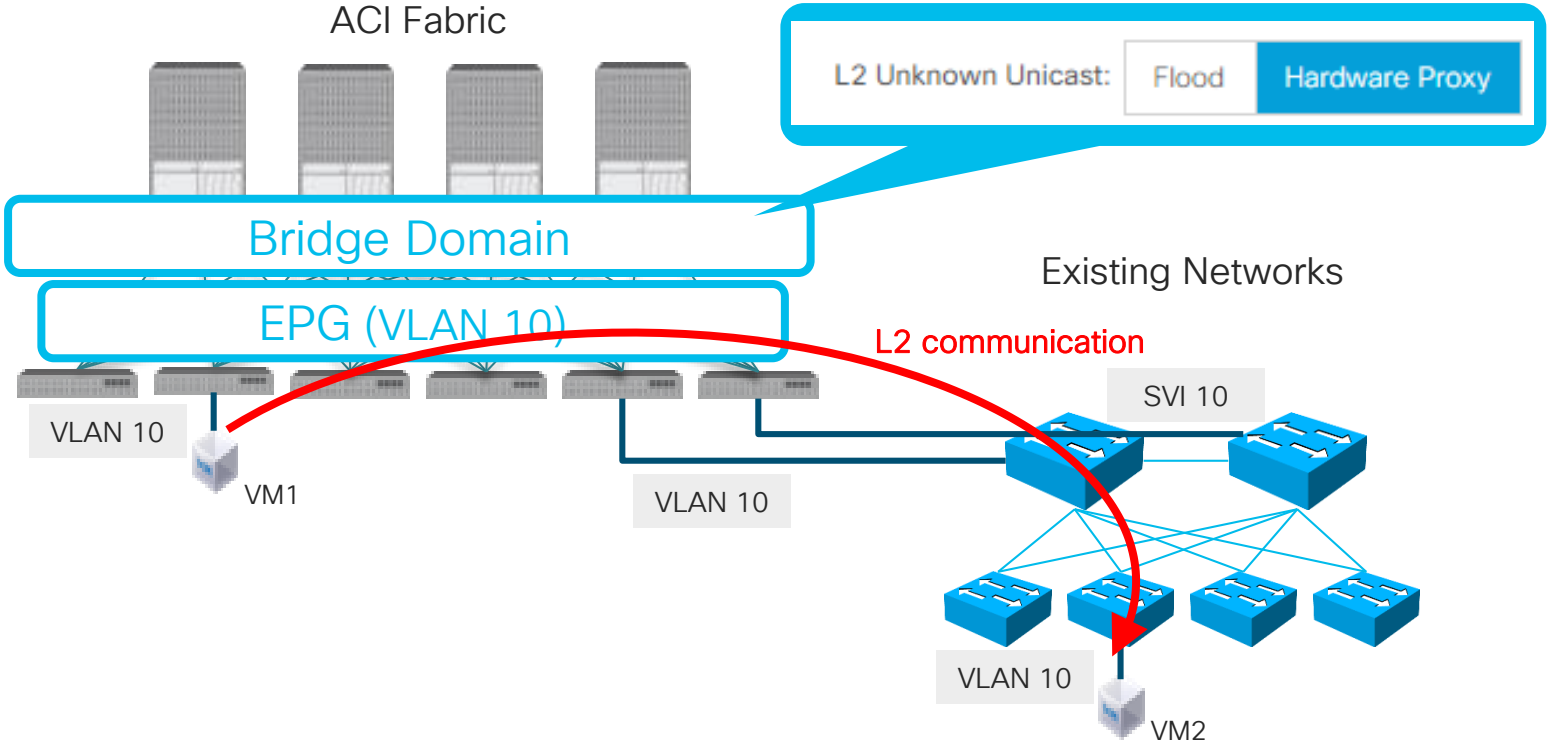
L2UU Hardware-Proxy drop example

Migration Scenario



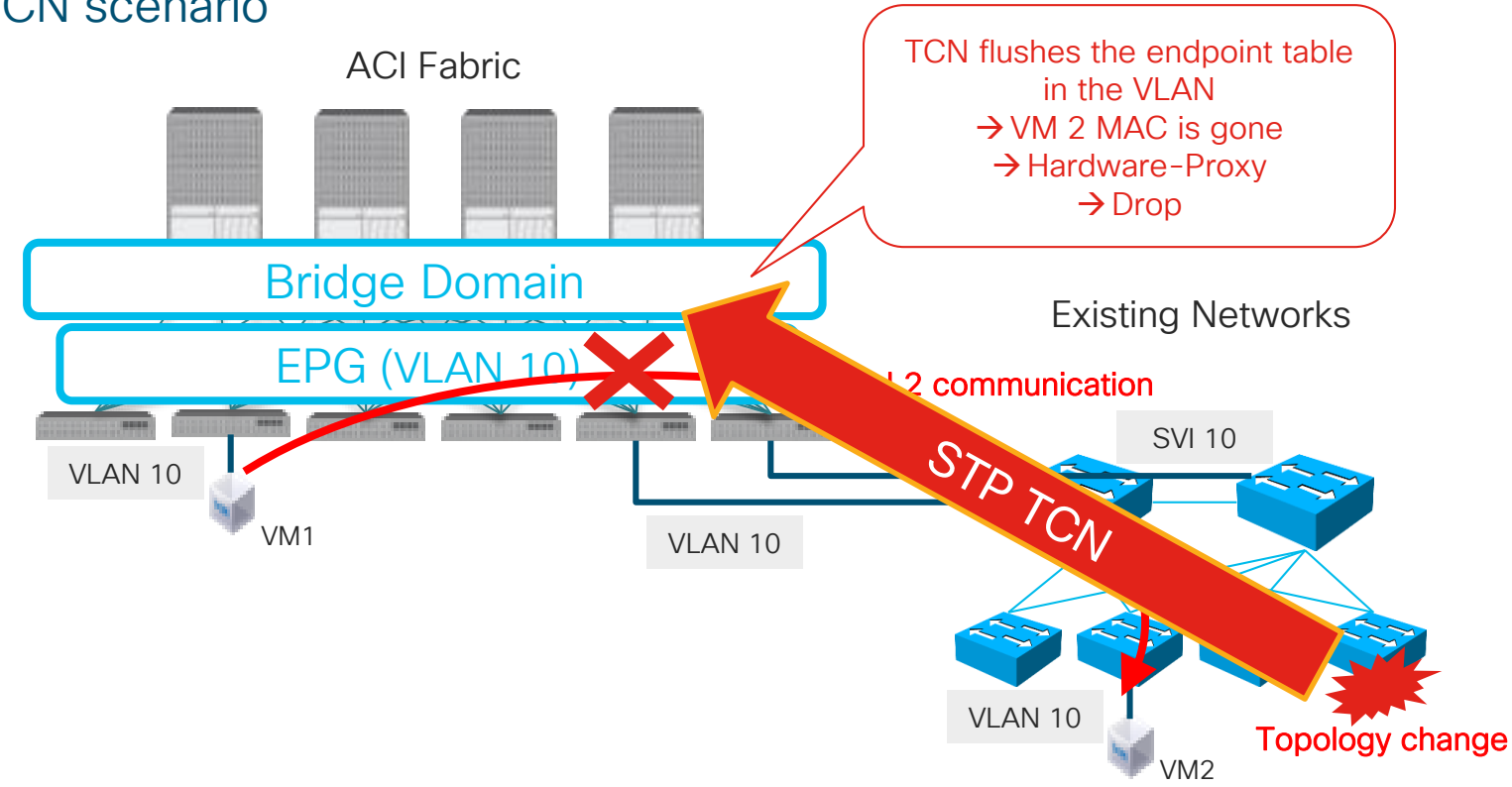
L2UU Hardware-Proxy drop example 2

STP TCN scenario



L2UU Hardware-Proxy drop example 2

STP TCN scenario



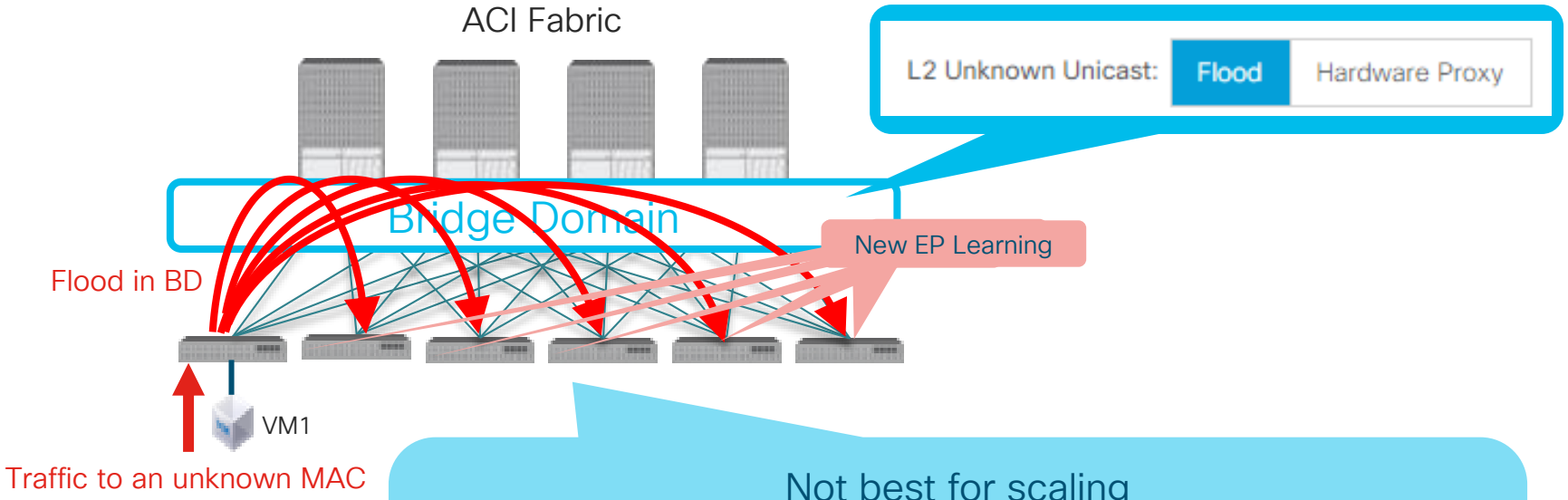
Rule of thumb (again)

Rule of thumb

Flood everything

- in L2 BD
- when connecting external networks

Down side of L2UU Flood



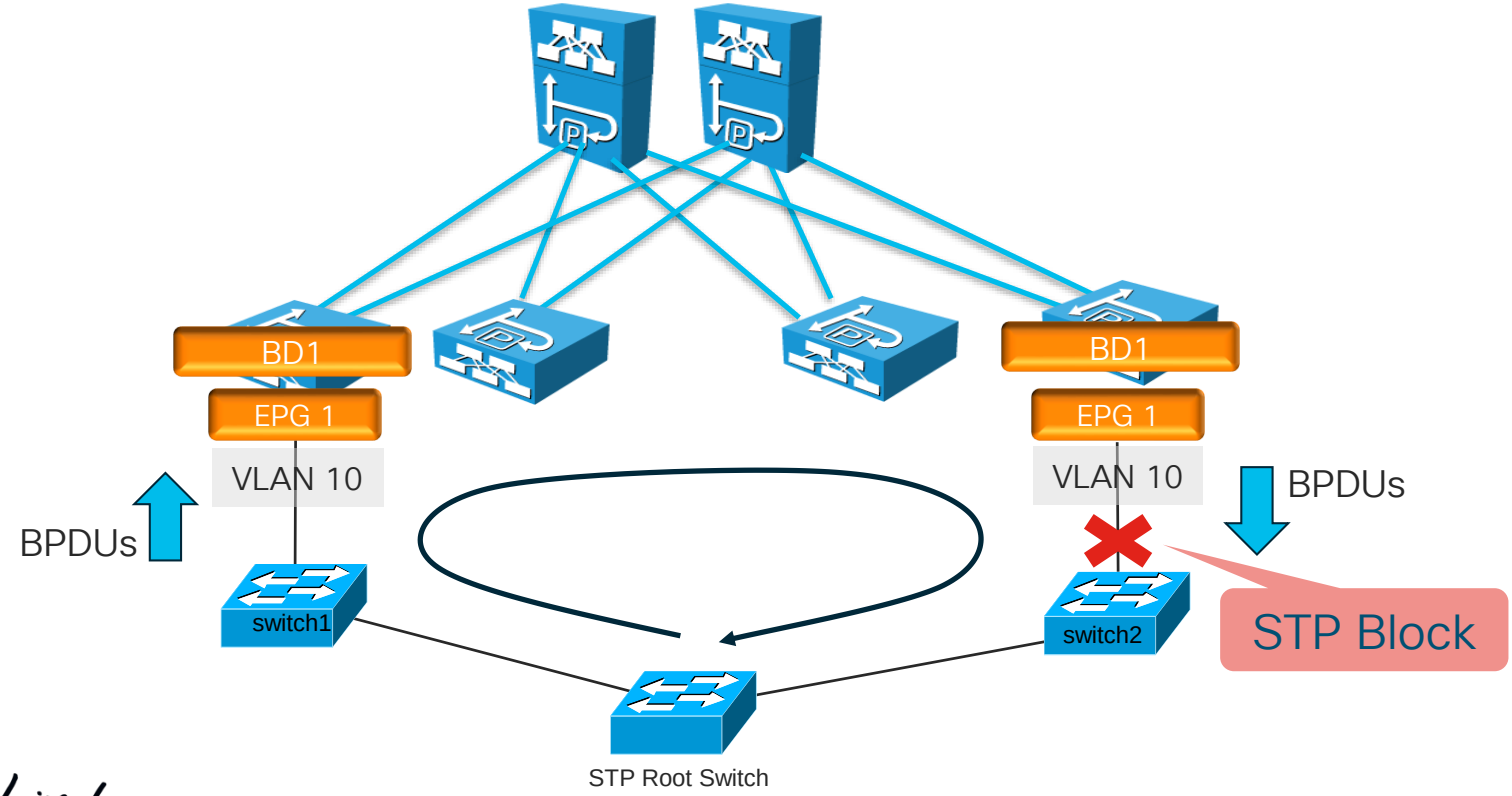
Not best for scaling

1. Non-optimal bandwidth usage
2. All leaf switches learn remote endpoint
   - Easier to fill up the endpoint table
   - No more new remote endpoints mean another flood

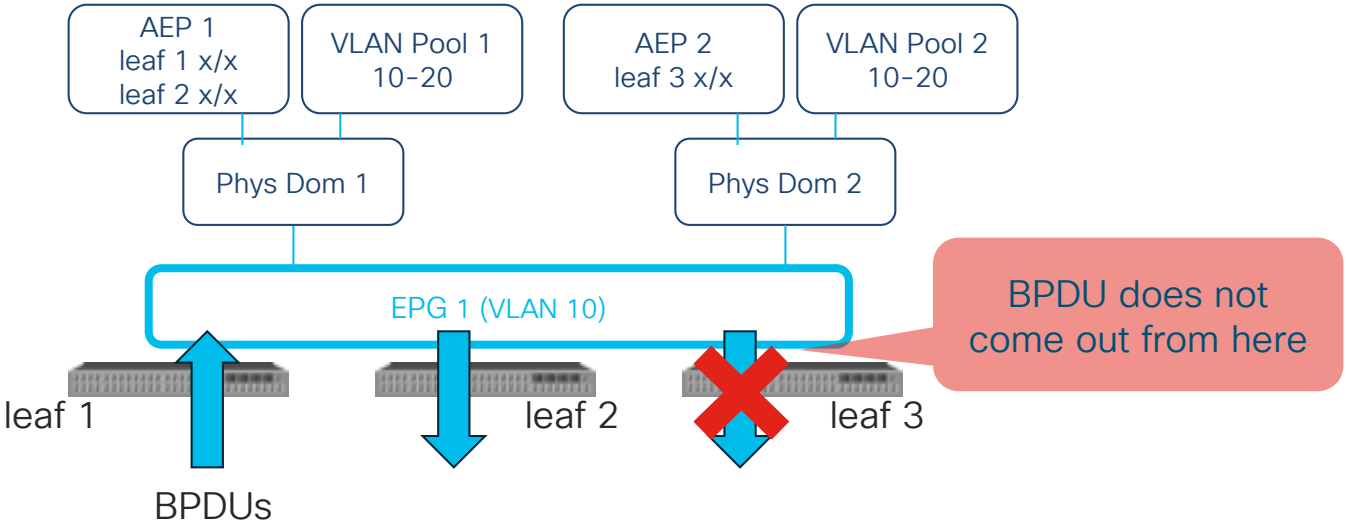
Agenda

- Single Fabric/Pod Topology Options
- ACI Design Considerations
  - ACI Fabric Bring Up
  - L2 Connectivity to Existing Networks
  - Loop prevention in ACI
  - Moving the default gateway to ACI
  - Connecting servers (Physical, VMM Integration, UCSM Integration)
  - Teaming Options
  - Allowing Traffic through ACI
  - Configuring L3 Connectivity to the Outside
- Application Centric Design
  - Network Centric & Application Centric
  - Contract Priority and Optimization

ACI forwarding BPDUs allows the external switches which run STP to prevent loops



STP flooding domain and ACI VLAN



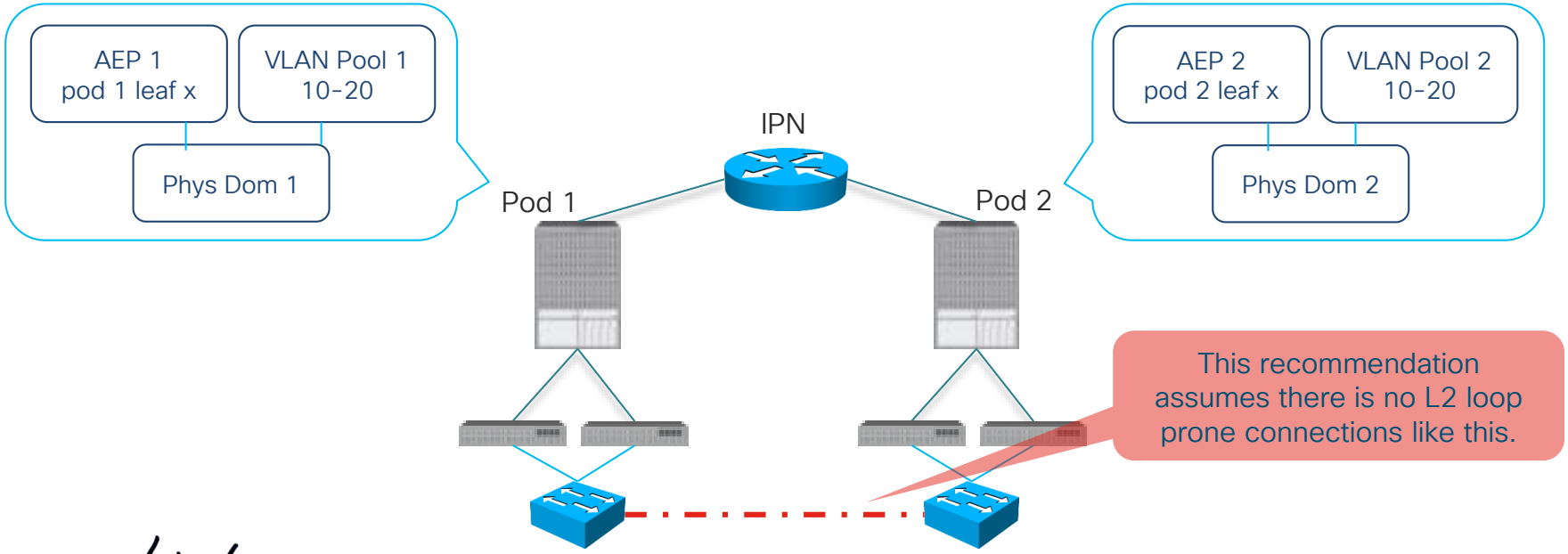
Each encap VLAN is assigned a VxLAN ID based on a VLAN Pool. This VxLAN ID is used to flood STP BPDUs.

- Different VLAN Pool = different STP flooding domain

STP flooding domain and ACI VLAN (cont.)

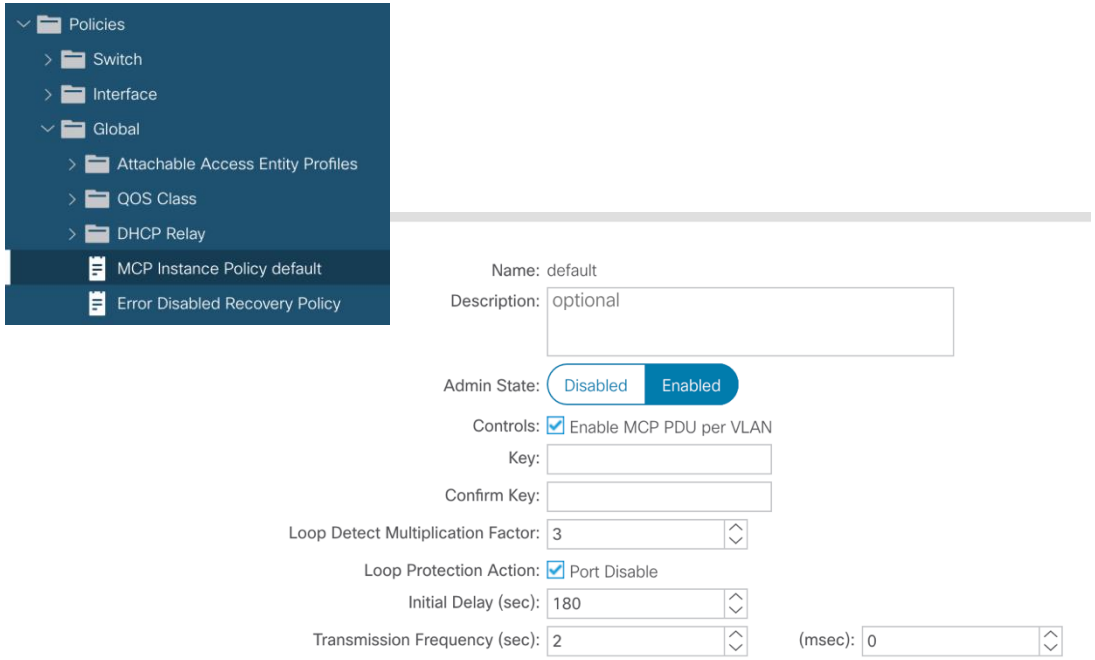
Recommended to create a domain and VLAN pool per pod

- Isolate STP domain per pod
- Prevent unnecessary BPDU, TCN propagations to other pods



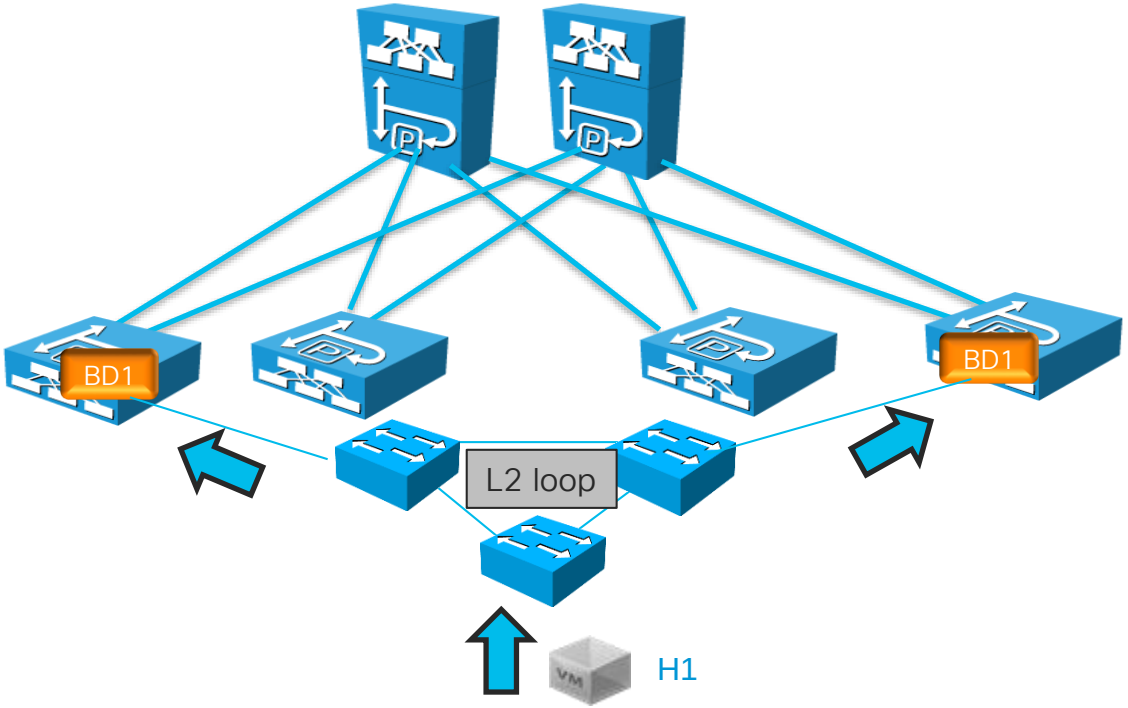
Miscabling Protocol (MCP)

- After ports go up, MCP waits before sending MCP PDUs
- This is so that Spanning-Tree IF present can converge
- *If during that time there is a temporary loop*, Rogue EP will quarantine the IP/MAC of the hosts on the BD that are experiencing the loop.
- This protects the fabric from the effects of the temporary loop



If a loop is occurring endpoints would be continuously flapping

- If a Loop occur like in this picture the host MAC keeps flapping between leafs like leaf1 and leaf4
- ACI has multiple ways to detect loops with this type of scenario
- ACI can detect loops from EP moves and it can disable BD learning when too many EP moves occur, or it can disable the ports where the move is happening or it can quarantine the specific EP that are flapping.



EP move dampening, EP Loop Protection, Rogue EP Detection

	EP move dampening	EP Loop Protection	Rogue EP Detection
Scope	per BD	Global	Global
Detection	aggregate numer of all moves per BD in a second	number of moves of an individual endpoint between two ports	number of moves of an individual endpoint in the specified interval
Detects MAC and/or IP move	Detects MAC moves, IP moves	Detects MAC moves	Detects MAC moves, IP moves
Possible actions	BD learn disable per leaf	port disable or BD learn disable per leaf	Programs static entry to disable learning for the specific entry

Hardening the ACI fabric to reduce the chance and/or impact of Layer 2 Loops

MCP globally enabled

MCP Instance Policy default

Admin State: Disabled Enabled

Rogue EP Control

Properties

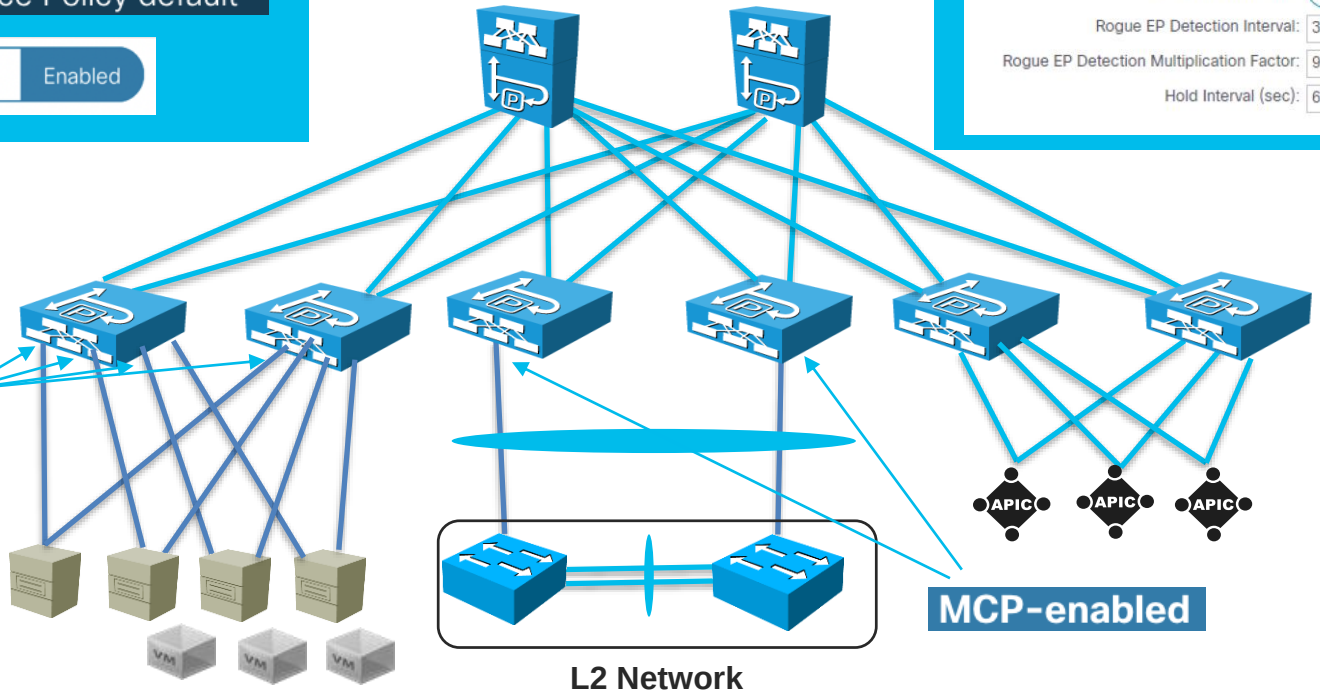
Administrative State: Disabled Enabled

Rogue EP Detection Interval: 30

Rogue EP Detection Multiplication Factor: 9

Hold Interval (sec): 600

BPDU guard

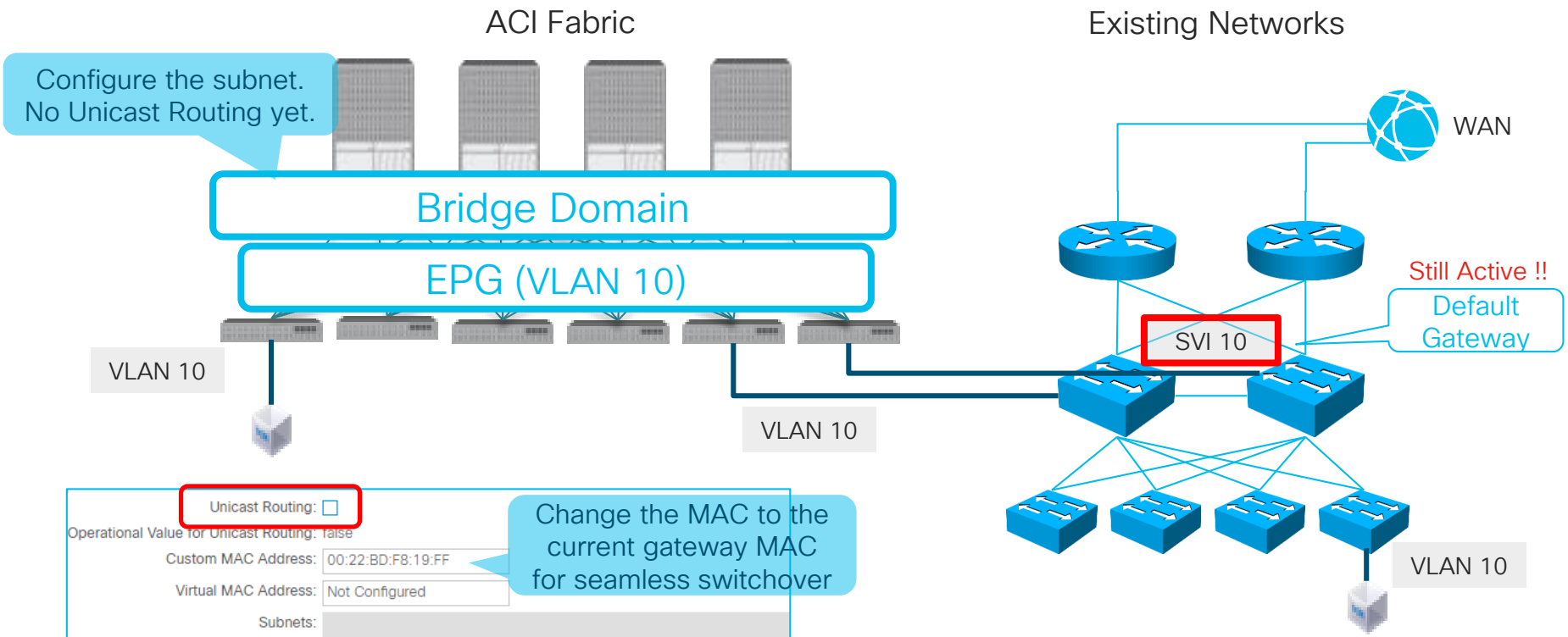


Agenda

- Single Fabric/Pod Topology Options
- ACI Design Considerations
  - ACI Fabric Bring Up
  - L2 Connectivity to Existing Networks
  - Loop prevention in ACI
  - Moving the default gateway to ACI
  - Connecting servers (Physical, VMM Integration, UCSM Integration)
  - Teaming Options
  - Allowing Traffic through ACI
  - Configuring L3 Connectivity to the Outside
- Application Centric Design
  - Network Centric & Application Centric
  - Contract Priority and Optimization

Migrate the default gateway to ACI

with Network Centric Design (1 BD = 1 VLAN)

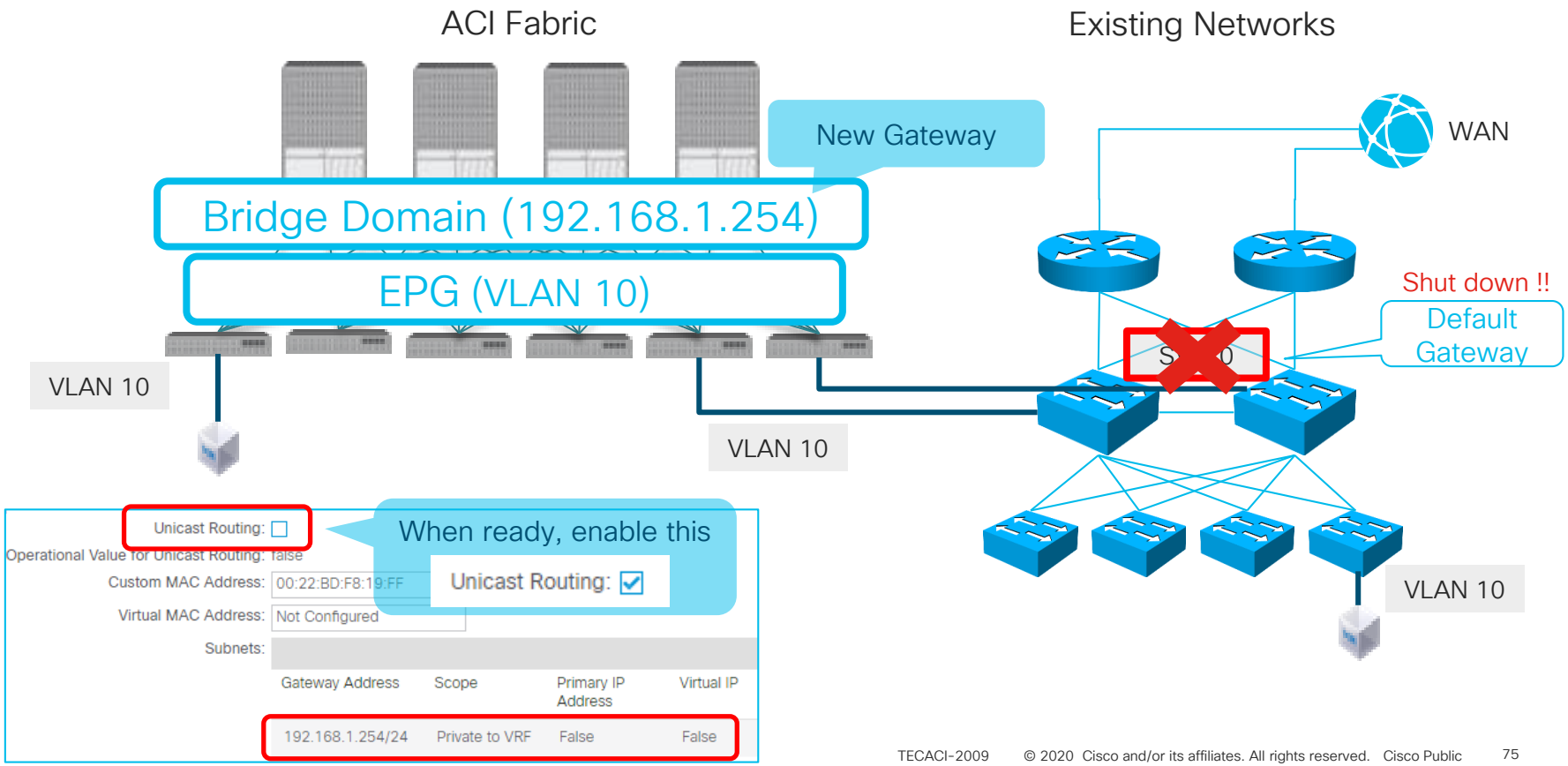


Unicast Routing: ☐
Operational Value for Unicast Routing: false
Custom MAC Address: 00:22:BD:F8:19:FF
Virtual MAC Address: Not Configured
Subnets:

Gateway Address	Scope	Primary IP Address	Virtual IP
192.168.1.254/24	Private to VRF	False	False

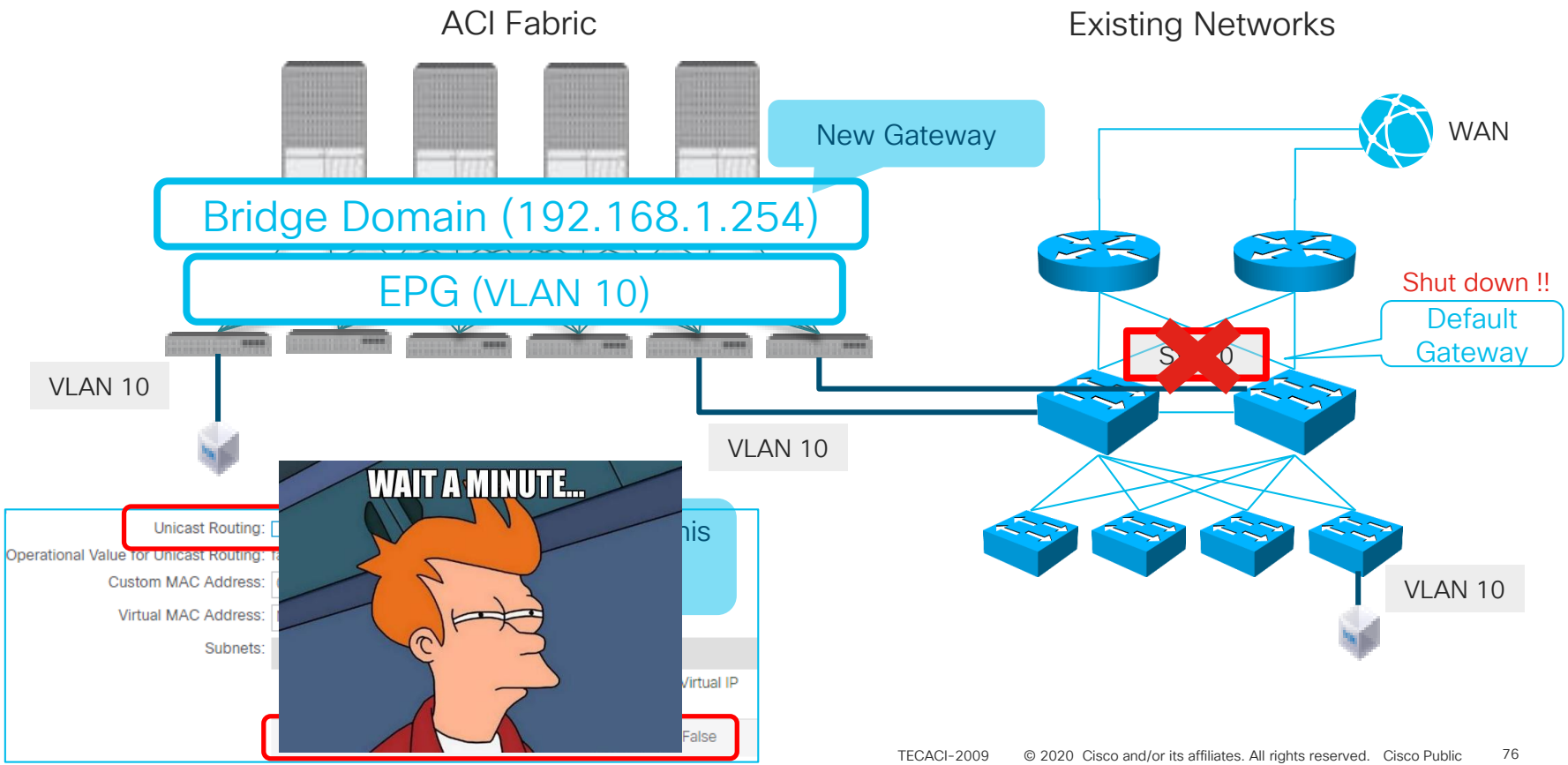
Migrate the default gateway to ACI

with Network Centric Design (1 BD = 1 VLAN)



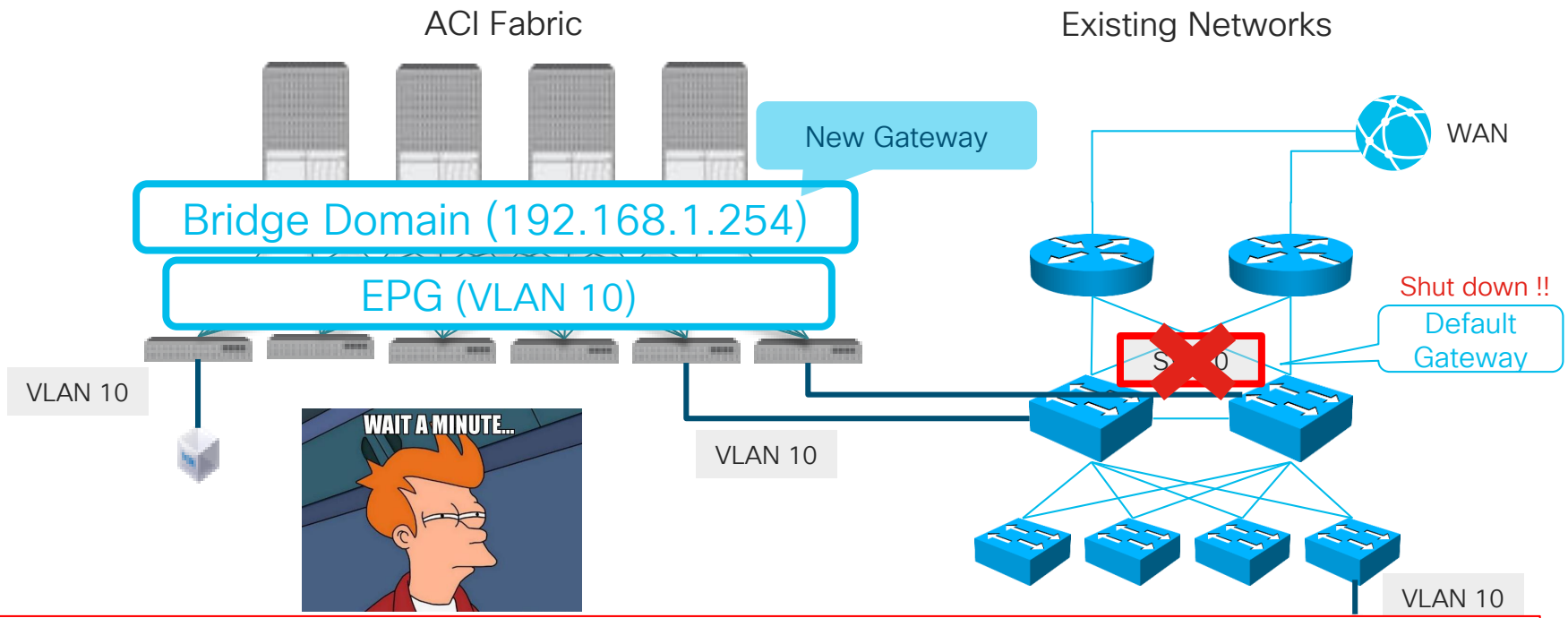
Migrate the default gateway to ACI

with Network Centric Design (1 BD = 1 VLAN)



Migrate the default gateway to ACI

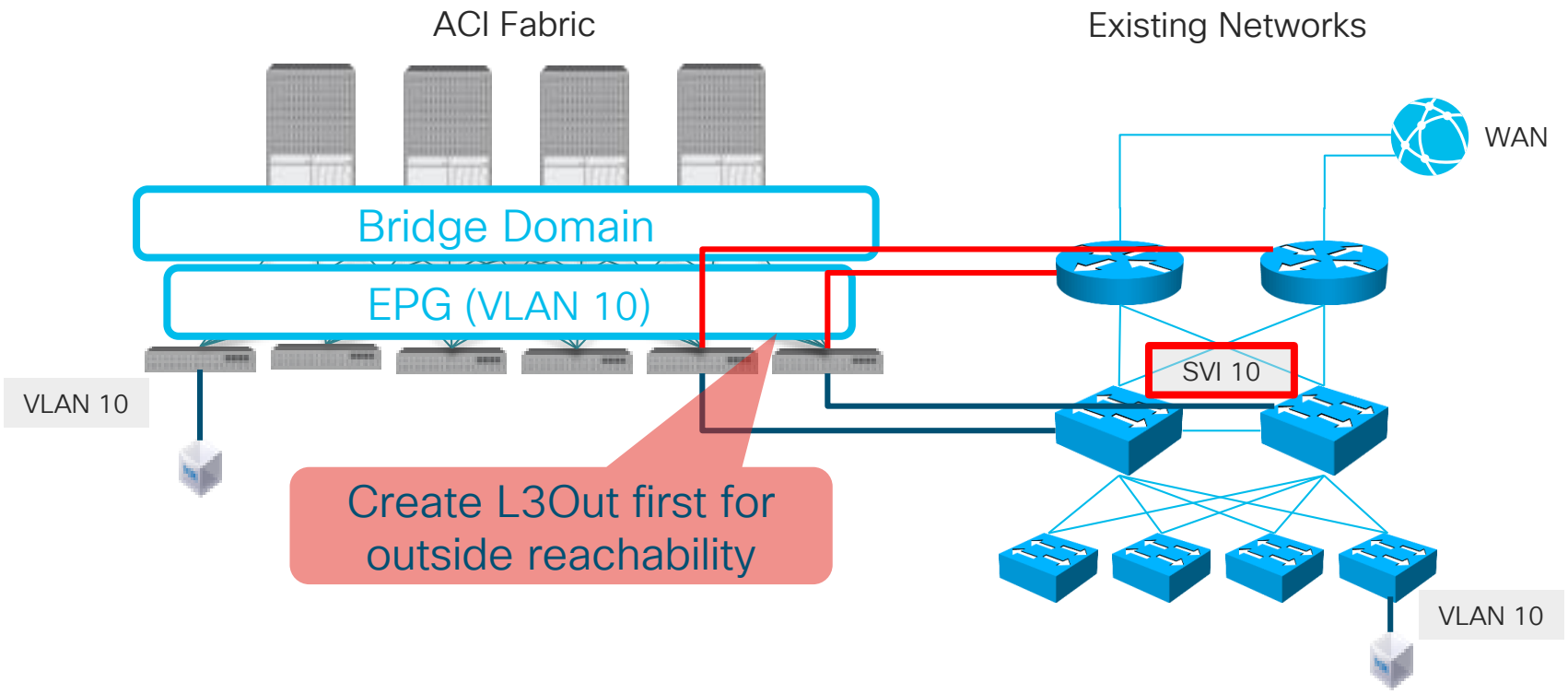
with Network Centric Design (1 BD = 1 VLAN)



how am I supposed to go out to WAN now?

Migrate the default gateway to ACI

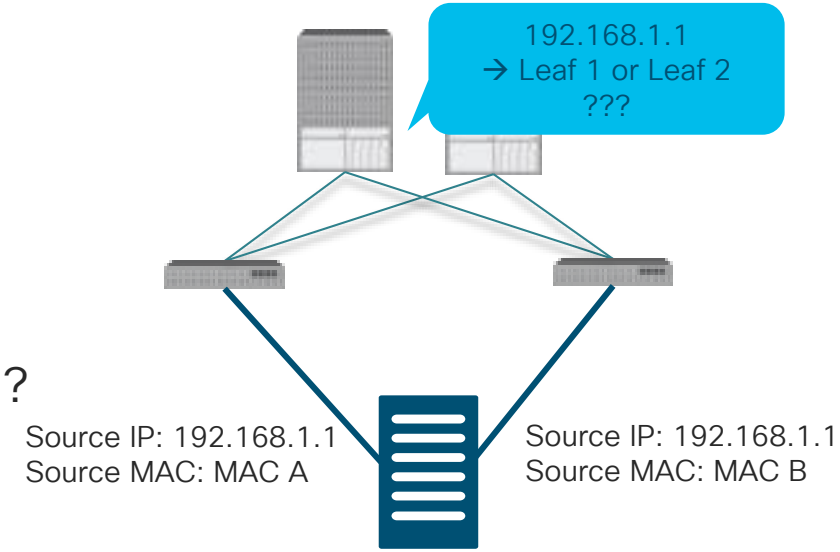
with Network Centric Design (1 BD = 1 VLAN)



Enabling Unicast Routing enables Endpoint IP learning as well

Keep in mind:
"ACI does data-plane learning for IP as well"

Bridging Traffic → MAC learning
Routing Traffic → MAC & IP learning
ARP Request → MAC & IP learning

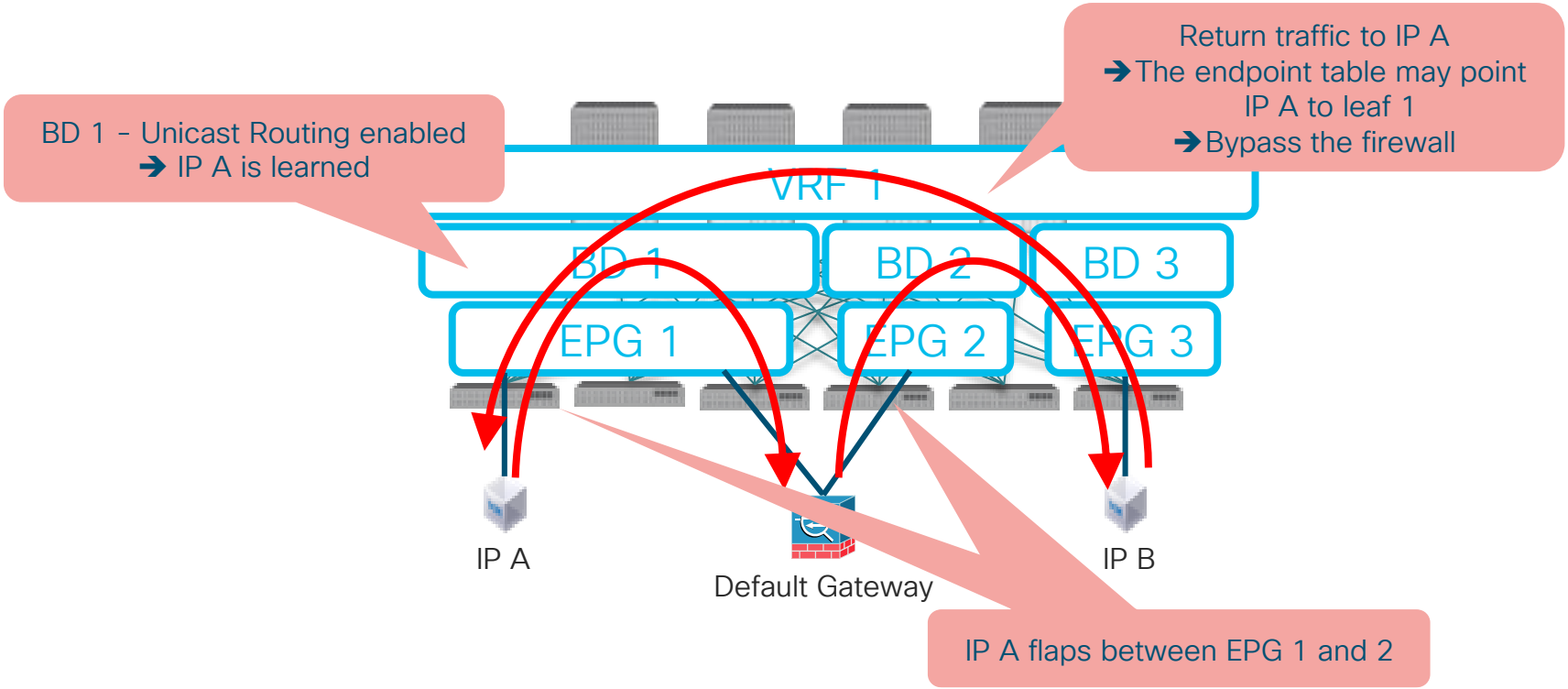


Does your NIC teaming work well with ACI?

If you have servers with Active/Active teaming
=> Tuning may be required => more on this later

Make sure you understand how endpoint learning works:
https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-739989.html

Why not Unicast Routing on L2 BD?



Caution - Just one example of many other corner cases

Agenda

- Single Fabric/Pod Topology Options
- ACI Design Considerations
  - ACI Fabric Bring Up
  - L2 Connectivity to Existing Networks
  - Loop prevention in ACI
  - Moving the default gateway to ACI
  - Connecting servers (Physical, VMM Integration, UCSM Integration)
  - Teaming Options
  - Allowing Traffic through ACI
  - Configuring L3 Connectivity to the Outside
- Application Centric Design
  - Network Centric & Application Centric
  - Contract Priority and Optimization

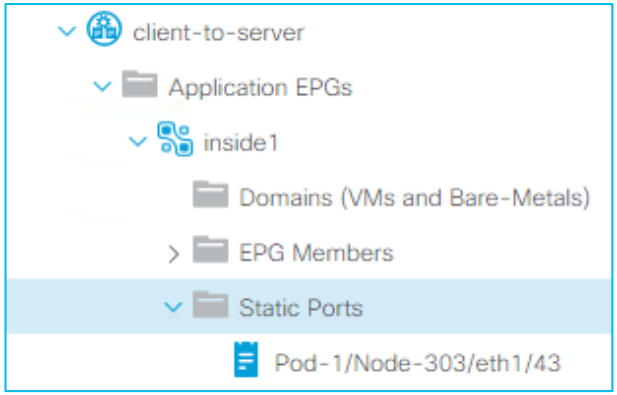
Two Server Connectivity options

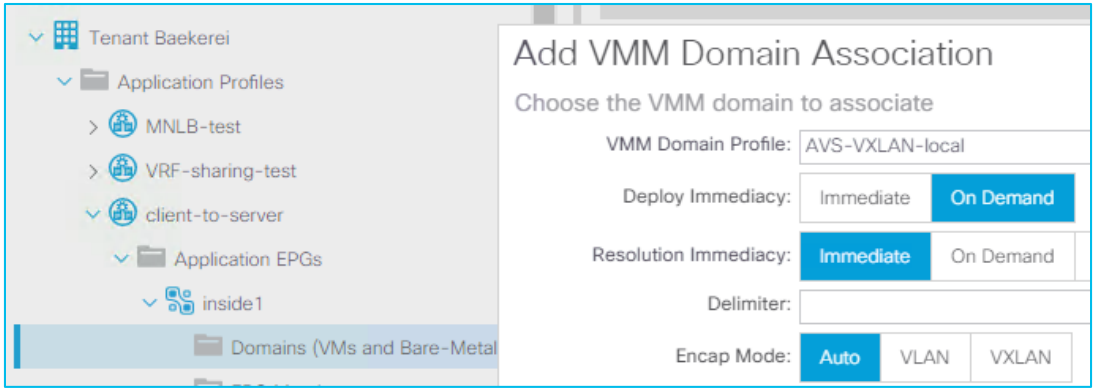
Manual (Static Path Binding)

- Associate a physical domain to an EPG
- Manually assign a VLAN (EPG) on a leaf interface

VMM Integration (Dynamic Path Binding)

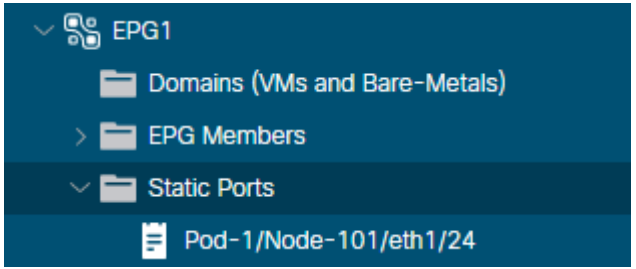
- Associate a VMM domain to an EPG.
- ACI-VMM dynamically figures out a leaf interface and VLAN via CDP/LLDP/opflex



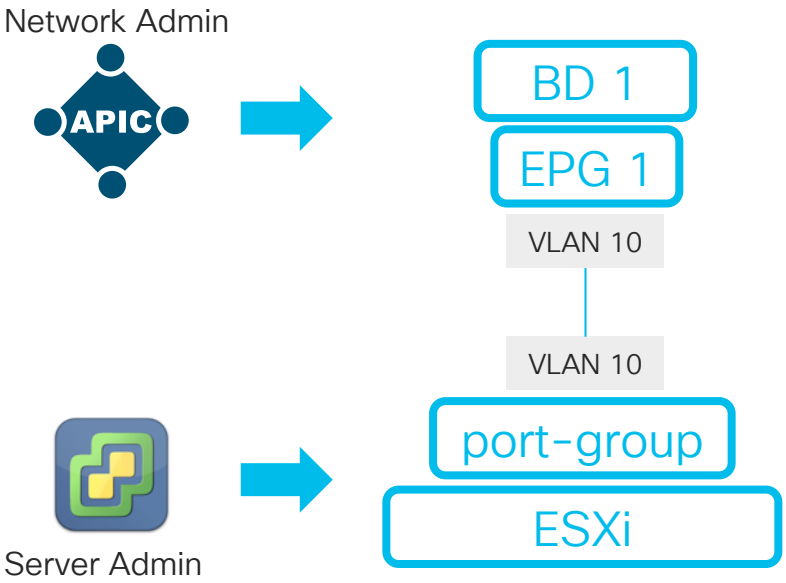


Physical Domain (physdom)

Can be used for both physical and virtual servers.



Configure Manually



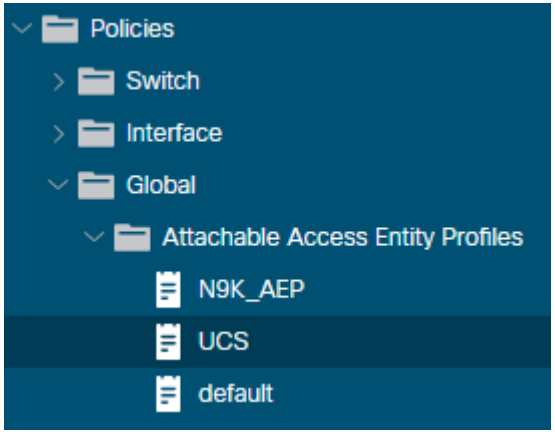
This can be tedious and error prone...

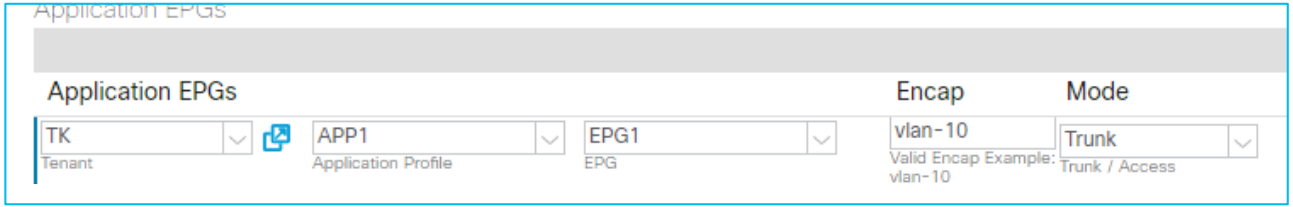
Bulk Static Path Binding via AEP

Assign VLAN 10 as EPG 1 on all interfaces in AEP UCS

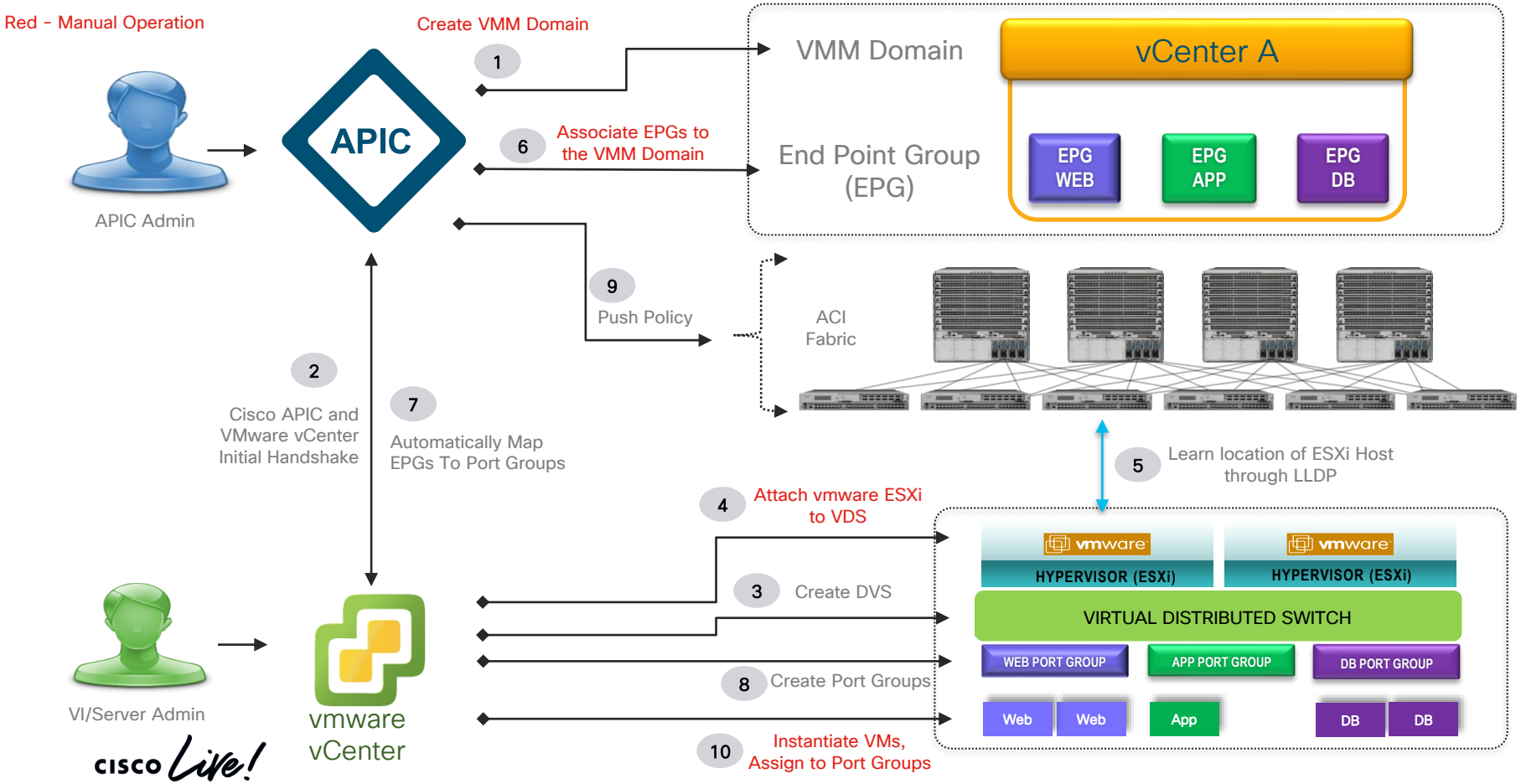
- Design AEP (group of interfaces) carefully

AEP | Tenant | Application profile | EPG | VLAN

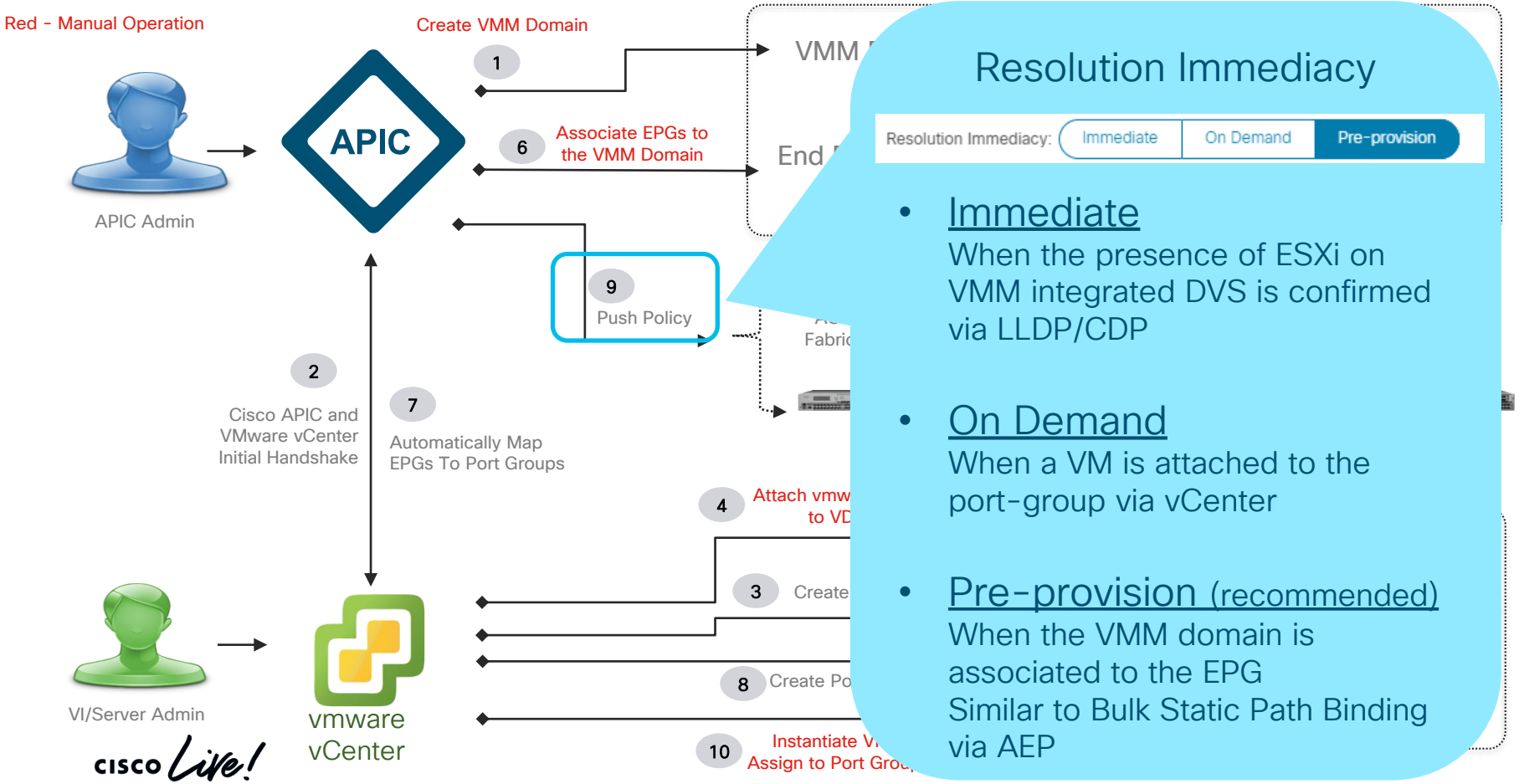




VMM Domain (vmware vCenter example)



VMM Domain (vmware vCenter example)

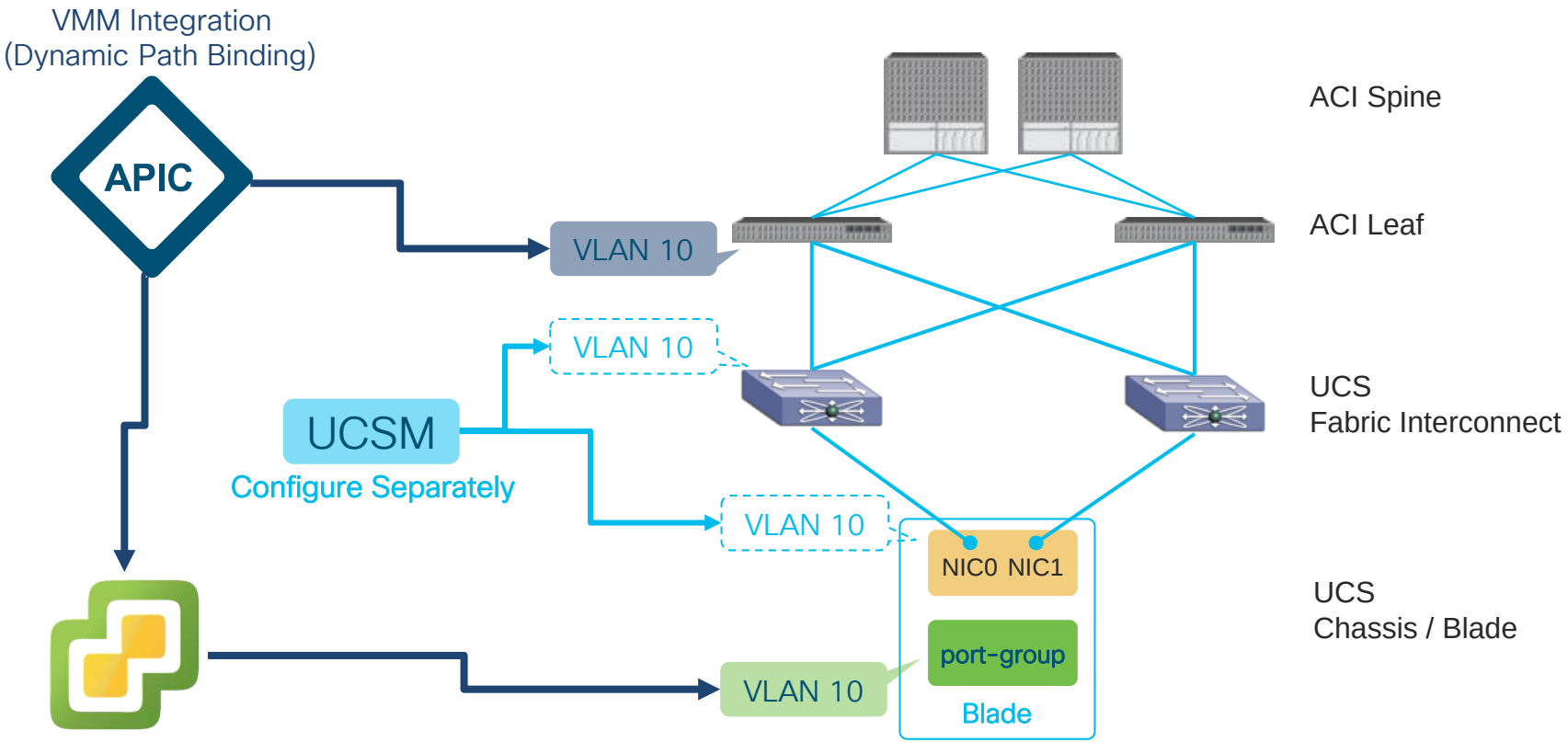


Resolution Immediacy

Resolution Immediacy: Immediate | On Demand | Pre-provision

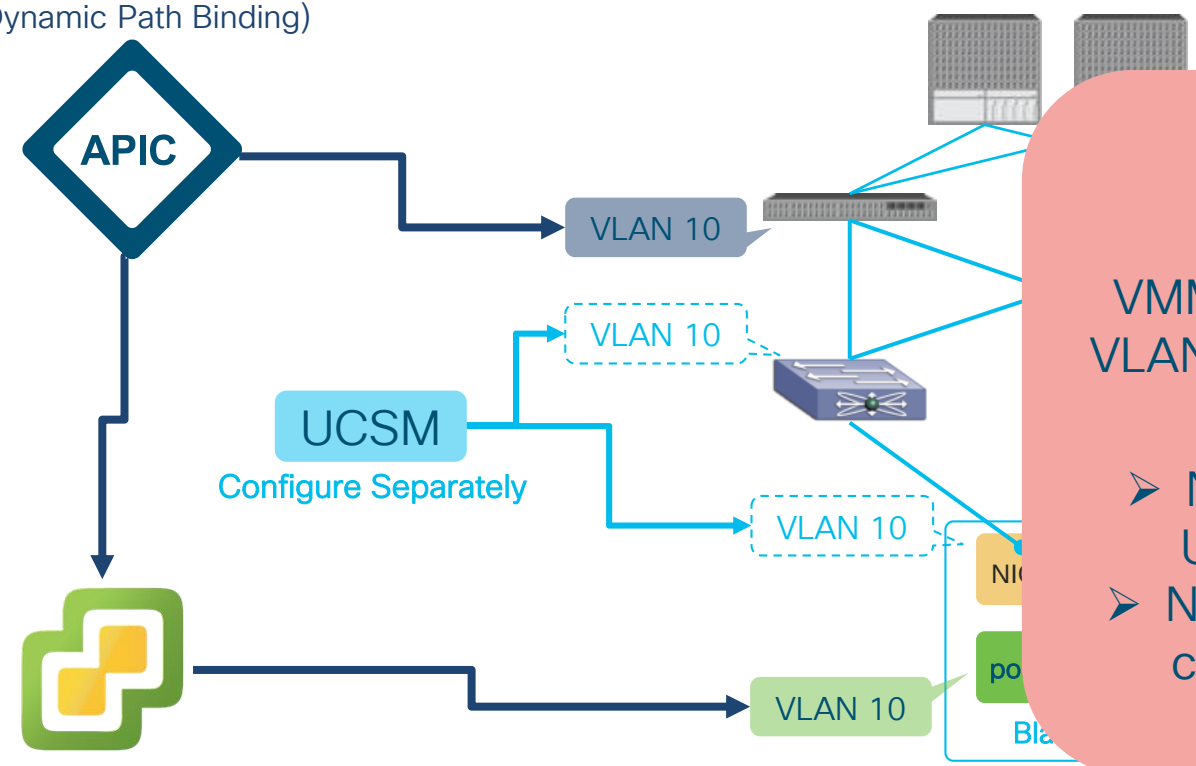
- **Immediate**
  When the presence of ESXi on VMM integrated DVS is confirmed via LLDP/CDP
- **On Demand**
  When a VM is attached to the port-group via vCenter
- **Pre-provision (recommended)**
  When the VMM domain is associated to the EPG
  Similar to Bulk Static Path Binding via AEP

ACI and UCS Fabric Interconnect (UCSM)



ACI and UCS Fabric Interconnect (UCSM)

VMM Integration
(Dynamic Path Binding)



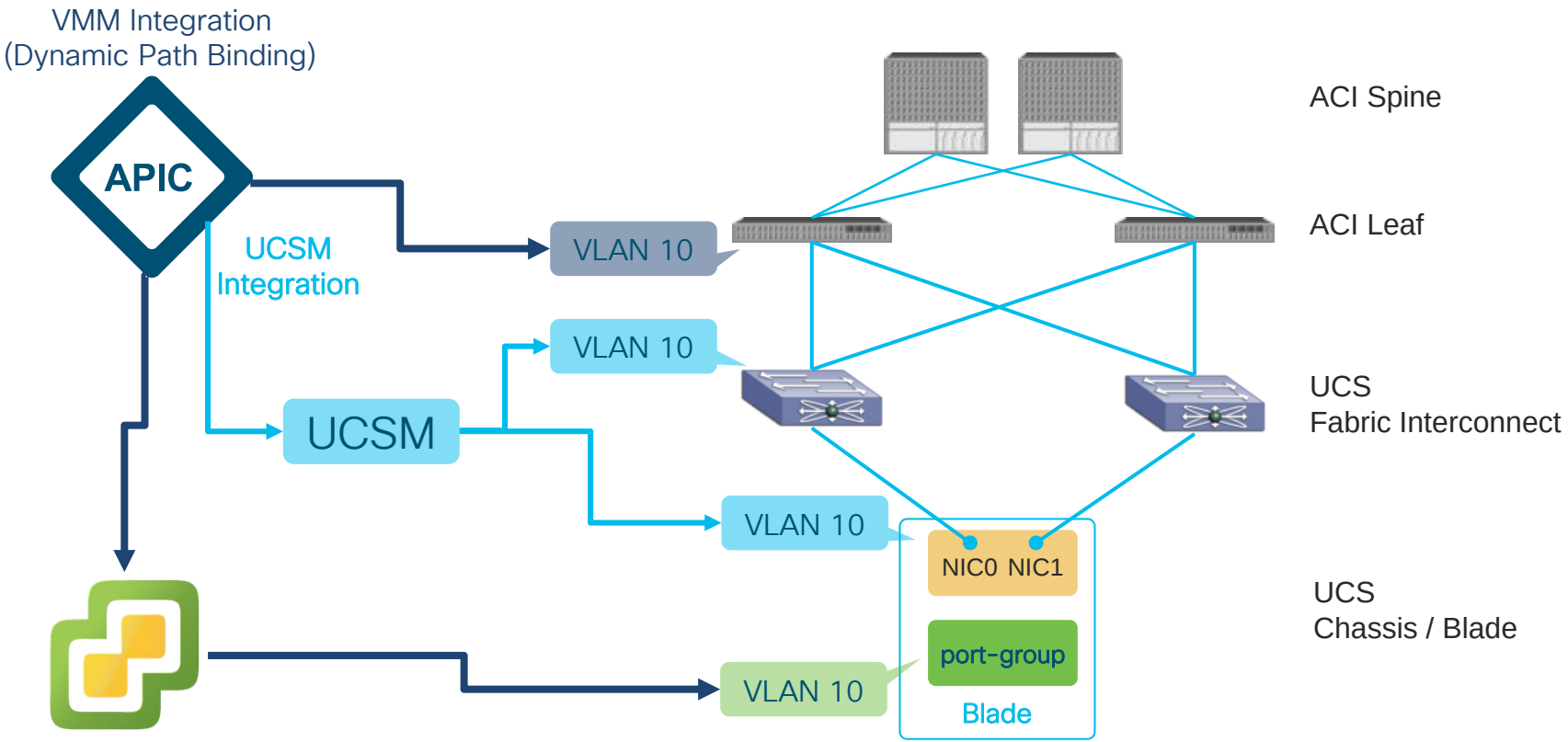
Problem

VMM integration picks VLANs dynamically from a pool.

- Not easy to match UCS configuration.
- Not scalable to pre-configure all VLANs

ACI and UCSM Integration

From ACI 4.1



UCS FI has a range of reserved VLANs

Quoting the UCSM Network Management Guide:

“”“

You cannot create VLANs with IDs from 4030 to 4047 or from 4093 to 4095. These ranges of VLAN IDs are reserved.

“”“

https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/ucs-manager/GUI-User-Guides/Network-Mgmt/3-2/b_UCSM_Network_Mgmt_Guide_3_2/b_UCSM_Network_Mgmt_Guide_3_2_chapter_0101.html

ACI (UCSM Integration) does not verify if the VMM domain uses those reserved VLANs.

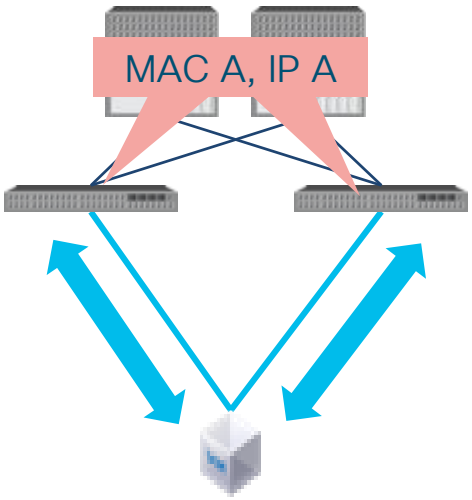
➔ Ensure not to use such VLANs in your pool.

Agenda

- Single Fabric/Pod Topology Options
- ACI Design Considerations
  - ACI Fabric Bring Up
  - L2 Connectivity to Existing Networks
  - Loop prevention in ACI
  - Moving the default gateway to ACI
  - Connecting servers (Physical, VMM Integration, UCSM Integration)
  - Teaming Options
  - Allowing Traffic through ACI
  - Configuring L3 Connectivity to the Outside
- Application Centric Design
  - Network Centric & Application Centric
  - Contract Priority and Optimization

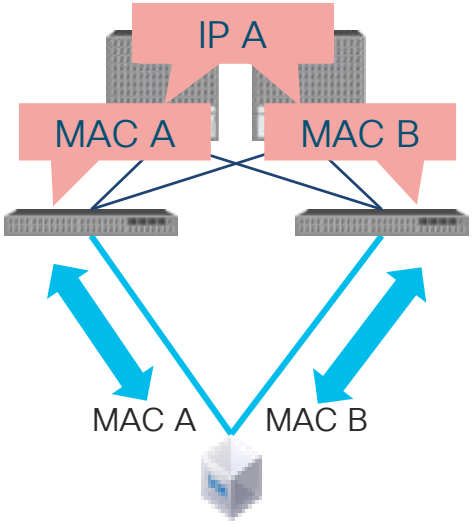
NIC Teaming Issues with ACI

MAC/IP flap between nodes/ports



Same problem as any other switch

IP flap between MAC



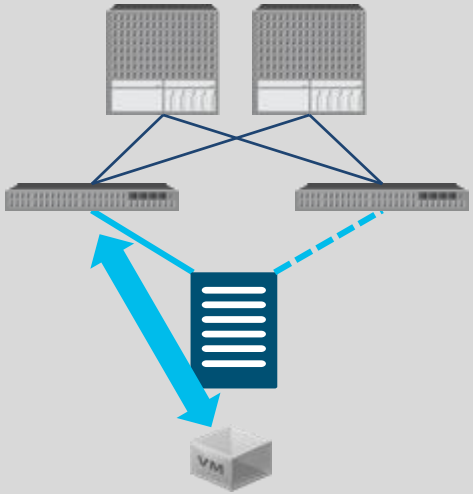
ACI specific due to IP data-plane learning

General NIC Teaming Types

All three work well with ACI

Microsoft - Switch Independent Hyper-V
VMware - Route Based on Originating Port ID

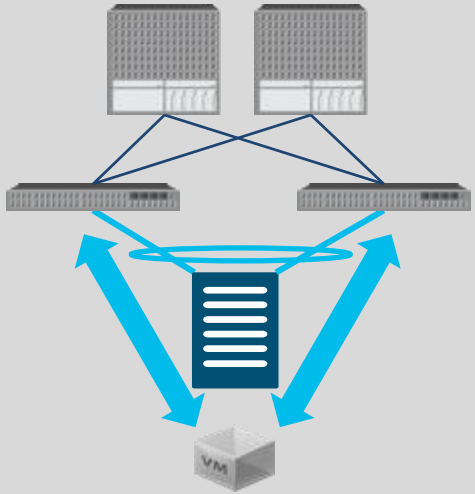
Active / Standby



One uplink is used for all VMs
Another uplink is only for failover

Active / Active

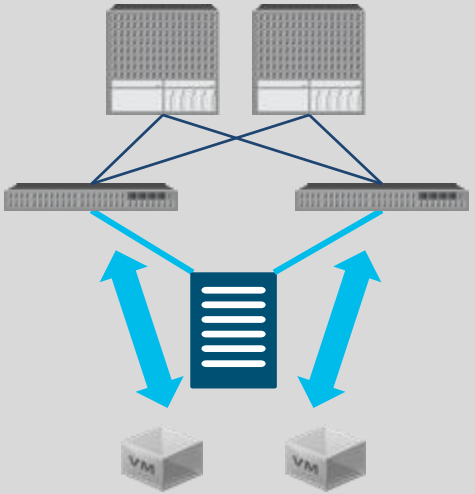
(Link Aggregation (LAG))



All VMs use all uplinks as logically one uplink

Active / Active

(MAC Pinning)

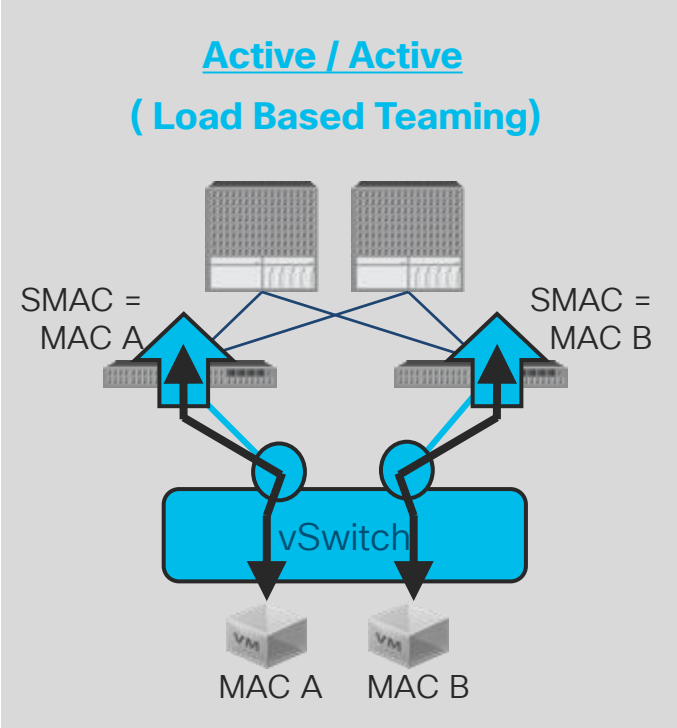


Each VM is assigned (pinned) to a dedicated uplink for all traffic

NIC Teaming Types to avoid

VMware – Load Based Teaming (LBT)

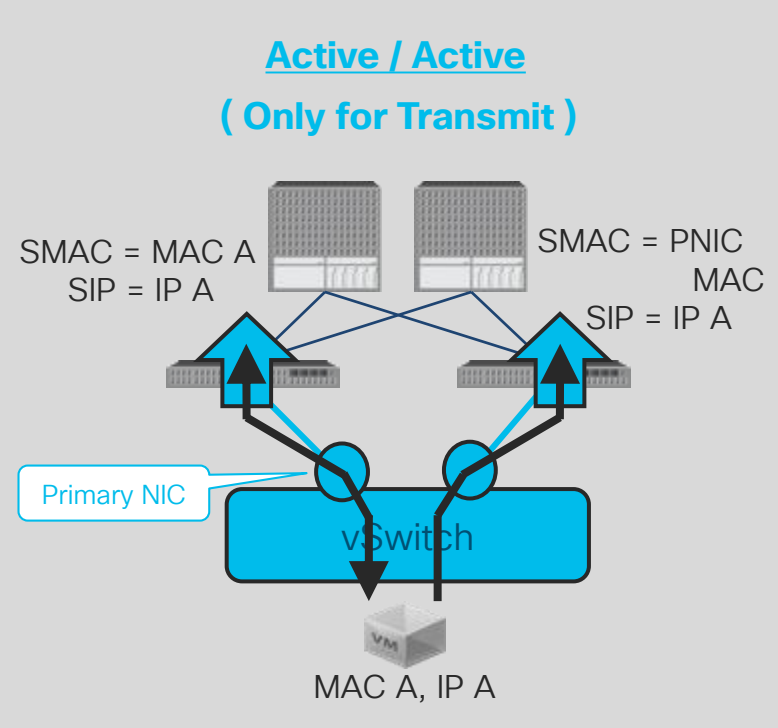
- How it works:
  - Similar to MAC Pinning (Route Based on Originating Port ID)
  - Re-pin every 30 sec based on the link utilization
- Reasons why not to use it:
  - MAC/IP flap between nodes/ports
  - 30 sec interval could be acceptable (not preferred just like any other switches)
- Solution:
  - LAG such as LACP or static port-channel



NIC Teaming Types to avoid

Microsoft – Switch Independent with Address Hash or Dynamic Load

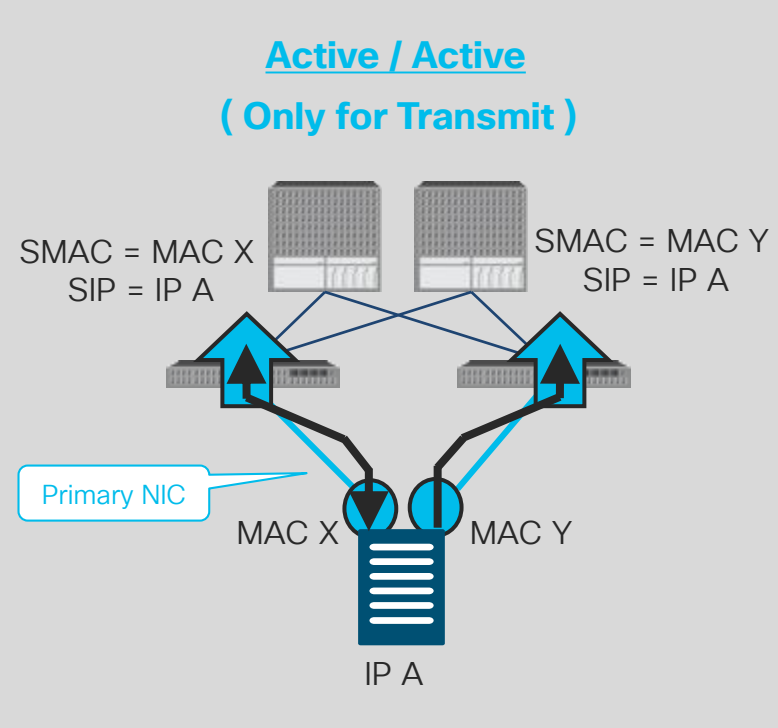
- How it works:
  - Outgoing traffic on the primary NIC
    sMAC is original MAC (could be primary NIC's MAC)
  - Outgoing traffic on non-primary NIC
    sMAC is each NIC's MAC
  - Return traffic is only on the primary NIC
    Controlling it via ARP (IP A -> MAC A)
- Reasons why not to use it:
  - IP flaps between MACs
  - No load balancing for return traffic
- Solution:
  - LAG such as LACP or static port-channel
  - Switch Independent with Hyper-V
  - Disable VRF IP data-plane learning on ACI



NIC Teaming Types to avoid

HP – Transmit Load Balancing Teaming

- How it works:
  - Outgoing traffic
    sMAC is each NIC's MAC
  - Return traffic is only on the primary NIC
    Controlling it via ARP (IP A -> MAC X)
- Reasons why not to use it:
  - IP flaps between MACs
  - No load balancing for return traffic
- Solution:
  - LAG such as LACP or static port-channel
  - Active / Standby (HP – Network Fault Tolerance, NFT)
  - Disable VRF IP data-plane learning on ACI



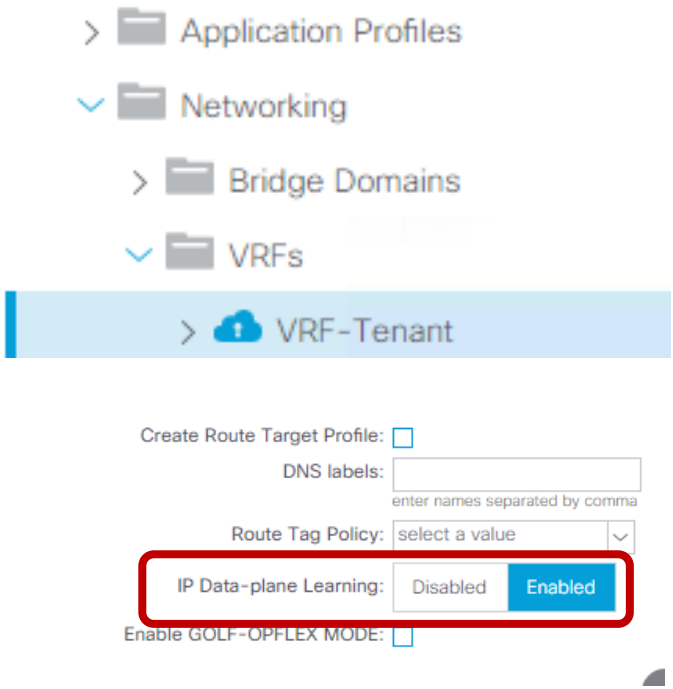
https://support.hpe.com/hpsc/doc/public/display?docId=c01984706

VRF IP Data-Plane Learning knob on ACI

From ACI 4.0

- Enabling VRF IP Data-Plane Learning (default)
  - ✓ Quick detection of endpoint move without using CPU resource
  - ✓ Traffic optimization with IP remote endpoint (less flood, etc.)
  - × IP flaps between MACs with some NIC teaming options

- Disabling VRF IP Data-Plane Learning
  - × Relying on CPU and ARP for IP learning (the same as traditional switch)
  - × Non optimal forwarding with no IP remote endpoint learning
  - ✓ No IP flaps between MACs with some NIC teaming options



https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-739989.html#_Toc18440064

Which Teaming Options to Use

	Server uses all pNICs for client-to-server traffic	Server uses all the pNICs for server-to-client traffic	One VM uses all pNICs for client-to-server traffic	One VM uses all pNICs for server-to-client traffic	ACI configuration required
Port-channeling/LACP	Yes	Yes	Yes	Yes	vPC
Active/Standby Teaming	No	No	No	No	Nothing special
MAC pinning	Yes	Yes	No	No	Nothing special
Hyper-V Wwitch Independent with Hyper-V port	Yes	Yes	No	No	Nothing special
Hyper-V Switch Independent with Dynamic Load	Yes	Yes	No	Yes	dataplane learning tuning
A/A TLB teaming	No	Yes	No	Yes	dataplane learning tuning
Vmware LBT Teaming	Yes	Yes	No	No	Don't configure too aggressive timers with Rogue EP detection

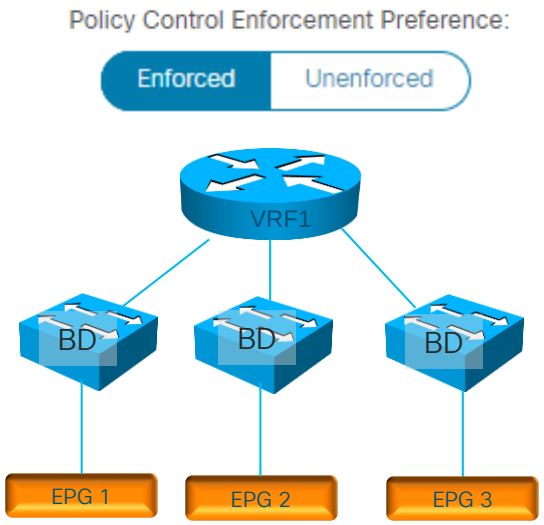
Agenda

- Single Fabric/Pod Topology Options
- ACI Design Considerations
  - ACI Fabric Bring Up
  - L2 Connectivity to Existing Networks
  - Loop prevention in ACI
  - Moving the default gateway to ACI
  - Connecting servers (Physical, VMM Integration, UCSM Integration)
  - Teaming Options
  - Allowing Traffic through ACI
  - Configuring L3 Connectivity to the Outside
- Application Centric Design
  - Network Centric & Application Centric
  - Contract Priority and Optimization

How to allow all traffic

Many start migration to ACI without enforcing security in ACI first

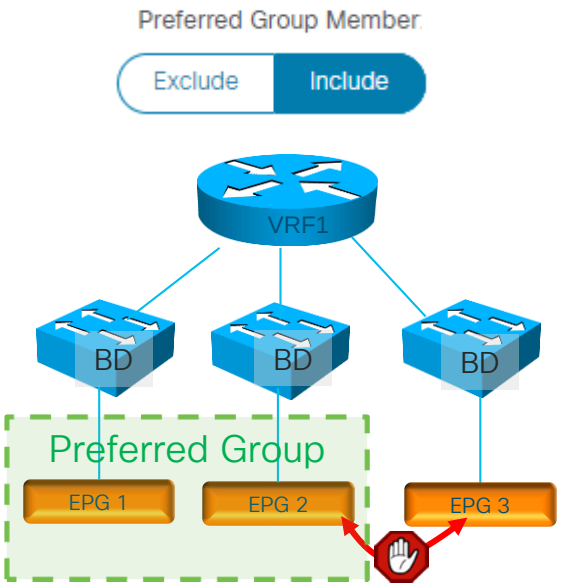
VRF unenforced



No contract at all.

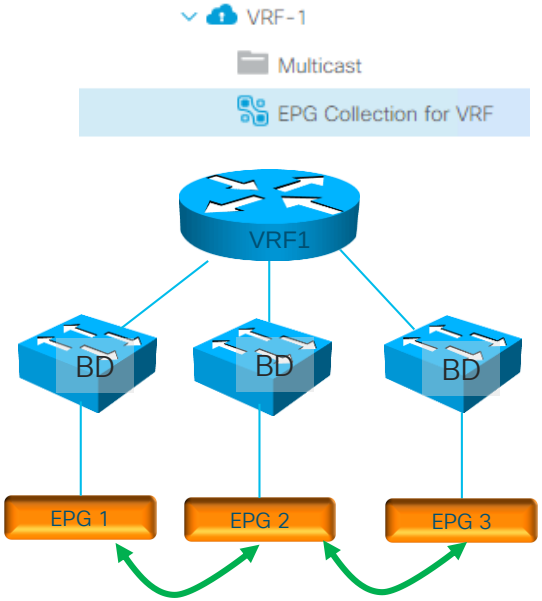
- EPG means nothing but VLAN(s)

Preferred Group



Implicit allow for all EPGs in the preferred Group (recommended)

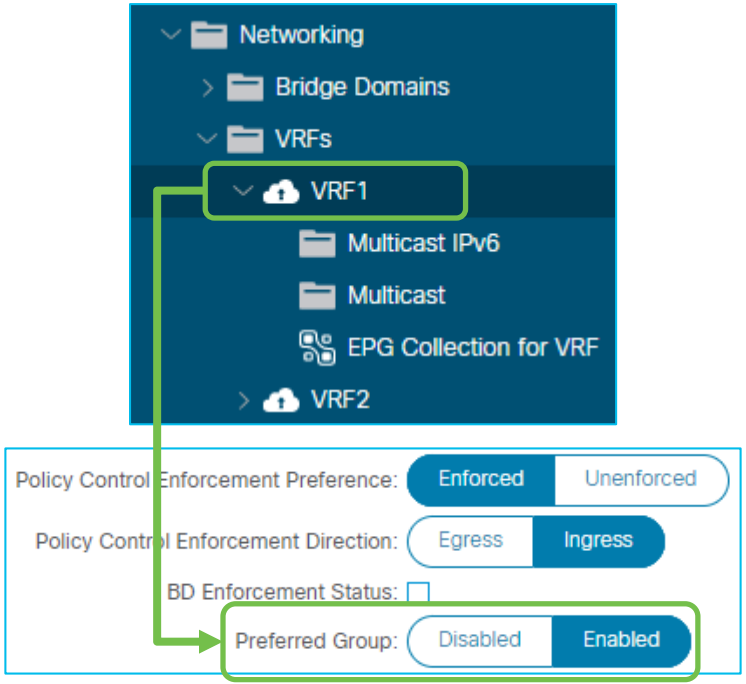
vzAny with permit



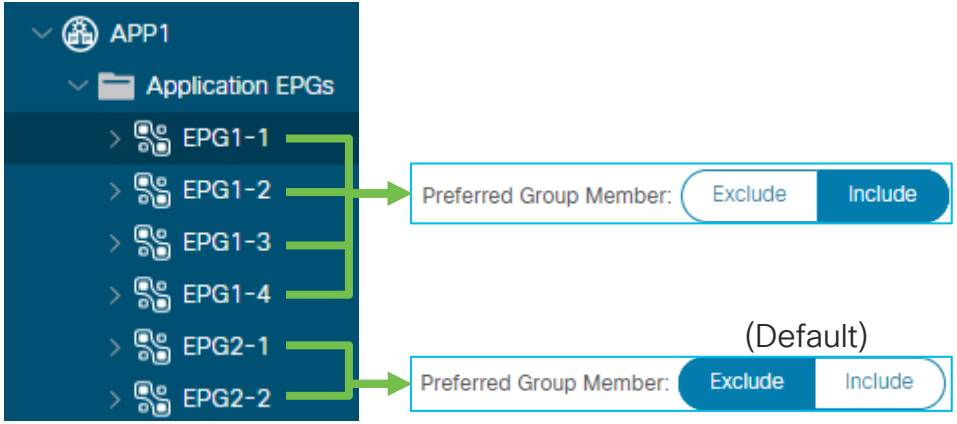
Manually creating permit any any

Preferred Group Configuration

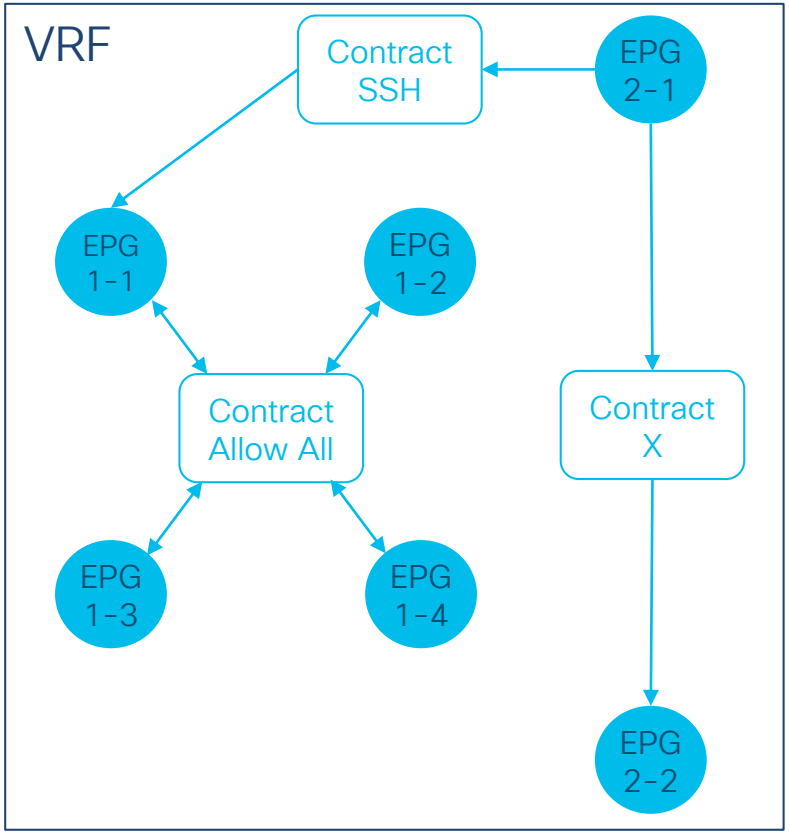
1. Enable Preferred Group in the VRF



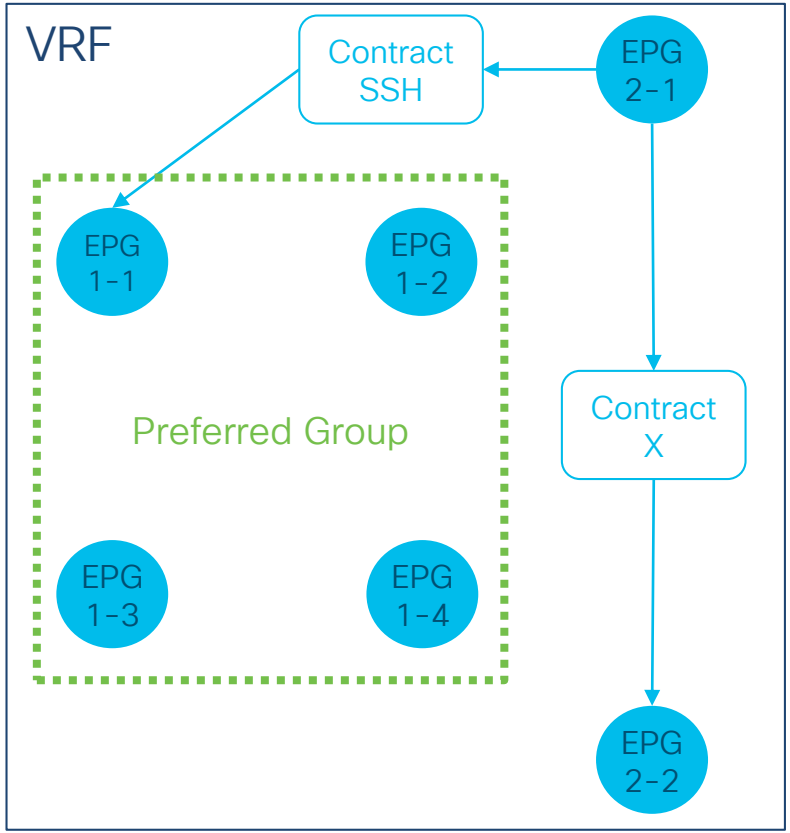
2. Include each EPG in the Preferred Group



Without Preferred Group

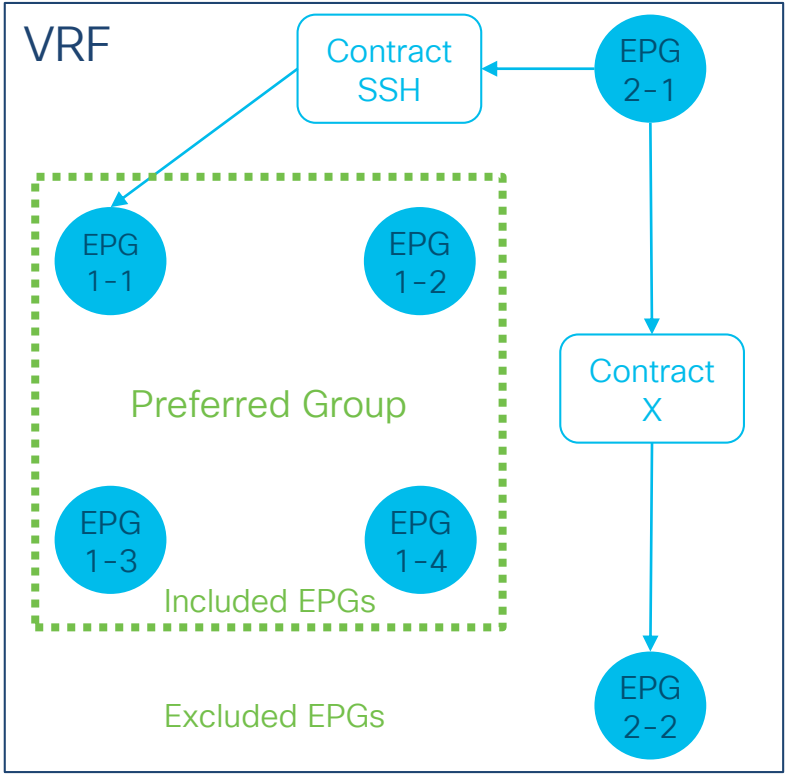


With Preferred Group



Preferred Group

How does the zoning-rule look like?



1. Your contracts (priority 7)
   Excluded EPG ←→ Excluded EPG
   Excluded EPG ←→ Included EPG

Source	Destination	Filter	Action
EPG-2-1	EPG-1-1	SSH	permit
EPG-1-1	EPG-2-1	SSH-r	permit
EPG-2-1	EPG-2-2	X	permit
EPG-2-2	EPG-2-1	X-r	permit

2. Implicit deny for all Excluded EPGs (priority 18)
   Excluded EPGs → Any
   Any → Excluded EPGs

Source	Destination	Filter	Action
EPG-2-1	any	implicit	deny
any	EPG-2-1	implicit	deny
EPG-2-2	any	implicit	deny
any	EPG-2-2	implicit	deny

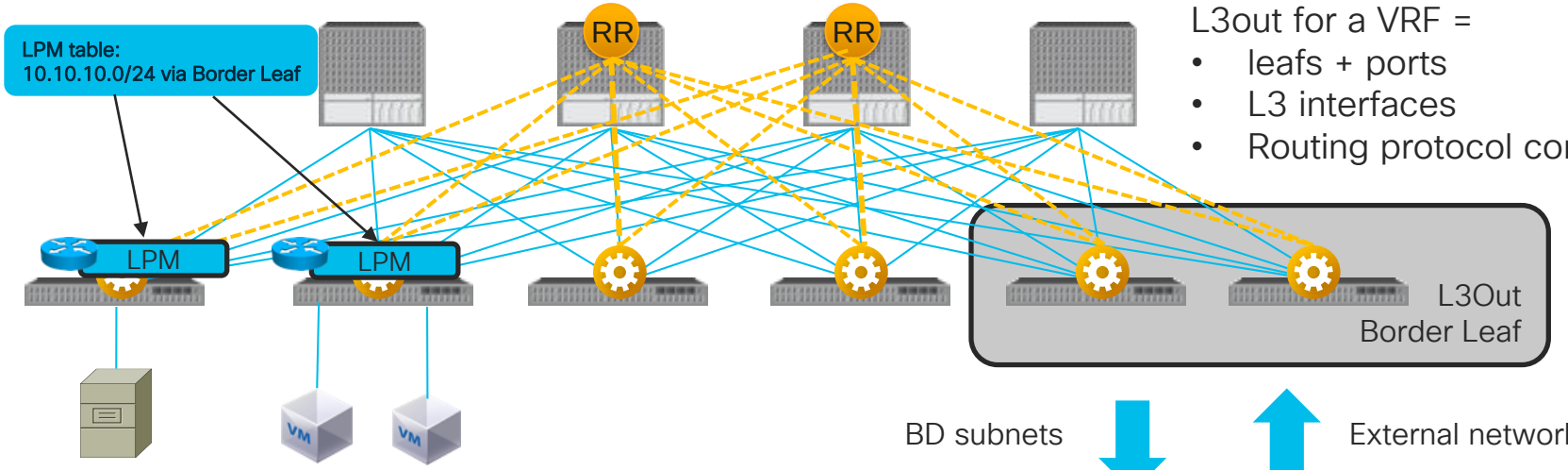
3. Implicit permit all Included EPGs (priority 20)

Source	Destination	Filter	Action
any	any	implicit	permit

Agenda

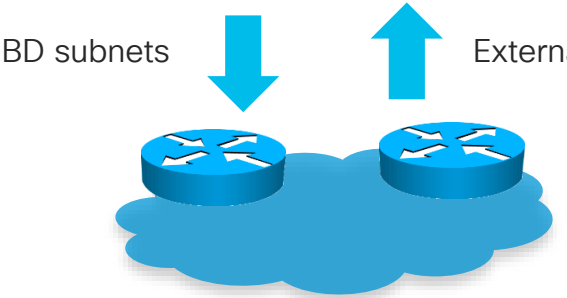
- Single Fabric/Pod Topology Options
- ACI Design Considerations
  - ACI Fabric Bring Up
  - L2 Connectivity to Existing Networks
  - Loop prevention in ACI
  - Moving the default gateway to ACI
  - Connecting servers (Physical, VMM Integration, UCSM Integration)
  - Teaming Options
  - Allowing Traffic through ACI
  - Configuring L3 Connectivity to the Outside
- Application Centric Design
  - Network Centric & Application Centric
  - Contract Priority and Optimization

If ACI is the default gateway, L3Out is required for outside reachability



L3out for a VRF =

- leafs + ports
- L3 interfaces
- Routing protocol configuration



E.g. 10.10.10.0/24

L3Out Interface Types

Supported L3Out Interface Types:

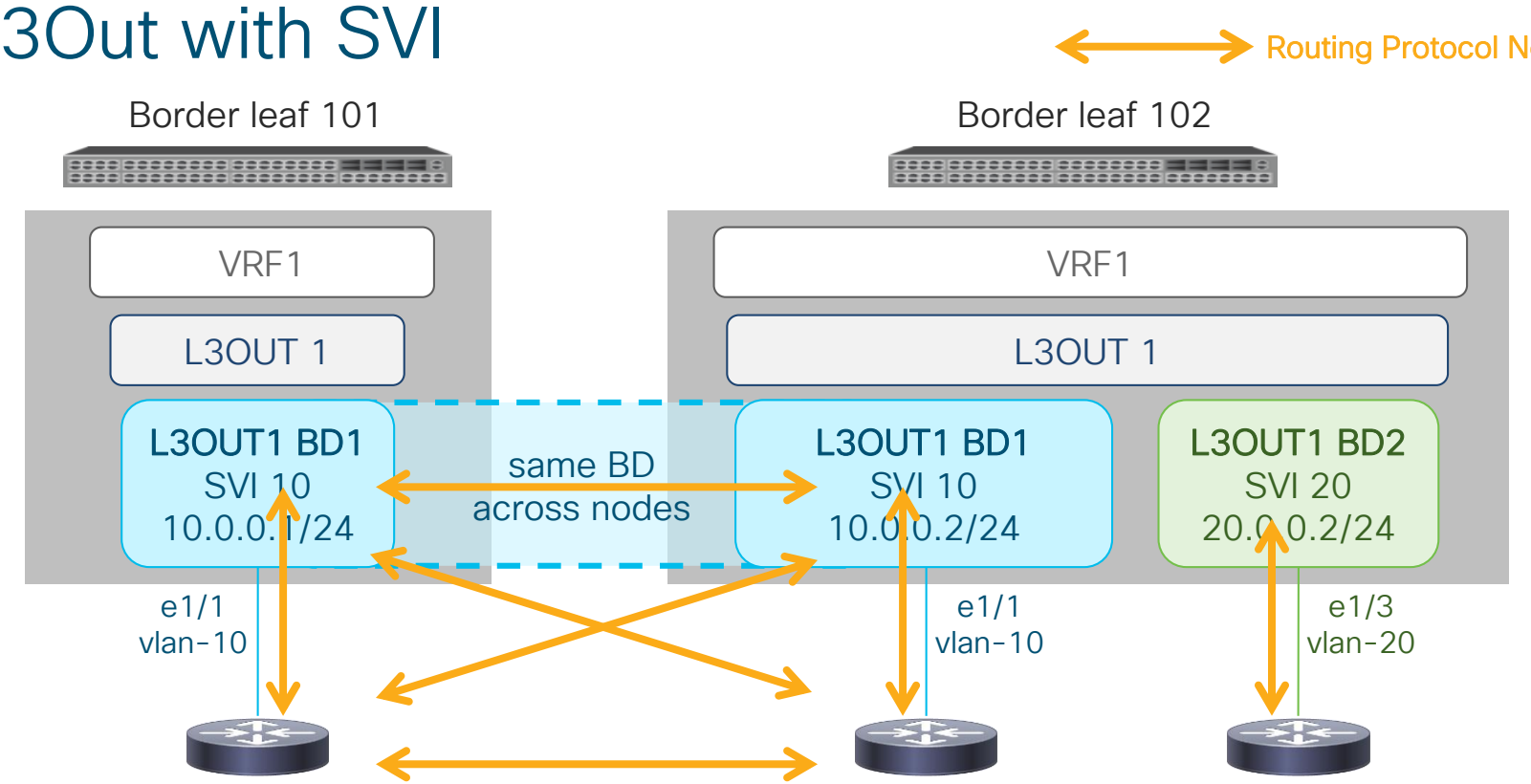
- Routed Interface
- Routed Sub-Interface
- SVI

Rule of thumb

Use Routed (Sub-)Interface if possible
SVI is typically required only for FW, LB, etc.

Why? fast routing, next-hop convergence/detection is 101 for routing, right?

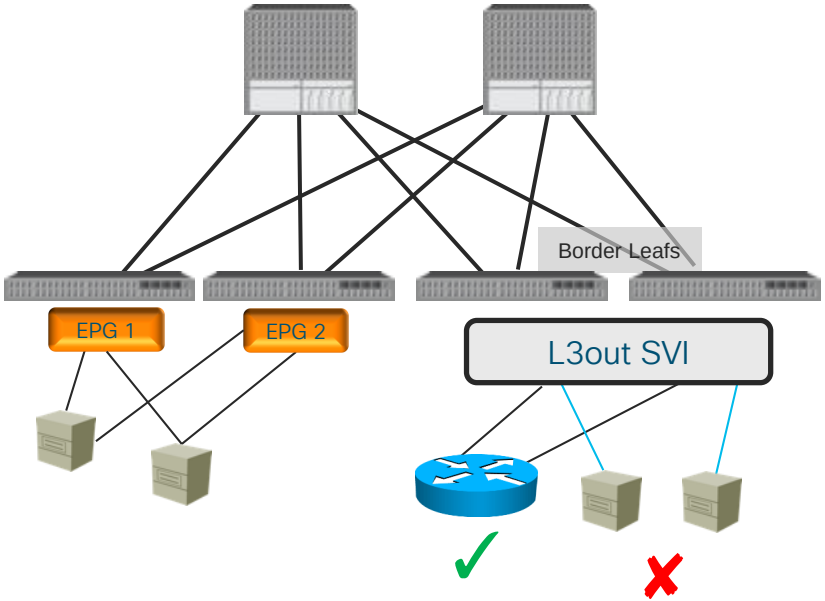
L3Out with SVI



L3Out exception:
The same access VLAN means the same broadcast domain

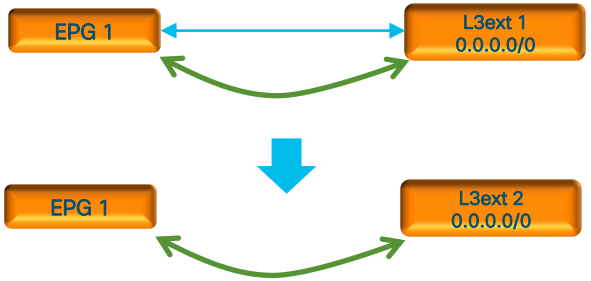
The L3Out is for routing devices (routers, FW, etc.)

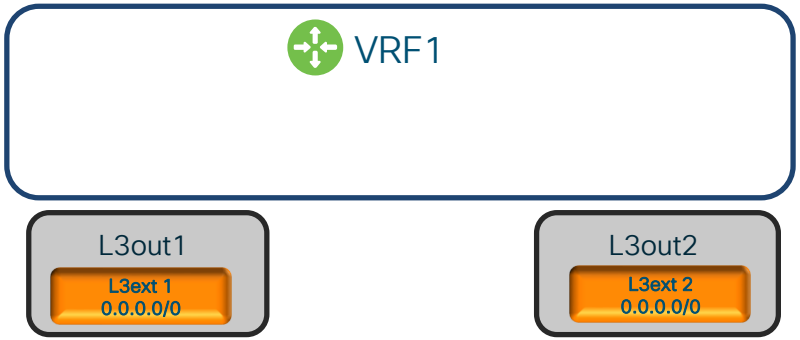
- The L3out is meant to attach Routing devices
- It is not meant to attach servers directly on the SVI of a L3Out
- Servers should be attached to EPGs and Bridge Domains
- There are multiple reasons for this:
  - The L2 domain created by a L3out with SVIs is not equivalent to a regular Bridge Domain
  - The L3ext classification is designed for hosts multiple hops away (more on this later)



Avoid using 0.0.0.0/0 as a L3external if possible

- Classifying traffic from the L3Out with a L3Out subnet 0.0.0.0/0 is possible but it can lead to misconfigurations
- If two L3exts in the same VRF have a 0.0.0.0/0
- If EPG1 is allowed to talk to L3ext1 => EPG1 can also talk to L3ext2

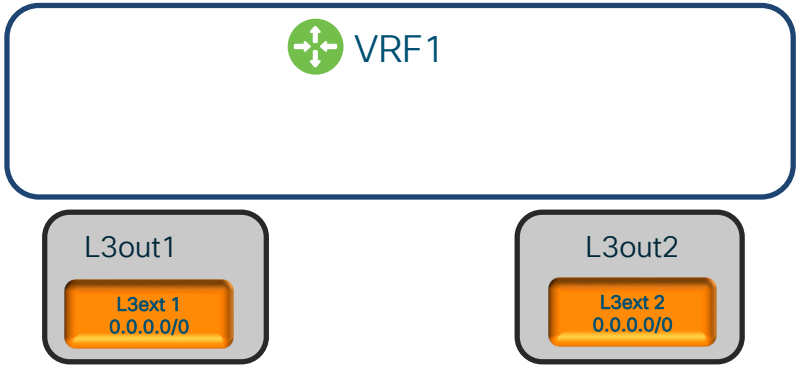




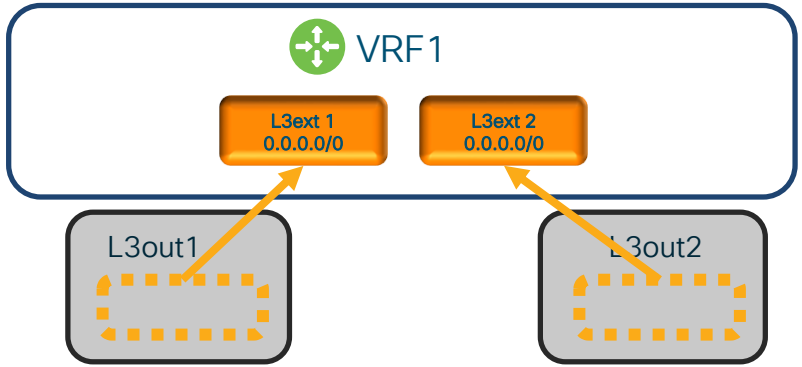
Note – This is specifically for L3Out subnet with “External Subnets with the External EPG” scope

L3Out subnet design

How you configure



How you should imagine



Rule of thumb

L3Out subnet is per VRF

L3Out Interface subnet as L3Out Subnets?

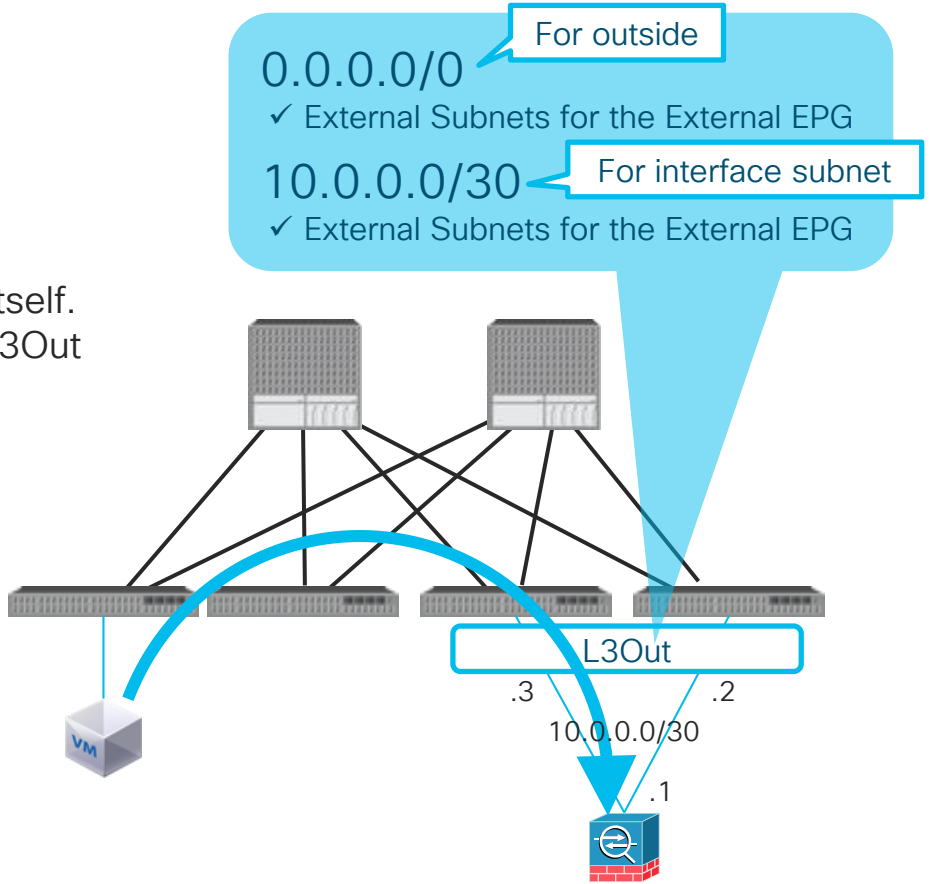
L3Out for L4-L7 devices (FW, LB)

Typically traffic is destined to the IP of the FW or LB itself.

- L3Out subnet should include such subnet (L3Out interface subnet)

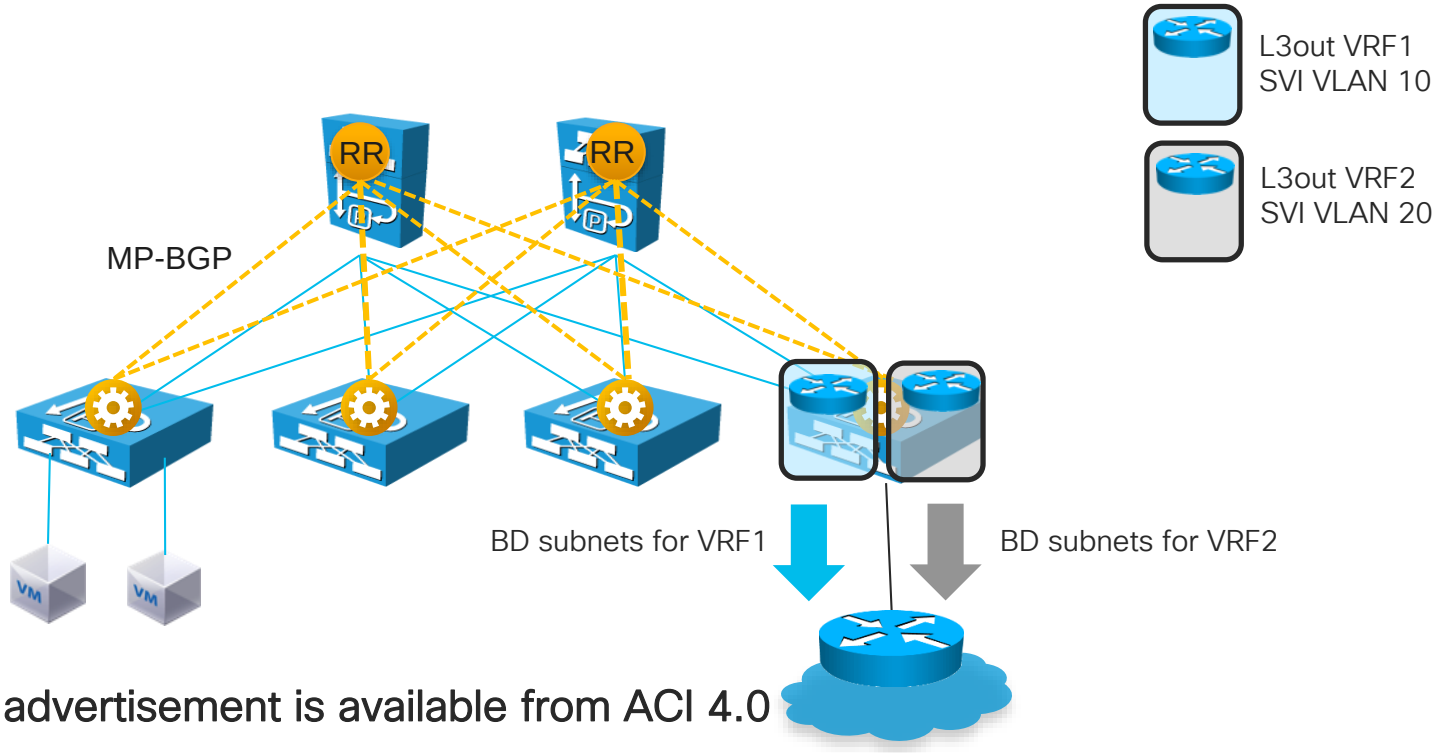
Why?

- By default, L3Out Interface Subnet uses a special pcTag (class ID) 1.
  This may lead an unintended result.
  (See CSCuz12913 for details if needed)



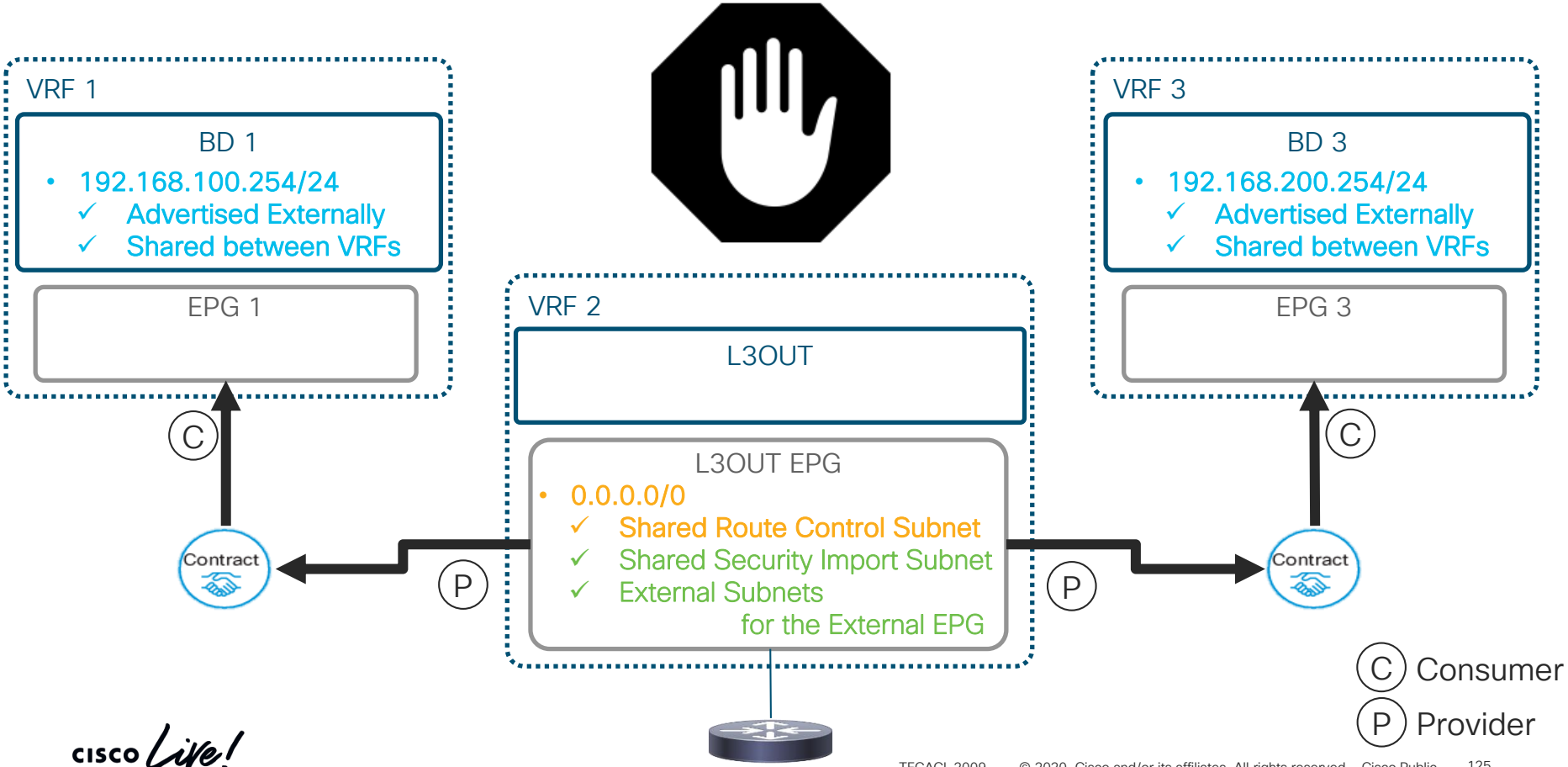
Multi Tenancy Design and L3Out:

One L3Out per VRF = VRF-Lite



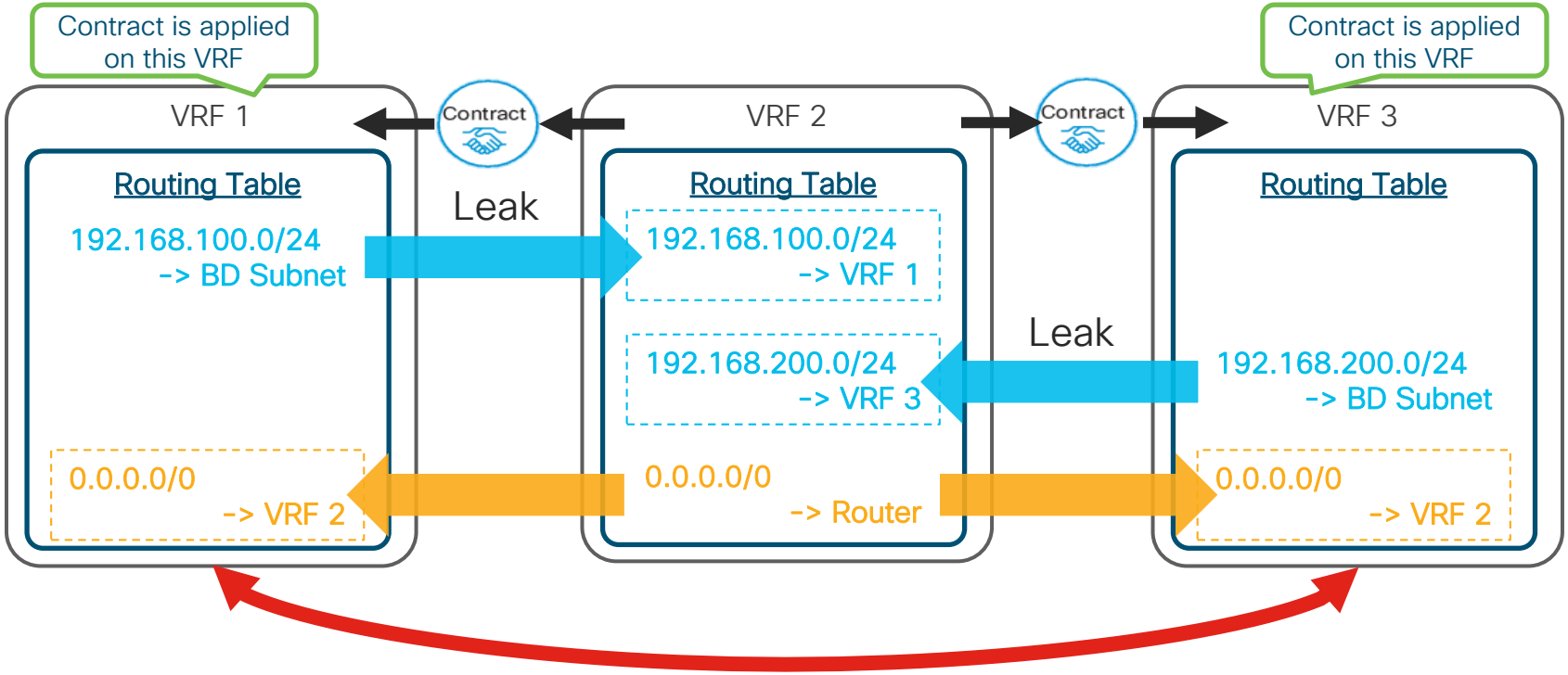
Host route advertisement is available from ACI 4.0

Multi Tenancy with Shared L3Out



Multi Tenancy with Shared L3Out

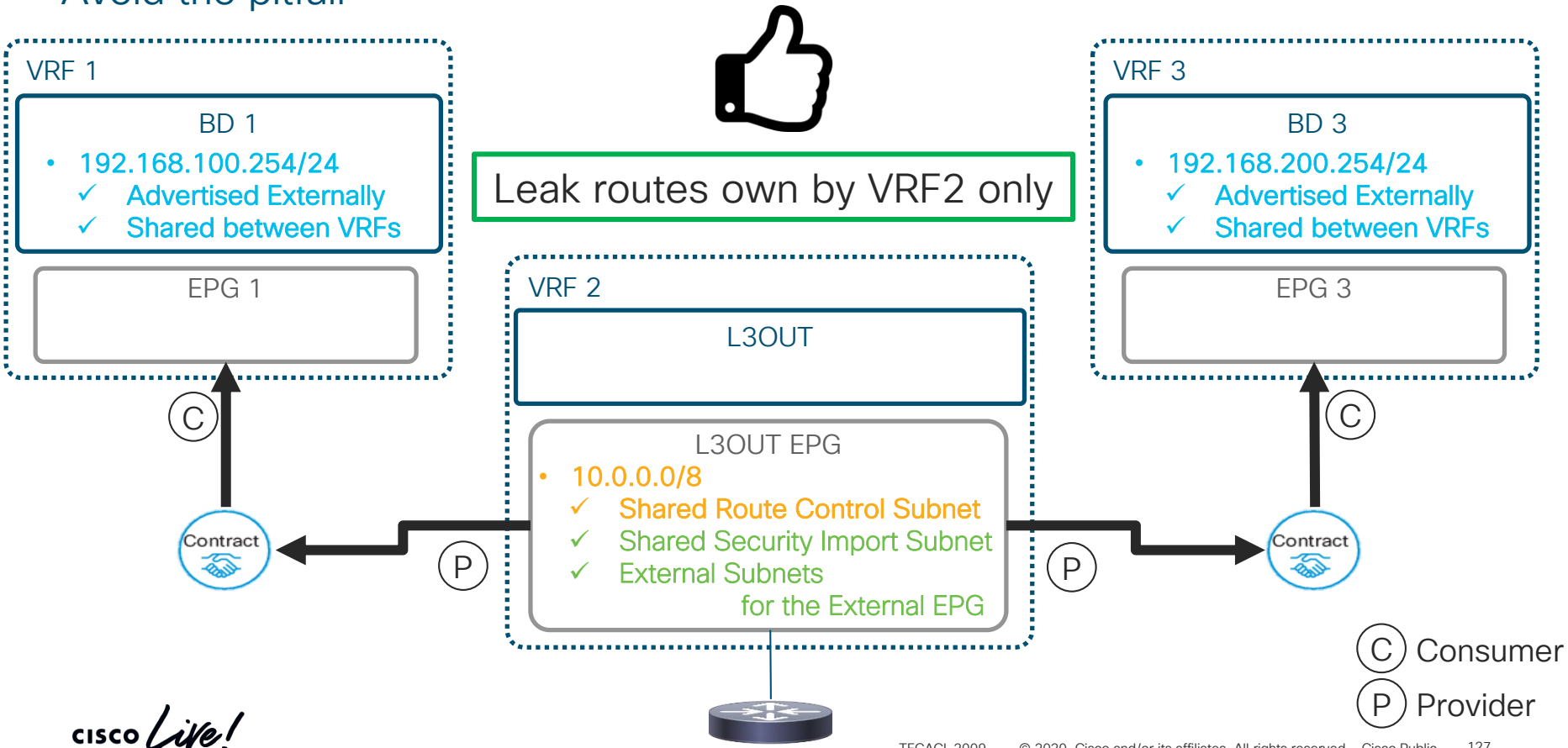
And its pitfall



VRF 1 and VRF 3 can talk to each other

Multi Tenancy with Shared L3Out

Avoid the pitfall



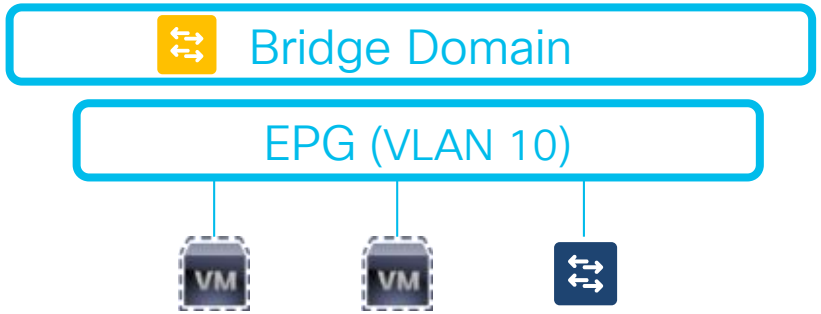
Agenda

- Single Fabric/Pod Topology Options
- ACI Design Considerations
  - ACI Fabric Bring Up
  - L2 Connectivity to Existing Networks
  - Loop prevention in ACI
  - Moving the default gateway to ACI
  - Connecting servers (Physical, VMM Integration, UCSM Integration)
  - Teaming Options
  - Allowing Traffic through ACI
  - Configuring L3 Connectivity to the Outside
- Application Centric Design
  - Network Centric & Application Centric
  - Contract Priority and Optimization

Network Centric & Application Centric

Network Centric

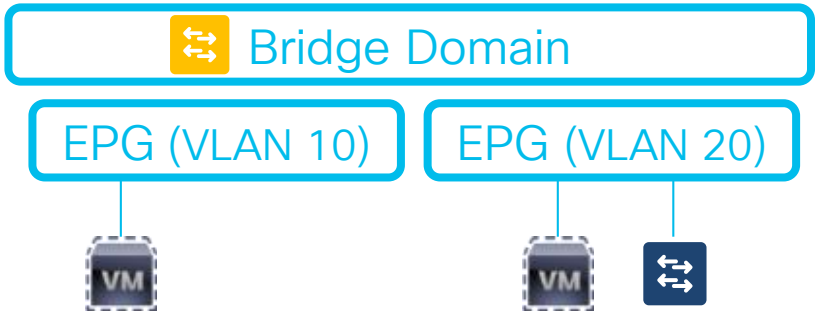
- 1 VLAN = 1 EPG = 1 BD



- Similar to traditional network
  - VLAN as a broadcast domain
  - Easy to connect to legacy networks
- Typically simple or no contracts (no ACLs)

Application Centric

- Multiple EPGs per BD



- Multiple security domains (EPGs) in one broadcast domain
- Flexible network and security design

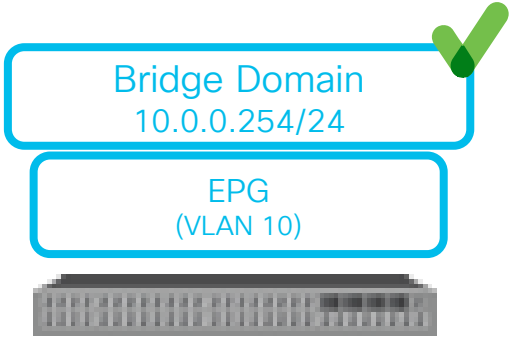
Now is the time for details

Network Centric & Application Centric

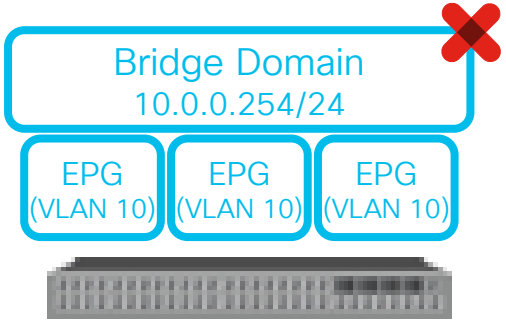
Could I re-use the same VLAN with Application Centric Design?

✔ Supported

✘ Not Supported



Need granular security groups



Not Supported

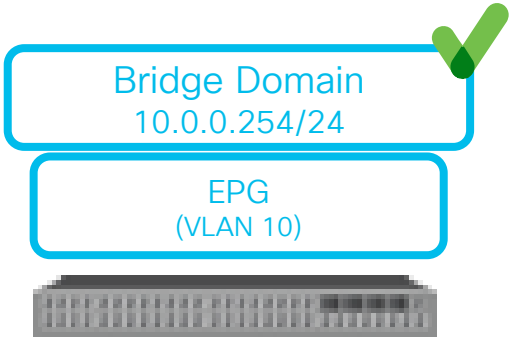
The same VLAN can be used for only one EPG in the same leaf

Network Centric & Application Centric Design

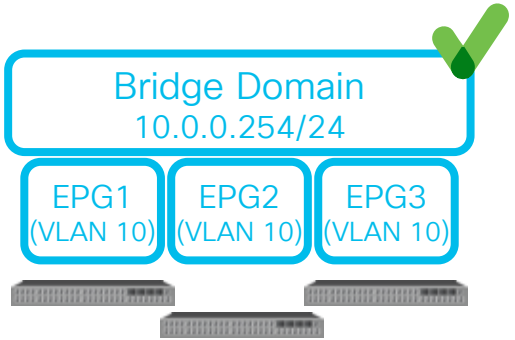
Could I re-use the same VLAN with Application Centric Design? (cont.)

✔ Supported

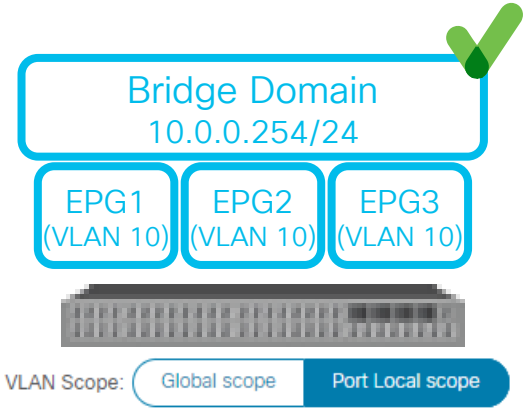
✘ Not Supported



Need granular security groups



The same VLAN can be re-used for any EPGs on **different** leaf.
EPG (security group) deployment is physically limited.

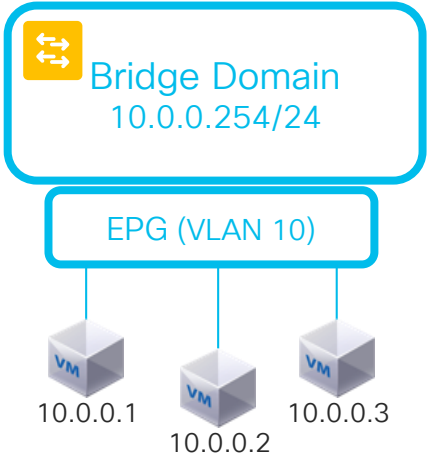


Local VLAN scope allows the same VLAN to be re-used on the same leaf.
Lower scalability
Complex Domain, VLAN Pool design

Network Centric & Application Centric Design

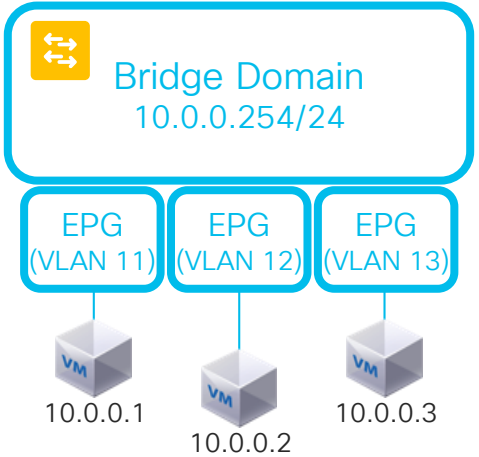
Network Centric

A security group in 1 subnet



More granular security?

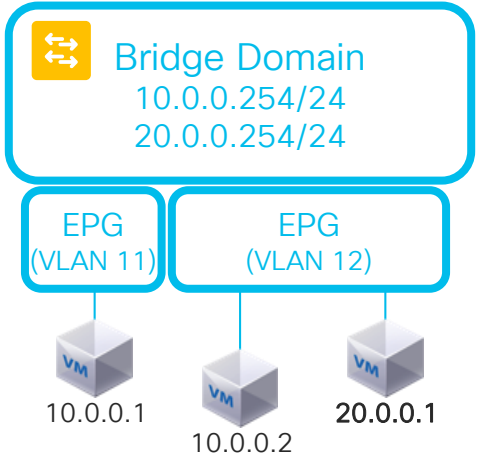
Multiple security groups in 1 subnet



What if multiple subnets need to share the same security rules?

Application Centric

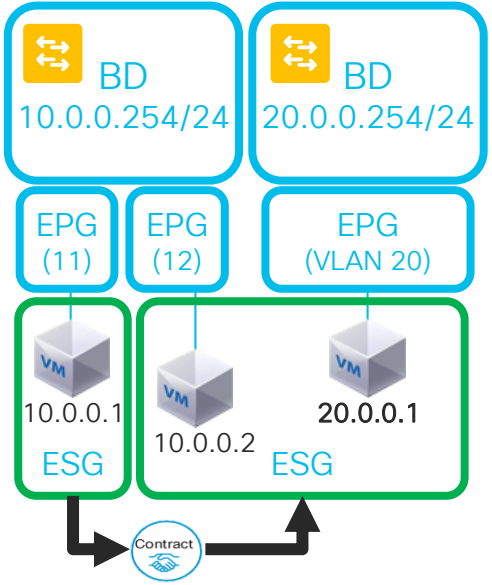
Security groups across subnets



Sharing a broadcast domain leads another security concern

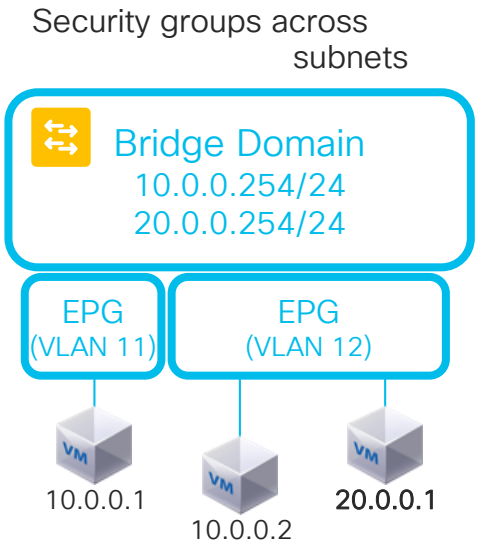
== Future ==
Endpoint Security Group (ESG)

Security groups across broadcast domains



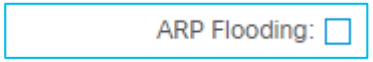
Network Centric & Application Centric Design

Considerations for multiple subnets in one BD



- Spine-Proxy to minimize the impact of flood

Be aware of the risk for L2 communication drop.



Silent Host Detection (ARP glean) can handle unknown targets. Unicast Routing and BD subnet are required.

OR

- Flood in Encap to minimize the impact of flood

From 3.1 and on 2nd (or newer) leaf, all traffic is flooded only within each encap VLAN (not EPG).

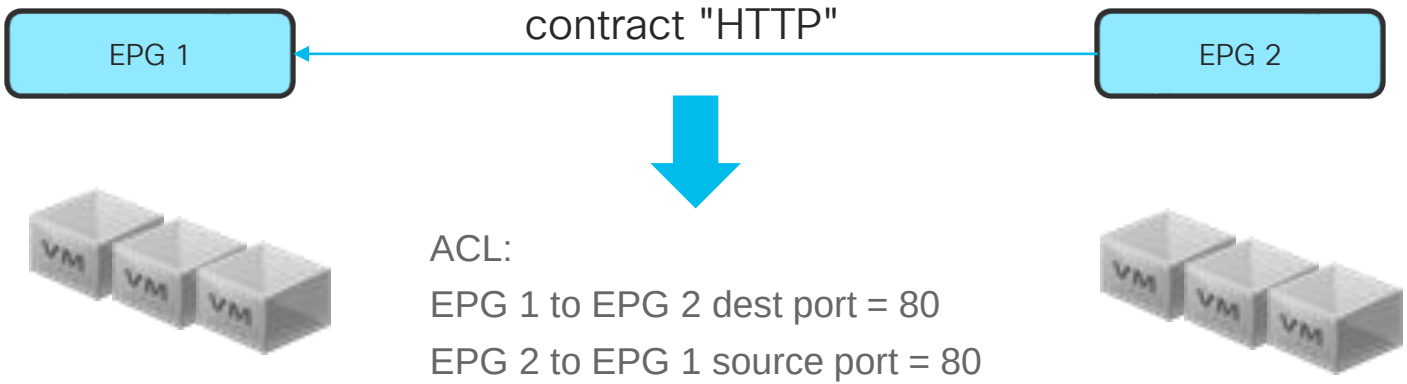
Agenda

- Single Fabric/Pod Topology Options
- ACI Design Considerations
  - ACI Fabric Bring Up
  - L2 Connectivity to Existing Networks
  - Loop prevention in ACI
  - Moving the default gateway to ACI
  - Connecting servers (Physical, VMM Integration, UCSM Integration)
  - Teaming Options
  - Allowing Traffic through ACI
  - Configuring L3 Connectivity to the Outside
- Application Centric Design
  - Network Centric & Application Centric
  - Contract Priority and Optimization

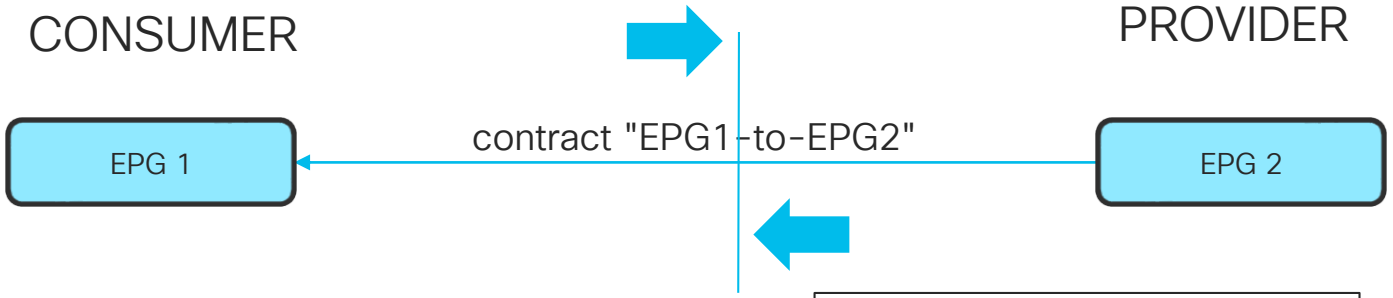
Communication between EPGs is configured via "contracts"

EPG 1 consumes "HTTP"

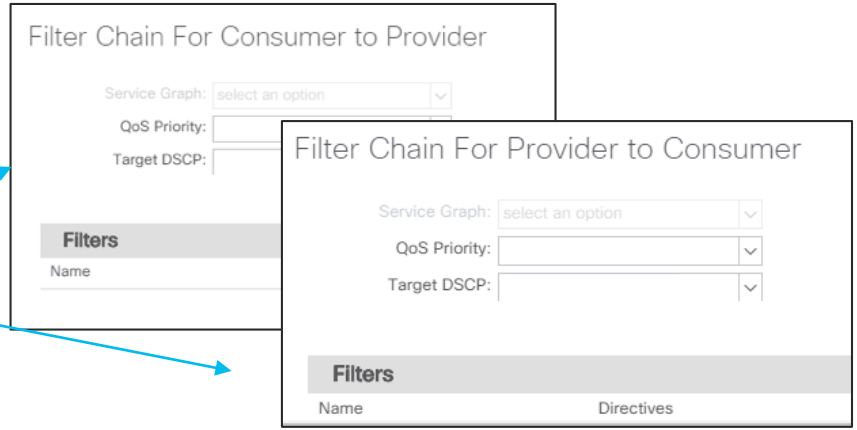
EPG 2 provides "HTTP"



You can then add subjects to the contract and define the filters for each direction

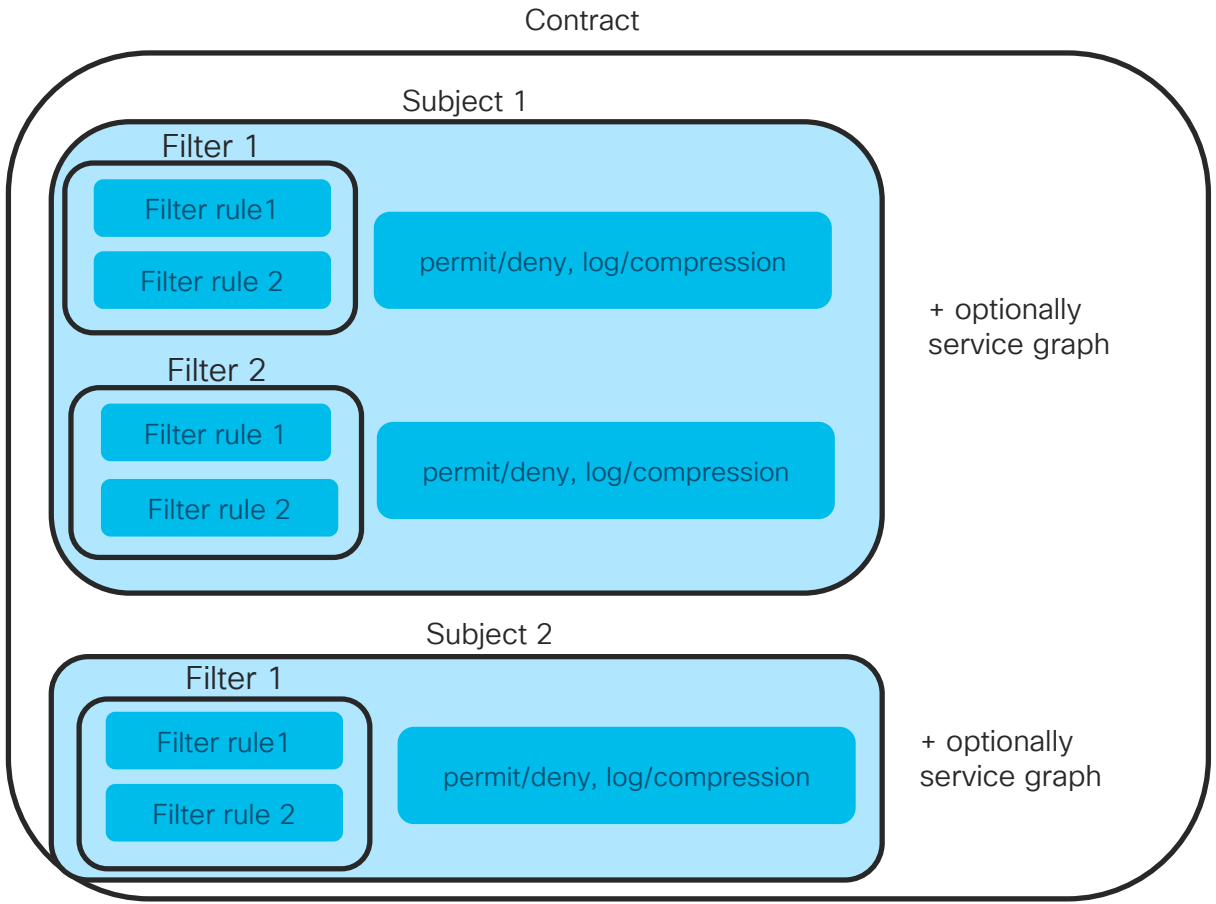


subject1: HTTP permit
subject2: Various other filters
subject3: etc..

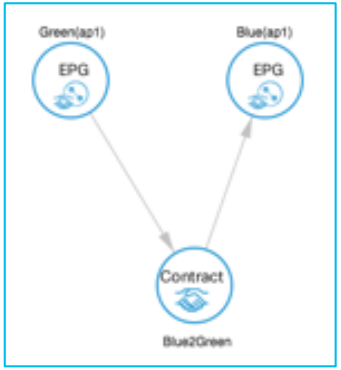


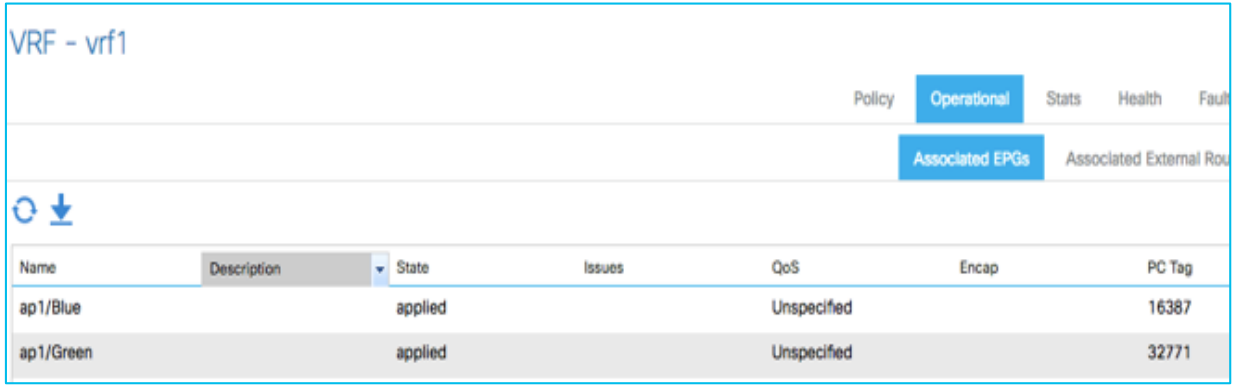
A Contract defines a set of filter rules in each "direction"

- Contract subjects contain a list of filters
- Each Filter consits of multiple filter rules
- Filter rules match protocols, src and dst ports
- Filters are attached to subjects as permit or deny rules
- Filters are attached to subjects with a directive (log, none)



The contract filters are programmed in the Policy Cam on the Leaf





```
leaf1# show zoning-rule scope 2162697 | egrep -E "Scope|32771|16387"
Rule ID  SrcEPG   DstEPG   FilterID operSt   Scope    Action   Priority
4616     16387    32771    5        enabled  2162697  permit   src_dst_any(8)
4617     32771    16387    5        enabled  2162697  permit   src_dst_any(8)
```

Contract Priority 101

When a packet matches multiple contracts:

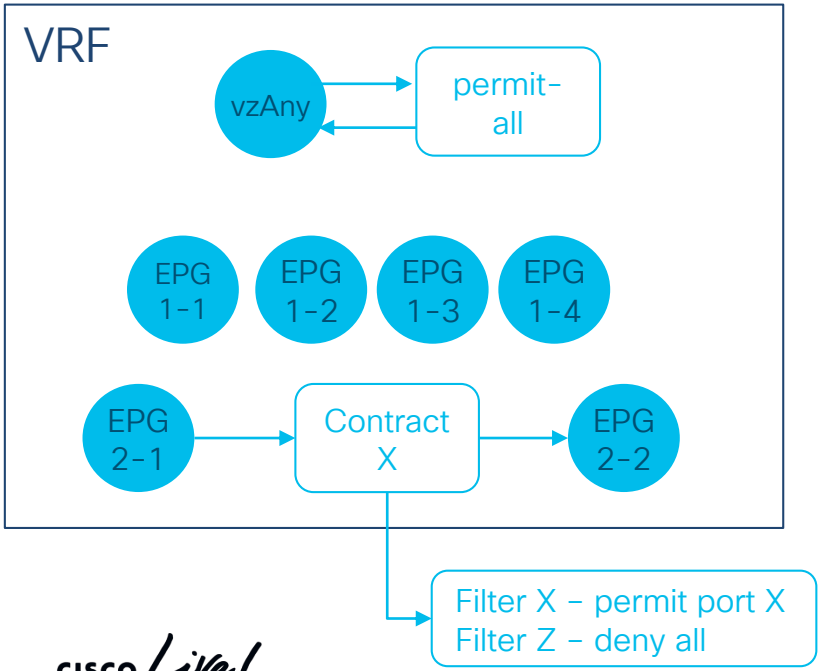
1. Specific EPG > vzAny
2. Specific match parameter wins
3. Deny > Redirect > Permit

Note - “Redirect” is used for Service Graph PBR

Contract Priority 101

Requirements:

1. Allow EPG2-1 to EPG2-2 on port X
2. All other communications are allowed



Src	Dst	Filter	Action
EPG2-1	EPG2-2	X	permit
EPG2-1	EPG2-2	all	deny
any	any	all	permit
any	any	implicit	deny

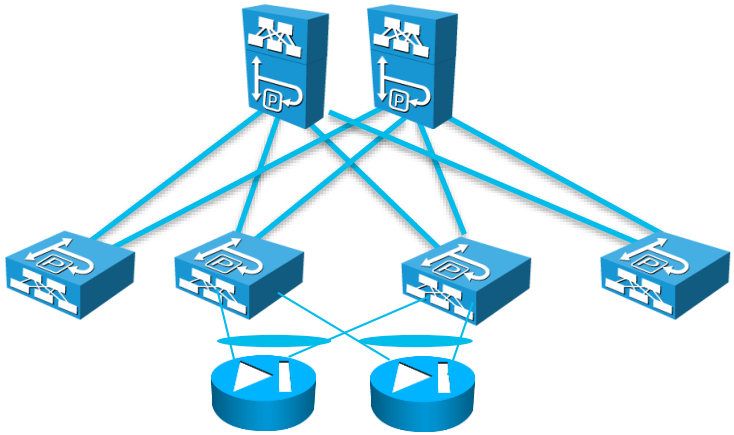
Higher Priority

Redirect all traffic to a FW

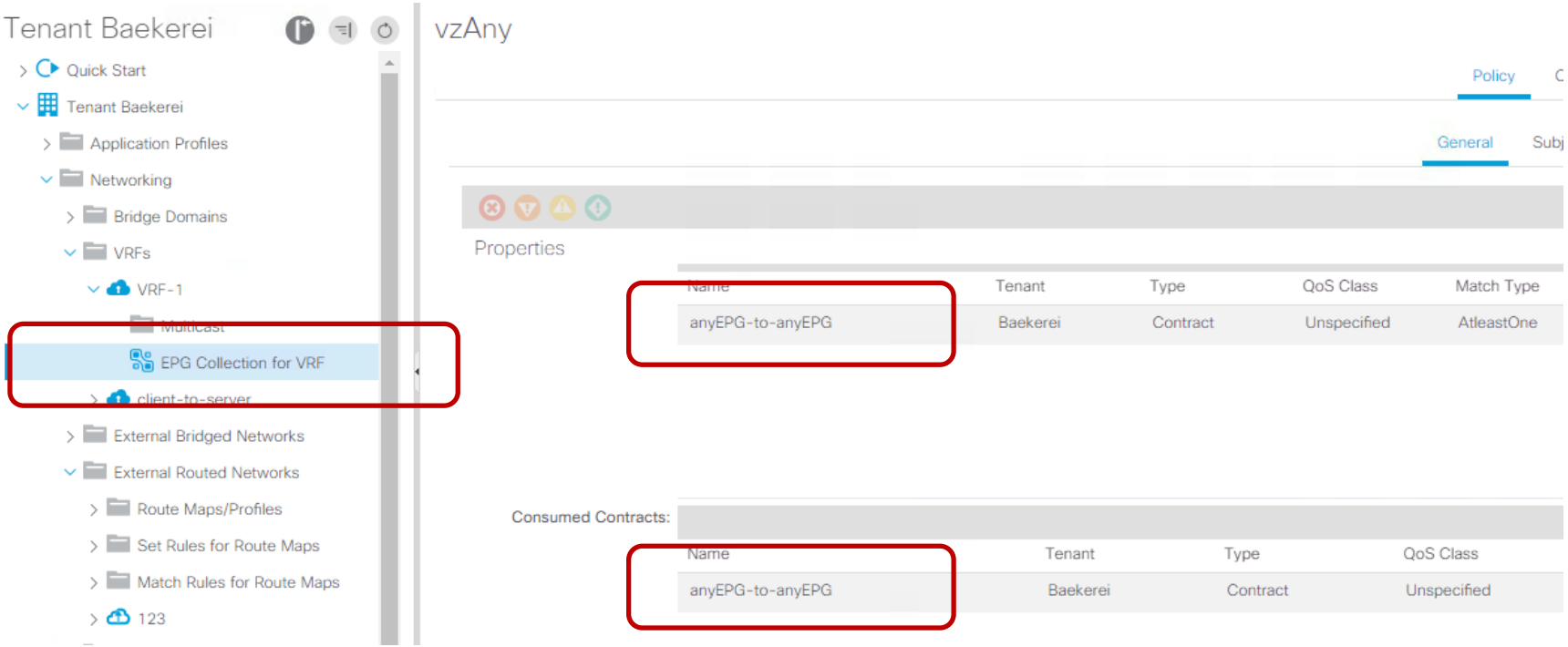
With Service Graph PBR and vzAny

1. Create a PBR Service Graph that points to the FW
2. Create a permit all contract
3. Apply the PBR Service Graph to the permit all contract
4. vzAny provides and consume the permit all contract

Note - The servers' default gateway needs to the BD



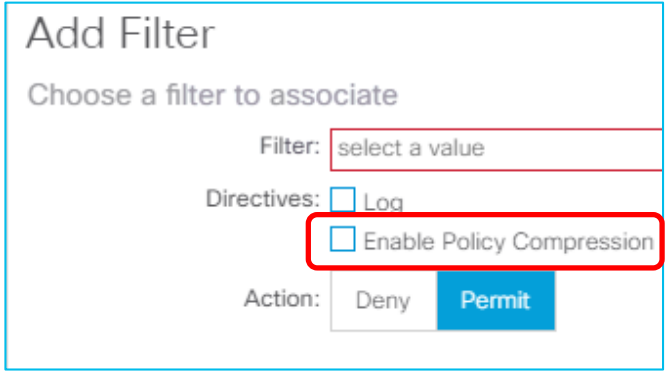
The contract is provided and consumed under vzAny



Contract Optimization

- Contracts (policy TCAM) Usage becomes a bottleneck easier in Application Centric Design

- There are many optimizations (depending on the leaf hw):
  - Change the ASIC profile to "policy TCAM" intensive profiles
  - Use L4 port range operations. It uses one entry only in TCAM.
  - Enable Policy Compression (at the cost of granularity of statistics)
    - Bidirectional rules are combined into one (from 3.2).
    - Rules used by the same EPG are combined into one. This is called indirection (from 4.0).



Design Whitepapers

- https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-737592.pdf
- https://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/sw/2-x/multicast/b_Using_Layer3_Multicast/b_Using_Layer3_Multicast_chapter_00.html

For More Reading

- Design Guide: https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-737909.pdf
- About endpoint learning and BD settings: https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-739989.html
- About the L3Out:
- https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/guide-c07-743150.html
- About Migration: https://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/sw/migration_guides/migrating_existing_networks_to_aci.html
  - Getting Started Guide: https://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/sw/4-x/getting-started/Cisco-APIC-Getting-Started-Guide-401.html
  - Step by Step ACI deployment: https://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/white_papers/Cisco-ACI-Initial-Deployment-Cookbook.html
  - Virtualization Guide: https://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/sw/4-x/virtualization/Cisco-ACI-Virtualization-Guide-411/Cisco-ACI-Virtualization-Guide-411_chapter_010.html
  - Verified Scalability Guide for ACI 4.1: https://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/sw/4-x/verified-scalability/Cisco-ACI-Verified-Scalability-Guide-411.html

ACI Anywhere, Extending the ACI Fabric

Agenda

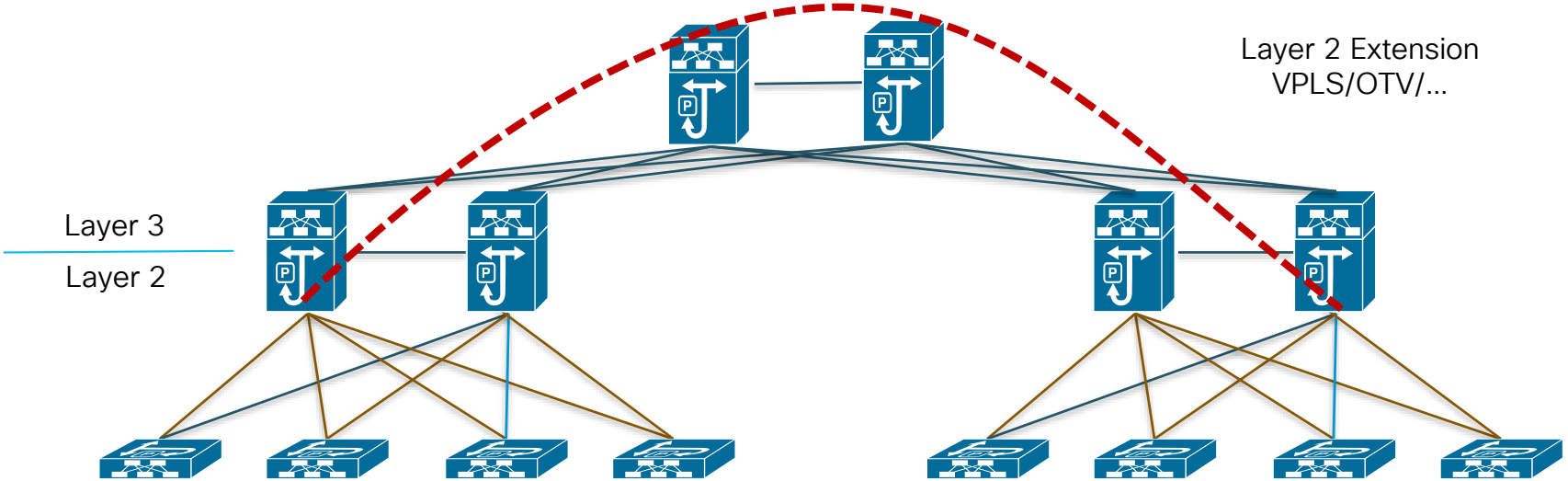
- ACI Anywhere, Extending the ACI Fabric
  - Overall Design Principles (AZs and Regions)
  - Mapping use cases to the proper solutions
    - Active/Active DC → Multi-Pod
    - Disaster Recovery → Multi-Site
    - Migration/Coexistence with Legacy DC Networks and 'Disaggregated DCs' Model → Physical Remote Leaf
    - Baremetal Cloud Integration → Virtual Pod (vPod)
  - Extending ACI to the Cloud
  - Connecting the users to the Multi-Cloud DC
    - ACI and SDA Integration
    - ACI and SDWAN Integration

Overall Design Principles

Systems Availability

Best of Breed prior to 2014

- Distinct Network Domains for availability
  - Extension of Layer 2 works but complicates change and fault isolation
  - Sizing of each domain is a balance between need, risk and cost



Network Fault/Change/Reachability Domain

Network Fault/Change/Reachability Domain

Data Center Interconnect Solutions

A Tale from the Past

Yesterday

Ethernet

Over dark fiber or protected D-WDM

- *vPC double-sided (caution with DWDM SLA)*
  - Dual site interconnection
- *OTV or VXLAN EVPN*
  - Dual/Multiple sites interconnection

MPLS

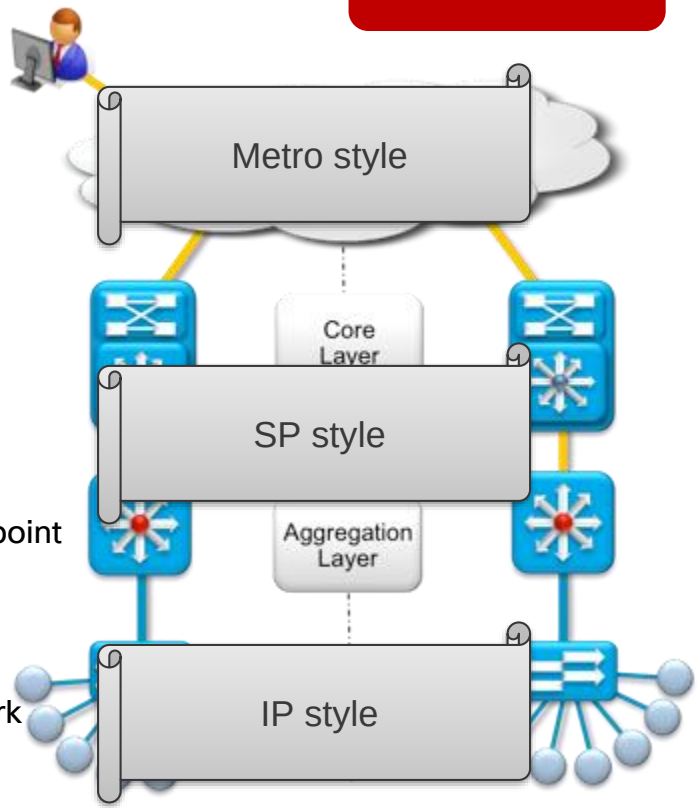
MPLS Transport

- *MPLS-EVPN*
  - SP, Point to Multipoint
- *PBB-EVPN*
  - Large scale & Multi-tenants, Point to Multipoint

IP

IP Transport

- *OTV*
  - Interconnect Traditional-based DC Network
- *VXLAN EVPN*
  - interconnect VXLAN-based Fabric
  - Layer 2 Ext. only and/or Layer 3 Ext. (multitenancy)
- *LISP*
  - For Subnet extension and Path Optimization



Does ACI Change Anything?

YES

Reachability is now Decoupled from Fault and Change

Reachability is Decoupled from Topology as well

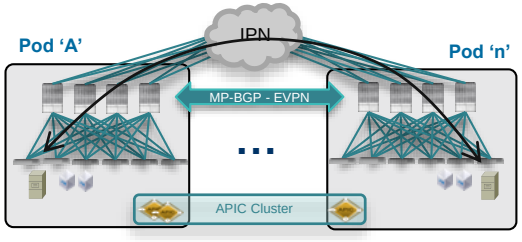
Data Center Interconnect Solutions

ACI Simplifies the Deployment of DCI

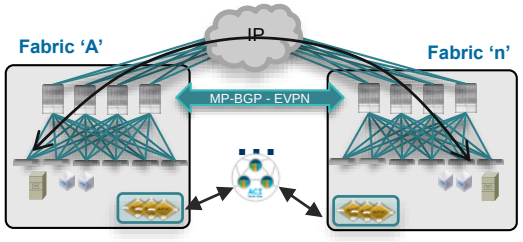
Today

- Common Control/Data Plane options used across different architectures
- Consistent security policies end-to-end

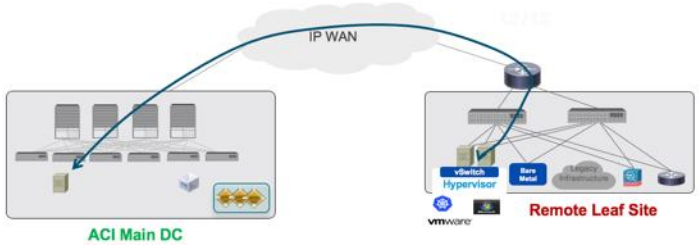
ACI Multi-Pod Fabric



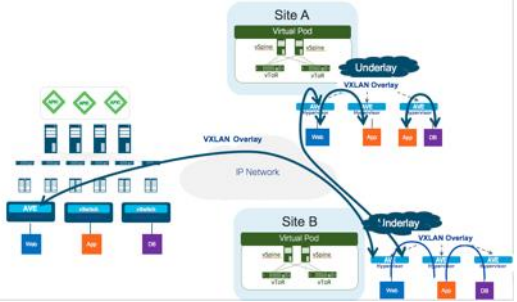
ACI Multi-Site



ACI Physical Remote Leaf



ACI Virtual Remote Leaf (vPod)



Multi-Pod or Multi-Site?

That is the question…

And the answer is...

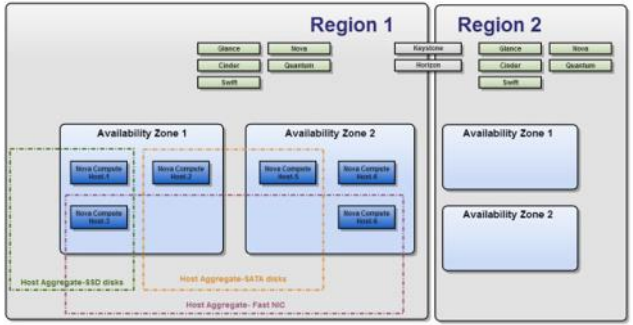
BOTH!

Framework for Multi-Cloud High Availability Design

Regions and Availability Zones

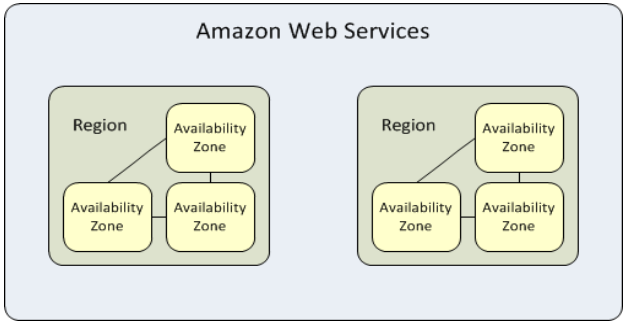
- Regions – Each Region has its own full OpenStack deployment, including its own API endpoints, networks and compute resources
- Availability Zones – Inside a Region, compute nodes can be logically grouped into Availability Zones, when launchir new VM instance, we can specify AZ or even a specific node in a AZ to run the VM instance

OpenStack



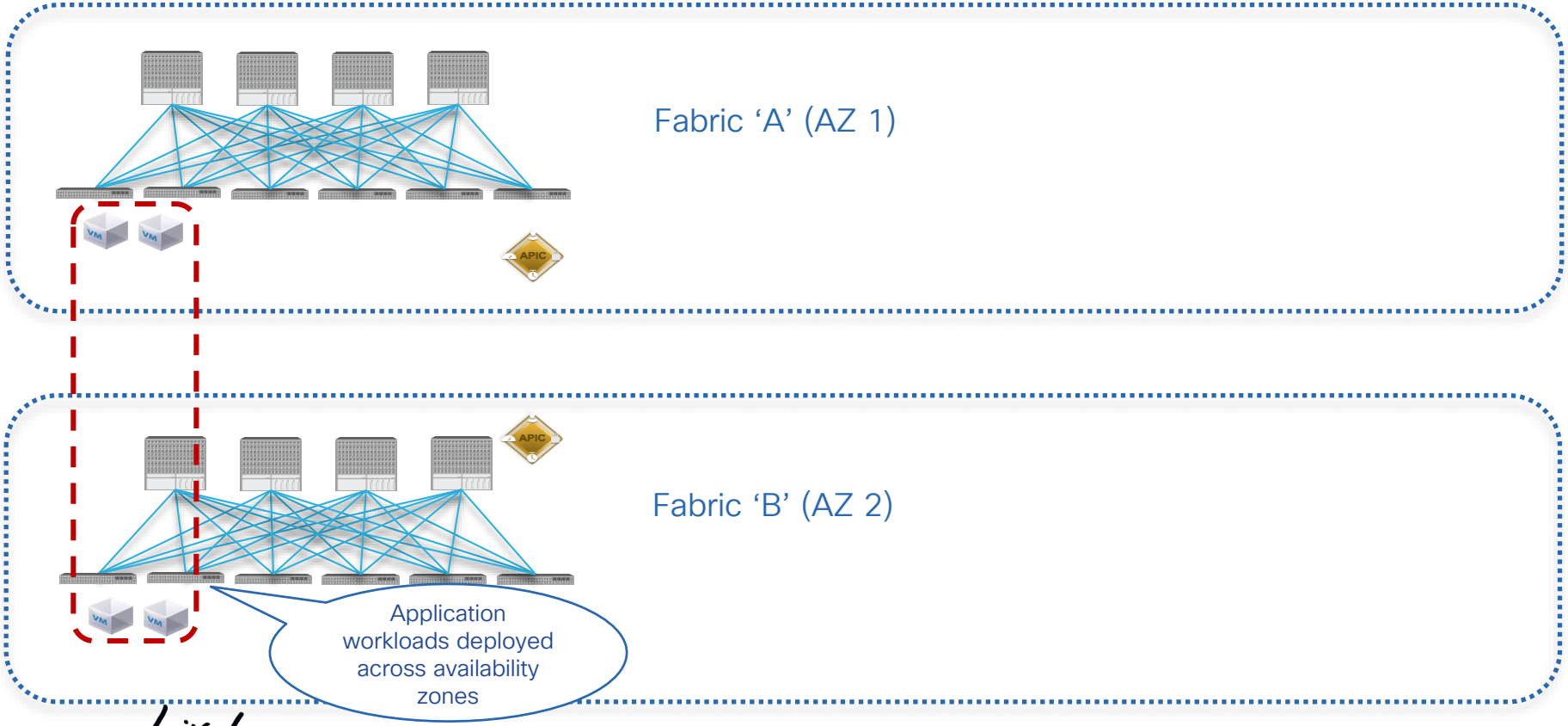
- Regions – Separate large geographical areas, each composed of multiple, isolated locations known as Availability Zones
- Availability Zones – Distinct locations within a region that are engineered to be isolated from failures in other Availability Zones and provide inexpensive, low latency network connectivity to other Availability Zones in the same region

Amazon Web Services



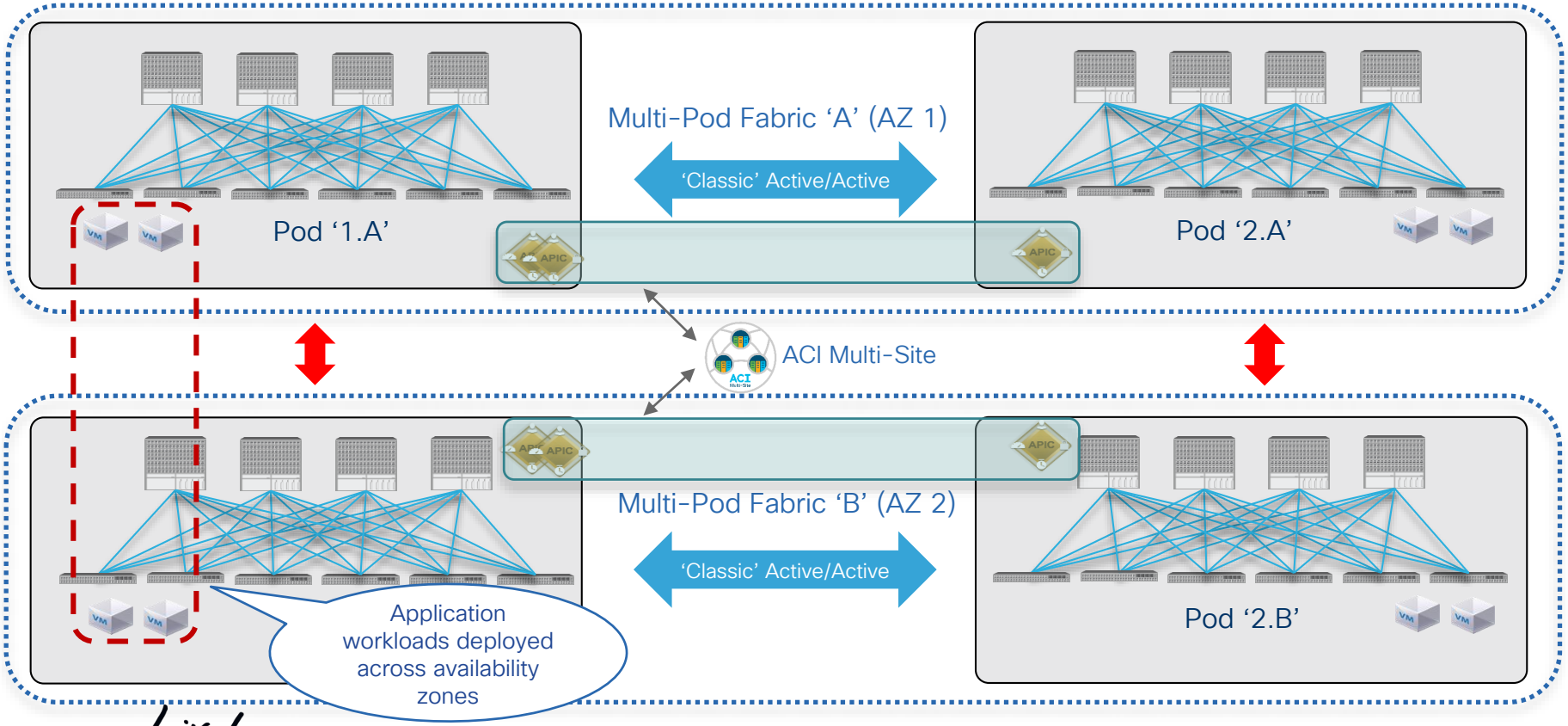
Typical Requirement

Creation of Two Independent Fabrics/AZs



Typical Requirement

Creation of Two Independent Fabrics/AZs



Agenda

- ACI Anywhere, Extending the ACI Fabric
  - Overall Design Principles (AZs and Regions)
  - Mapping use cases to the proper solutions
    - Active/Active DC → Multi-Pod
    - Disaster Recovery → Multi-Site
    - Migration/Coexistence with Legacy DC Networks and 'Disaggregated DCs' Model → Physical Remote Leaf
    - Baremetal Cloud Integration → Virtual Pod (vPod)
  - Extending ACI to the Cloud
  - Connecting the users to the Multi-Cloud DC
    - ACI and SDA Integration
    - ACI and SDWAN Integration

Active/Active DC Deployment

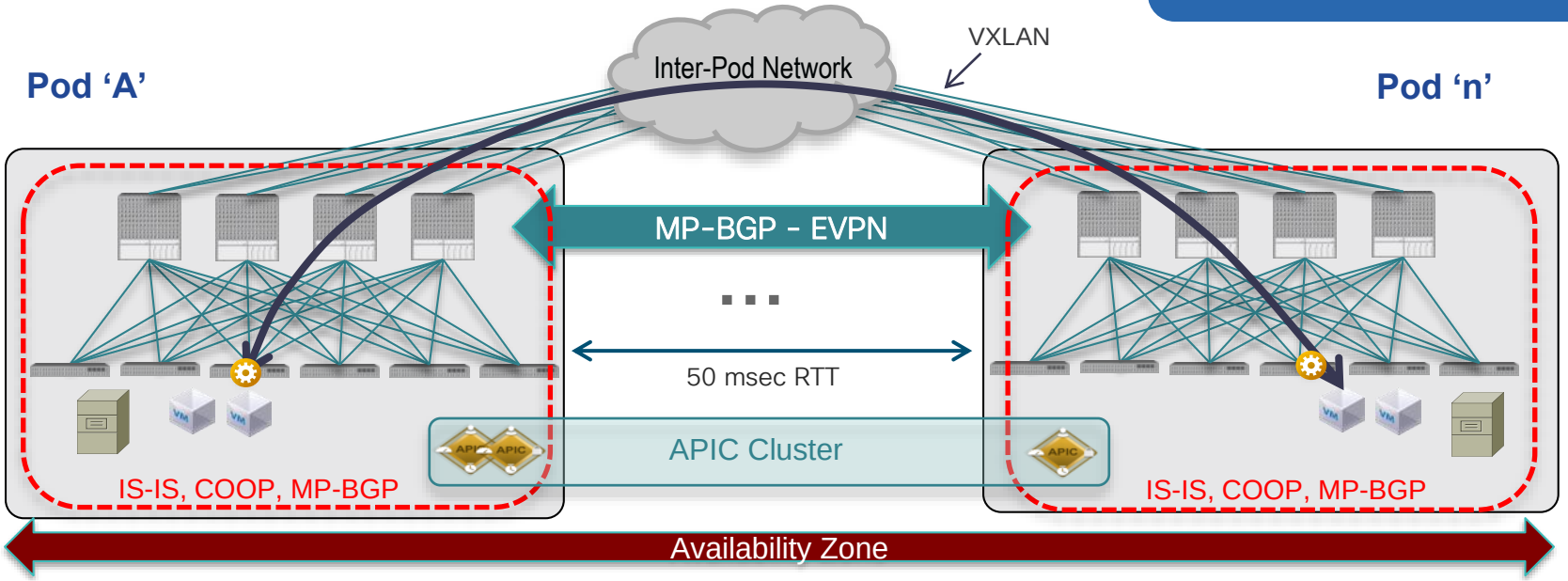
Metro Virtual Data Centre

- High-availability application and data solution architecture which leverages a dual data centre physical infrastructure (tightly coupled DC for short distances)
- Management and interaction of applications in a paired data centre environment
- Disaster Avoidance and Prevention by pro-actively migrates seamlessly Virtual Machines with no interruption (vMotion for example)
- Active-active capability and workload rotation to accelerate incident response time and increase confidence
  - Deployment of an ESXi Metro Cluster with vSphere HA, Fault Tolerance (FT), DRS
  - Service Nodes (FW, SLB) clustered across DCs (Active/Standby, Active/Active)

ACI Multi-Pod

The Ideal Architecture for Active/Active DC Deployments

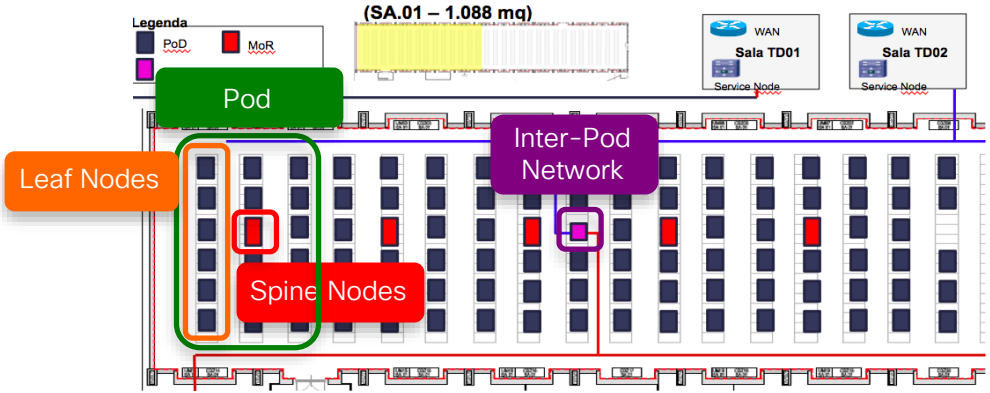
For More Information on ACI Multi-Pod: BRKACI-2003



- Multiple ACI Pods connected by an IP Inter-Pod L3 network, each Pod consists of leaf and spine nodes
- Managed by a single APIC Cluster
- Single Management and Policy Domain
- Forwarding control plane (IS-IS, COOP) fault isolation
- Data Plane VXLAN encapsulation between Pods
- End-to-end policy enforcement

ACI Multi-Pod
Most Common Use Cases

- Need to scale up a single ACI fabric above 200 leaf nodes supported in a single Pod
- Handling 3-tiers physical cabling layout (for example traditional N7K/N5K/N2K deployments)

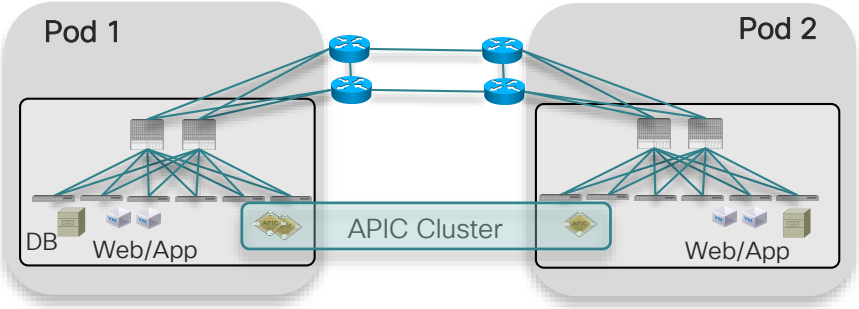


- True Active/Active DC deployments

  Single VMM domain across DCs (stretched ESXi Metro Cluster, vSphere HA/FT, DRS initiated workload mobility,...)

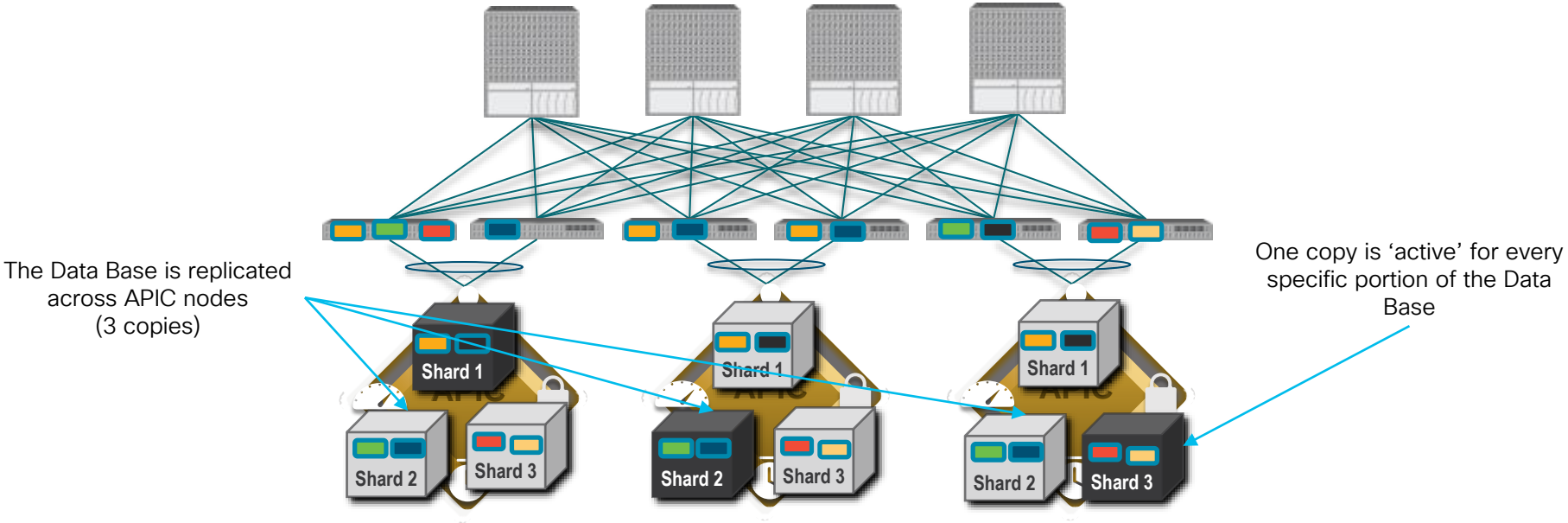
  Deployment of Active/Standby or Active/Active clustered network services (FWs, SLBs) across DCs

  Application clustering (L2 BUM extension across Pods)



ACI Multi-Pod
APIC Cluster Deployment

APIC – Distributed Multi-Active Data Base



- All Services in ACI run against their own portions of a Database
- Services and Database Processes are active on all nodes (not active/standby)
- The Data Base is distributed as active + 2 backup instances (shards) for every attribute

APIC Cluster Deployment with Multi-Pod

Deployment Recommendations

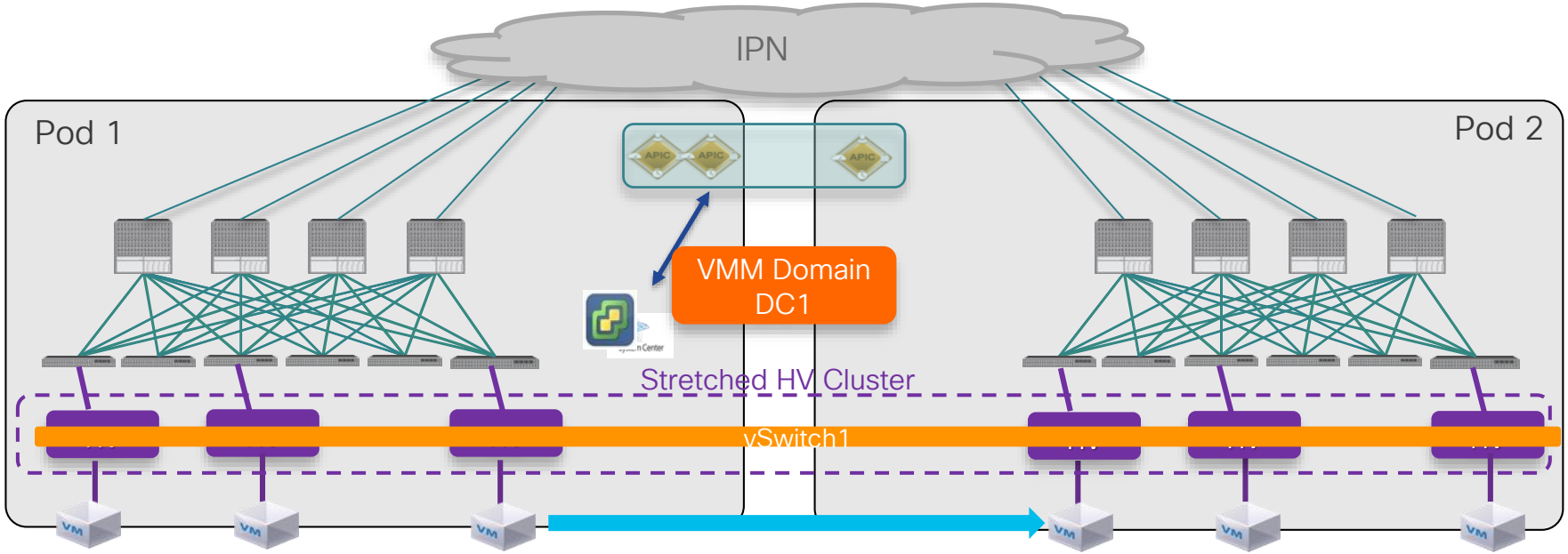
- Main recommendation: deploy a 3 nodes APIC cluster when less than 80 leaf nodes are deployed across Pods
- From ACI release 4.1(1) can deploy 4 nodes to support up to 200 leaf nodes across Pods
- When 5 (or 7) nodes are really needed for scalability reasons, follow the rule of thumb of never placing more than two APIC nodes in the same Pod (when possible):

	Pod1	Pod2	Pod3	Pod4	Pod5	Pod6
2 Pods*	APIC APIC APIC	APIC APIC				
3 Pods	APIC APIC	APIC APIC	APIC			
4 Pods	APIC APIC	APIC	APIC	APIC		
5 Pods	APIC	APIC	APIC	APIC	APIC	
6+ Pods	APIC	APIC	APIC	APIC	APIC	

*'ID Recovery' procedure possible for recovering of lost information

Multi-Pod and Virtual Machine Manager (VMM) Integration

ACI Multi-Pod and VMM Integration

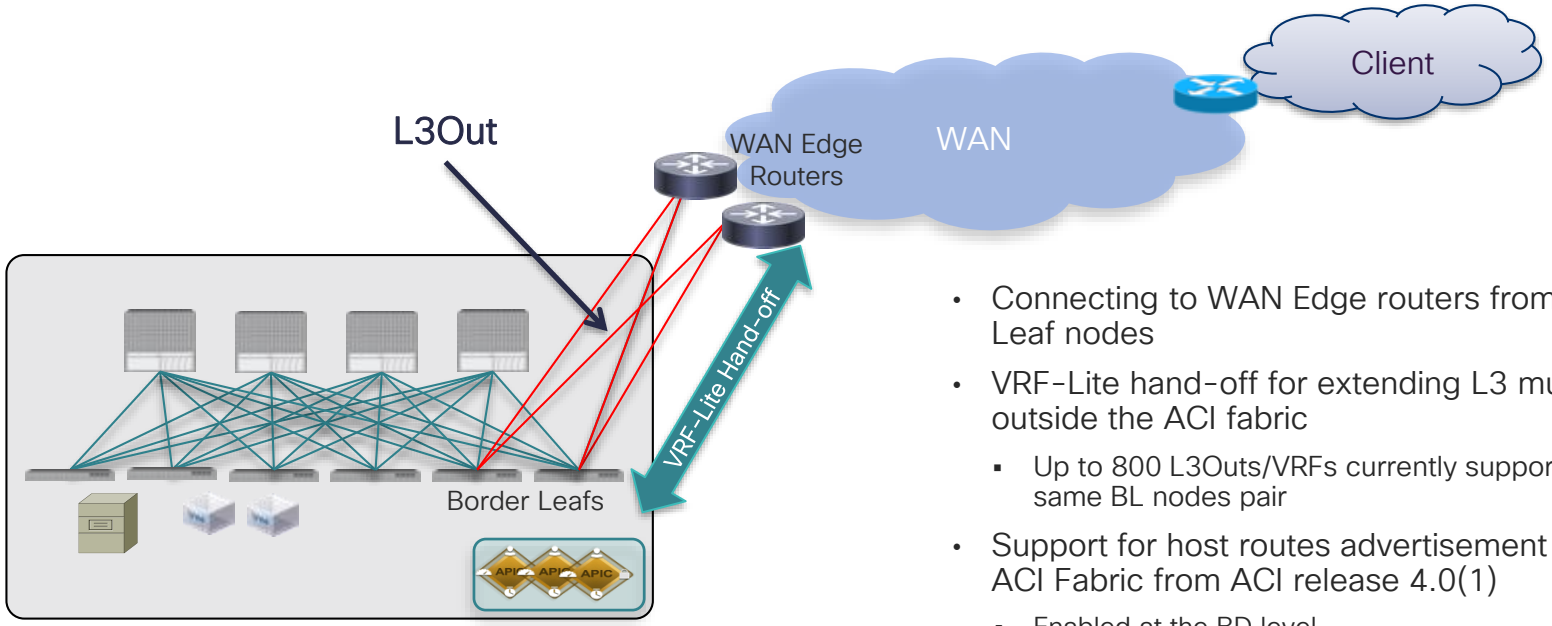


- Cluster of Hypervisors stretched across Pods
  - Single VMM domain created across Pods
- Logical switch extended across the hypervisors part of the same stretched cluster
- Support for all intra-cluster functions (vSphere HA/FT, DRS, etc.)

Multi-Pod Connectivity to the External L3 Domain

Connecting to the External Layer 3 Domain

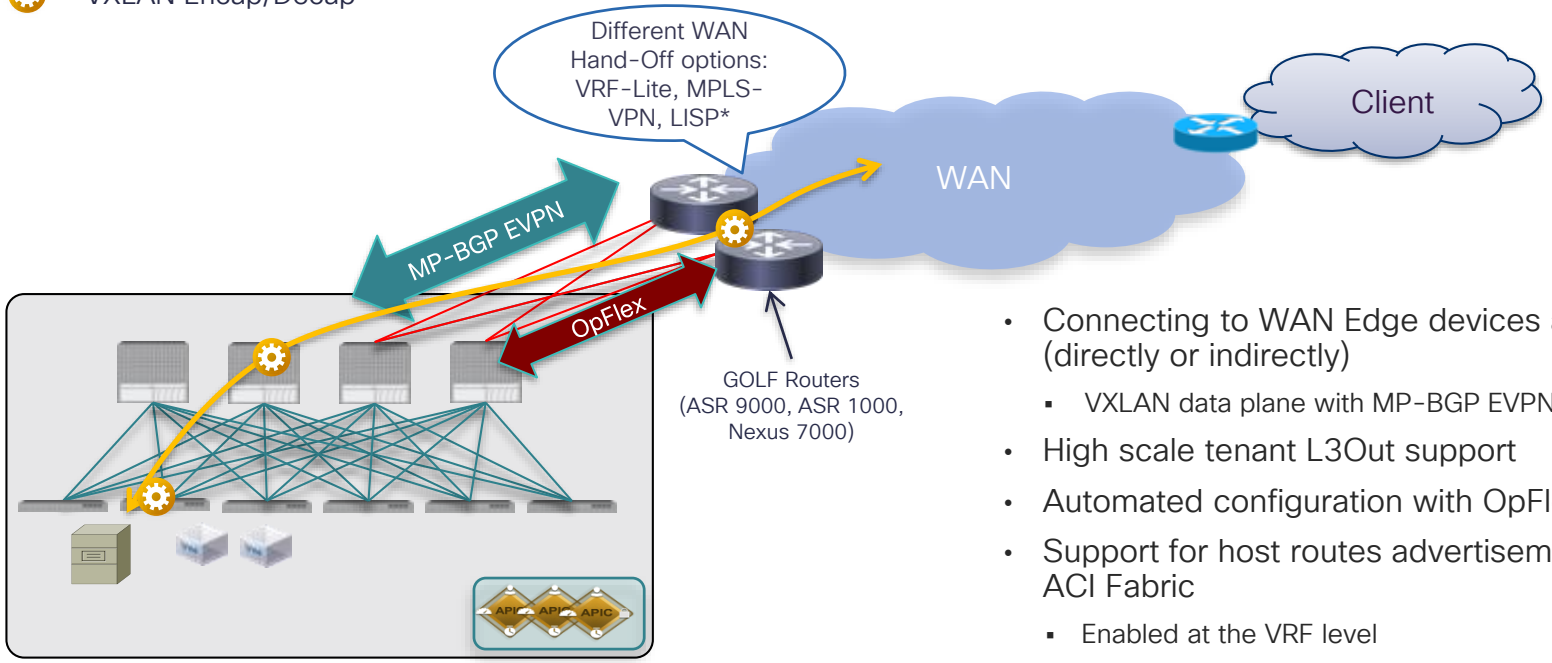
‘Traditional’ L3Outs on the BL Nodes (Recommended Option)



- Connecting to WAN Edge routers from Border Leaf nodes
- VRF-Lite hand-off for extending L3 multi-tenancy outside the ACI fabric
  - Up to 800 L3Outs/VRFs currently supported on the same BL nodes pair
- Support for host routes advertisement out of the ACI Fabric from ACI release 4.0(1)
  - Enabled at the BD level
- Support for L3 Multicast and Shared L3Out

Connecting to the External Layer 3 Domain
‘GOLF’ L3Outs (VRF High Scale Use Cases)

= VXLAN Encap/Decap

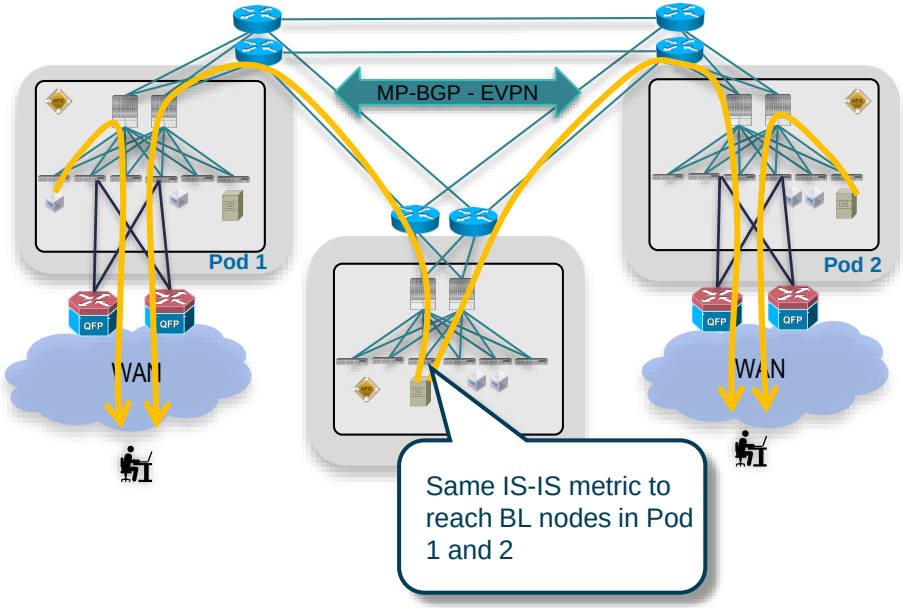


- Connecting to WAN Edge devices at Spine nodes (directly or indirectly)
  - VXLAN data plane with MP-BGP EVPN control plane
- High scale tenant L3Out support
- Automated configuration with OpFlex
- Support for host routes advertisement out of the ACI Fabric
  - Enabled at the VRF level
- No support for L3 Multicast or Shared L3Out (every tenant VRF requires its own L3Out)

Connecting Multi-Pod to Layer 3 Domain

Sharing L3Out Across Pods

- A Pod <u>does not need</u> to have a dedicated WAN connection (i.e. can offer transit services to other Pods)
- Multiple WAN connections can be deployed across Pods
- Outbound traffic: by default VTEPs always select WAN connection in the local Pod based on preferred metric
- Leaf nodes in Pods without local L3Outs will load-balance traffic between L3Outs in remote Pods

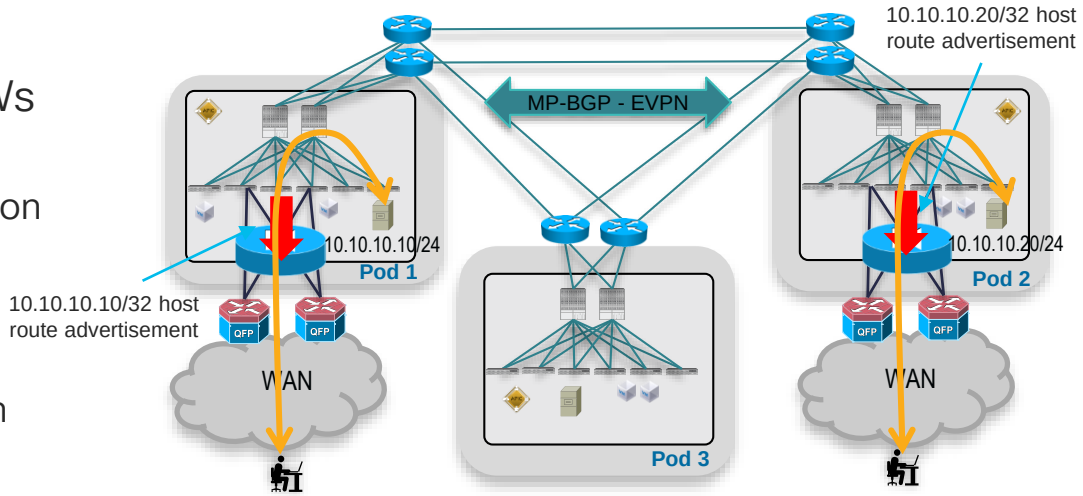


Connecting Multi-Pod to Layer 3 Domain

Use of Host-Route Advertisement on BL L3Outs

ACI 4.0(1) Release

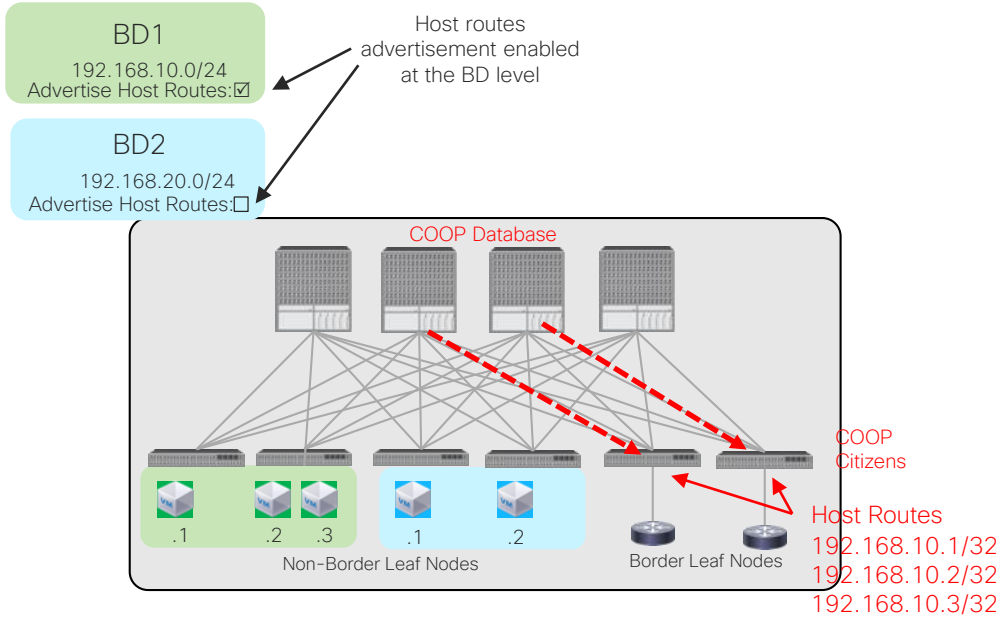
- Host routes advertisement is a best option to ensure all the deployed FWs are actively utilized
  - Support for host route advertisement on BL nodes available from ACI release 4.0(1)
  - Enabled at the BD level
  - Requires an L3Out connection in each Pod
  - Allows to keep symmetric inbound and outbound traffic paths



Host Routes Advertisement

Downloading the Host Routes to the Border Leaf Nodes

ACI 4.0(1) Release



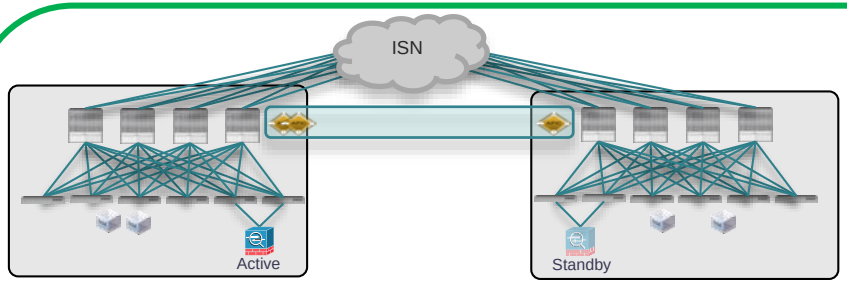
- Endpoint information is stored on the COOP database of the spines in the fabric
- When a BD is enabled with Host Routing the border leaf nodes download the host routes from the spines
- Enabling Host Routing on the Bridge Domain does not automatically advertise the host routes out of an L3Out, it must be explicitly configured (see a following slide)

Multi-Pod Network Services Integration Models

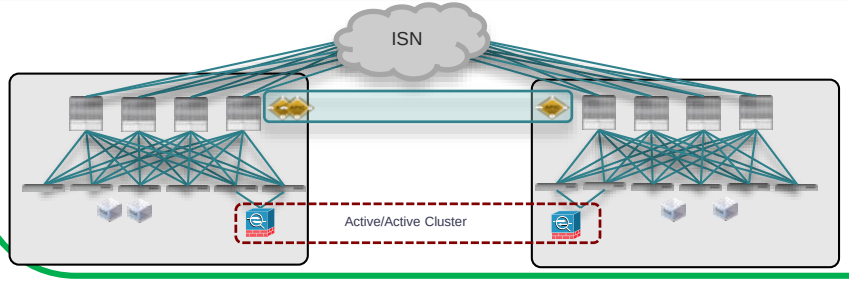
Multi-Pod and Network Services

Integration Models

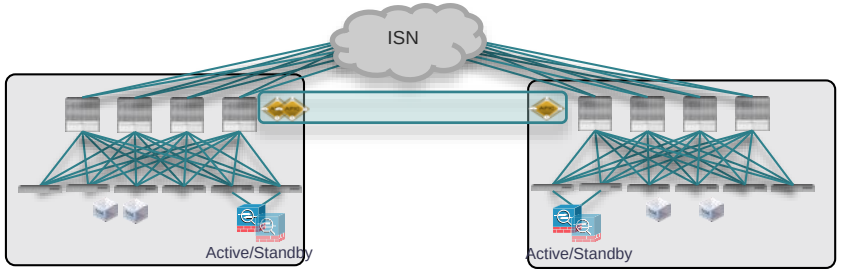
Typical options for an Active/Active DC use case



- Active and Standby pair deployed across Pods
- No issues with asymmetric flows



- Active/Active FW cluster nodes stretched across Sites (single logical FW)
- Requires the ability of discovering the same MAC/IP info in separate sites at the same time
- Supported from ACI release 3.2(4d) with the use of Service-Graph with PBR



- Independent Active/Standby pairs deployed in separate Pods
- Use of Symmetric PBR to avoid the creation of asymmetric paths crossing different active FW nodes

ACI Multi-Pod

Where to Go for More Information

- ✓ ACI Multi-Pod White Paper

  http://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-737855.html?cachemode=refresh

- ✓ ACI Multi-Pod Configuration Paper

  https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-739714.html

- ✓ ACI Multi-Pod and Service Node Integration White Paper

  https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-739571.html

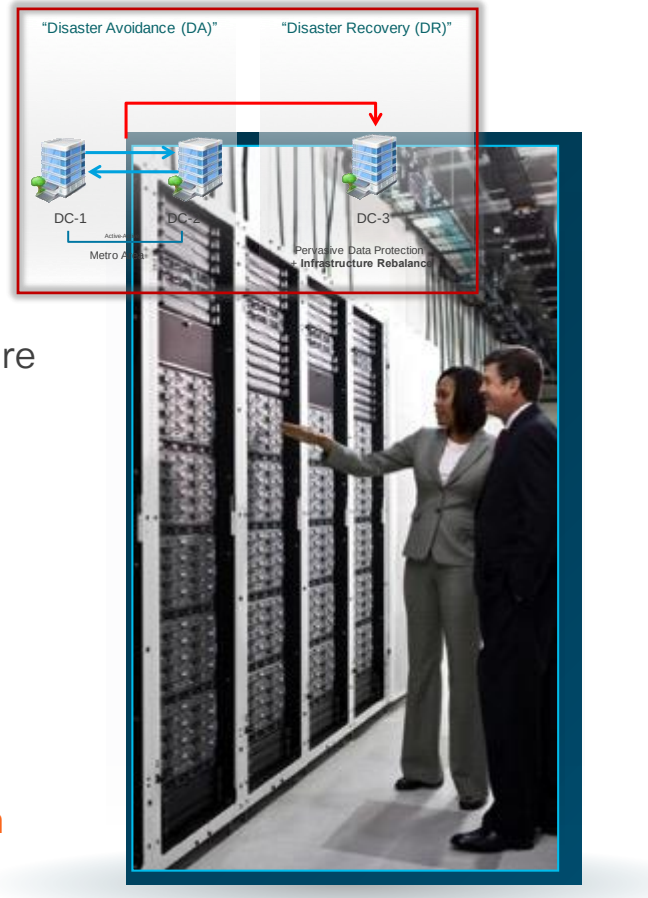
- ✓ BRKACI-2003

Agenda

- ACI Anywhere, Extending the ACI Fabric
  - Overall Design Principles (AZs and Regions)
  - Mapping use cases to the proper solutions
    - Active/Active DC → Multi-Pod
    - Disaster Recovery → Multi-Site
    - Migration/Coexistence with Legacy DC Networks and 'Disaggregated DCs' Model → Physical Remote Leaf
    - Baremetal Cloud Integration → Virtual Pod (vPod)
  - Extending ACI to the Cloud
  - Connecting the users to the Multi-Cloud DC
    - ACI and SDA Integration
    - ACI and SDWAN Integration

Disaster Recovery Use Case

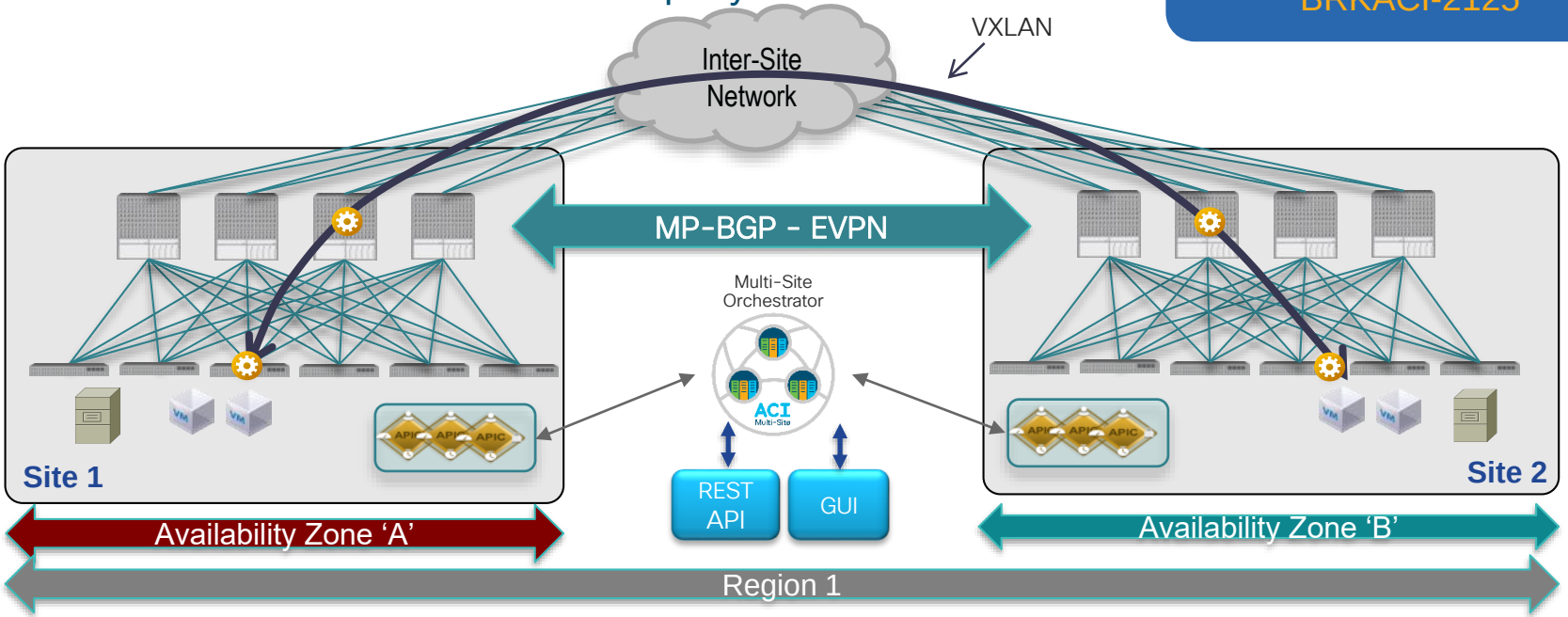
- Drive cost efficiencies through re-use of infrastructure and processes
- Integrate Disaster Recovery into day to day operations
- Make capacity growth sustainable through repurposed infrastructure and shared resources leveraging virtualization
- Provide a DR capability that is acceptable for any framework to all subsidiaries (multi-tenancy)
- Supports failover of passive services
  - VMware Site Recovery Manager (SRM)
  - Microsoft Cluster Services (MSCS)
  - IBM HA Clustering Multi-proc (HACMP) / (PowerHA)
  - Etc..
- Global Server Load Balancing, Route Health Injection or LISP (Path Redirection)



ACI Multi-Site

The Ideal Architecture for DR Deployments

For More Information on ACI Multi-Site: BRKACI-2125



- Separate ACI Fabrics with independent APIC clusters
- No latency limitation between Fabrics
- ACI Multi-Site Orchestrator pushes cross-fabric configuration to multiple APIC clusters providing scoping of all configuration changes

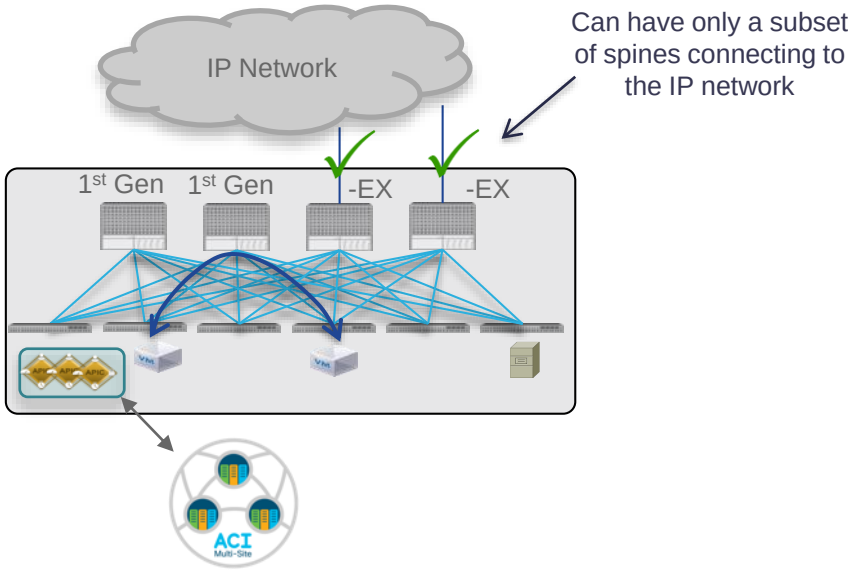
- MP-BGP EVPN control plane between sites
- Data Plane VXLAN encapsulation across sites
- End-to-end policy definition and enforcement

ACI Multi-Site

Hardware Requirements

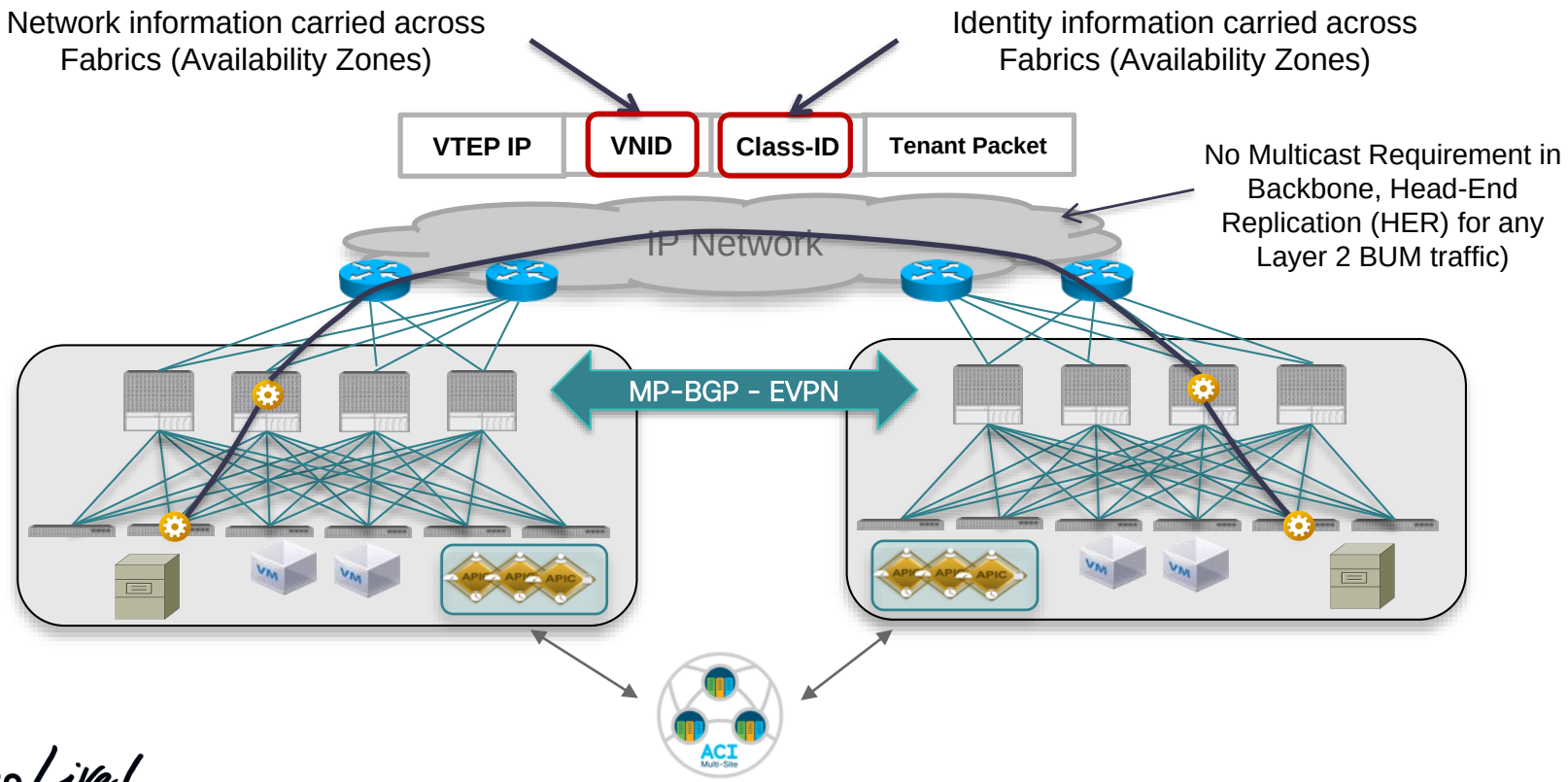
- Support all ACI leaf switches (1st Generation, -EX and -FX)
- Only –EX spine (or newer) to connect to the inter-site network
- New 9364C non modular spine (64x40G/100G ports) supported for Multi-Site from ACI 3.1 release (shipping)
- 1st generation spines (including 9336PQ) not supported

  Can still leverage those for intra-site leaf to leaf communication



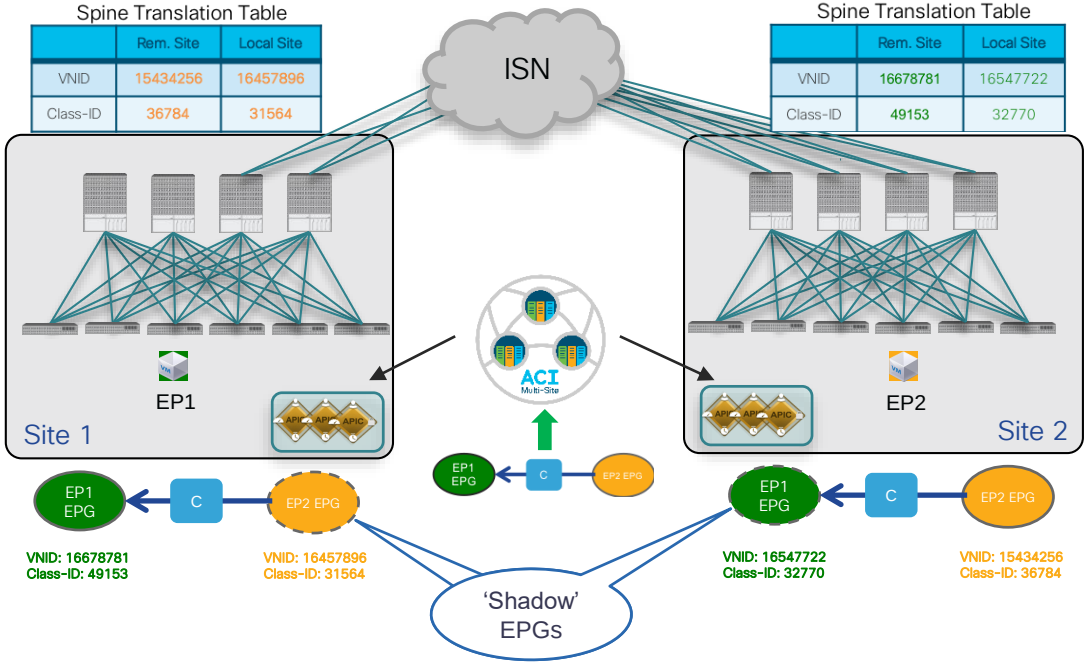
ACI Multi-Site

Network and Identity Extended between Fabrics



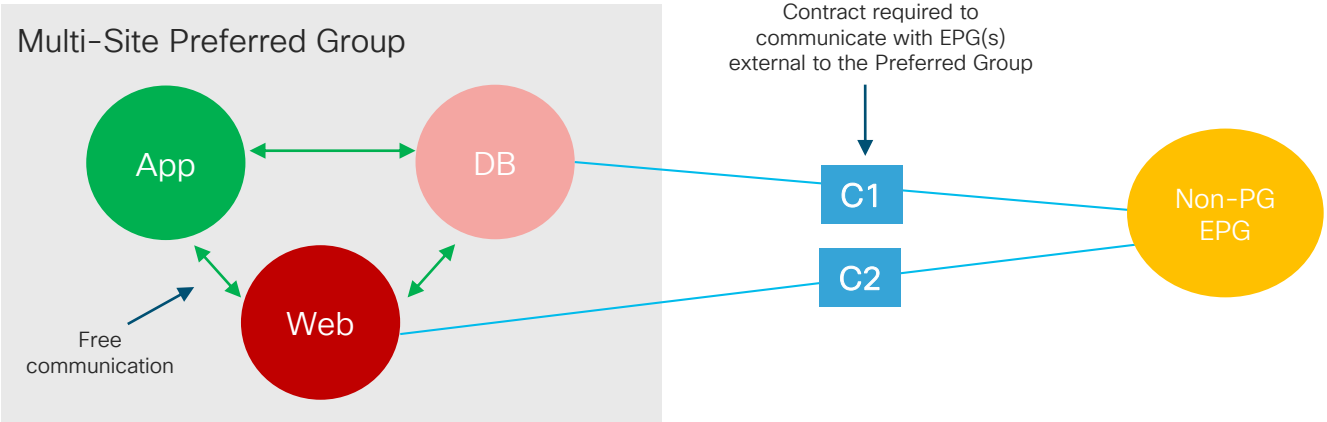
ACI Multi-Site
Inter-Site Policies and Spines' Translation Tables

- Inter-Site policies defined on the ACI Multi-Site Orchestrator are pushed to the respective APIC domains
  - End-to-end policy consistency
  - Creation of 'Shadow' EPGs to locally represent the policies
- Inter-site communication requires the installation of translation table entries on the spines (namespace normalization)
- Up to ACI release 4.0(1) translation entries are populated only in two cases:
  1. Stretched EPGs/BDs
  2. Creation of a contract between not stretched EPGs



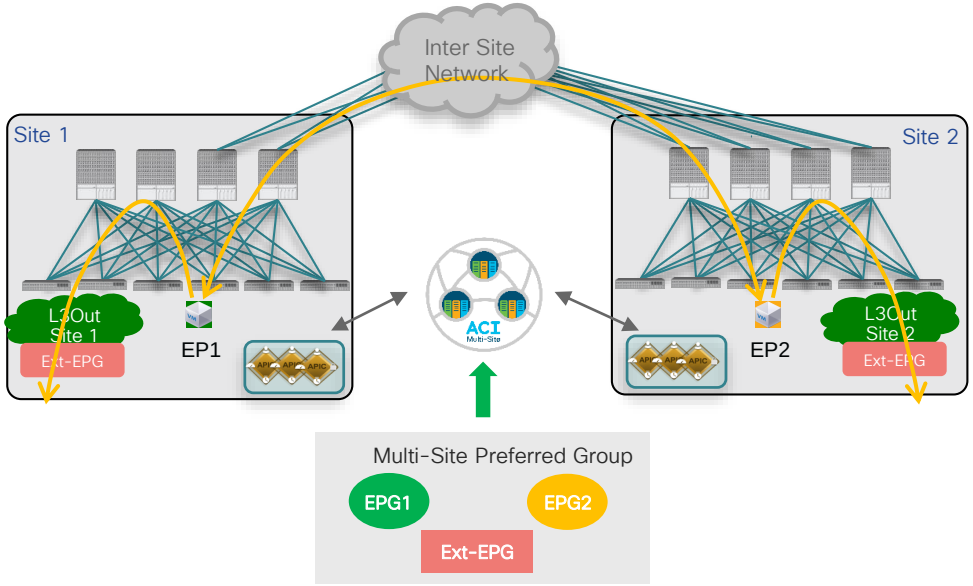
ACI Multi-Site
Removing Policy Enforcement: Preferred Groups

MSO 2.0(2) Release



- ”VRF unenforced” not supported with Multi-Site
- Multi-Site Preferred Group configuration from the Multi-Site Orchestrator is supported from MSO 2.0(2) release
  - Creates ‘shadow’ EPGs and translation table entries ‘under the hood’ to allow ‘free’ inter-site communication
  - 250 Preferred Groups supported as ACI release 4.1(1)
- Typically desired in legacy to ACI migration scenarios

Removing Policy Enforcement
Preferred Groups for E-W and N-S Flows



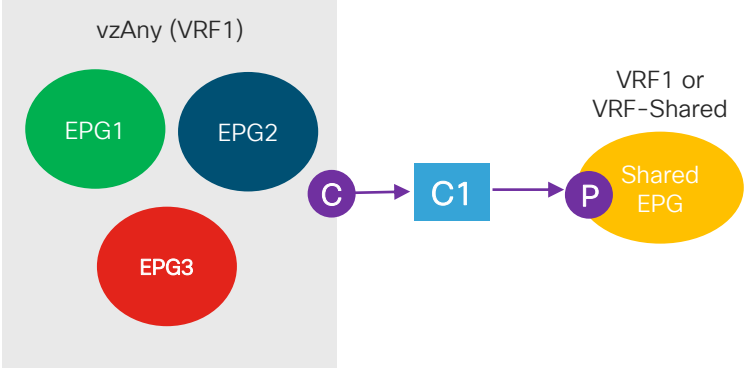
- Adding internal EPGs and External EPGs (associated to L3Outs) to the Preferred Group allows to enable free east-west and north-south connectivity
- When adding the Ext-EPG to the Preferred Group:
  - Can't use 0.0.0.0/0 for classification, needs more specific prefixes
  - As workaround it is possible to use 0.0.0.0/1 and 128.0.0.0/1 to achieve the same result
  - Must ensure Ext-EPG is a stretched object

ACI Multi-Site
vzAny Support (MSO 2.2(4) Release)

MSO 2.2(4) Release

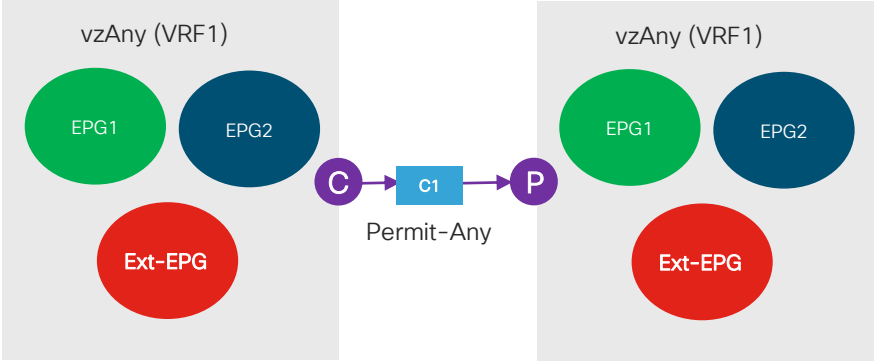
What is vzAny? Logical object representing all the EPGs in a VRF

Use case 1: Many-to-One communication (Shared Services)



- Multiple EPGs part of a specific VRF1 consume the services provided by a shared EPG (part of VRF1 or of a VRF-shared)
- VRF-shared can be part of the same tenant or of a different tenant

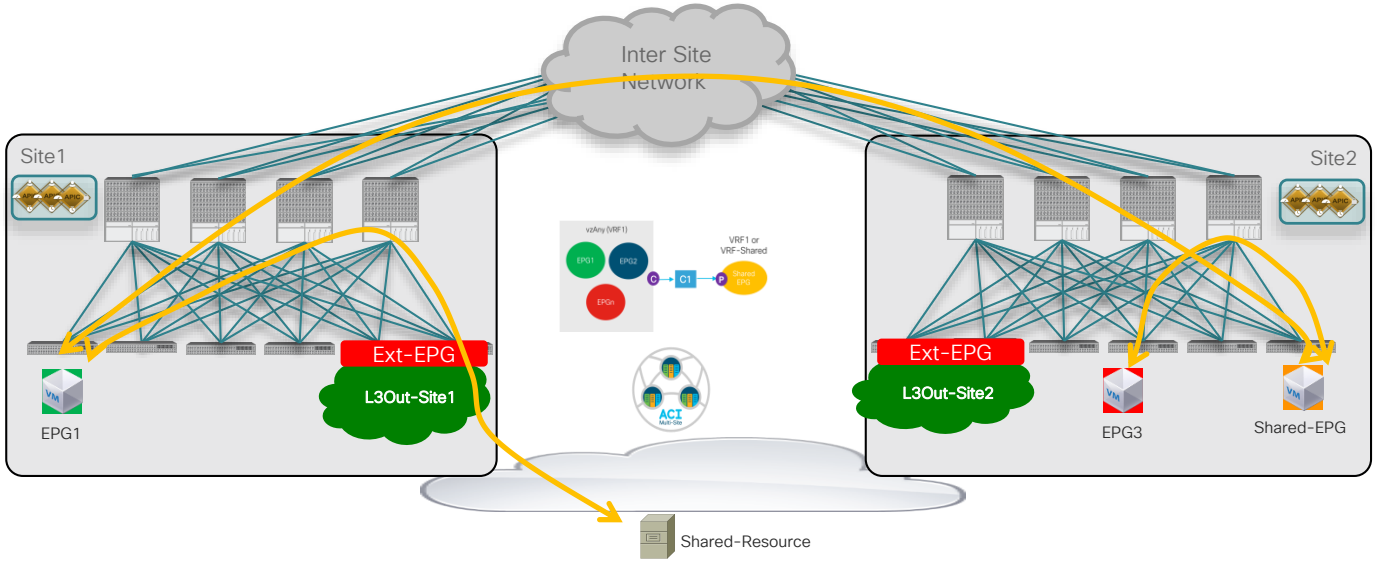
Use case 2: Enable free communication inside a VRF



- vzAny provides and consumes a contract with an associated “Permit-any” filter
- Use ACI fabric only for network connectivity without policy enforcement
- Equivalent to “VRF unenforced”

ACI Multi-Site and vzAny
Many-to-One Communication (Shared Services)

MSO 2.2(4) Release

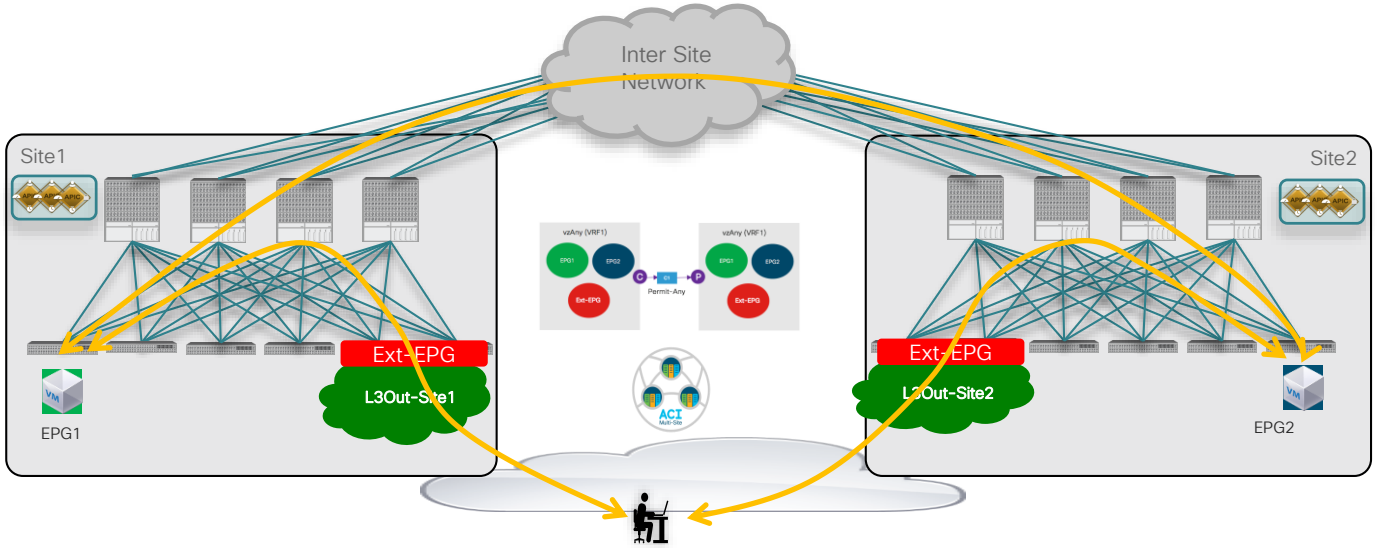


- Proper translation entries are created on the spines of both fabrics to enable east-west communication
- Supported also for Shared Services behind an L3Out

ACI Multi-Site and vzAny

Enable Inter-Site Free Communication Inside a VRF

MSO 2.2(4) Release



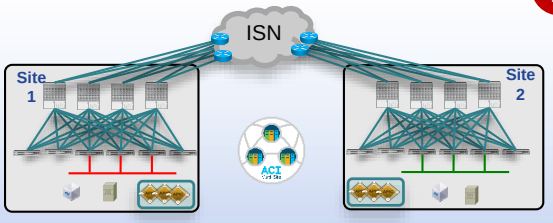
- Proper translation entries are created on the spines of both fabrics to enable east-west communication
- Supported also for connecting to the external Layer 3 domain

ACI Multi-Site
Per Bridge Domain Behavior

Should be the behavior for the majority of BDs with Multi-Site

Layer 3 only across sites

1

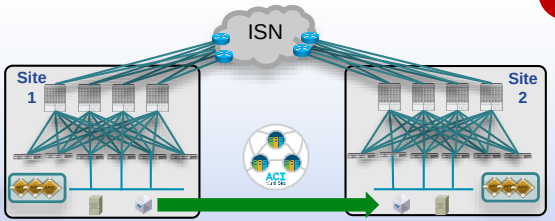


- Bridge Domains and subnets not extended across Sites
- Layer 3 Intra-VRF or Inter-VRF communication (shared services across VRFs/Tenants)

MSO GUI (BD)

IP Mobility without BUM flooding

2

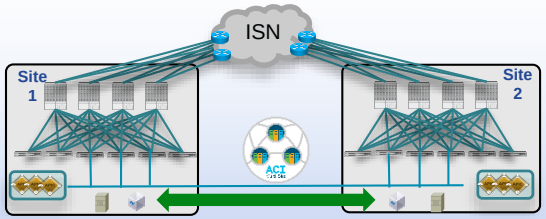


- Same IP subnet defined in separate Sites
- Support for IP Mobility ('cold' and 'live'* VM migration) and intra-subnet communication across sites
- **No Layer 2 BUM flooding across sites**

MSO GUI (BD)

Layer 2 adjacency across Sites

3



- Interconnecting separate sites for fault containment and scalability reasons
- Layer 2 domains stretched across Sites, support for <u>application clustering</u>
- **Layer 2 BUM flooding across sites**

MSO GUI (BD)

*'Live' migration officially supported from ACI release 3.2

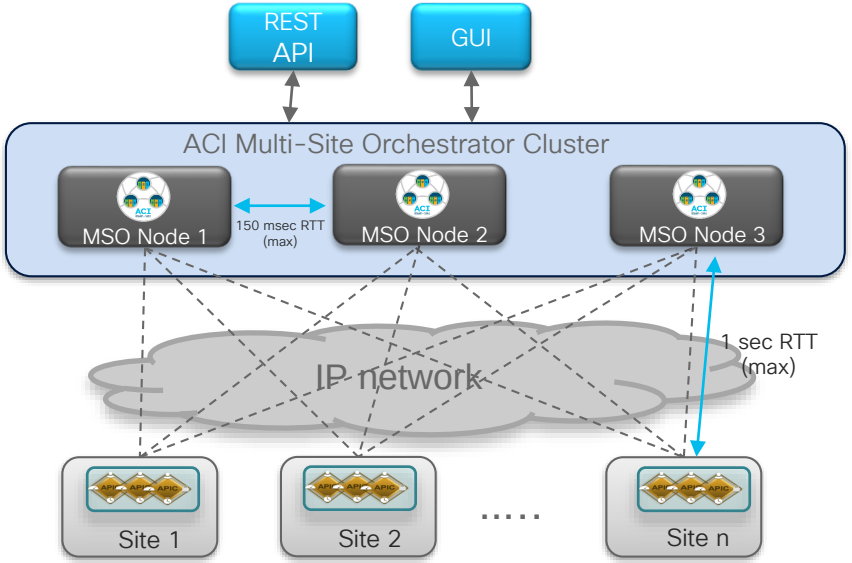
ACI Multi-Site
Continuous Scale Improvements

	ACI Release 3.0	ACI Release 3.1	ACI Release 3.2		ACI Release 4.2
	MSO 1.0	MSO 1.1	MSO 1.2		MSO 2.1
Number Of Sites	5	8	10	...	12
Max Leafs (across sites)	250	800	1200		1600
Tenants	100	200	300		400
VRF	400	400	800		1000
BD	800	2,000	3,000		4,000
EPGs	800	2,000	3,000		4,000
Contracts	1,000	2,000	3,000		4,000
L3Out External EPGs	500	500	500		500
Isolated EPGs	N/A	400	400		400

For more information please refer to:
https://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/sw/4-x/verified-scalability/Cisco-ACI-Verified-Scalability-Guide-422.pdf

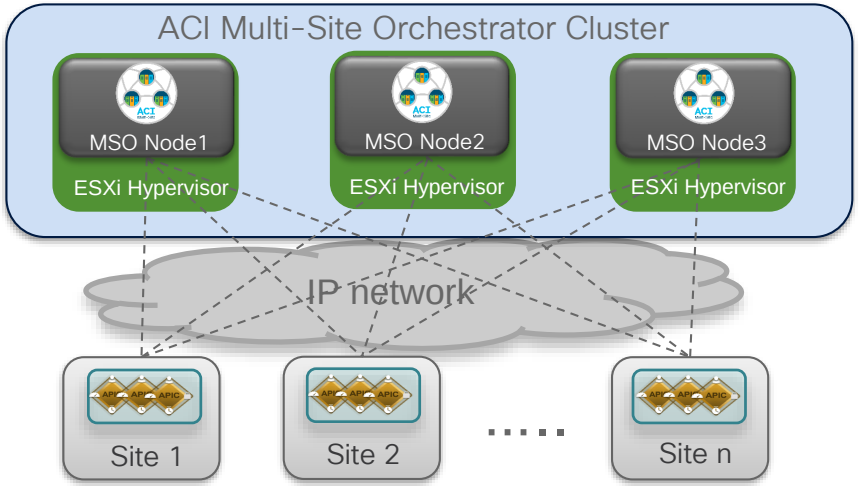
Multi-Site
Multi-Site Orchestrator, Schemas and Templates

ACI Multi-Site
Multi-Site Orchestrator (MSO)



- Three MSO nodes are clustered and run concurrently (active/active)
  - Typical database redundancy considerations (minority/majority rules)
  - Up to 150 msec RTT latency supported between MSO nodes
- OOB Mgmt connectivity to the APIC clusters deployed in separate sites
  - Up to 1 sec RTT latency between MSO and APIC nodes
- Main functions offered by MSO:
  - Monitoring the health-state of the different ACI Sites
  - Provisioning of day-0 infrastructure configuration to establish inter-site EVPN control plane and VXLAN data plane
  - Defining and provisioning tenant policies
  - Day-2 operation functionalities

ACI Multi-Site Orchestrator
VM Based MSO Cluster



- Supported from the beginning (MSO release 1.0(1))
- Each Cisco ACI Multi-Site Orchestrator node is packaged in a VMware vSphere virtual appliance
- For high availability, you should deploy each Cisco ACI Multi-Site Orchestrator virtual machine on its own VMware ESXi host
- Requirements for MSO Release 1.2(x) and above:
  - VMware ESXi 6.0 or later
  - Minimum of eight virtual CPUs (vCPUs), 24 Gbps of memory, and 100 GB of disk space

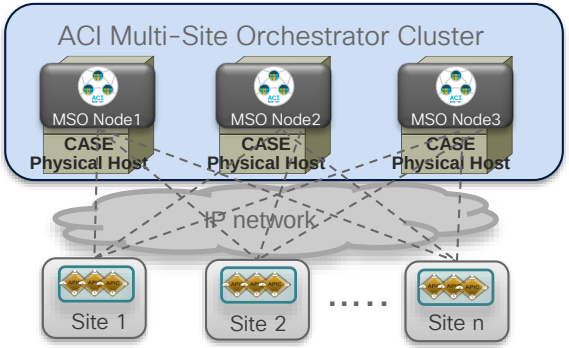
ACI Multi-Site Orchestrator

Cisco Application Service Engine (CASE) Based MSO Cluster

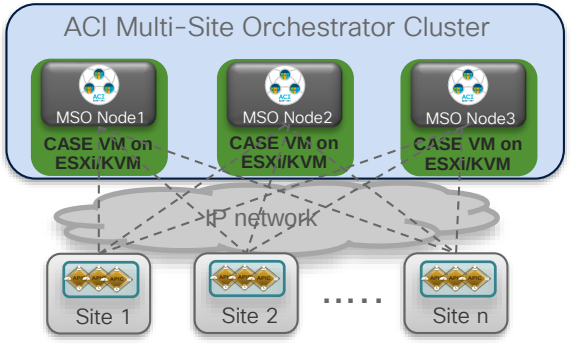
MSO 2.2(3) Release

- CASE cluster available in different form factors (physical, on-prem VM, AWS instance)
- MSO is installed as an App on the CASE cluster
- Recommended MSO cluster deployment option going forward

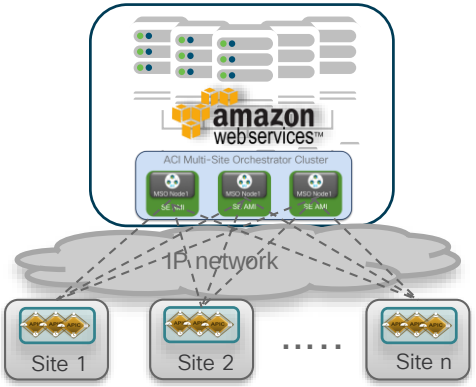
1 MSO App on CASE physical form factor



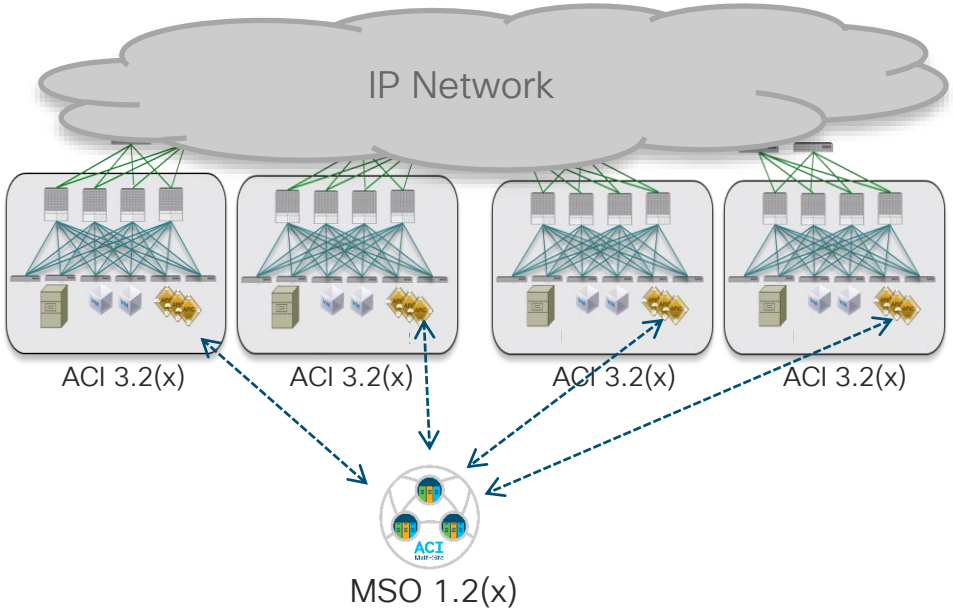
2 MSO App on CASE VM form factor (on premises)



3 MSO App on CASE VM form factor (on AWS)



ACI Multi-Site
MSO and APIC Release Dependency (Pre-MSO 2.2(1) Release)



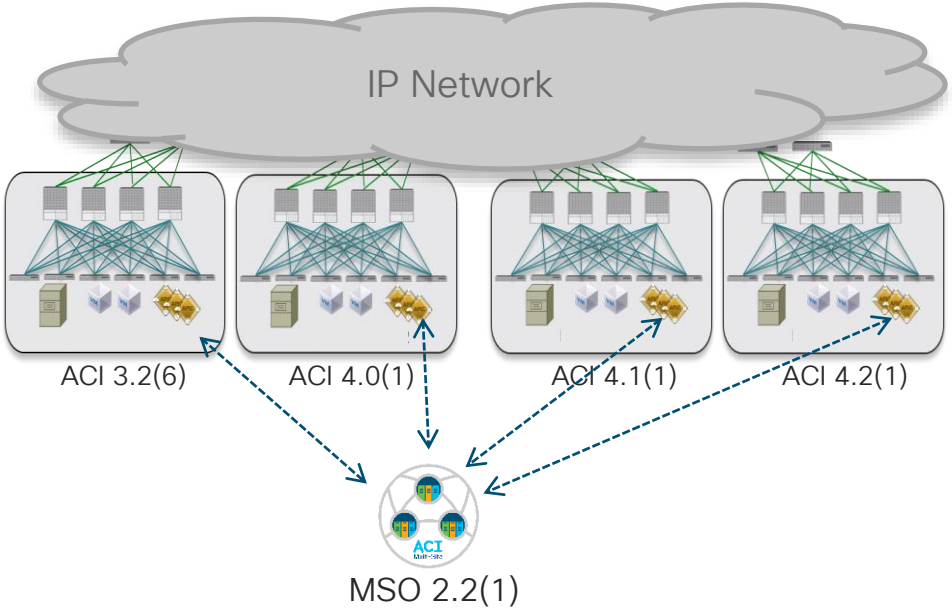
- Pre-MSO 2.2(1) release, MSO and ACI releases have to be aligned
  - For example MSO 1.2(x) is used with all sites running ACI release 3.2(x)
- Different ACI versions across sites are only supported during an ACI SW upgrade procedure
- The supported order of upgrade is:
  1. APIC firmware
  2. Switches firmware
  3. MSO

  For each Site (1–2)

  As last task (3)

ACI Multi-Site Interversion
Decoupling MSO and APIC Releases

MSO 2.2(1) Release



- MSO “interversion” support is available with MSO release 2.2(1)
- Different ACI versions across sites can be supported at steady state
- MSO gets visibility into what functionalities are supported in each fabric (based on the specific ACI releases)
  - Preventing the deployment of unsupported functionalities

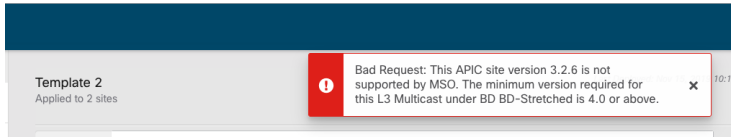
ACI Multi-Site Interversion
Decoupling MSO and APIC Releases

SW dependency for different ACI functionalities

Feature	Minimum APIC Version
ACI Multi-Pod Support	Release 3.2(1)
Service Graphs (L4-L7 Services)	Release 3.2(1)
External EPGs	Release 3.2(1)
ACI Virtual Edge VMM Support	Release 3.2(1)
DHCP Support	Release 3.2(1)
Consistency Checker	Release 3.2(1)
CloudSec Encryption	Release 4.0(1)
Layer 3 Multicast	Release 4.0(1)
MD5 Authentication for OSPF	Release 4.0(1)
EPG Preferred Group	Release 4.0(2)
Host Based Routing	Release 4.1(1)
Intersite L3Out	Release 4.2(1)

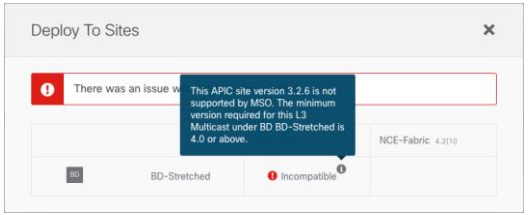
1

Version check during the “Save” operation of a template



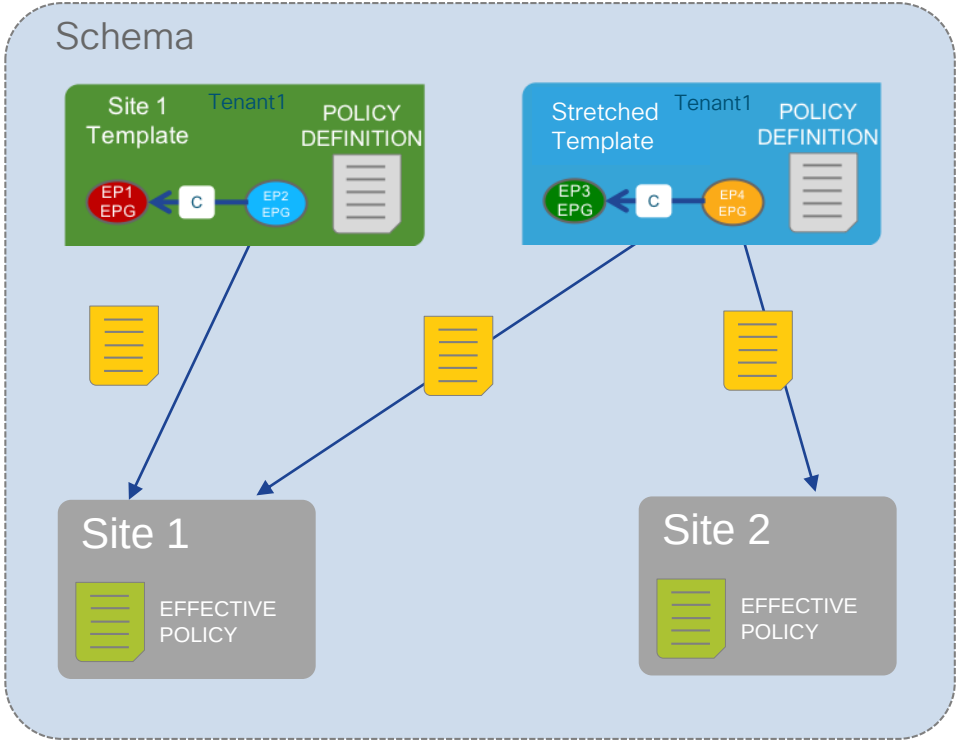
2

Version check at deployment time for a template



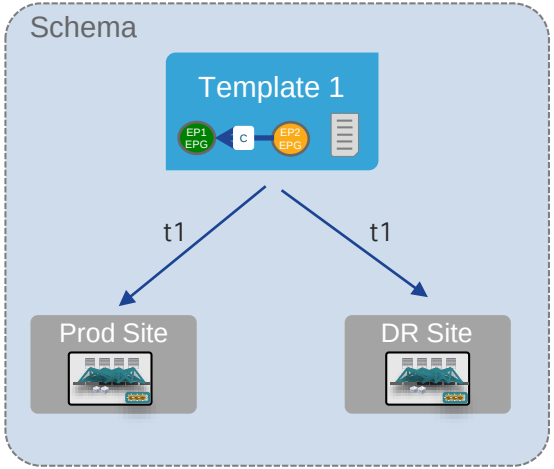
ACI Multi-Site
MSO Schema and Templates

- Template = ACI policy definition (ANP, EPGs, BDs, VRFs, etc.)
- Schema = container of Templates sharing a common use-case
  - As an example, a schema can be dedicated to a Tenant
- The template is currently the atomic unit of change for policies
  - Such policies are concurrently pushed to one or more sites
- Scope of change: policies in different templates can be pushed to separate sites at different times

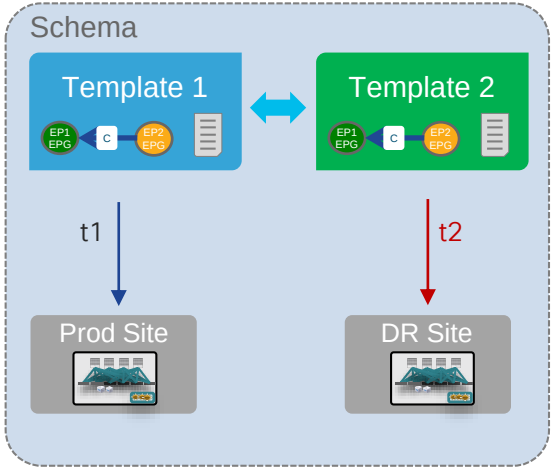


ACI Multi-Site

Schema and Templates Definition for the DR Use Case

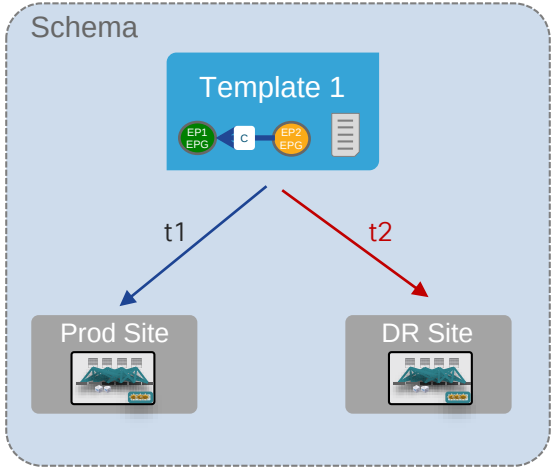


- Single Template associated to Prod and DR Sites
- Any change applied to the template is pushed to both sites simultaneously
- Easiest way to keep consistent policies deployed across sites



- Separate Template associated to Prod and DR Sites (can use cloning)
- Changes made to a template can be applied only to the mapped site
- Requires sync between the two templates (manual or performed by an higher level Orchestration tool)

Future

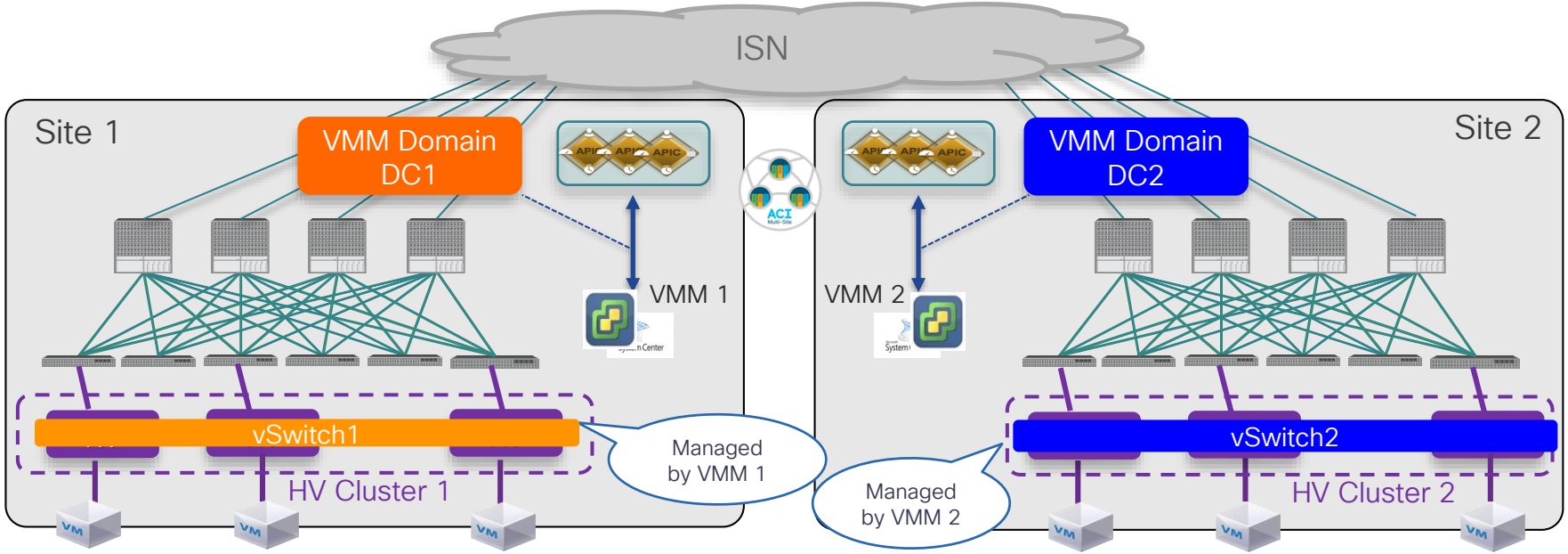


- Single Template associated to Prod and DR Sites
- Capability of independently apply changes to each site
- Brings together the advantages of the previous two options

Multi-Site and Virtual Machine Manager (VMM) Integration

ACI Multi-Site and VMM Integration

Option 1 – Separate VMM per Site

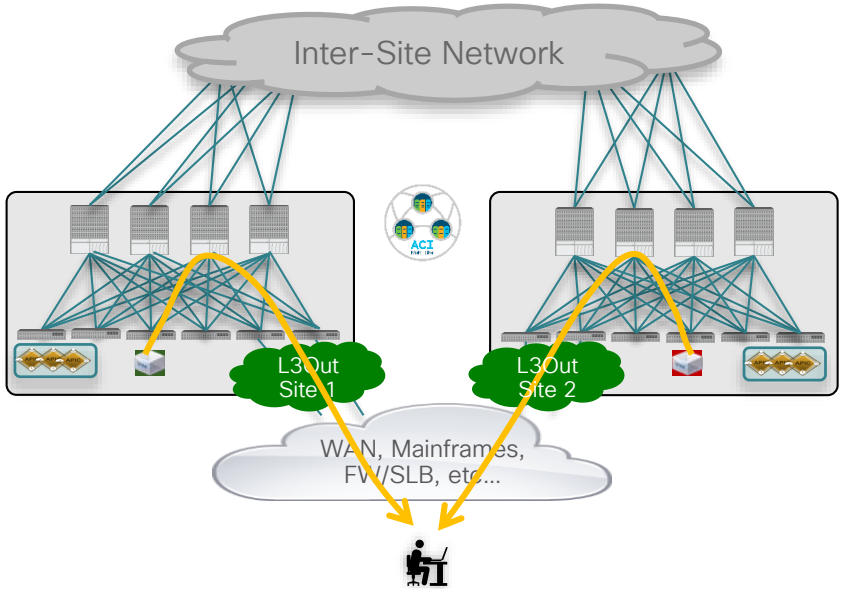


- Typical deployment model for an ACI Multi-Site
- Creation of separate VMM domains in each site, which are then exposed to the Multi-Site Orchestrator

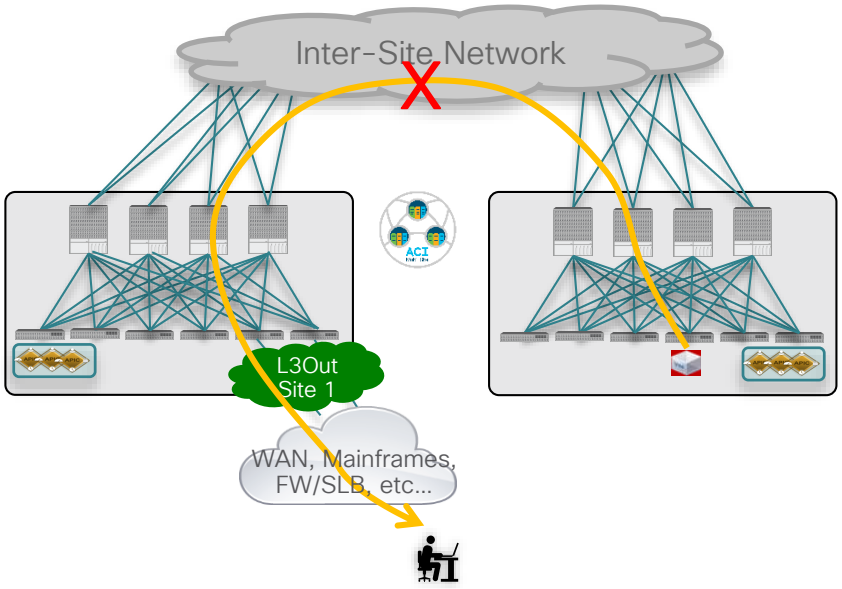
Multi-Site Connectivity to the External L3 Domain

Problem Statement
Behavior before ACI Release 4.2(1)

Supported Design



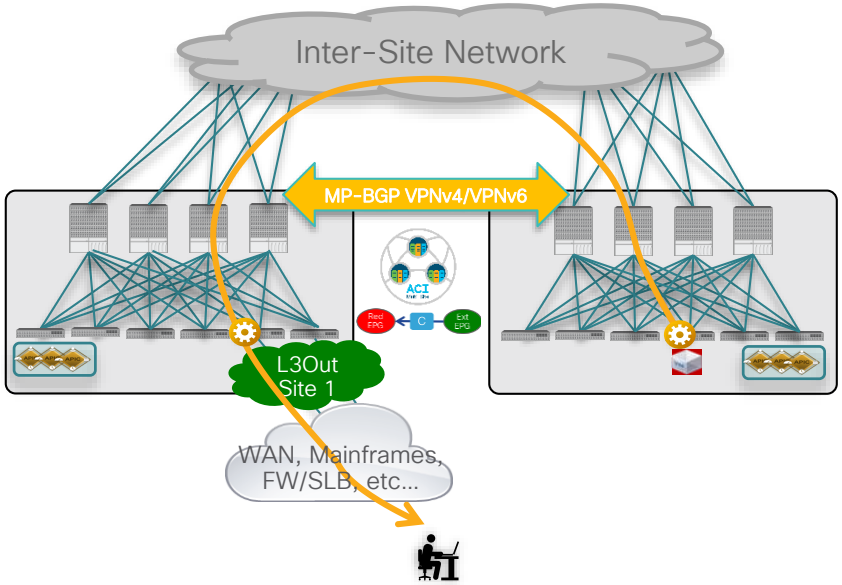
Not Supported Design



Note: the same consideration applies to both Border Leaf L3Outs and GOLF L3Outs

ACI Multi-Site and L3Out
Support of Intersite L3Out

ACI 4.2(1) Release

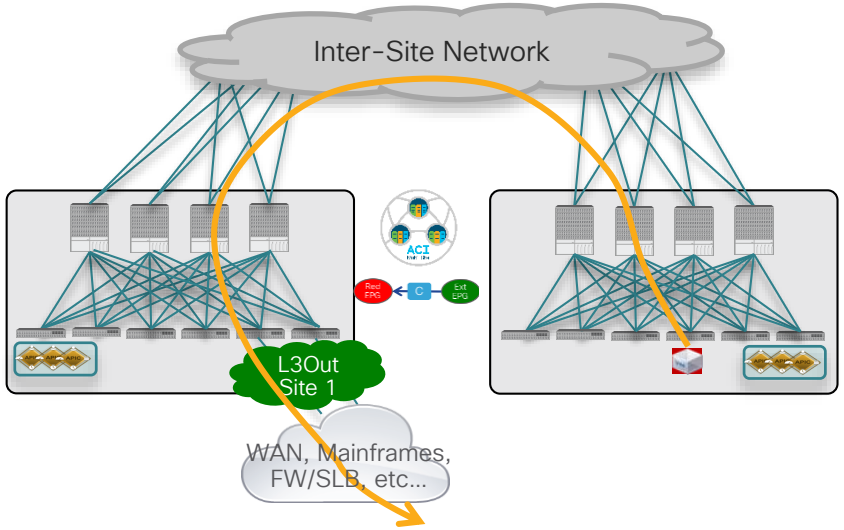


- Starting with ACI Release 4.2(1) it is possible for endpoints in a site to send traffic to resources (WAN, Mainframes, FWs/SLBs, etc.) accessible via a remote L3Out connection
- External prefixes are exchanged across sites via MP-BGP VPNV4/VPNv6 sessions between spines
- Traffic will be directly encapsulated to the TEP of the remote BL nodes
  - The BL nodes will get assigned an address part of an additional (configurable) prefix that must be routable across the ISN
  - This routable TEP pool can be configured on MSO or on APIC
- Same solution will also support transit routing across sites (L3Out to L3Out)

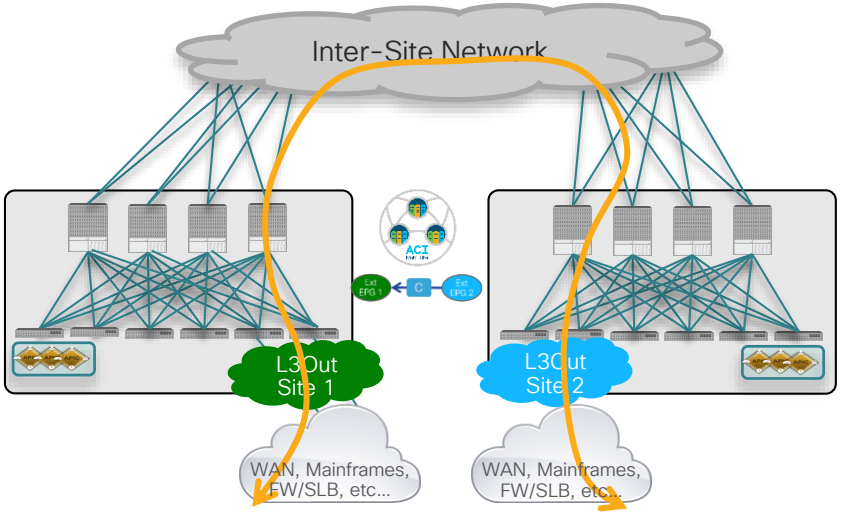
ACI Multi-Site and Intersite L3Out

Supported Scenarios

ACI 4.2(1) Release



- Endpoint to remote L3Out communication (intra-VRF)
- Endpoint to remote L3Out communication (inter-VRF)

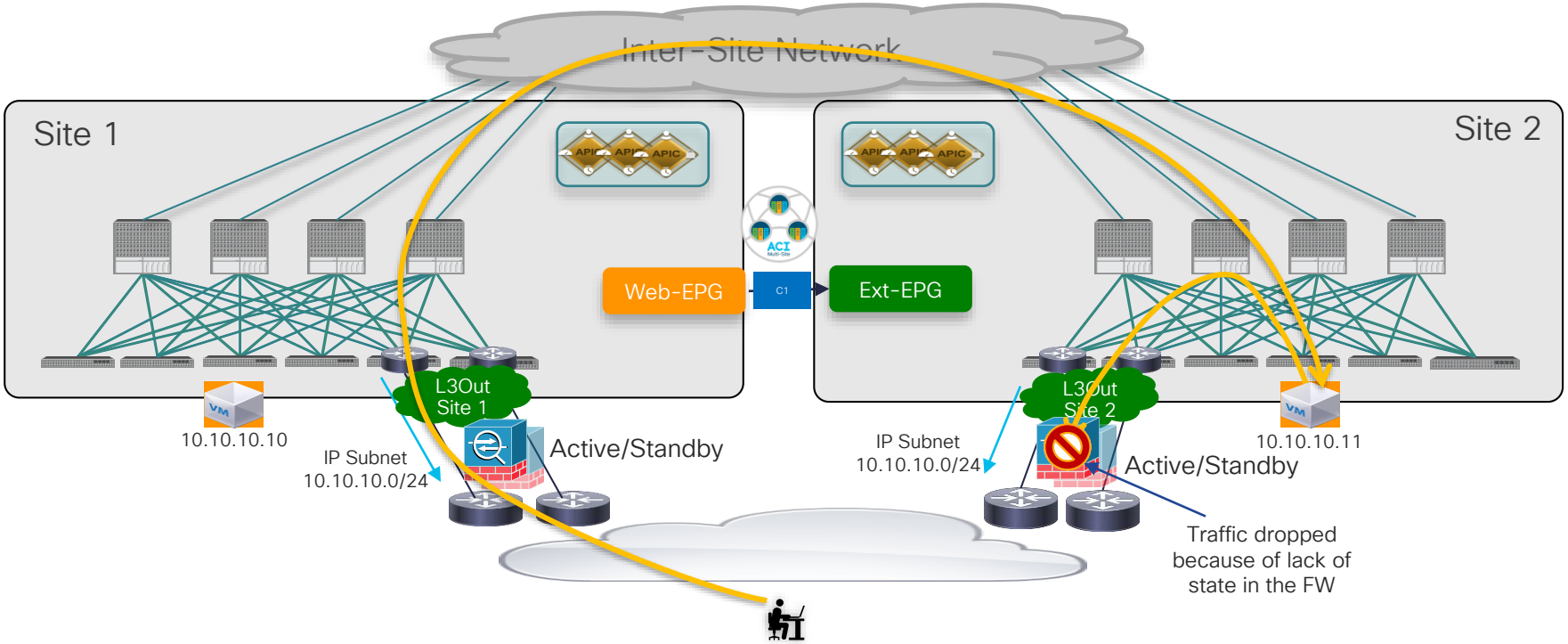


- Inter-site transit routing (intra-VRF)
- Inter-site transit routing (inter-VRF)

Solving Asymmetric Routing Issues with the External Network

Multi-Site and L3Out

Endpoints Normally Use Local L3Outs for Outbound Traffic



Solving Asymmetric Routing Issues

Use of Host-Routes Advertisement

ACI 4.0(1) Release

Inter-Site Network

Site 1

Site 2

Web-EPG

Ext-EPG

L3Out Site 1

L3Out Site 2

10.10.10.10

10.10.10.11

Host routes 10.10.10.10/32

Host routes 10.10.10.11/32

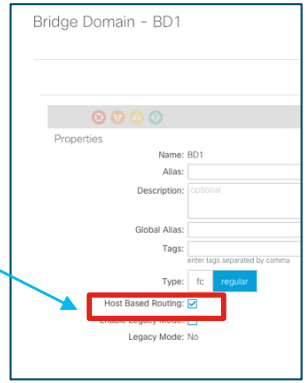
Active/Standby

Active/Standby

Host-routes injected into the WAN*

*Alternative could be running an overlay solution (LISP, GRE, etc.)

Enabled at the BD level

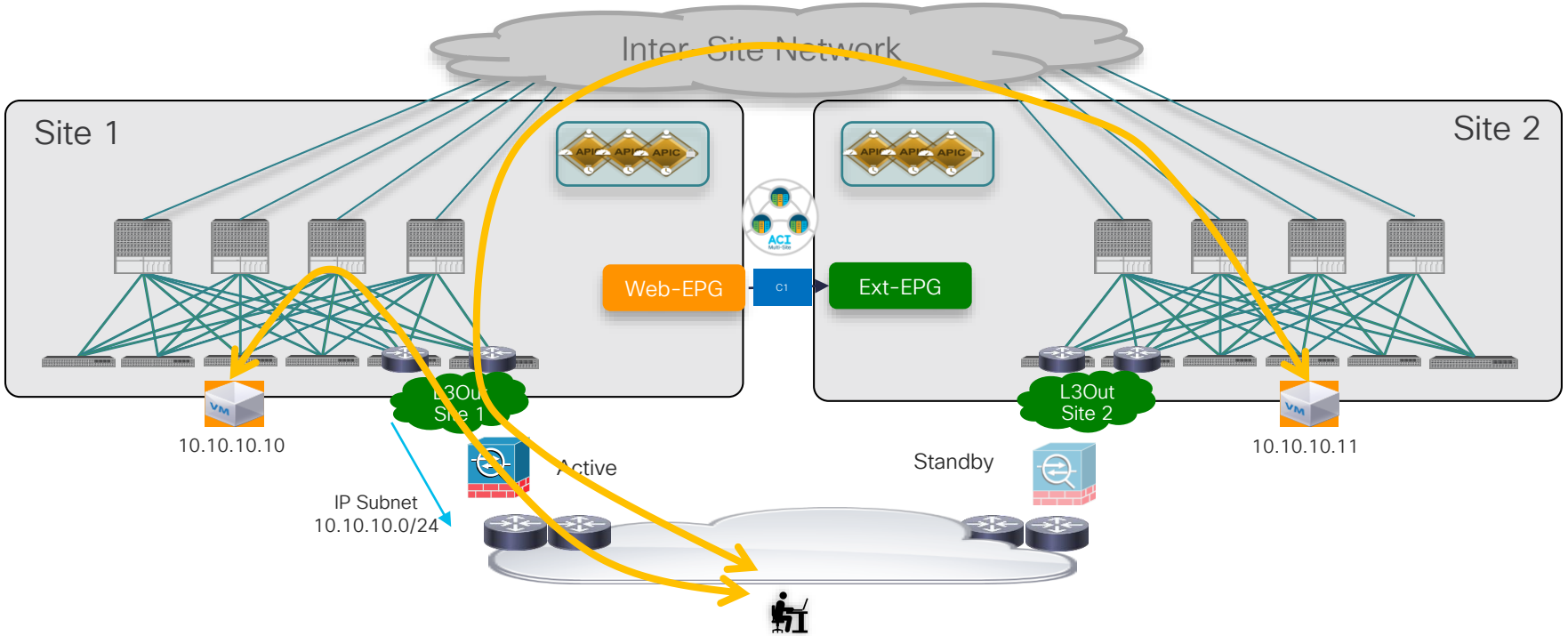


- Ingress optimization requires host-routes advertisement on the L3Out
  - Native support on ACI Border Leaf nodes available from ACI release 4.0(1)
  - Supported also on GOLF L3Outs (enabled at the VRF level)

Solving Asymmetric Routing Issues

Use of Active/Standby FW Pair Deployed across Sites

ACI 4.2(1) Release



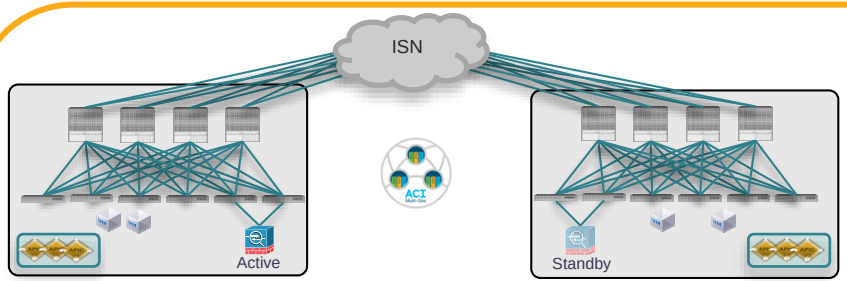
- Inbound and outbound flows are forced through the site with the active perimeter FW node
  - Common scenario in a Multi-Pod deployment, less desirable with Multi-Site
- <u>Requires Intersite L3Out support (ACI release 4.2(1))</u>

Multi-Site Network Services Integration

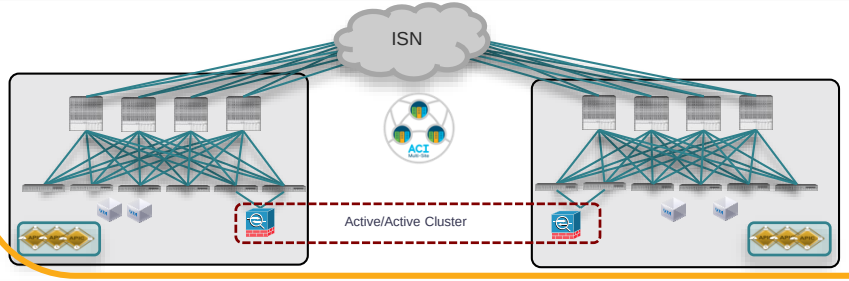
Multi-Site and Network Services

Integration Models

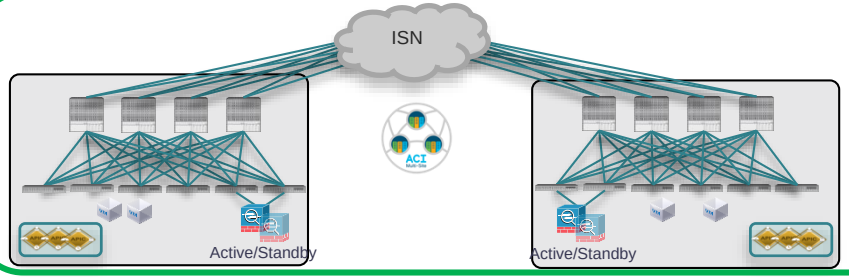
Deployment options fully supported with ACI Multi-Pod



- Active and Standby pair deployed across Pods
- Currently supported only if the FW is in L2 mode or in L3 mode but acting as default gateway for the endpoints
- From ACI 4.2(1) will be also supported as perimeter FW



- Active/Active FW cluster nodes stretched across Sites (single logical FW)
- Requires the ability of discovering the same MAC/IP info in separate sites at the same time
- Not supported



- **Recommended deployment model for ACI Multi-Site**
- Option 1: supported from 3.0 for N-S if the FW is connected in L3 mode to the fabric → mandates the deployment of traffic ingress optimization
- Option 2: supported from 3.2 release with the use of Service Graph with Policy Based Redirection (PBR)

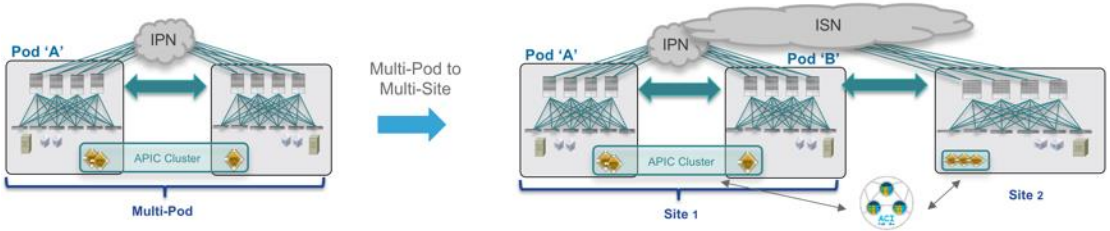
Multi-Site Integration with ACI Multi-Pod

ACI Multi-Pod and Multi-Site

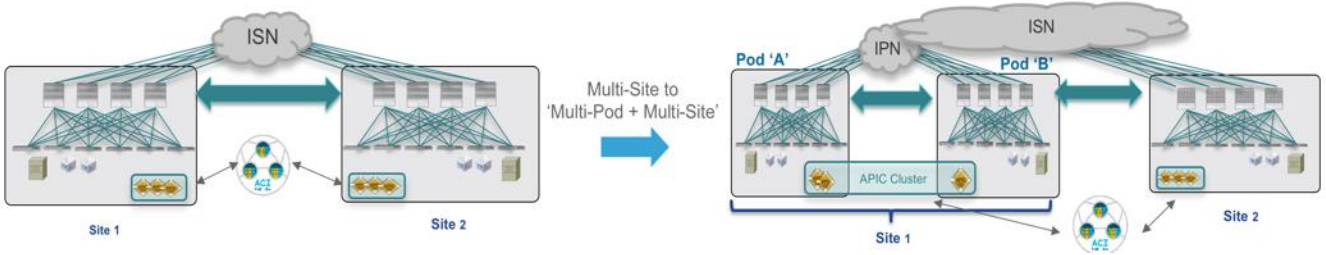
Main Use Cases

ACI 3.2(1) Release

- Adding a Multi-Pod Fabric as a 'Site' on the Multi-Site Orchestrator (MSO)

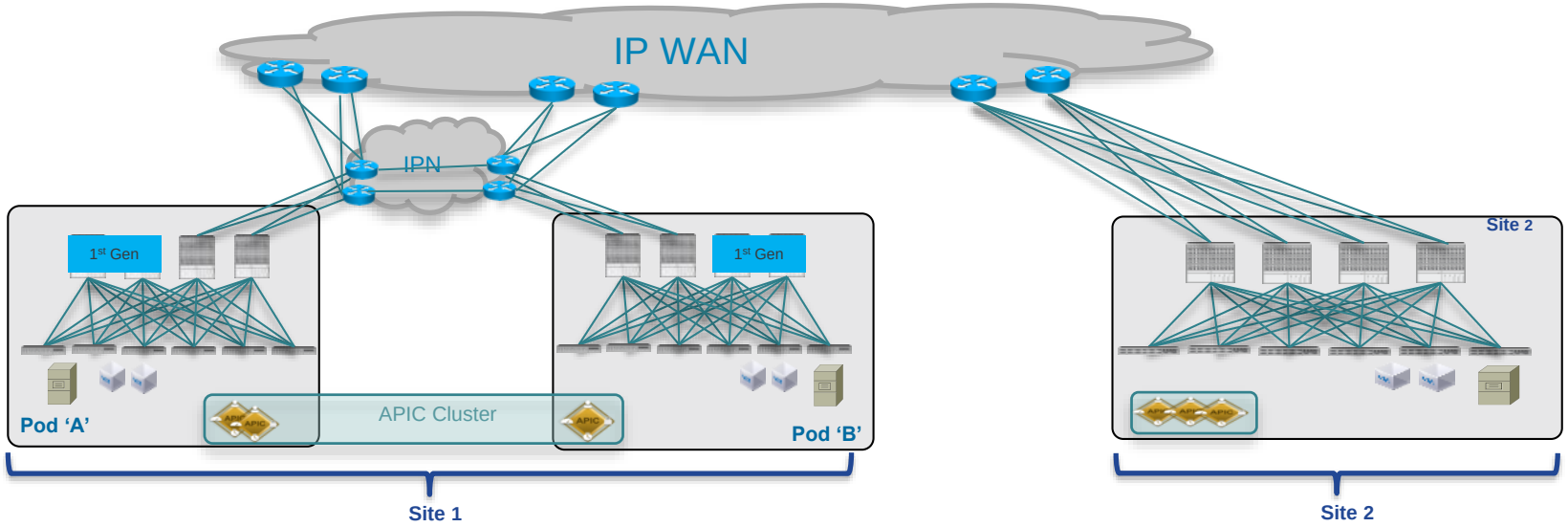


- Converting a single Pod Fabric (already added to MSO) to a Multi-Pod fabric



ACI Multi-Pod and Multi-Site

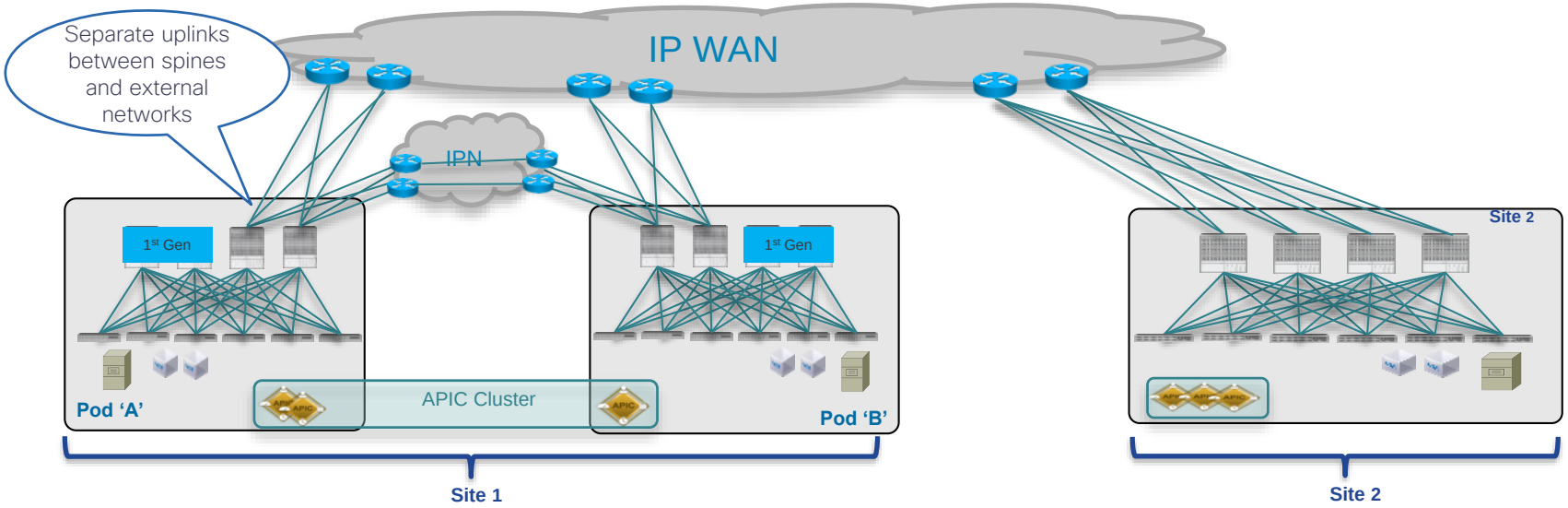
Connectivity between Pods and Sites



- Only 2nd generation spines must be connected to the external network
  - Need to add 2nd gen spines in each Pod (at least two per Pod) and migrate connections to the IPN from 1st gen spines to 2nd gen spines
- Single 'infra' L3Out and set of uplinks to carry both Multi-Pod and Multi-Site East-West traffic

Connectivity between Pods and Sites

Not Supported Topology



- Only 2nd generation spines must be connected to the external network
  - Need to add 2nd gen spines in each Pod (at least two per Pod) and migrate connections to the IPN from 1st gen spines to 2nd gen spines
- Single 'infra' L3Out and set of uplinks to carry both Multi-Pod and Multi-Site East-West traffic

ACI Multi-Site

Where to Go for More Information

- ✓ ACI Multi-Site White Paper

  https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-739609.html

- ✓ Deploying ACI Multi-Site from Scratch

  https://www.youtube.com/watch?v=HJJ8lznodN0

- ✓ BRKACI-2125

Agenda

- ACI Anywhere, Extending the ACI Fabric
  - Overall Design Principles (AZs and Regions)
  - Mapping use cases to the proper solutions
    - Active/Active DC → Multi-Pod
    - Disaster Recovery → Multi-Site
    - Migration/Coexistence with Legacy DC Networks and 'Disaggregated DCs' Model → Physical Remote Leaf
    - Baremetal Cloud Integration → Virtual Pod (vPod)
  - Extending ACI to the Cloud
  - Connecting the users to the Multi-Cloud DC
    - ACI and SDA Integration
    - ACI and SDWAN Integration

ACI Remote Physical Leaf

Business Value and Use Cases

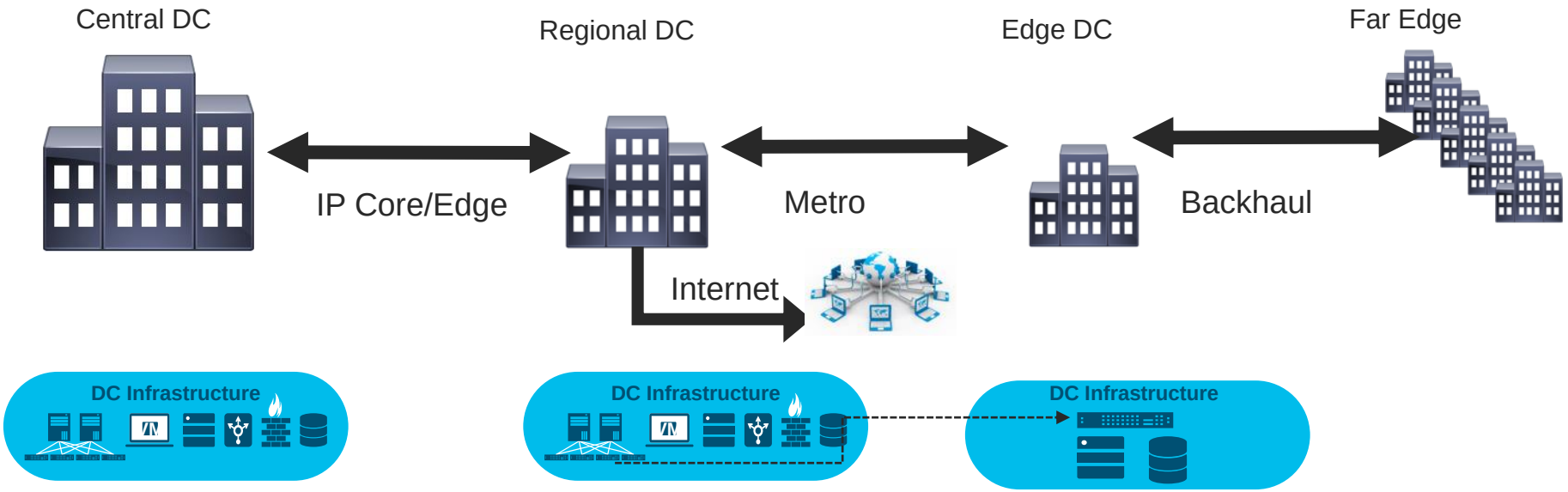
For More Information on ACI Remote Leaf: BRKACI-2387

- Extending the ACI policy model outside the main datacenter to remote sites distributed over IP Backbone (Telco DCs, CoLo locations, etc.)
- Extending ACI fabric policy and L2/L3 connectivity to a small DC site without requiring the deployment of a full-blown ACI Fabric or for migration/coexistence with legacy DC sites
- Centralized Policy Management and Control Plane for remote locations
- Small form factor solution at locations with space constraints

ACI Remote Physical Leaf

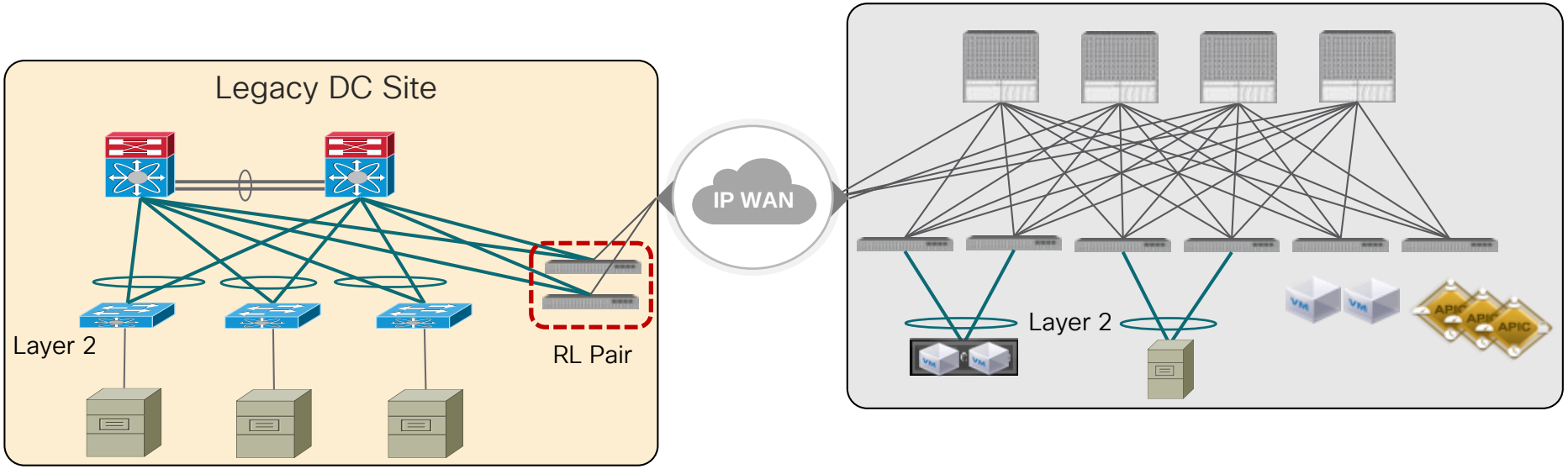
‘Disaggregated DC Model” for 5G Deployment

- Increased SP DC Footprint with edge transformation, vRAN
- Application/Services are distributed (CUPS, CU-DU split, etc.)



ACI Remote Physical Leaf

Migration/Coexistence Use Case



- Connecting a greenfield ACI fabric to a legacy DC location
  - Single point of management
  - Coexistence
  - Application migration

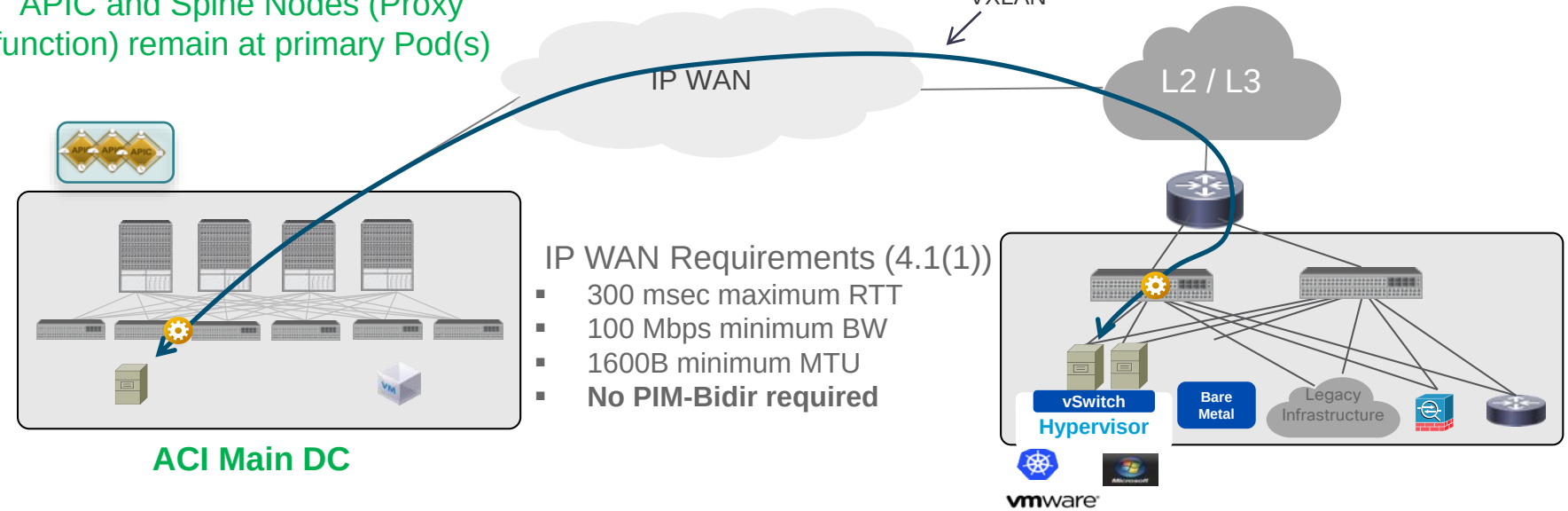
ACI Remote Physical Leaf

Conceptual Architecture

ACI 3.1(1) Release

A Remote Leaf 'Site' gets associated with the Spines of one specific Pod in Main DC

APIC and Spine Nodes (Proxy function) remain at primary Pod(s)



ACI Main DC

IP WAN Requirements (4.1(1))

- 300 msec maximum RTT
- 100 Mbps minimum BW
- 1600B minimum MTU
- **No PIM-Bidir required**

Remote Leaf Site: a pair of Nexus 9300 nodes connected to a L3 Network via uplink ports and fully managed by a centralized APIC cluster

ACI Remote Physical Leaf

Hardware/Software Support

ACI Main DC

Supported Spines

Fixed

- 9364C/9332C

Modular

- 9732C-EX
- 9736C-FX
- 9736Q-FX

Remote Location

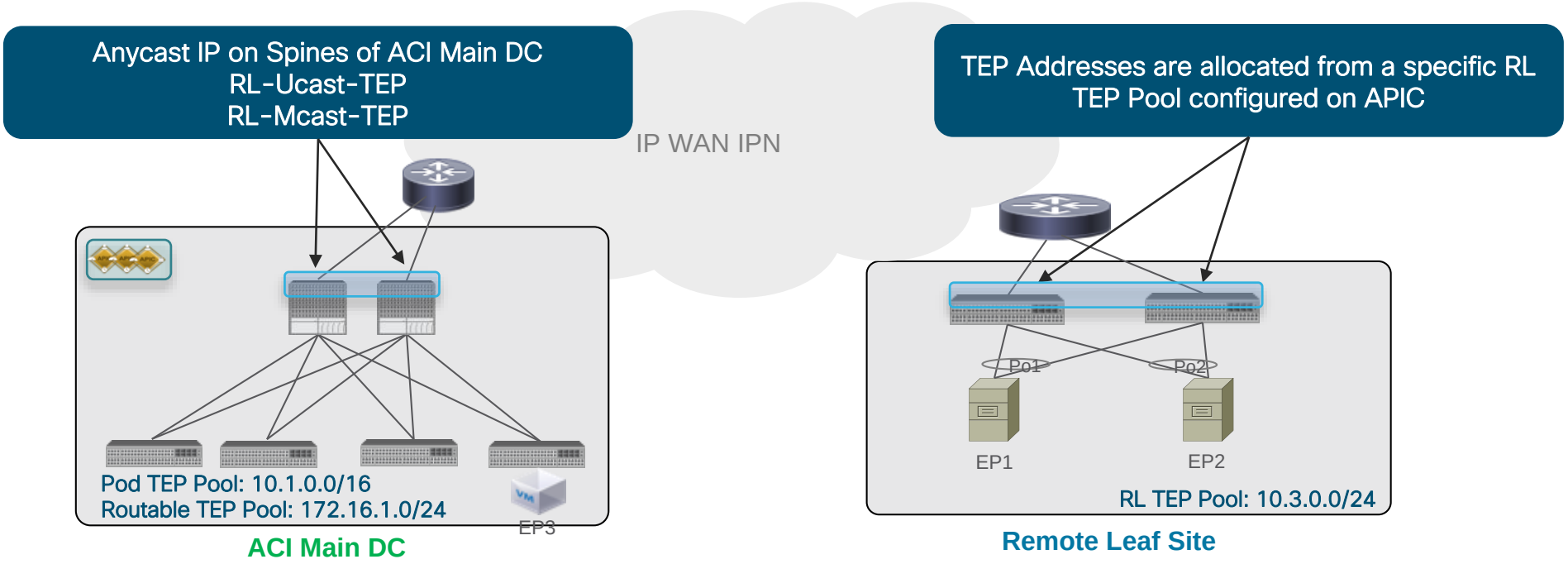
Supported Leaf

- N93180YC-EX
- N93108TC-EX
- N93180LC-EX
- N93180YC-FX
- N93180TC-FX
- N9348GC-FXP
- N9358GY-FXP
- N93240YC-FX2
- N9336C-FX2
- N93360YC-FX2
- N93216TC-FX2

All hardware from -EX onwards is supported

ACI Remote Physical Leaf

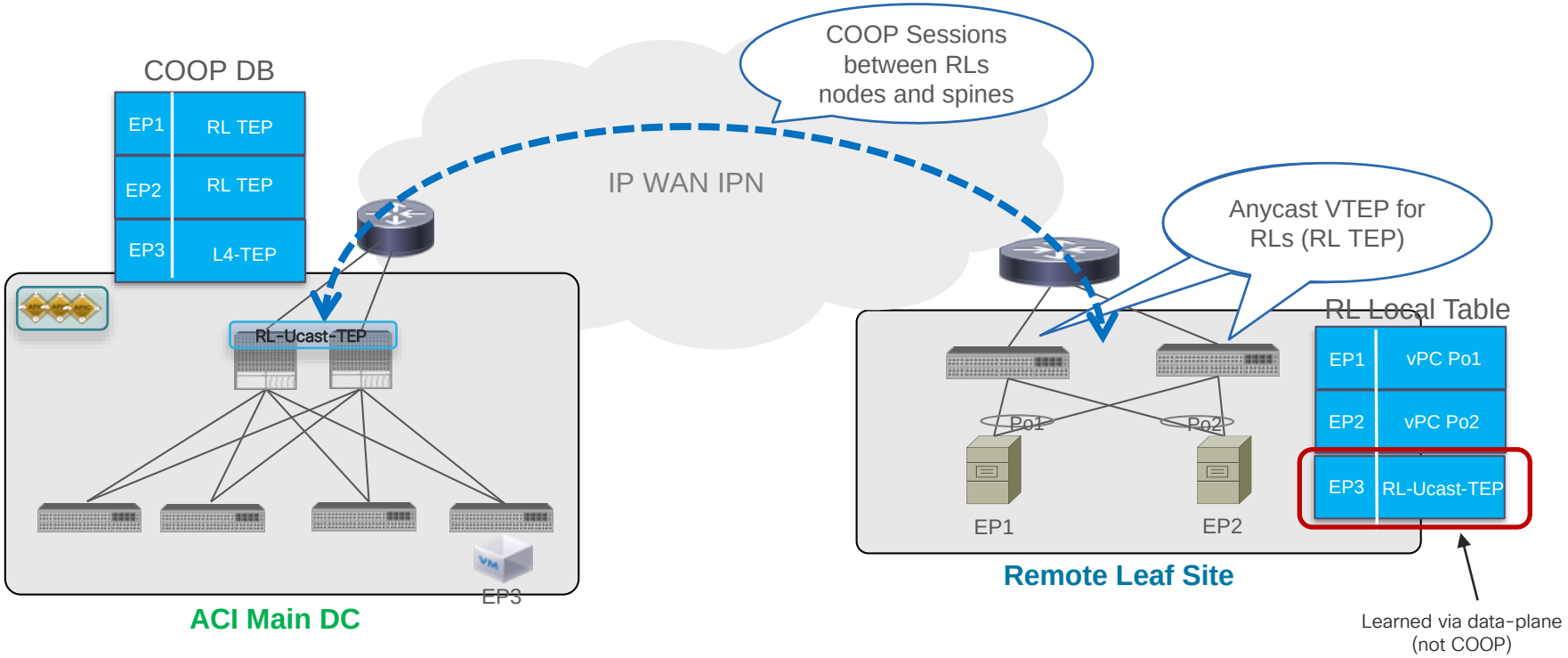
TEP Address Pools



RL-Ucast-TEP and RL-Mcast-TEP address are allocated from the routable TEP pool associated to the Pod

ACI Remote Physical Leaf

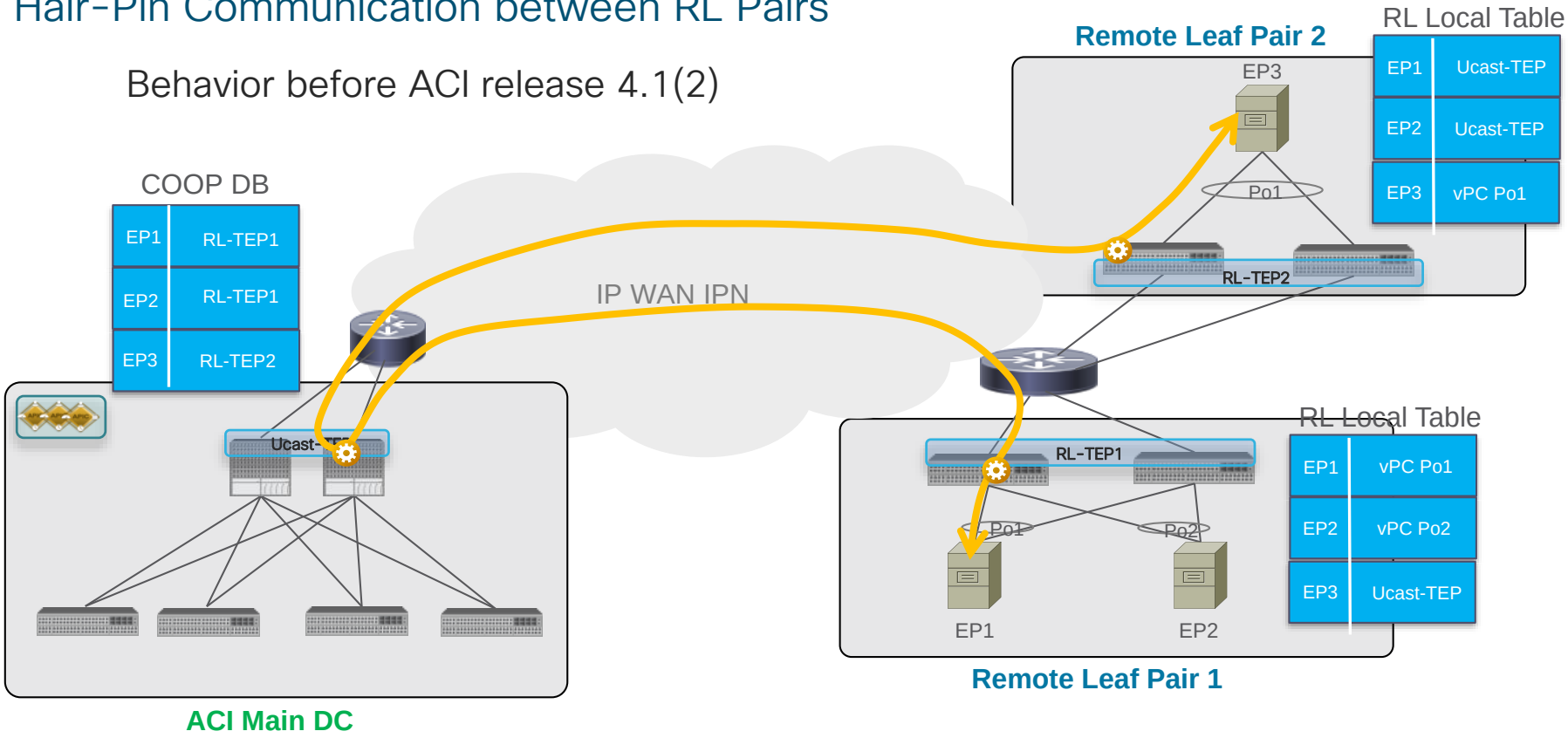
COOP for Announcing Remote Endpoint Information



ACI Remote Physical Leaf

Hair-Pin Communication between RL Pairs

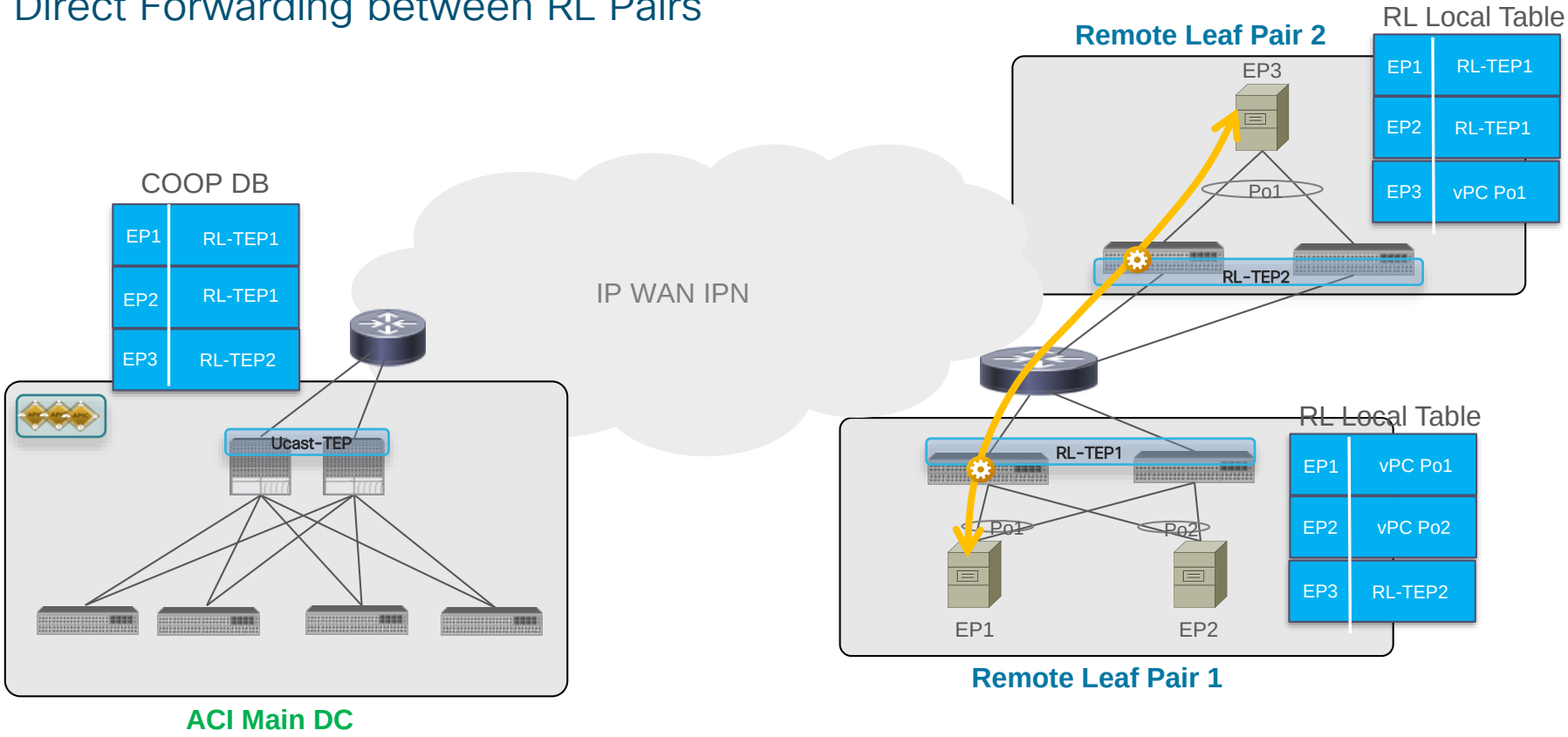
Behavior before ACI release 4.1(2)



ACI Remote Physical Leaf

Direct Forwarding between RL Pairs

ACI 4.1(2) Release



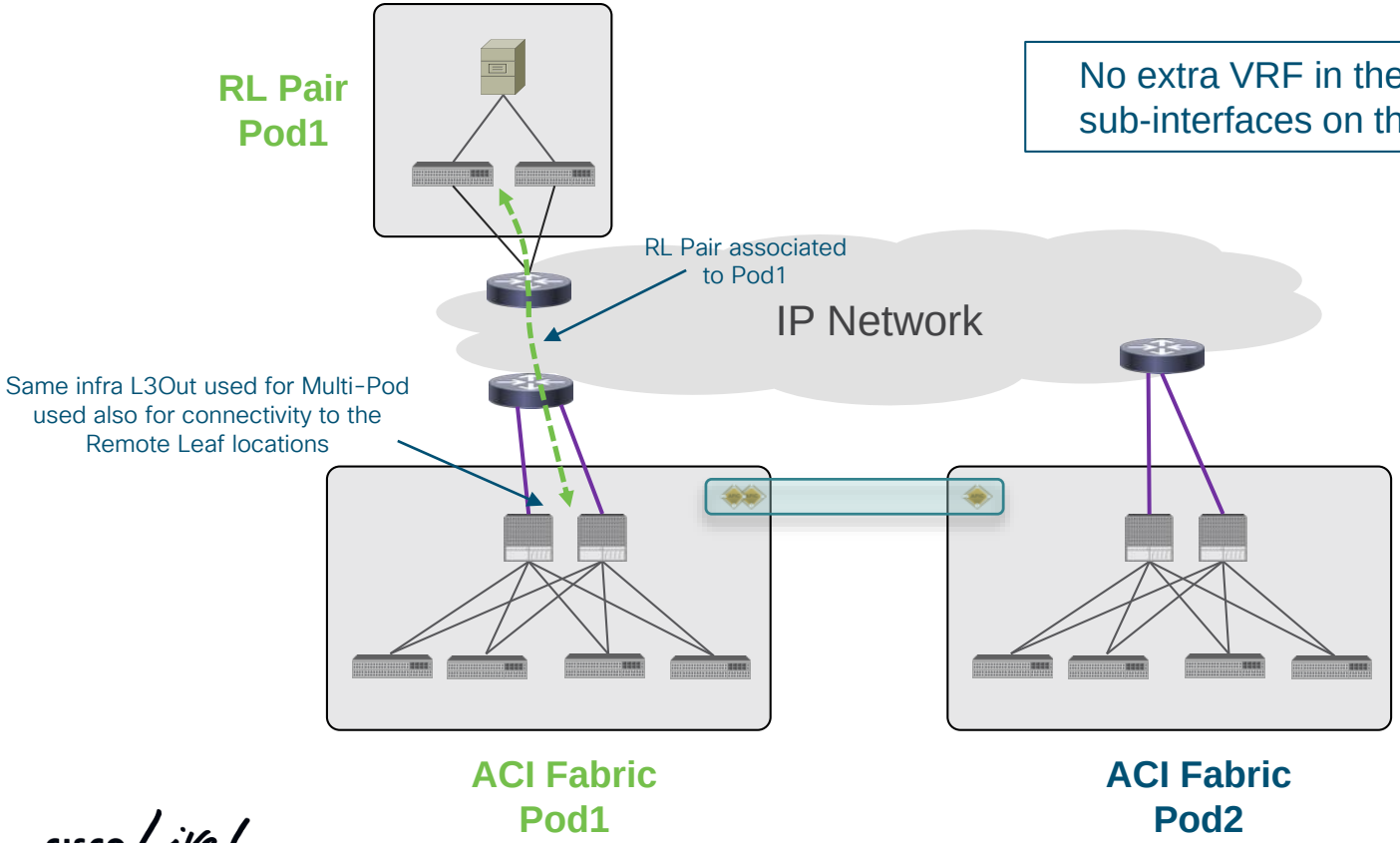
Integrating ACI Remote Leaf with Multi-Pod

ACI Remote Physical Leaf

Integration with Multi-Pod

ACI 4.1(2) Release

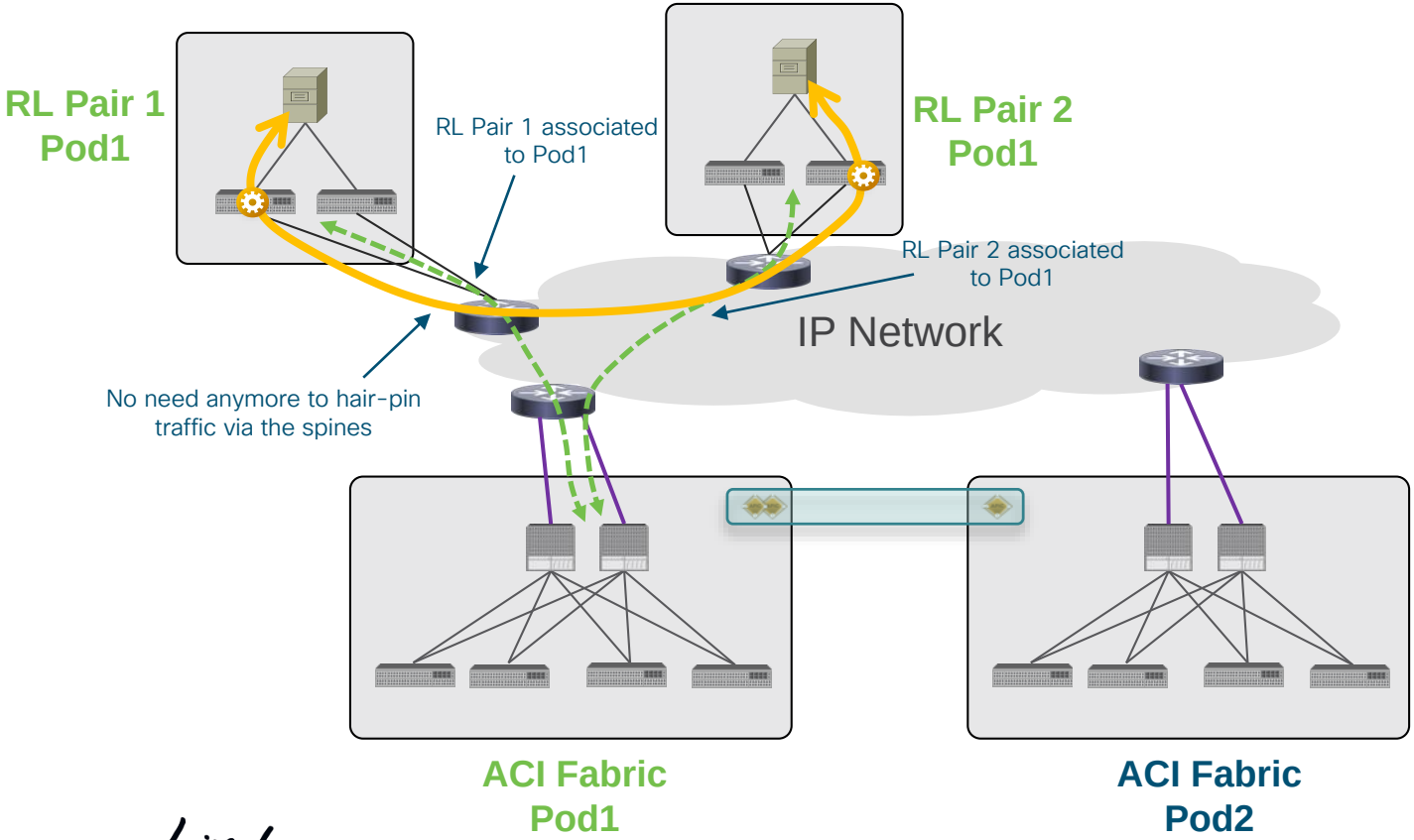
No extra VRF in the IPN and VLAN-5 sub-interfaces on the spines required



ACI Remote Leaf with Multi-Pod

Direct Forwarding between RL Pairs Part of the Same Pod

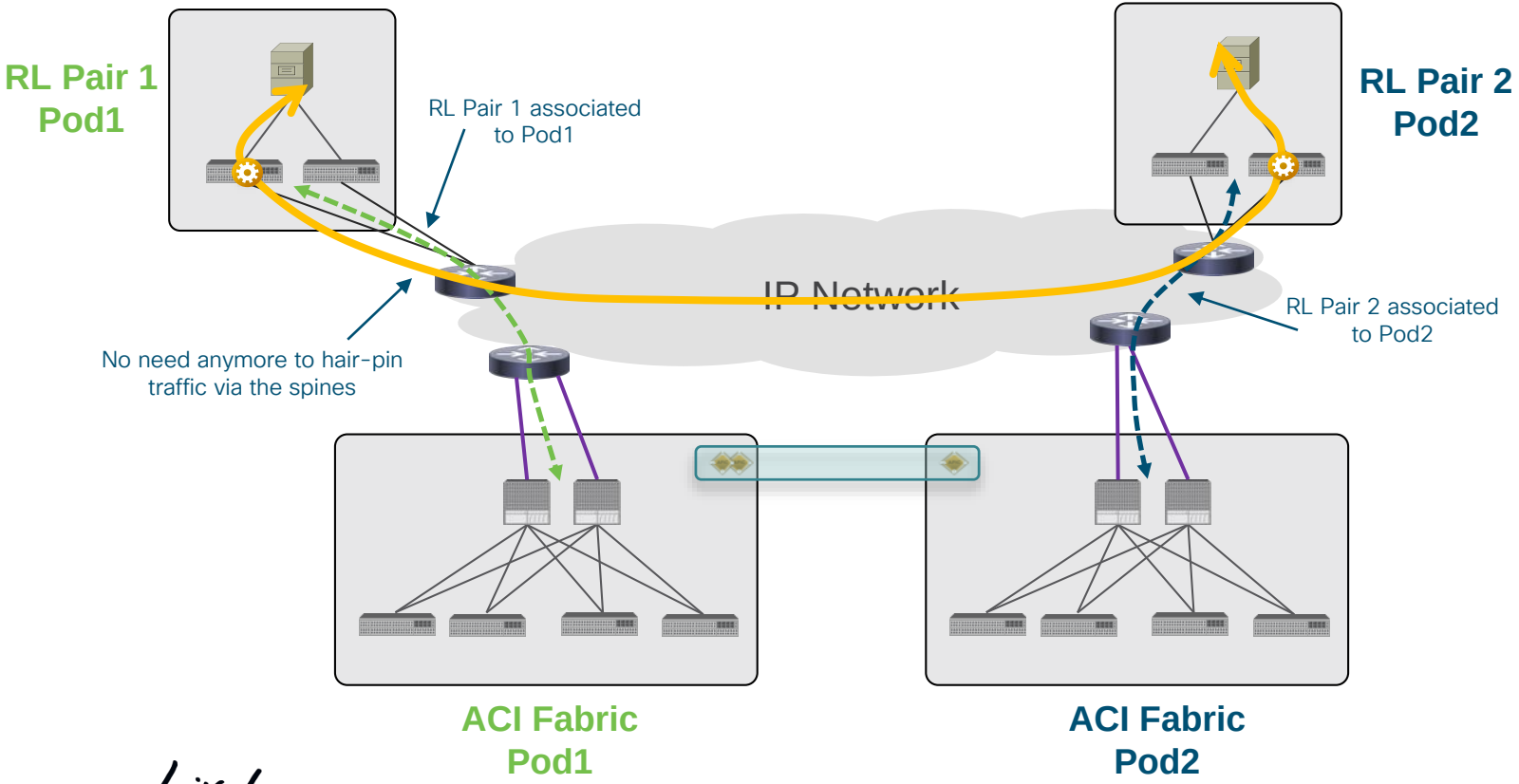
ACI 4.1(2) Release



ACI Remote Leaf with Multi-Pod

Direct Forwarding between RL Pairs Part of Different Pods

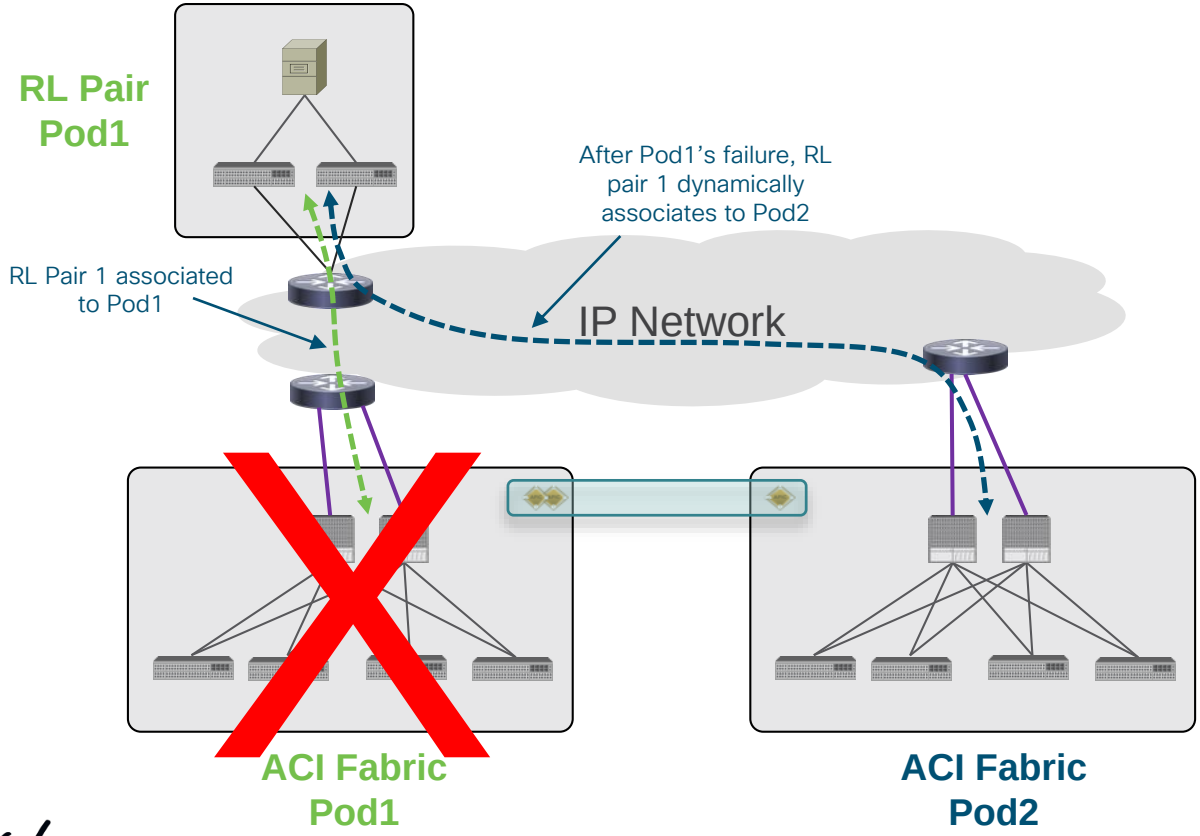
ACI 4.1(2) Release



ACI Remote Physical Leaf

RL Pair Resiliency in a Pod Failure Scenario

ACI 4.2(1) Release

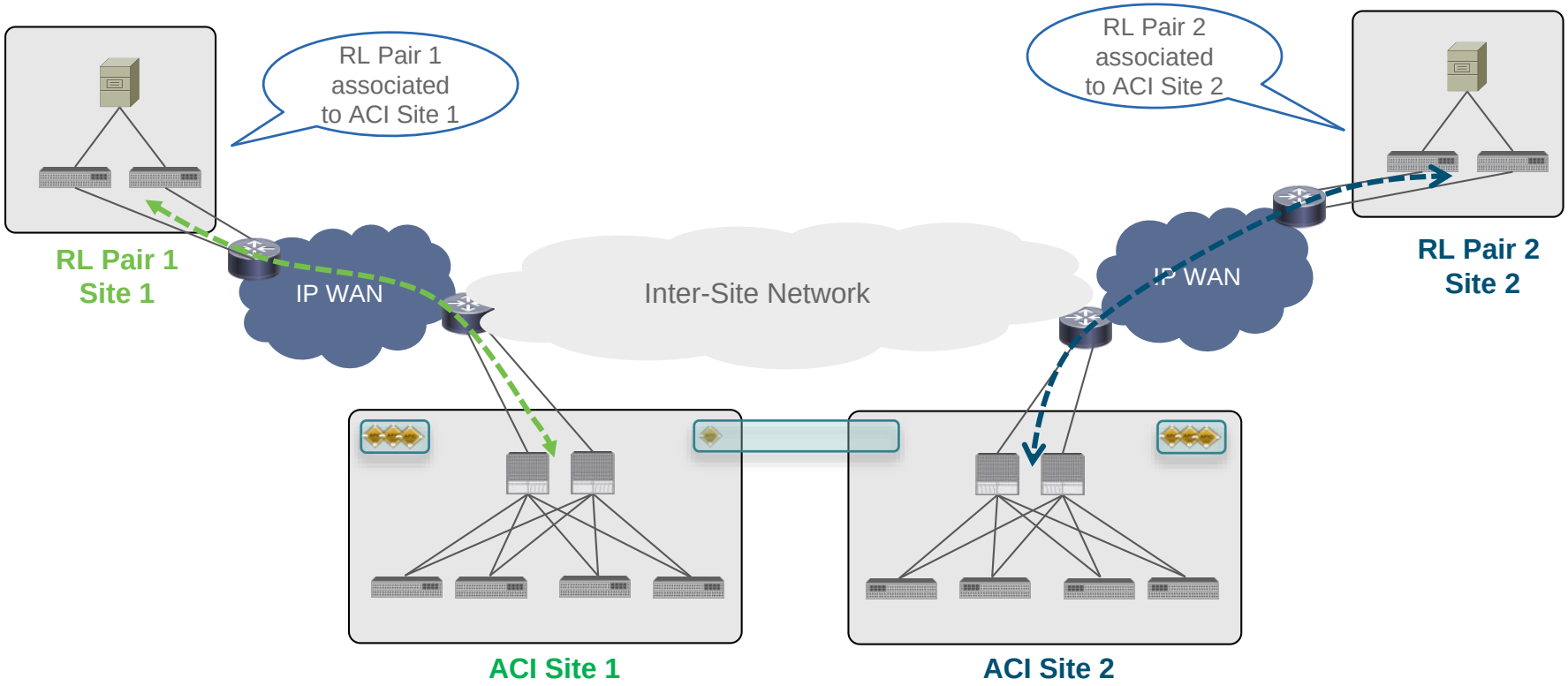


Integrating ACI Remote Leaf with Multi-Site

ACI Remote Physical Leaf and Multi-Site

RL Sites Can Be Associated to Separate Pods

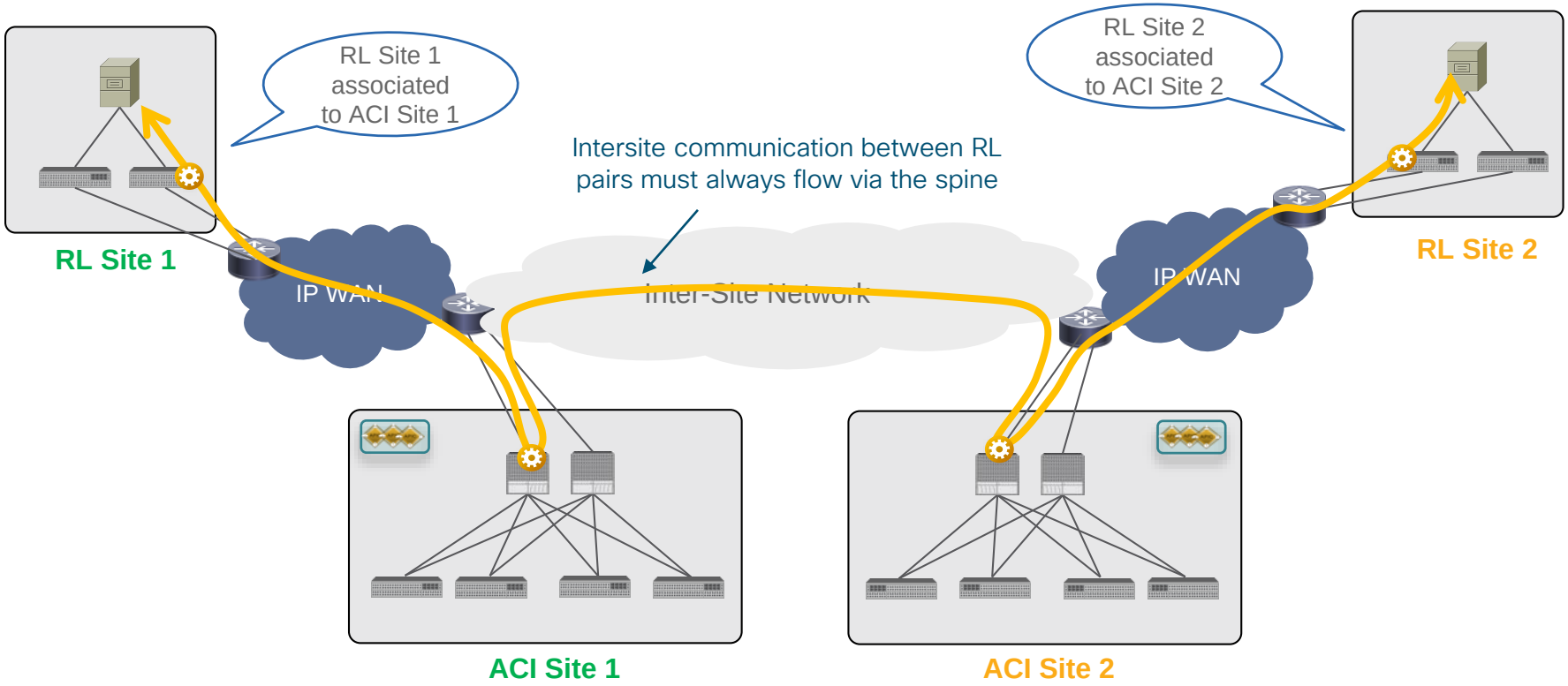
ACI 4.1(2) Release



ACI Remote Physical Leaf and Multi-Site

RL Sites Can Be Associated to Separate Pods

ACI 4.1(2) Release



ACI Physical Remote Leaf

Where to Go for More Information

- ✓ ACI Remote Physical Leaf White Paper

  https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-740861.html

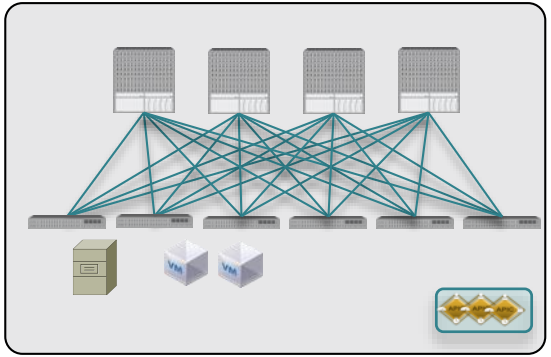
- ✓ BRKACI-2387

Agenda

- ACI Anywhere, Extending the ACI Fabric
  - Overall Design Principles (AZs and Regions)
  - Mapping use cases to the proper solutions
    - Active/Active DC → Multi-Pod
    - Disaster Recovery → Multi-Site
    - Migration/Coexistence with Legacy DC Networks and 'Disaggregated DCs' Model → Physical Remote Leaf
    - Baremetal Cloud Integration → Virtual Pod (vPod)
  - Extending ACI to the Cloud
  - Connecting the users to the Multi-Cloud DC
    - ACI and SDA Integration
    - ACI and SDWAN Integration

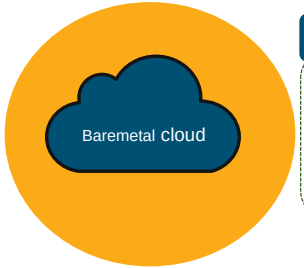
ACI Virtual Pod (vPod)

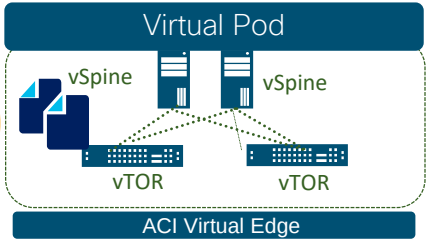
Main Use Cases



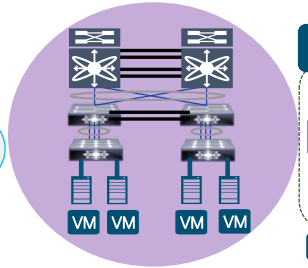
On-Premises ACI Data Center

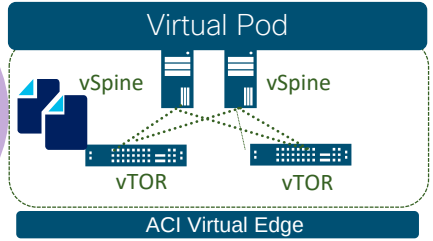
Baremetal Cloud



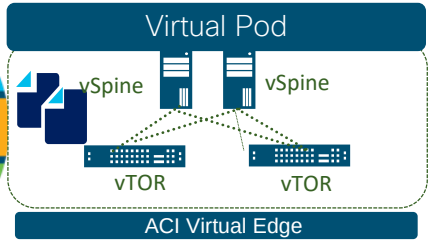


Brownfield



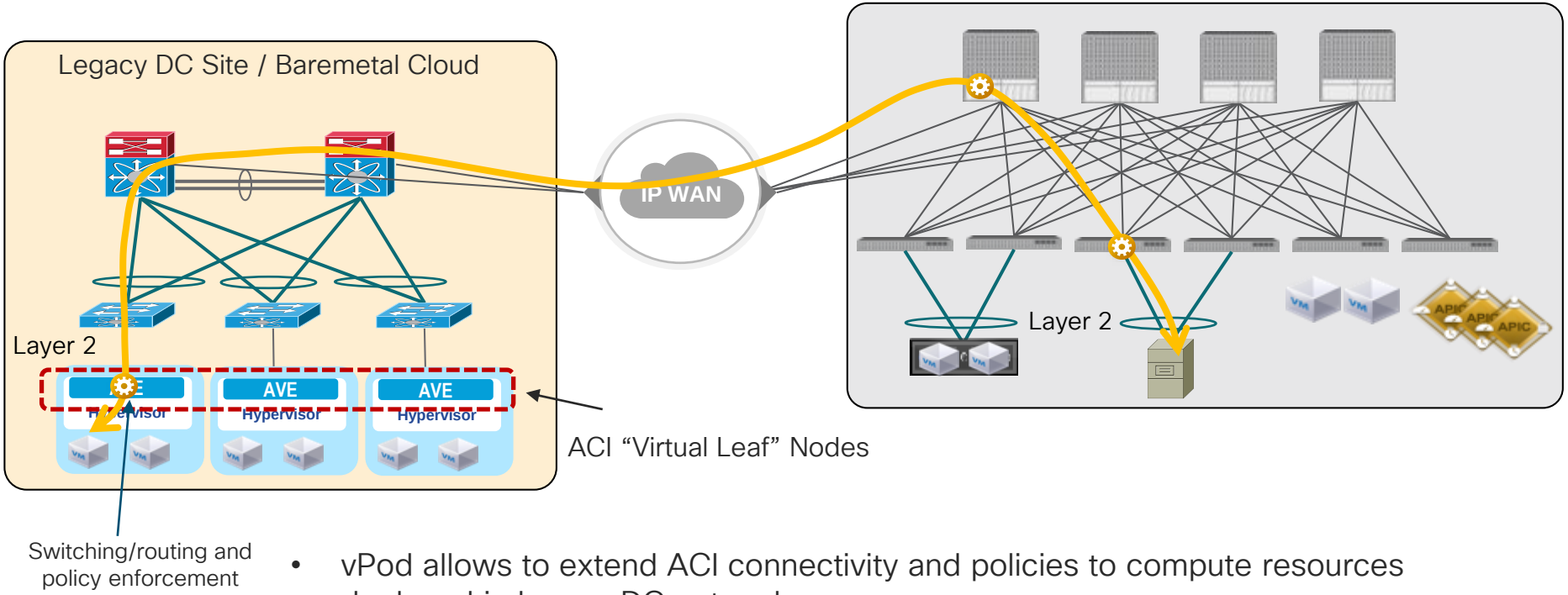


Colocation / Remote DC



ACI Virtual Pod

Extend ACI to Brownfield or Baremetal Cloud Locations

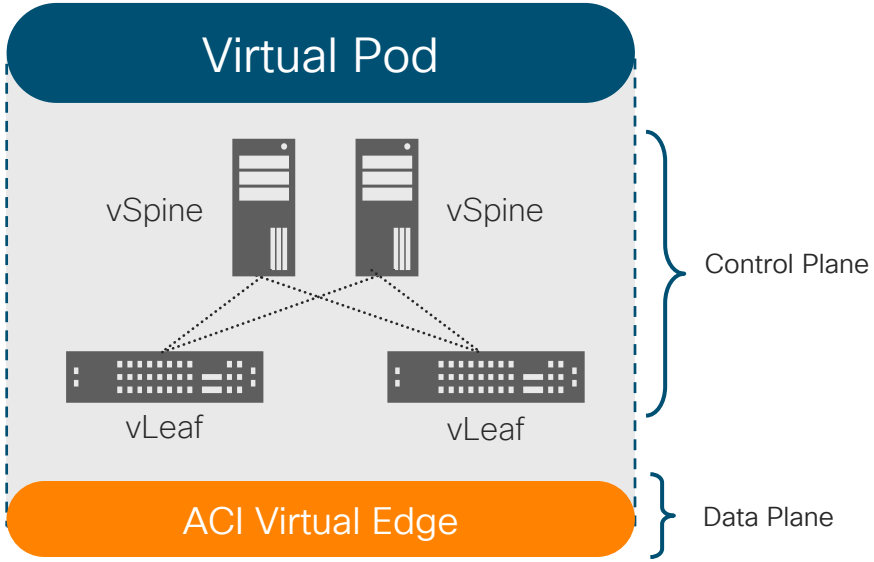


- vPod allows to extend ACI connectivity and policies to compute resources deployed in legacy DC networks
  - No need to deploy any ACI HW in the remote network

ACI Virtual Pod

Architectural Components

ACI 4.0(2) Release



Management Cluster (vSpine + vLeaf)

- vSpine nodes: centralized endpoint and LPM database (COOP and BGP)
- vLeaf nodes: distribute APIC policies to ACI Virtual Edges (DME/PE on vLeaf <-> Opflex on AVE)
- vSpine and vLeaf nodes are <u>not used</u> for data-plane forwarding

ACI Virtual Edge (vPod Mode)

- Implements ACI data plane function (switching and routing) <u>and</u> policy enforcement
- iVXLAN for communication within vPod and across Pods

ACI Virtual Pod

Software and Hardware Requirements

On-Premises Datacenter

Supported Spines

Fixed Spine:
- N9364C
- N9332C

Modular Spine LC: (C9504/C9508/C95016)
- N9732C-EX with FM N9K-C950x-FM-E(2)
- N9736C-FX with FM N9K-C950x-FM-E(2)

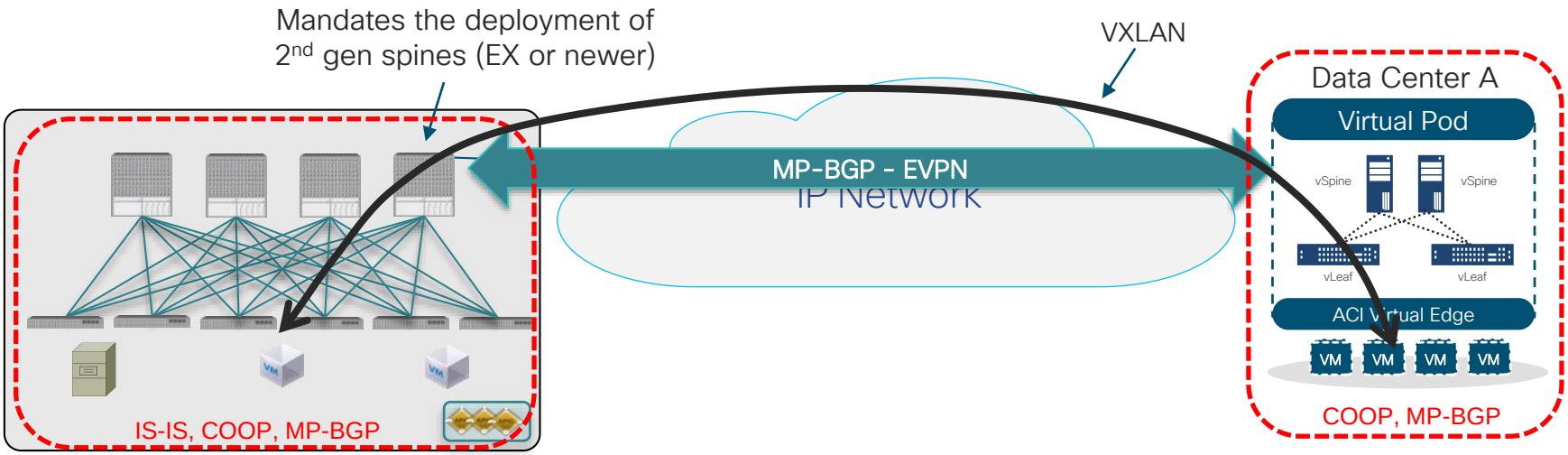
- APIC 4.0(2) onwards

vPod Datacenter

- VMware vCenter 6.0 or later
- 2 hosts for Management Cluster
- 2 hosts for Payload Cluster
- ESXi 6.0 or later
- ACI vCenter plugin or vPod python/PowerShell deployment scripts

ACI Virtual Pod

Control and Data Planes



On-Premises ACI Data Center

- Policies centrally defined on the APIC cluster deployed on-prem
- MP-BGP EVPN sessions established to exchange endpoint reachability information between Pods
- Ingress replication support on physical spines and AVEs to forward BUM traffic

vPod Management Plane

Fully Managed via APIC

Schedule Node Upgrade

Group type: Physical | Virtual

Upgrade Name: vFwGrp1
Select or create new

Target Firmware Version: 3.2(v1.3)

Upgrade start time: Now | Schedule for later

Ignore compatibility check: ☐

Graceful Maintenance: ☐

Run Mode: Pause upon upgrade failure

Node selection: Range | Manual

Only nodes that are not already in another Firmware Upgrade Group can be added in order to avoid a possible conflict in upgrades which may lead to unintentional outcome.

Selected	Node id	Node name	Model	Current Firmware	Status
Current Firmware: simsw-3.2(0.37) (5 Nodes)					
☐	101	vleaf1	N9K-C9396PX	simsw-3.2(0.37)	Not Scheduled
☐	102	vleaf2	N9K-C9396PX	simsw-3.2(0.37)	Not Scheduled

Cancel | Submit

ACI Virtual Pod

Where to Go for More Information

- ✓ Virtual Pod White Paper

  https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-742393.html

- ✓ BRKACI-2882 (Cisco Live San Diego 2019)

  https://www.ciscolive.com/global/on-demand-library.html?search=BRKACI-2882%20#/session/1542224297572001r8Mq

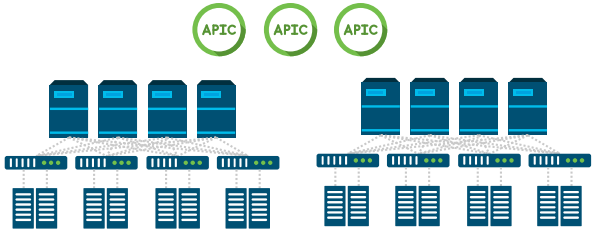
Agenda

- ACI Anywhere, Extending the ACI Fabric
  - Overall Design Principles (AZs and Regions)
  - Mapping use cases to the proper solutions
    - Active/Active DC → Multi-Pod
    - Disaster Recovery → Multi-Site
    - Migration/Coexistence with Legacy DC Networks and 'Disaggregated DCs' Model → Physical Remote Leaf
    - Baremetal Cloud Integration → Virtual Pod (vPod)
  - Extending ACI to the Cloud
  - Connecting the users to the Multi-Cloud DC
    - ACI and SDA Integration
    - ACI and SDWAN Integration

ACI Anywhere

Edge / Remote | **Core Data Centers** | **Multicloud**

Virtual ACI — IP WAN — ACI — IP WAN — Cloud ACI

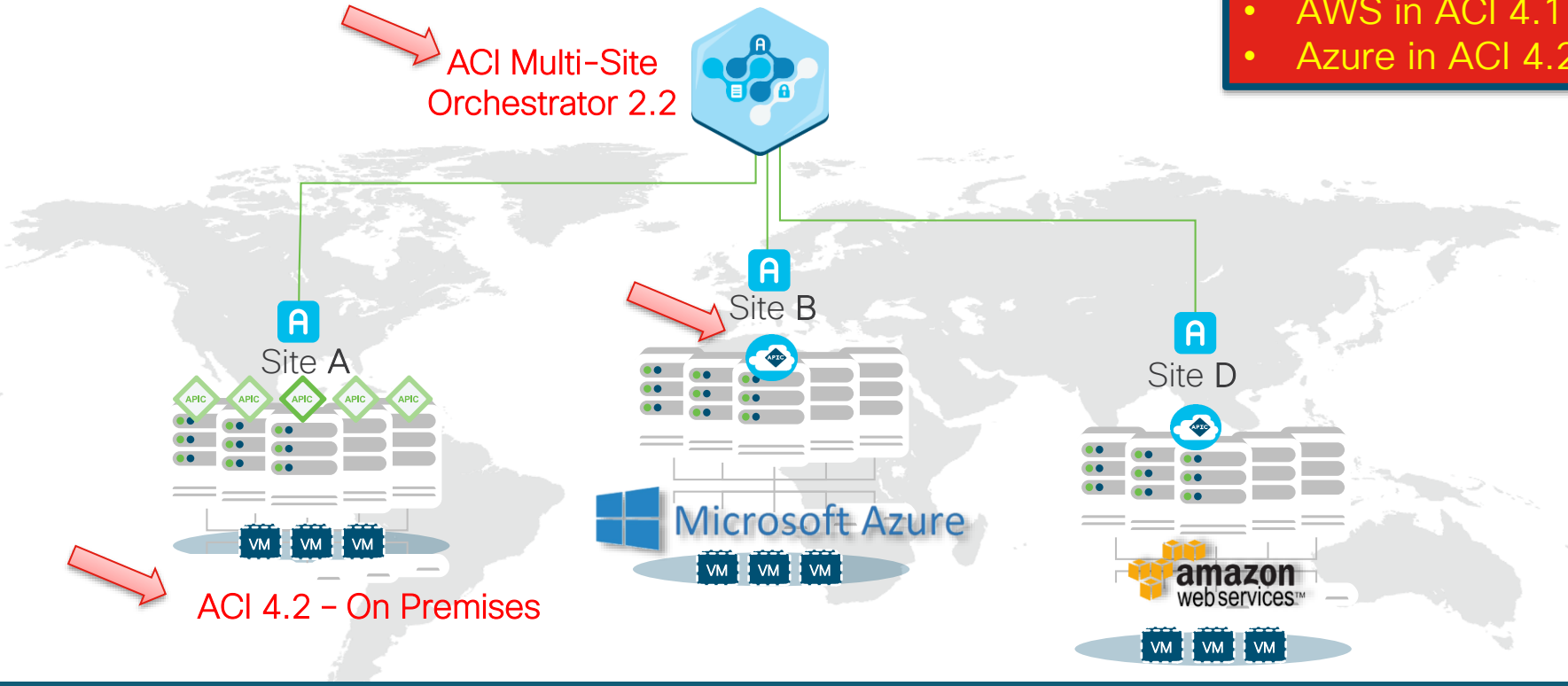


ACI **Multi-POD**	ACI **Multisite**	ACI **Remote Leaf**	**Virtual** ACI	**Cloud** ACI
ACI 2.0	**ACI 3.0**	**ACI 3.1**	**ACI 4.0**	**ACI 4.1 | ACI 4.2**

aws Azure

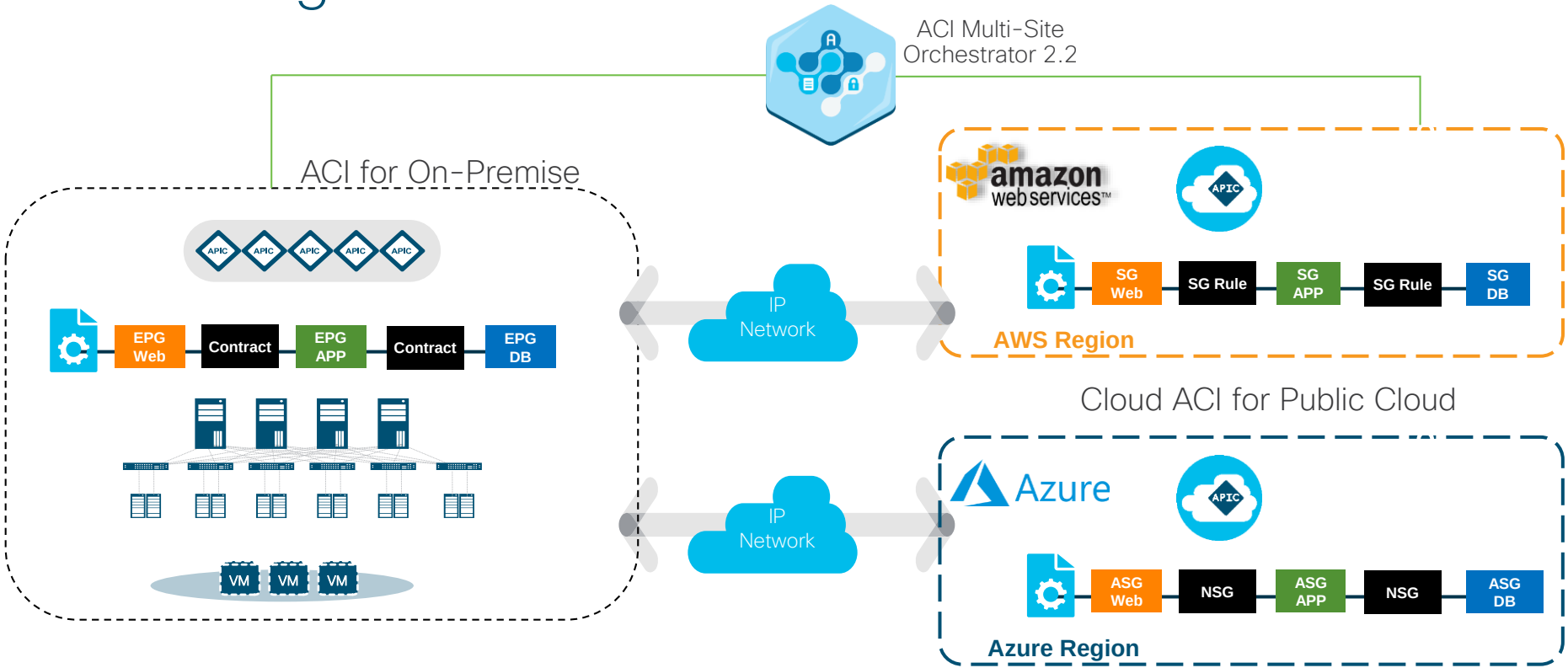
ACI Hybrid-Cloud Deployment Model

Hybrid Cloud:
- AWS in ACI 4.1
- Azure in ACI 4.2



Common Governance	Discovery & Visibility	Policy Translation	Monitoring & Troubleshooting	Single Point Of Orchestration	Operational Consistency

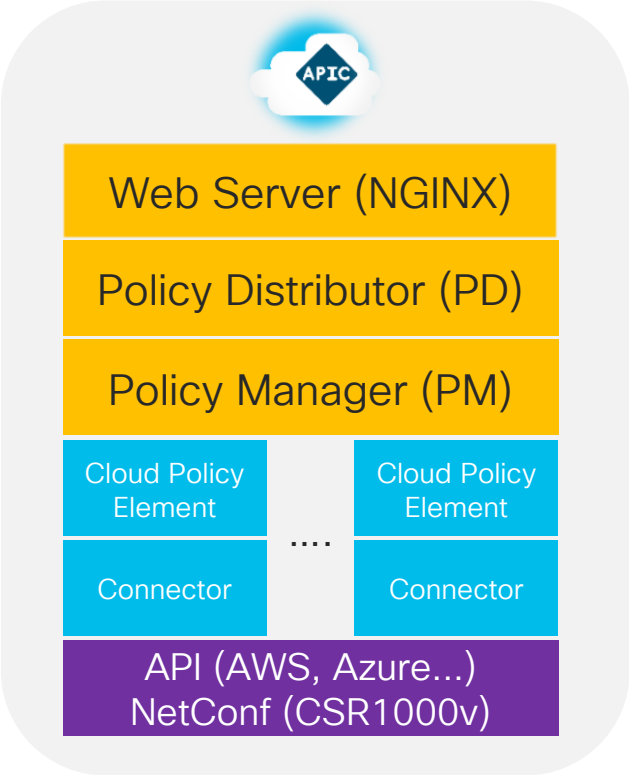
Extending ACI to the Cloud



Common Governance | Discovery & Visibility | Policy Translation | Monitoring & Troubleshooting | Single Point Of Orchestration | Operational Consistency

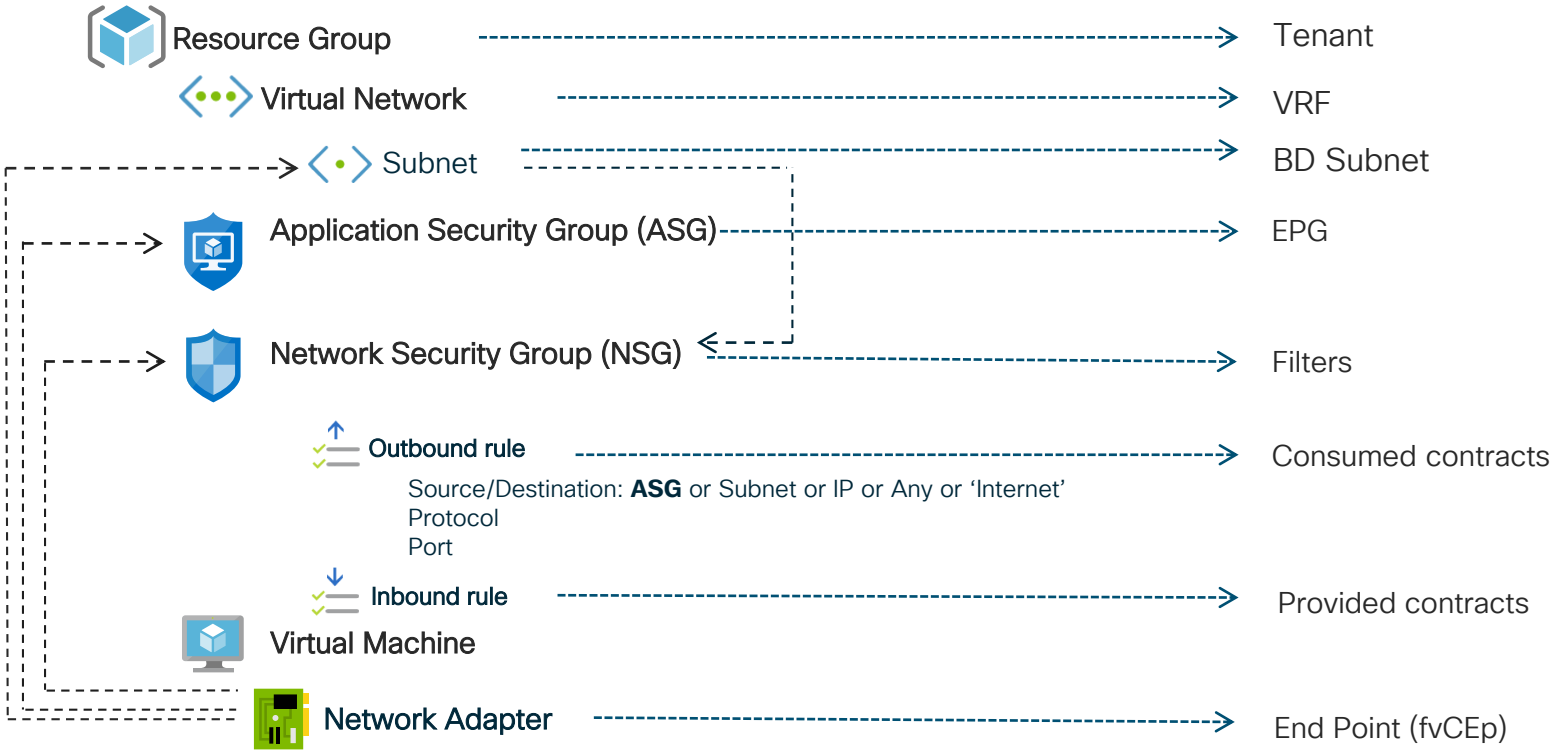
Cloud APIC (cAPIC)

Cloud APIC Architecture

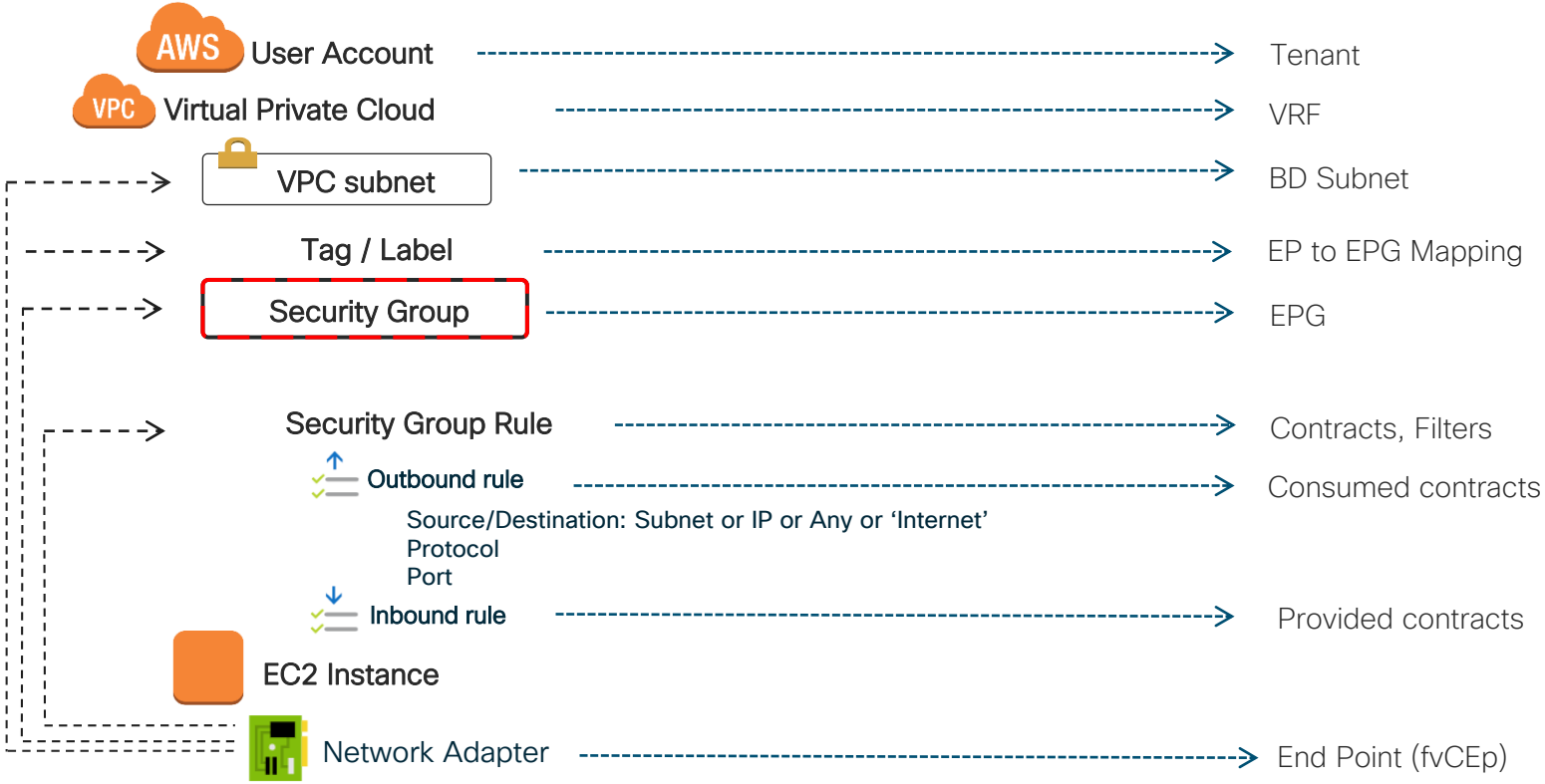


- Virtual Form Factor of APIC
- Automates / Manages Cloud Routers
- Translates ACI Policy to cloud native constructs
- Deploys cloud resources and infrastructure components
- Intuitive GUI and Similar ACI UI look and feel
- REST API North Bound Interface
- cAPIC manages 1 or more regions

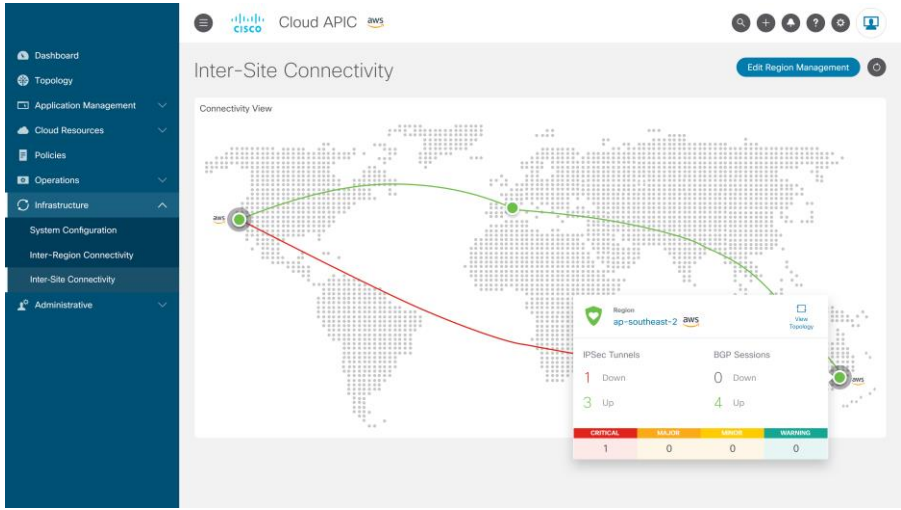
Policy Mapping – Azure

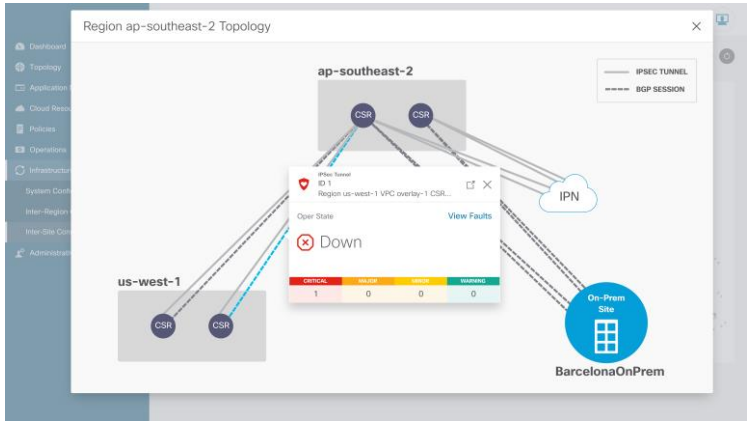


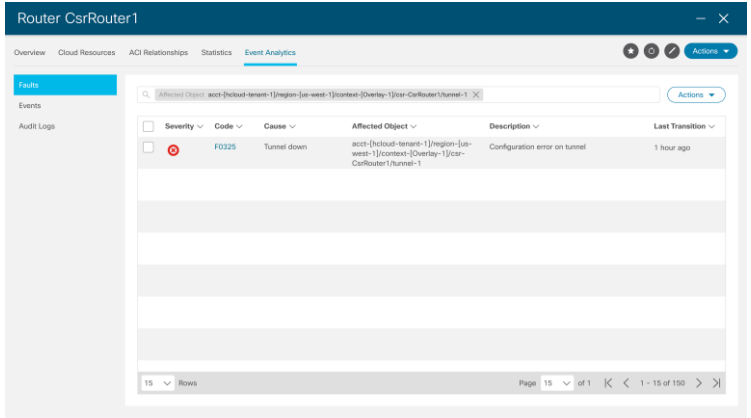
Policy Mapping – AWS



Topology Health







- Network connectivity and Health

Endpoints in an EPGs

EPG Web

Overview | Topology | Cloud Resources | Application Management | Statistics | Event Analytics

Actions

Virtual Networks
Endpoints
Security Groups

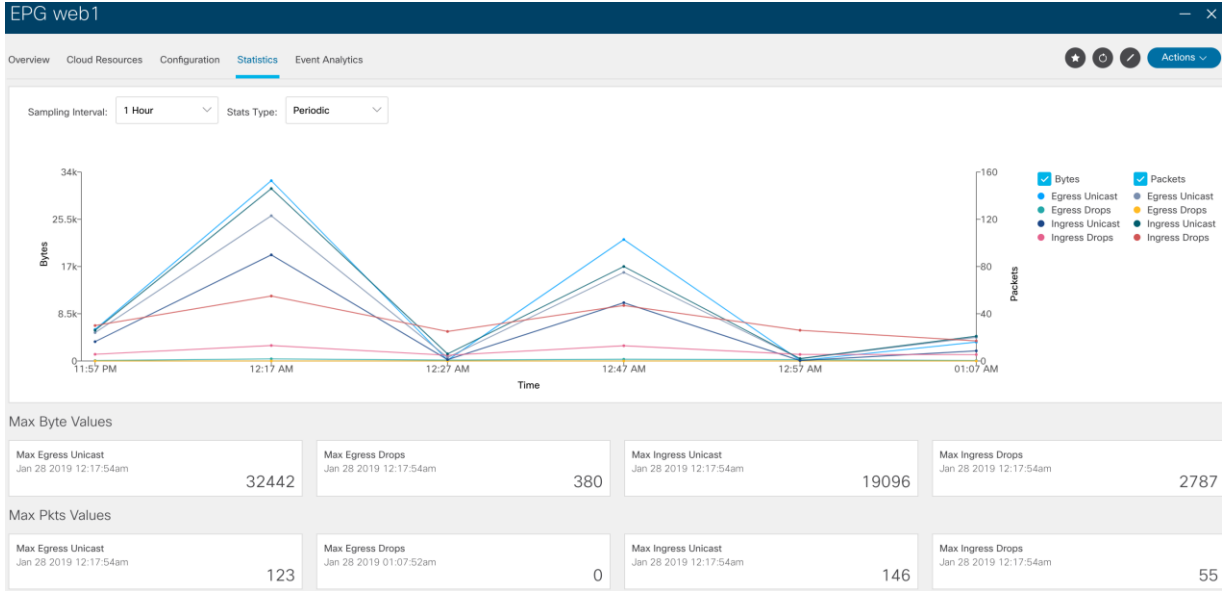
Filter by attributes

					Application Management	Cloud Resources	
Name	Oper State	Private IPv4 Addr.	Public IPv4 Addr.	EP Type	EPGs	Security Groups	Virtual Machines
wos-wordpress946 WoS > westus > WoS-VRF 10.101.200.0/24 > 10.101.200.0/24 > 10.101.200.128/25	in-use	10.101.200.132	13.64.103.72	vm	1	1	1

10 Rows

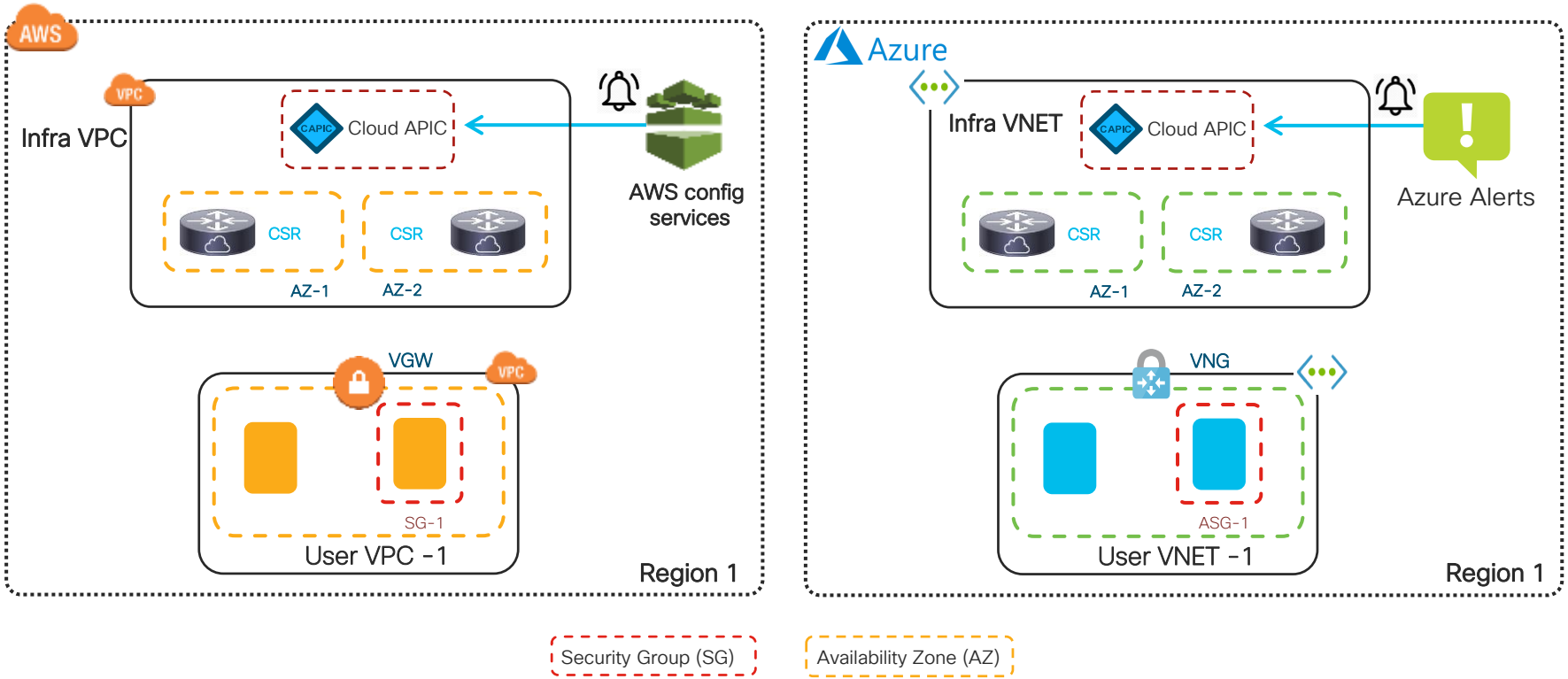
Page 1 of 1 1-1 of 1

Statistics



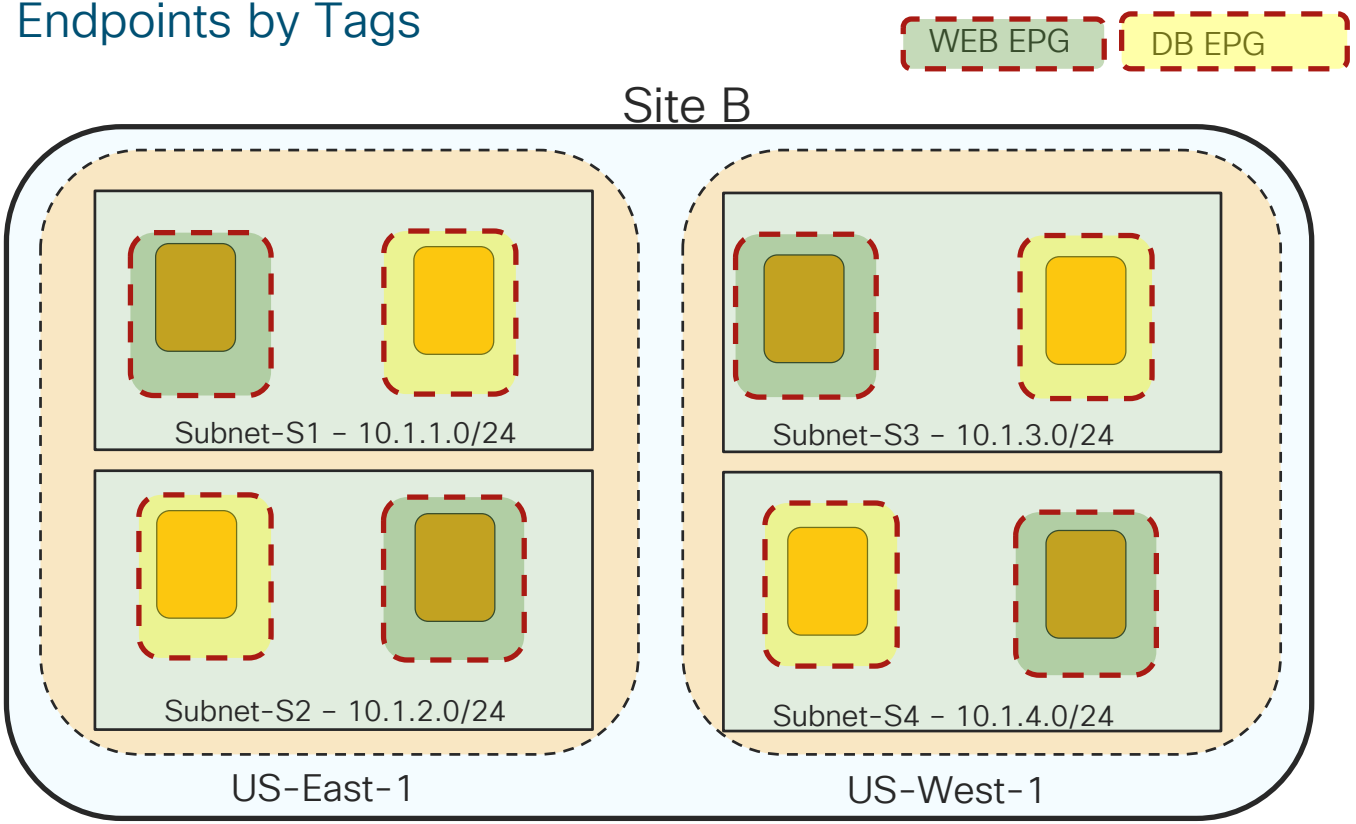
- We will show multiple statistics:
  - Inter-site
  - Inter-region
  - Inter-VPC
  - Cloud EPG
  - Cloud Routers

End Point Learning in Cloud



Cloud EPG

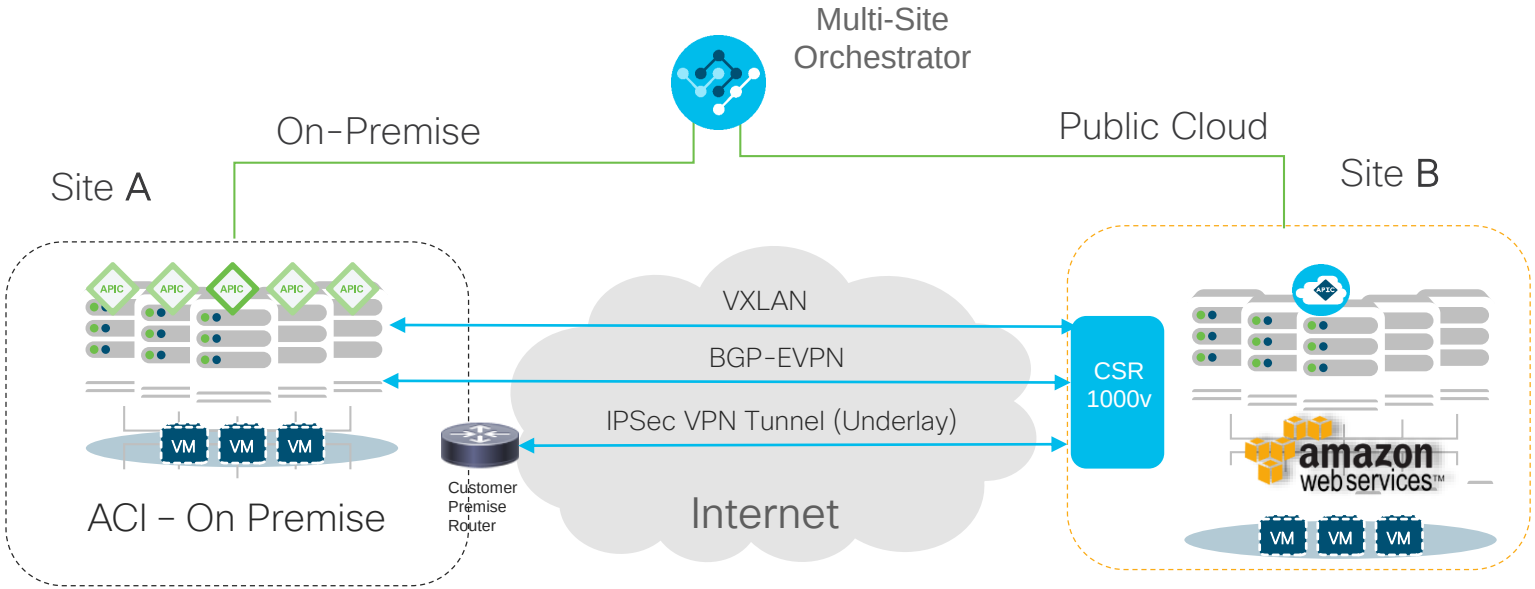
Mapping Endpoints by Tags



Cloud Connectivity (Shh, it’s just Multi-Site)

ACI Multi-Site Extension to Cloud

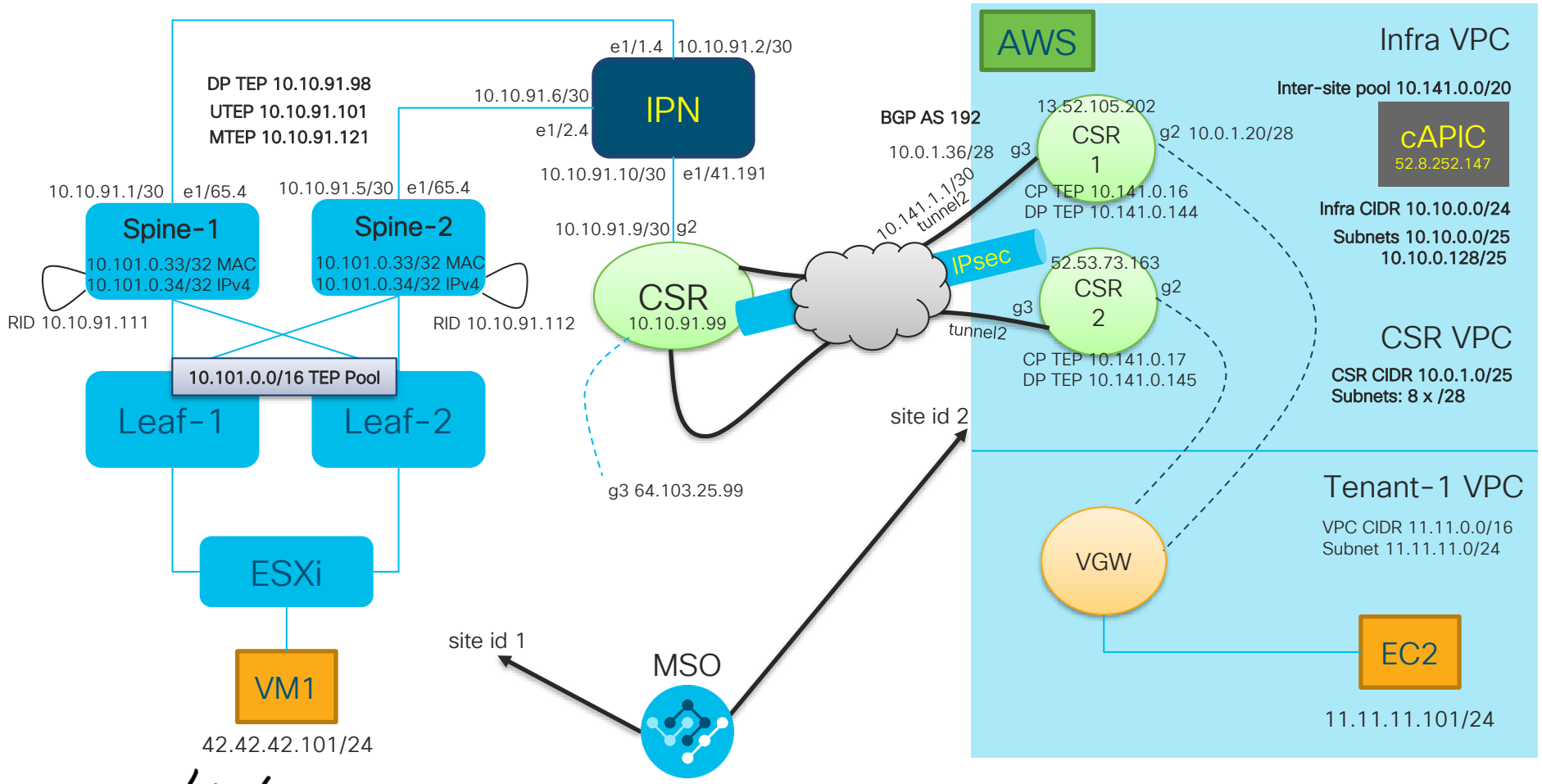
Simplest Mode through IPSec VPN

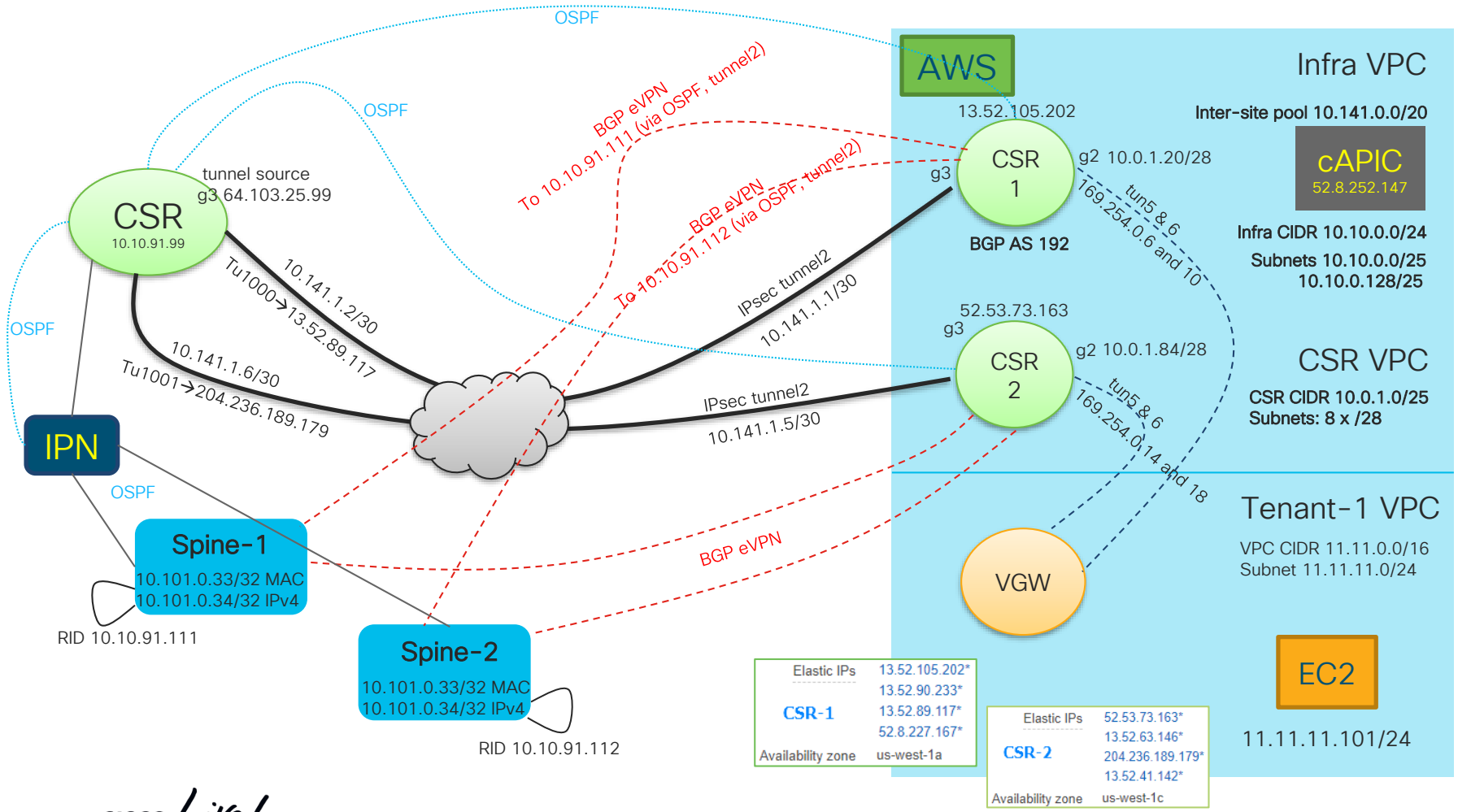


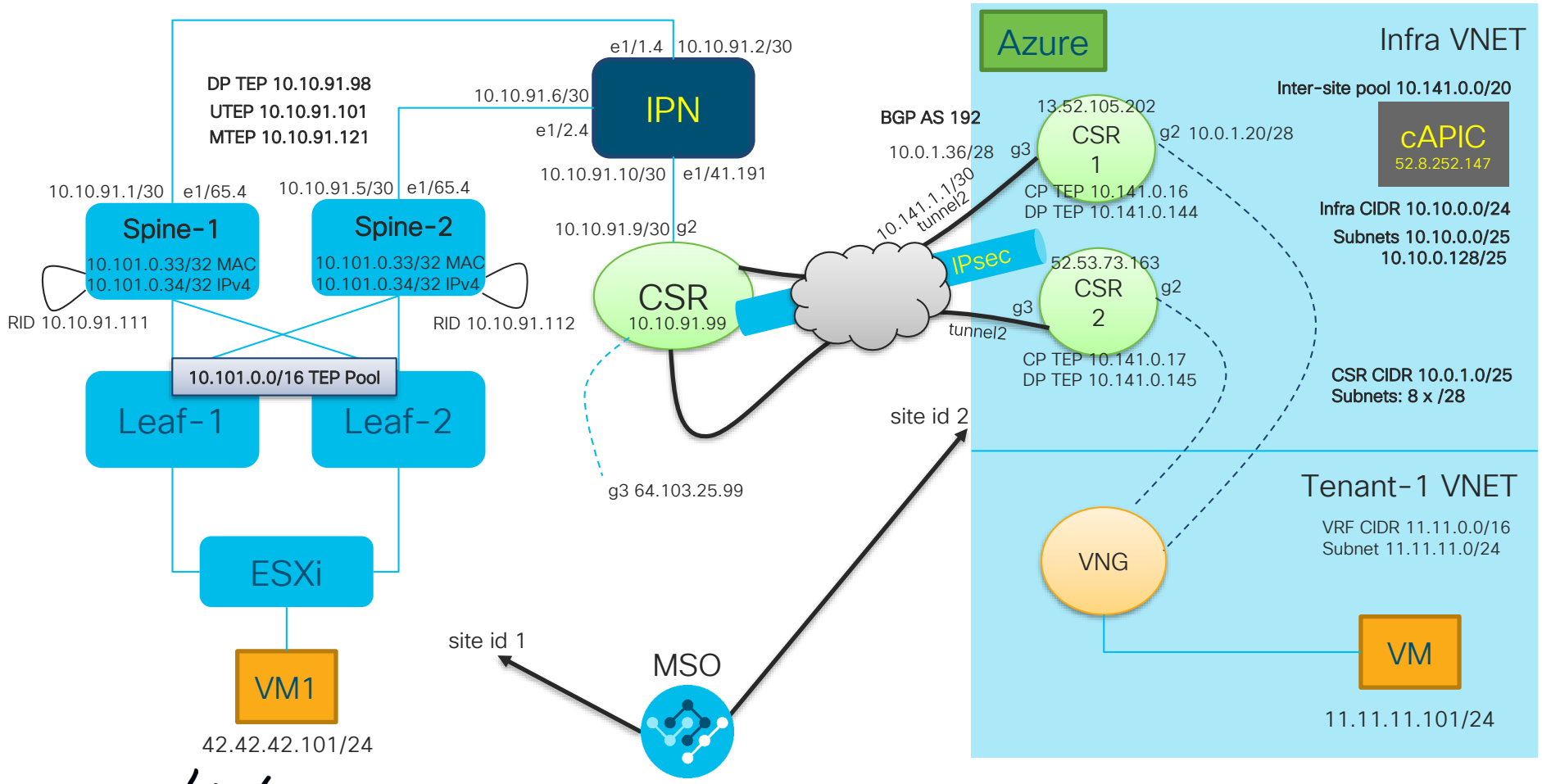
- VXLAN data-plane connects ACI fabric and Cloud site
- BGP-EVPN routing reachability between ACI fabric and Cloud Site
- IPSec VPN connection between customer Premise Router before ACI fabric and CSR1kv

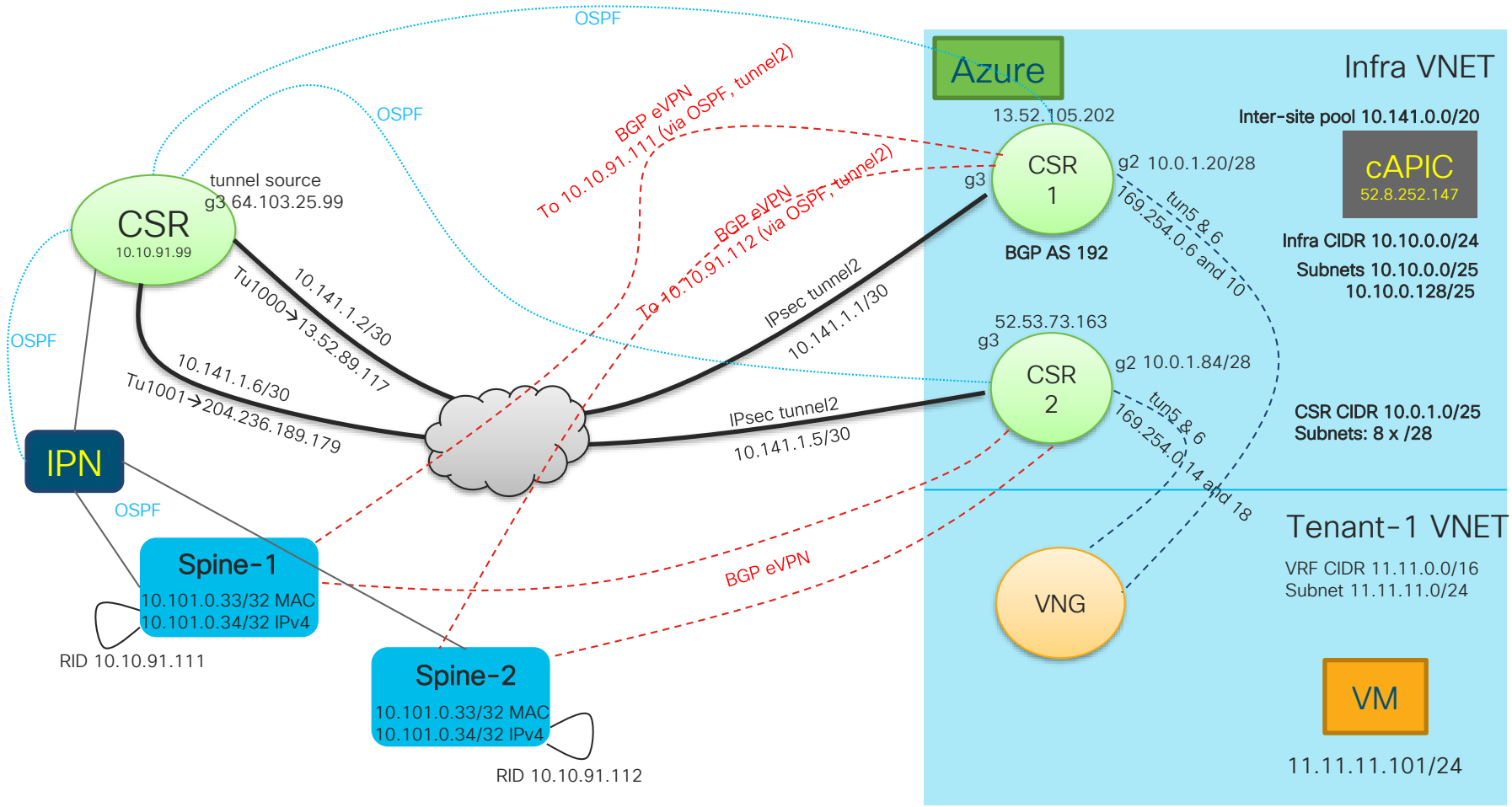
Cloud Multi-Site Components

- On-prem ACI
  - AWS requires software version 4.1
  - Azure requires software version 4.2
  - 2nd generation spines required (like regular multi-site)
- MSO
  - 2.1 supports AWS (hosted on-prem)
  - 2.2 supports AWS, Azure and cloud first
- Same IPN connectivity as traditional multi-site
  - You need an IPN layer (OSPF w/Spines)
  - And an on-prem CSR1Kv or ASR to terminate IPsec
  - MP-BGP eVPN between Spines and CSRs in the cloud
- Azure and/or AWS site(s) with a cloud APIC deployed
  - AWS - Infra VPC (where cAPIC and CSRs reside) & Tenant VPC (only AWS components)
  - Azure - Infra VNET (where cAPIC and CSRs reside) & Tenant VNETs (only Azure components)





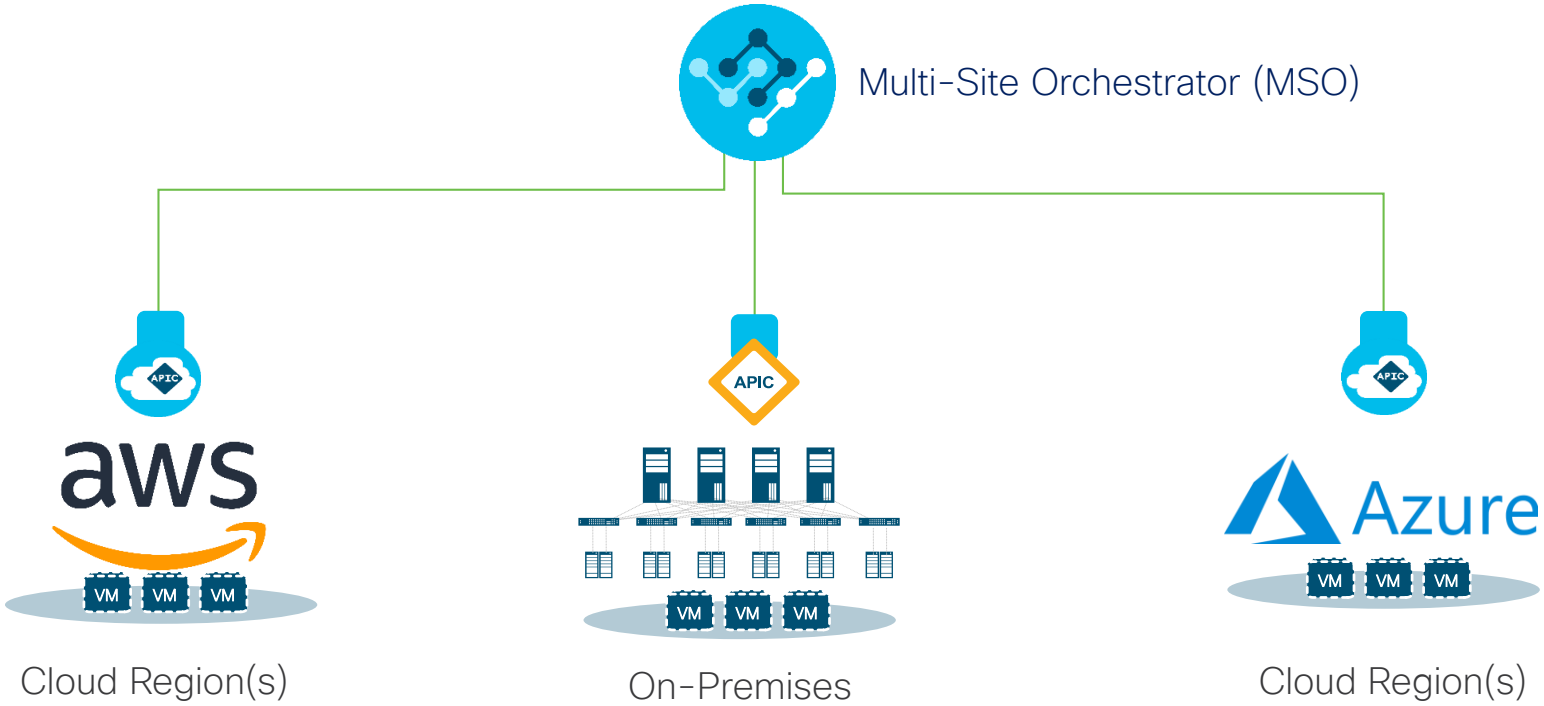




Wait, that looks quite complex
Actually, its just another instance of multi-site

- You don't configure half of what you just saw
- You spin up a cAPIC using a Cloud Formation (AWS) or ARM (Azure) Template
- Pair on-prem with your IPN just like regular multi-site
- MSO and cAPIC take care of most configuration aspects
  - You provide high-level config parameters
  - You get a ready-to-use IPsec configuration to copy/paste in your CSRs
  - All AWS configuration aspects (VGW BGP and IPsec, routing, security groups) is fully automated and abstracted
  - All Azure configuration aspects (GW BGP and IPsec, routing, ASG, NSG) is fully automated and abstracted

Cloud ACI
Its Multi-Site



MSO Form Factor

Hardware Appliance
(based on SE)

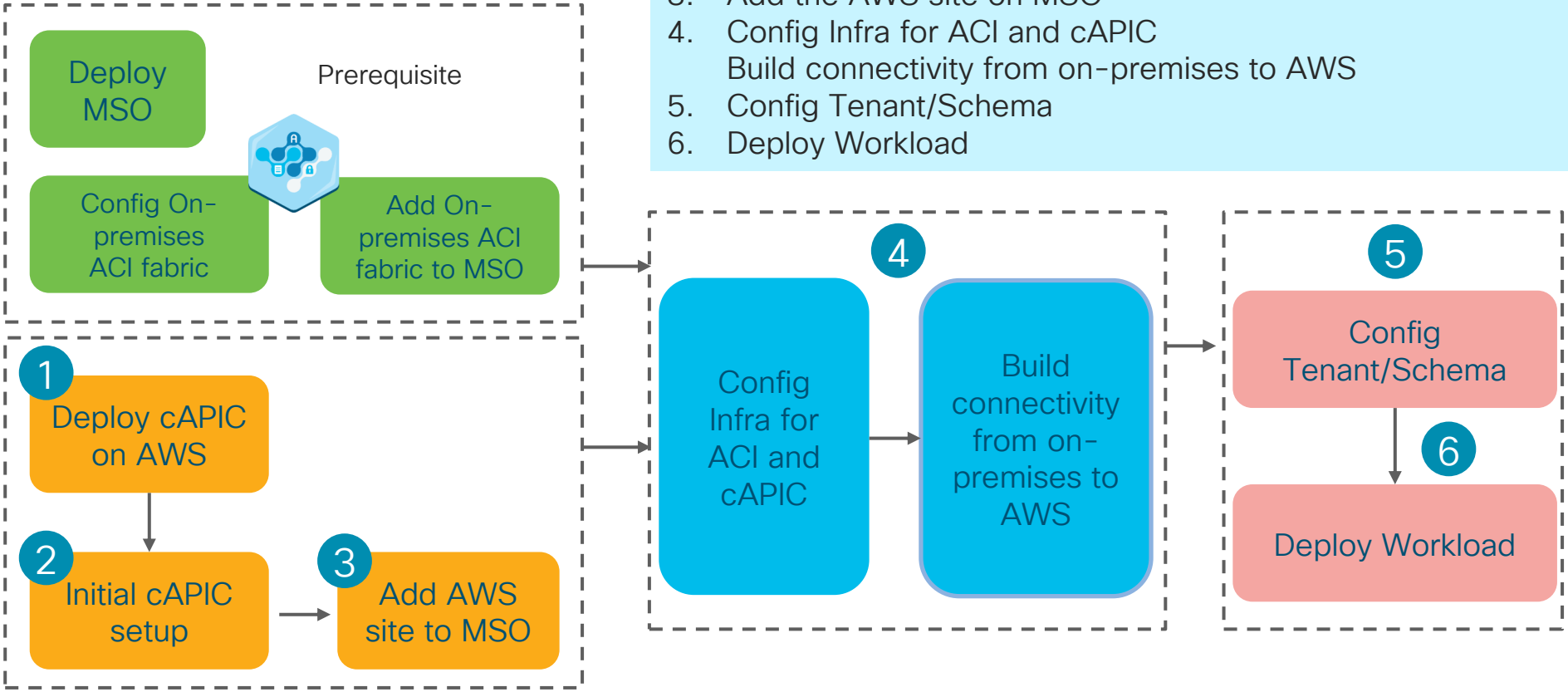
VMware OVA

Cloud MSO for AWS

Demo 1: ACI Cloud Connectivity

High Level steps

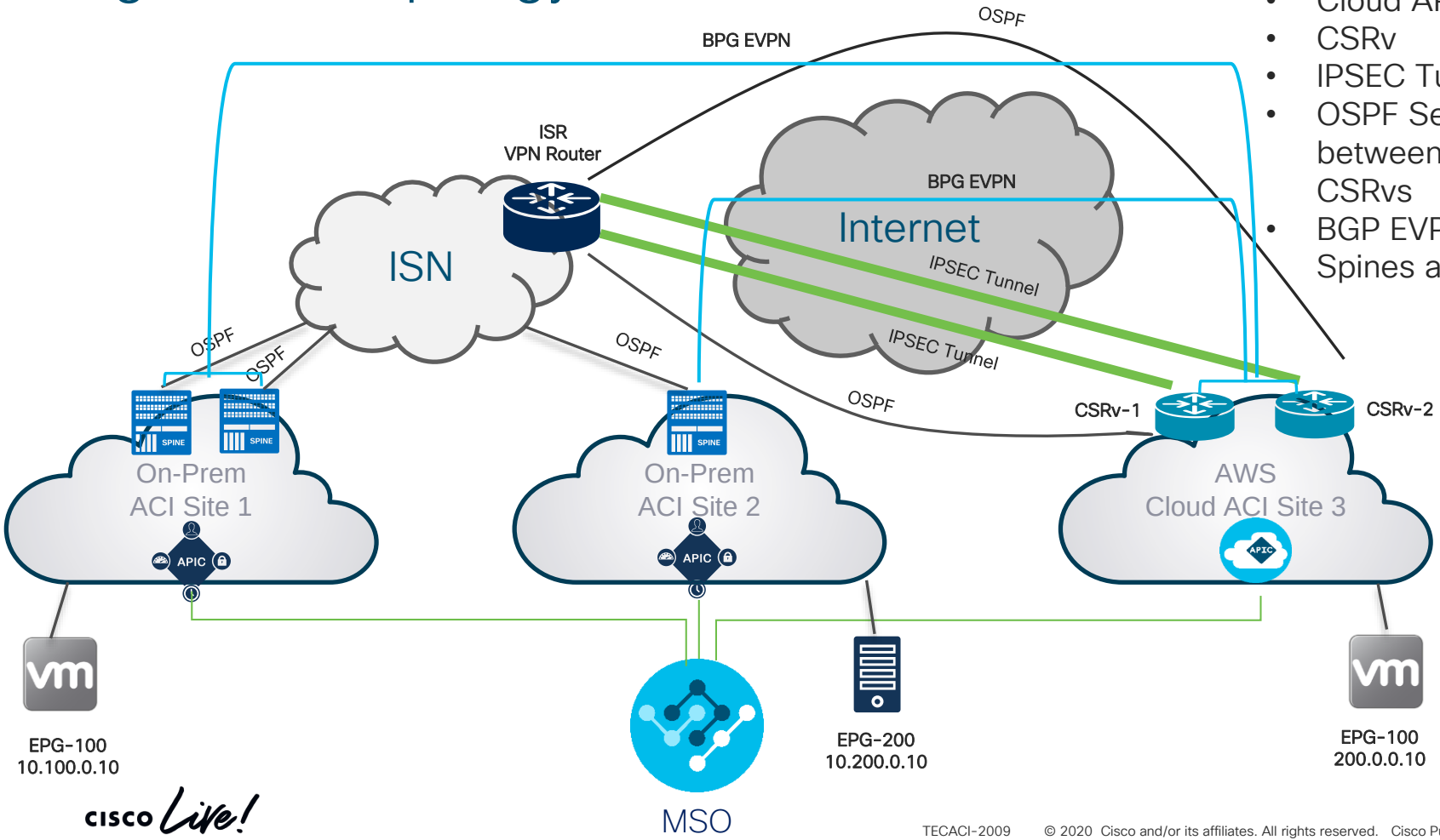
1. Deploy cAPIC on AWS
2. Initial cAPIC setup
3. Add the AWS site on MSO
4. Config Infra for ACI and cAPIC
   Build connectivity from on-premises to AWS
5. Config Tenant/Schema
6. Deploy Workload



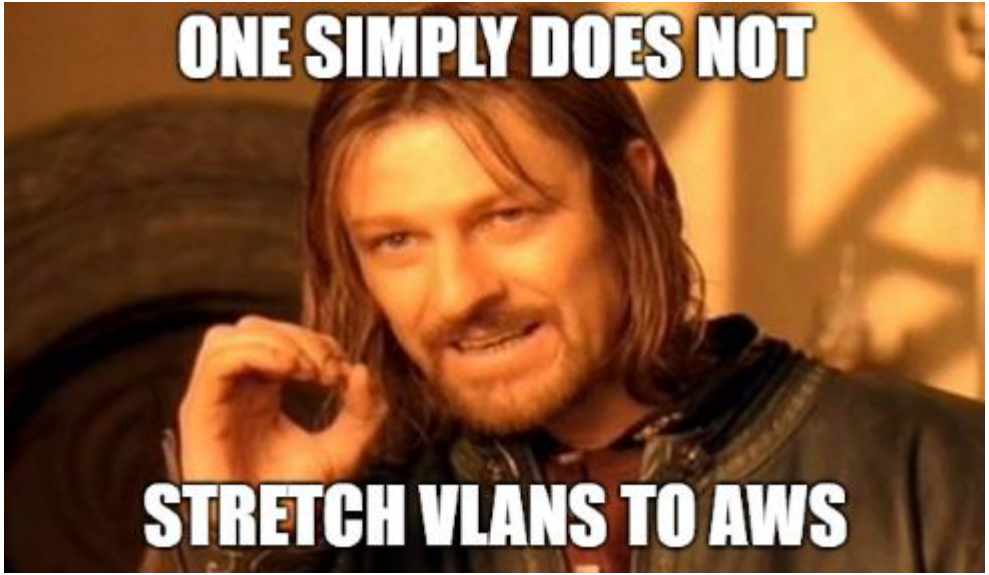
High Level Topology

Fully Automated:

- Cloud APIC
- CSRv
- IPSEC Tunnel
- OSPF Setup between ISR and CSRvs
- BGP EVPN between Spines and CSRvs



So I have Multi-Site to Cloud
You mean I can stretch VLANs???



No really, can I stretch VLANs to the cloud??

- **Absolutely not!**
  - Stretching an EPG does not mean stretching a VLAN or broadcast domain.
- Stretched EPG Foo uses subnet X on-prem and subnet Y in the cloud
  - You tell MSO what criteria(s) should be used to join EPG Foo
    - Could be tags, could be an IP prefix, could be a region or an AZ
    - cAPIC informs MSO when an EC2 instance matches the selector
    - EC2 instance appears as a /32 in external EPG on-prem, shadow EPG takes care of contracts to make it look like a stretched EPG
    - cAPIC programs a security group in AWS to match on-prem contracts
- Anyway, public cloud vendors do not allow broadcast or multicast.
  - There is no good use case for this. Don't let friends run NSX Cloud.

Agenda

- ACI Anywhere, Extending the ACI Fabric
  - Overall Design Principles (AZs and Regions)
  - Mapping use cases to the proper solutions
    - Active/Active DC → Multi-Pod
    - Disaster Recovery → Multi-Site
    - Migration/Coexistence with Legacy DC Networks and 'Disaggregated DCs' Model → Physical Remote Leaf
    - Baremetal Cloud Integration → Virtual Pod (vPod)
  - Extending ACI to the Cloud
  - Connecting the users to the Multi-Cloud DC
    - ACI and SDA Integration
    - ACI and SDWAN Integration

ACI and Multi-Domain Integration

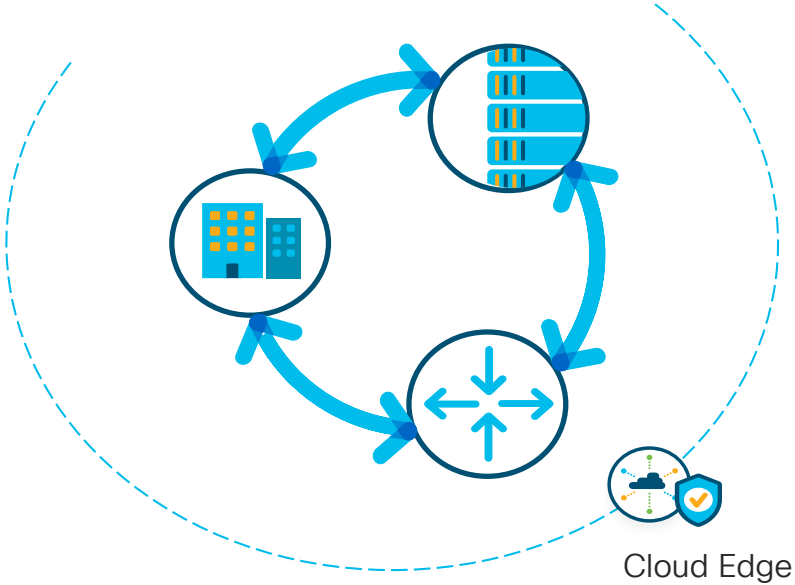
Our direction

Policy

Automation

Telemetry, Analytics and Assurance

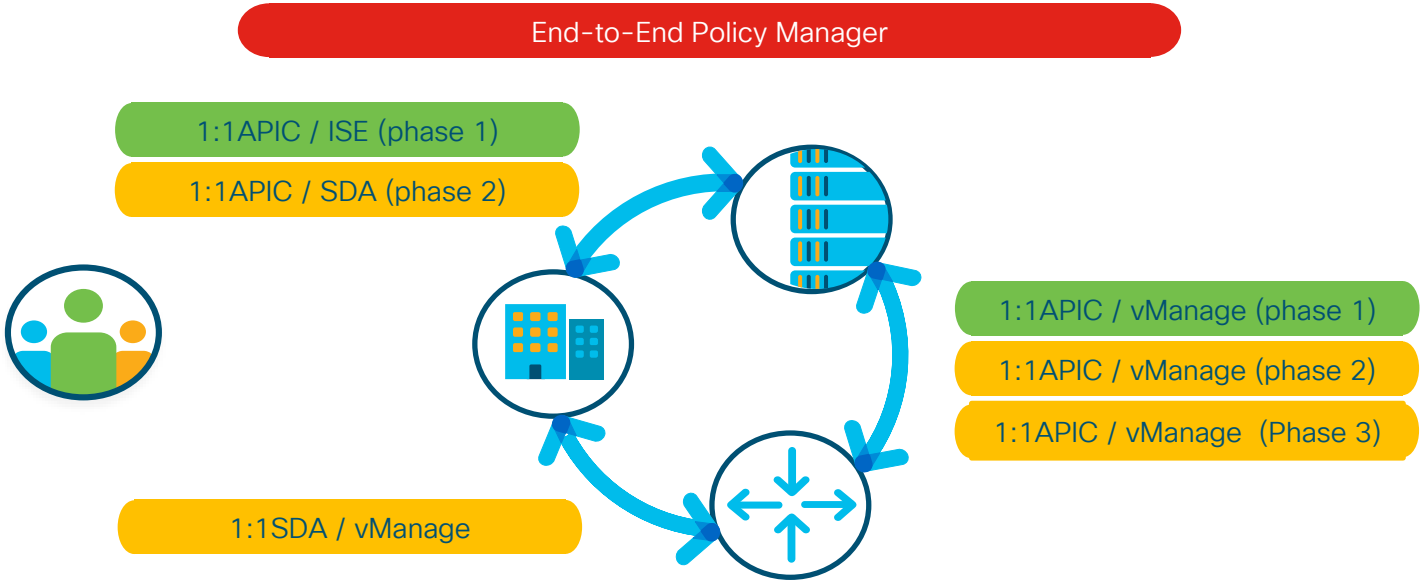
Security, Identity and Segmentation



Cloud Edge
Trust boundary

Domain specific integrations

with end-to-end policy



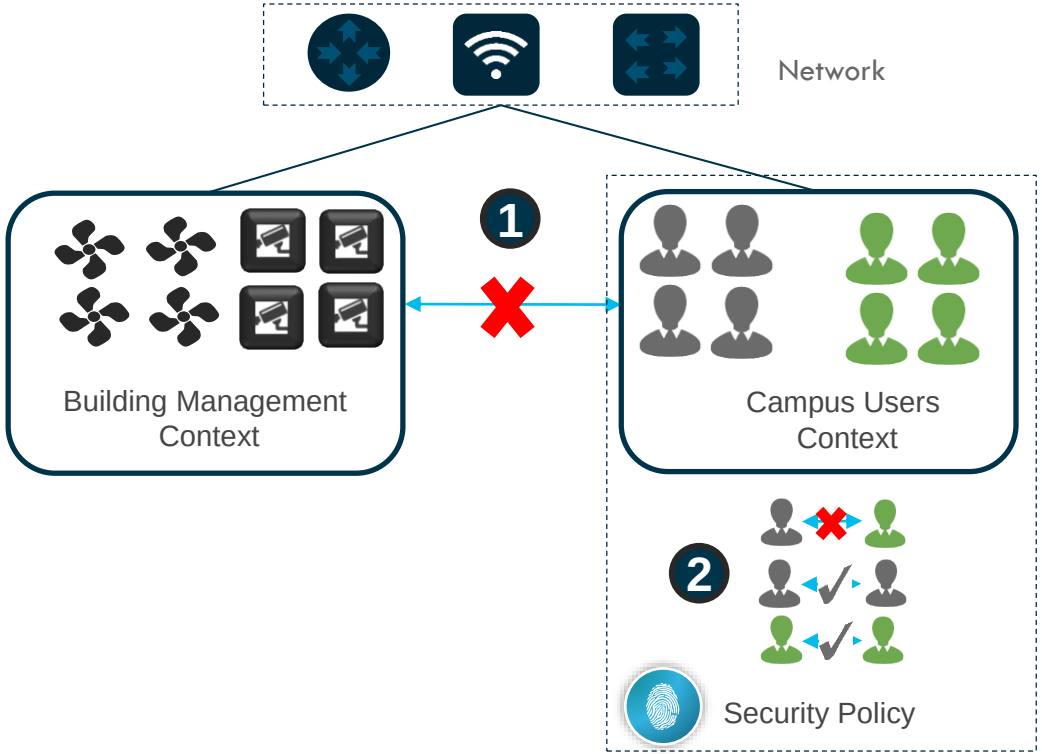
Shipping

1HCY20

2HCY20

Two Level Segmentation

Two level segmentation/label model



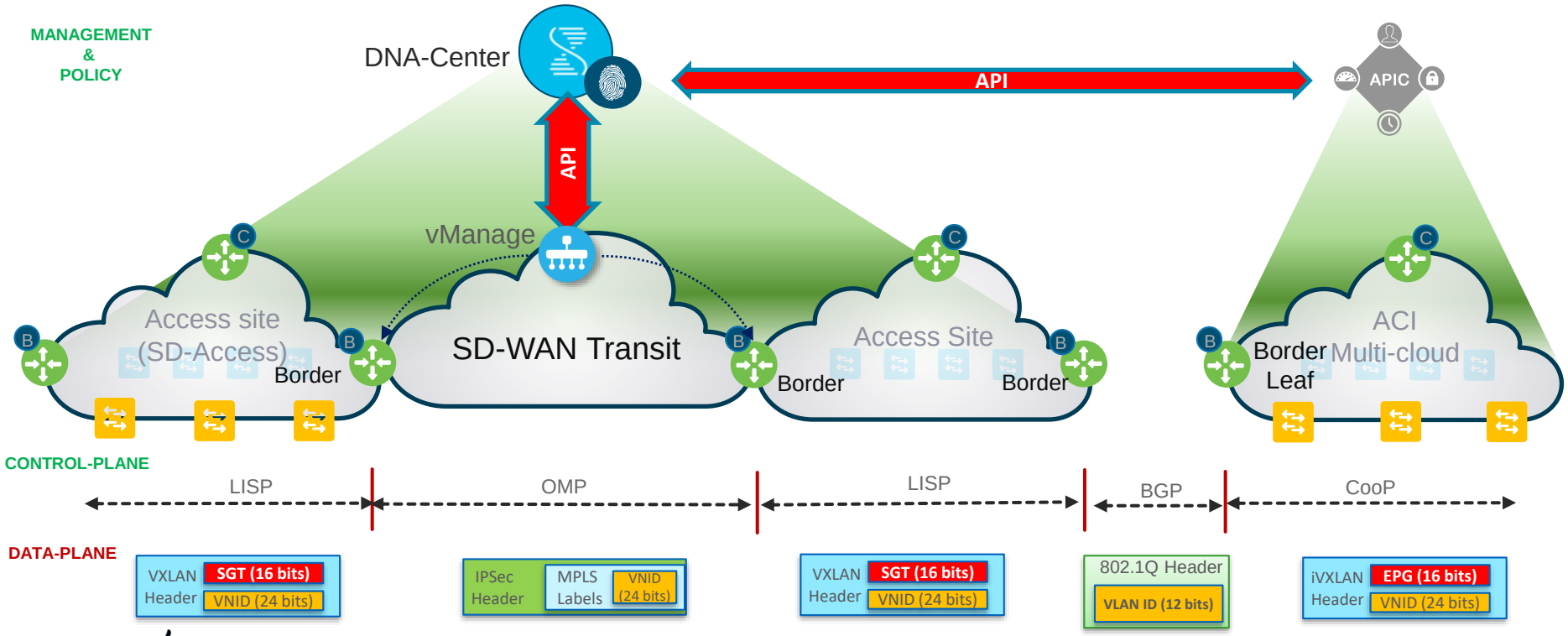
1 **Virtual Network (VN)**

First level Segmentation that ensures zero Communication between Building systems and Users

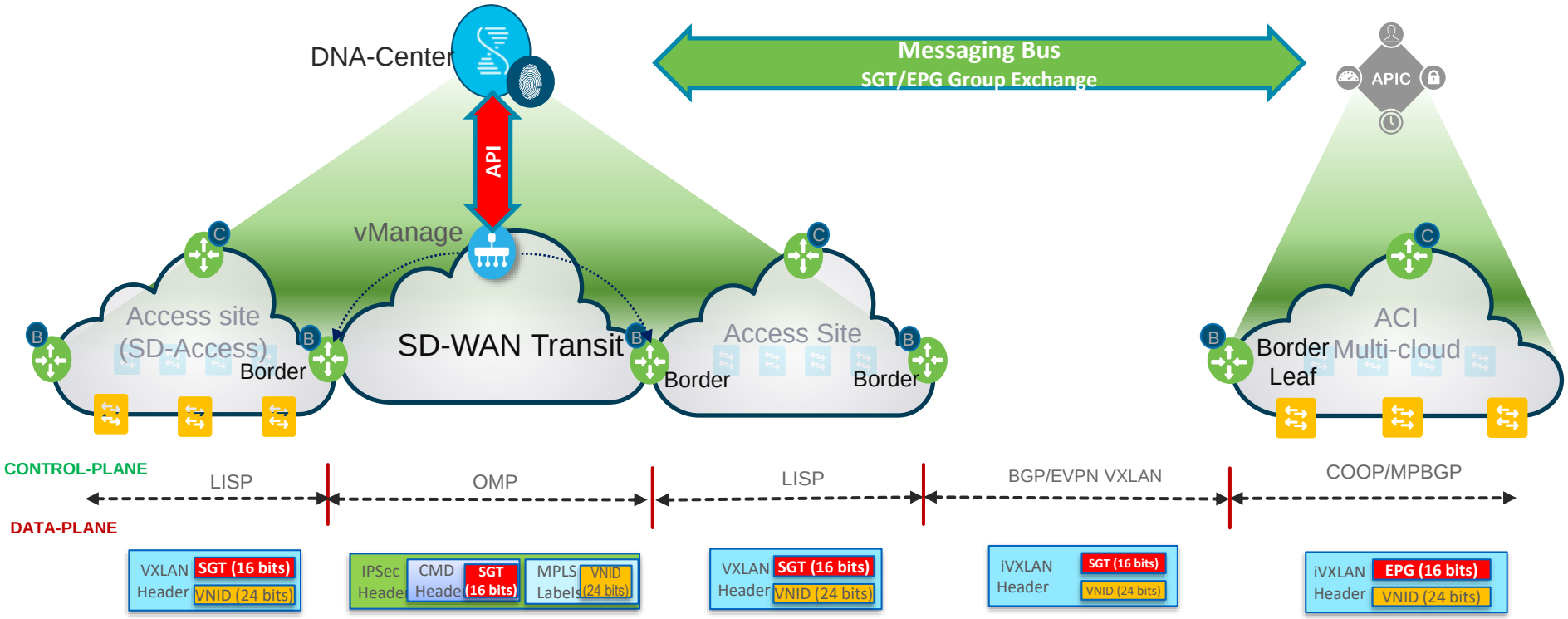
2 **Scalable / End-point Group**

Second level Segmentation within a VN that ensures role based access control between Blue-group and Red-group

End-to-end segmentation and cross domain interworking Phase 1 (Current)



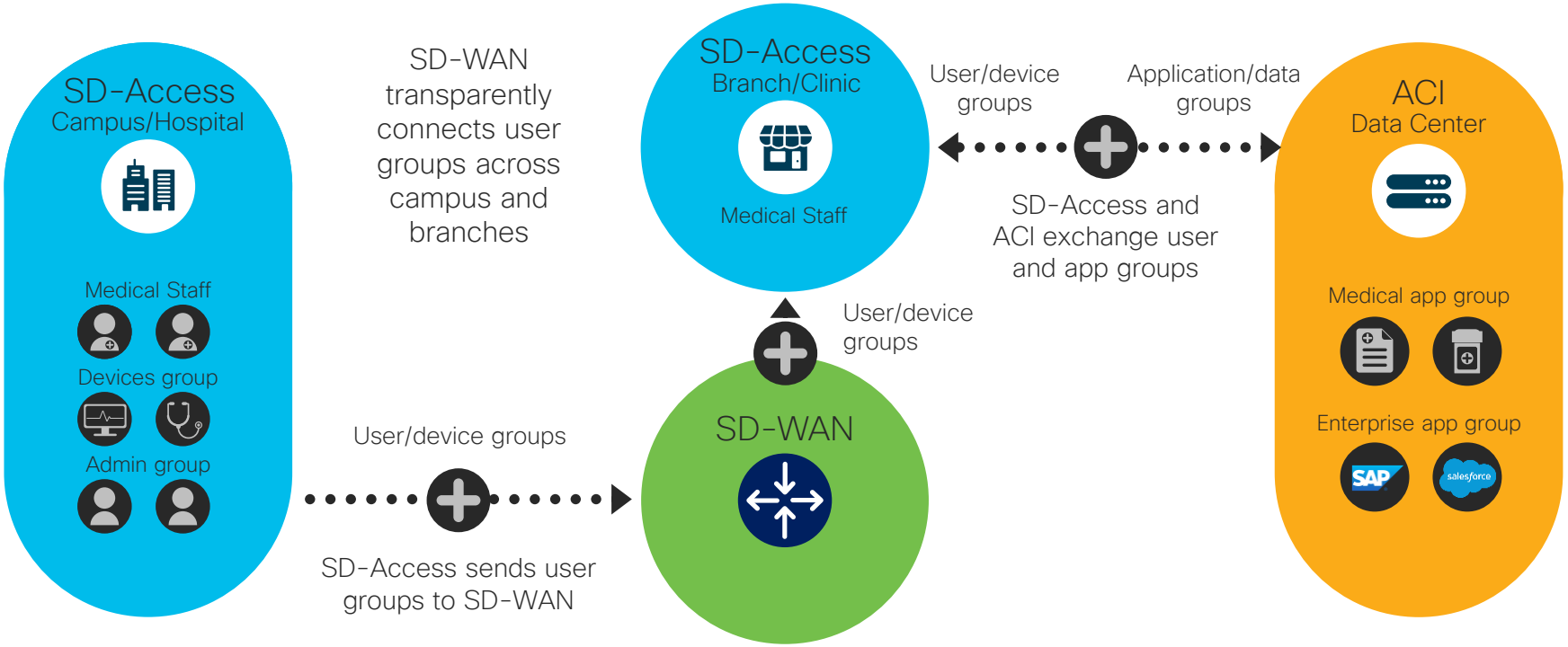
End-to-end segmentation and cross domain interworking Phase 2



SDA and SDWAN Pairwise Integration

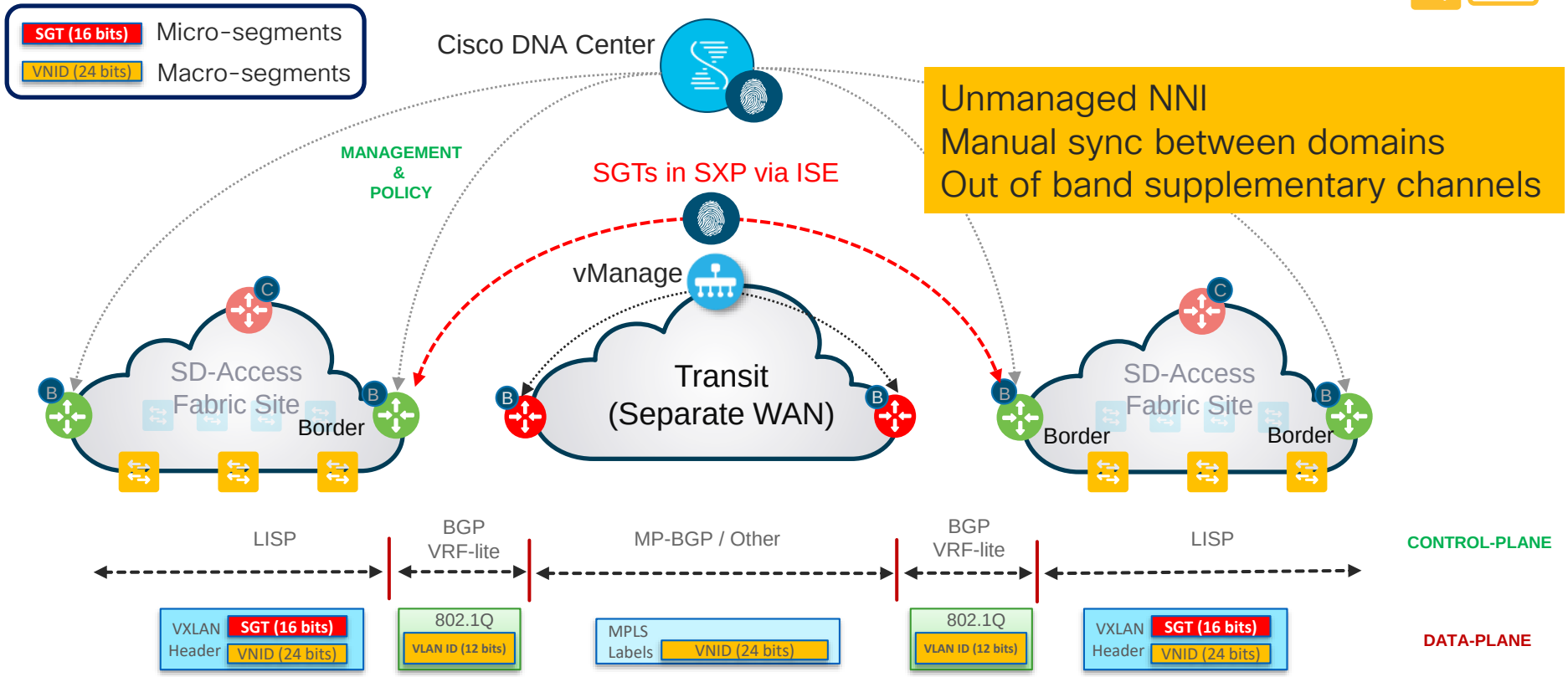
Segmentation Policy Follows the User

No reconfiguration required if users move

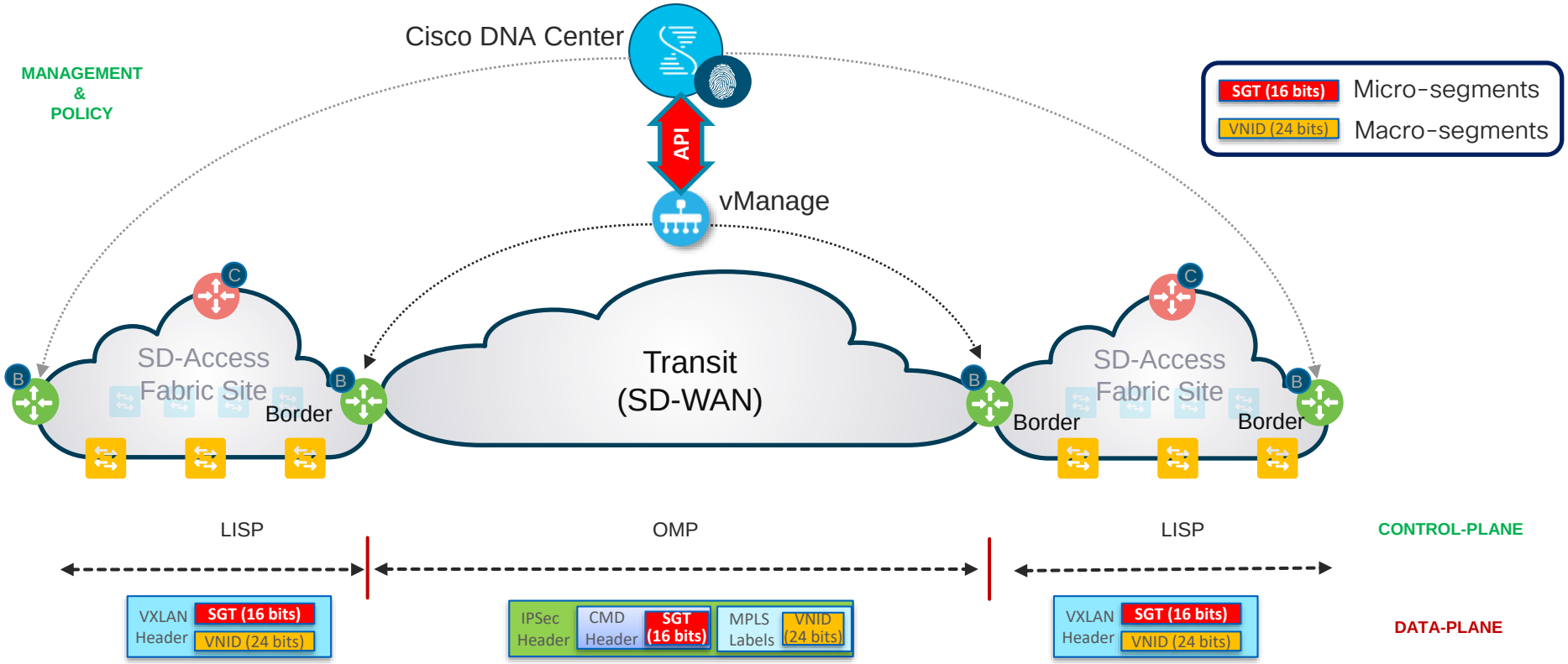


Segmentation maintained between users and apps at any location

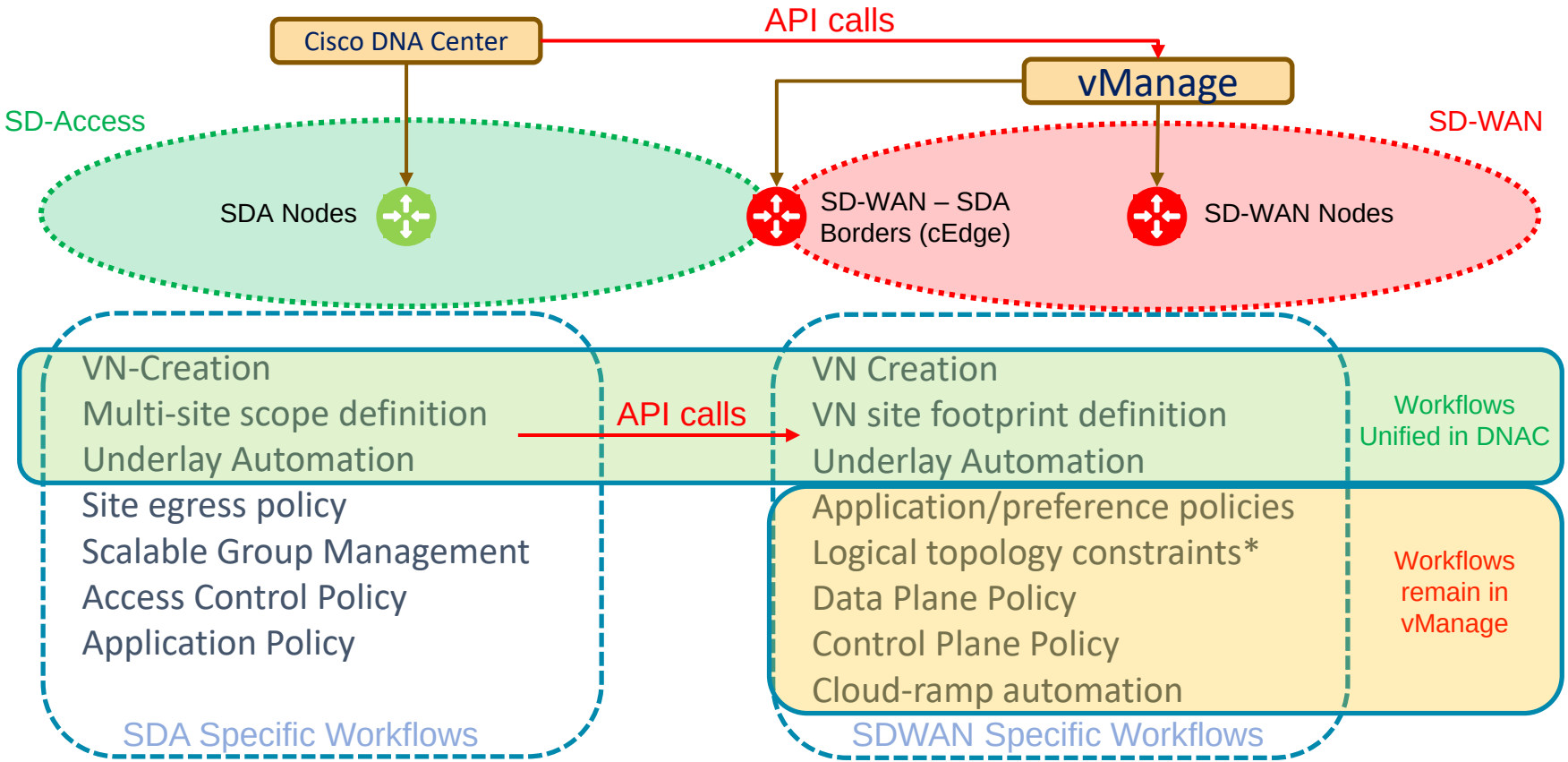
Before: Campus/WAN w/o Integration



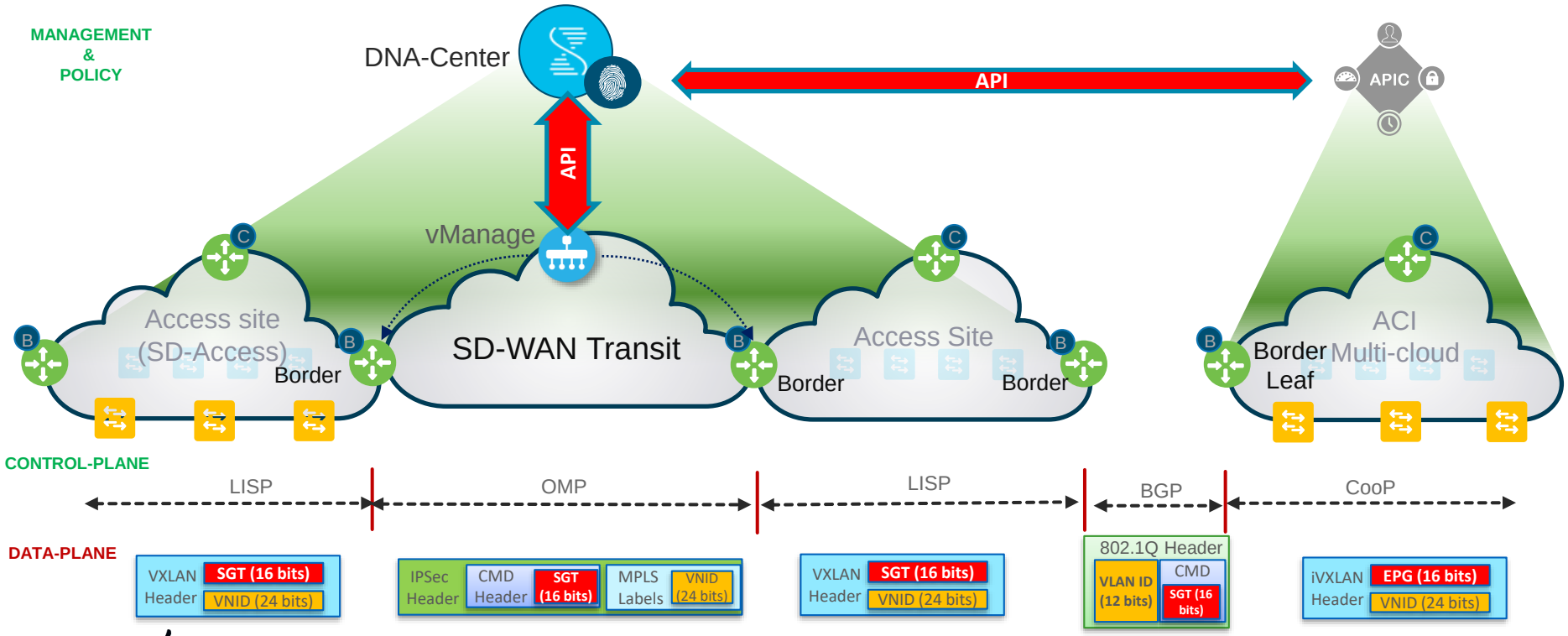
Integrated Multi-domain SD-Access/SD-WAN



Management of Integrated SDA+SD-WAN

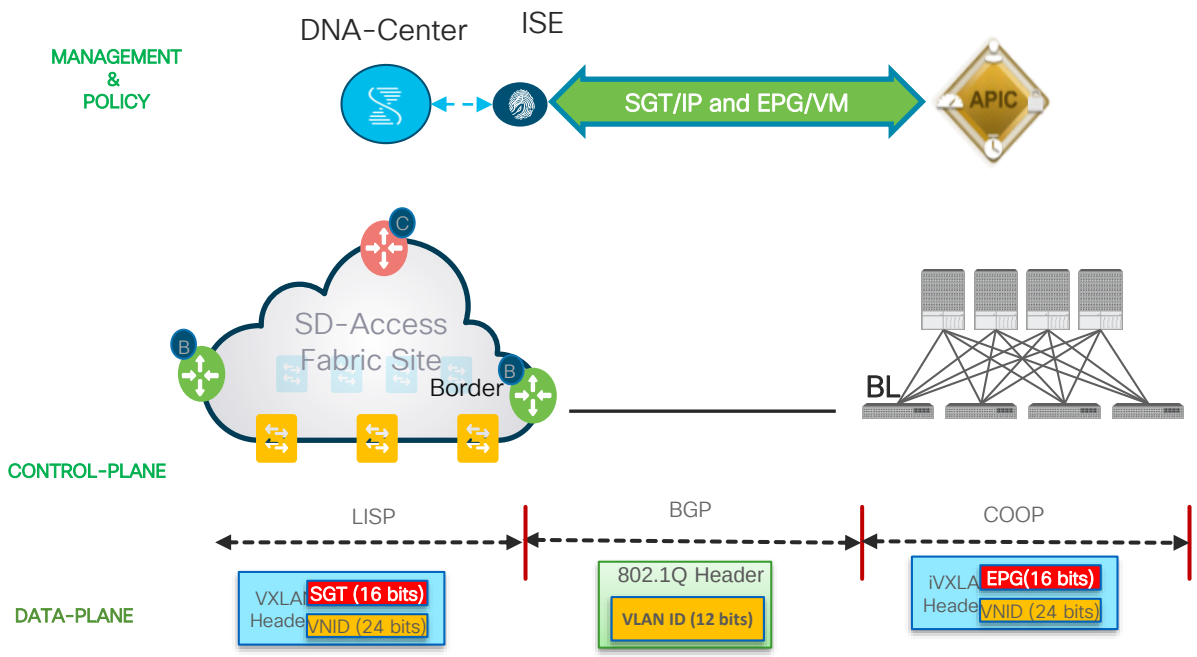


End-to-end segmentation and cross domain interworking



ACI and SDA Pairwise Integration – Phase 1

ACI and SDA Pairwise Integration
Phase 1 - SDA-ACI: Group/Identity Mapping



ACI 3.2 Scale
EX, FX and FX2 Hardware

SGT -> External EPG	250
Number of Mappings	64k
Mappings per External EPG	8k
Transaction rate (target)	100/s

Phase1: Solution Testing completed and supports:

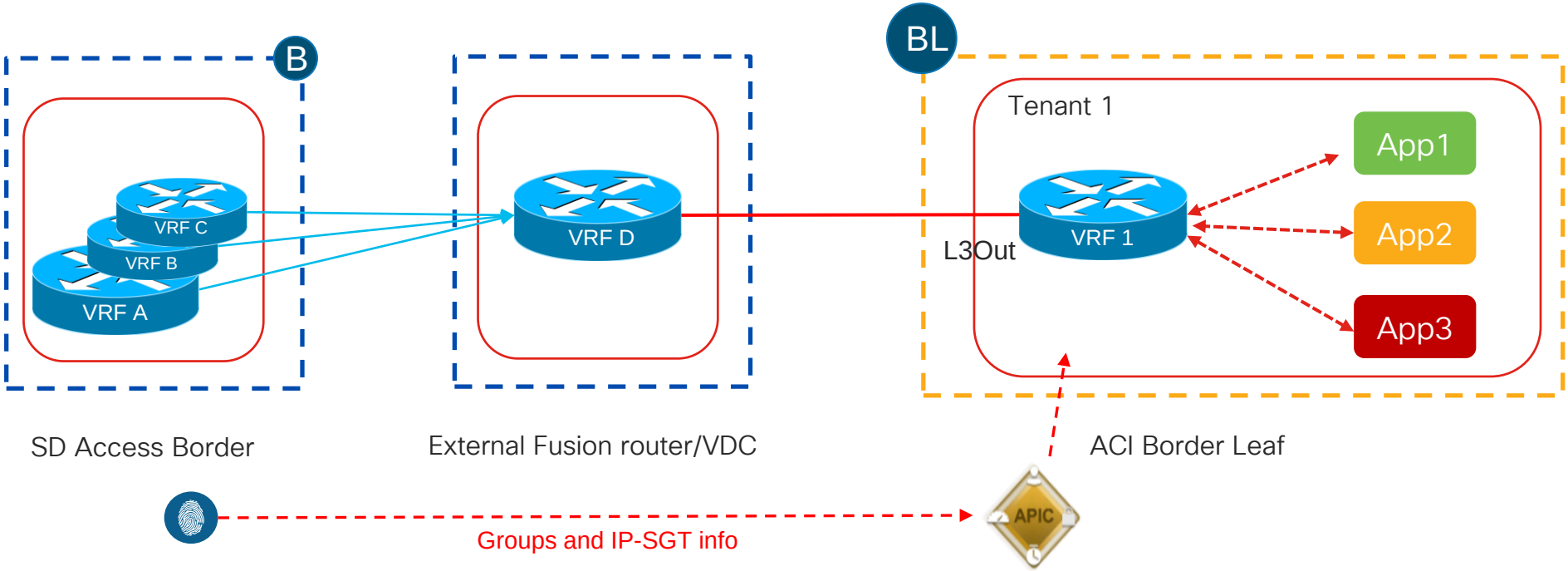
a) IP Data path

b) Exchange of SGT and EPG at the control plane layer

c) IP-SGT/EPG bindings in both directions

- ACI 3.2.x above
- ISE Version 2.4 Patch 6
- DNA Version 1.2.10

Phase 1 SDA-ACI
Current Solution: Single VRF, Single Tenant

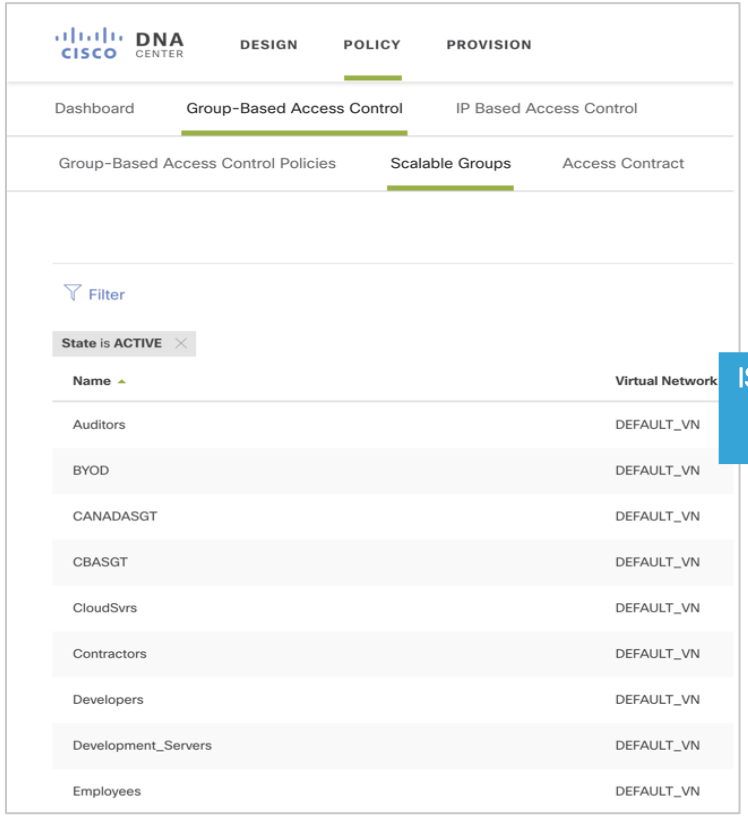


Note: Shared L3Out on ACI BL nodes currently not supported

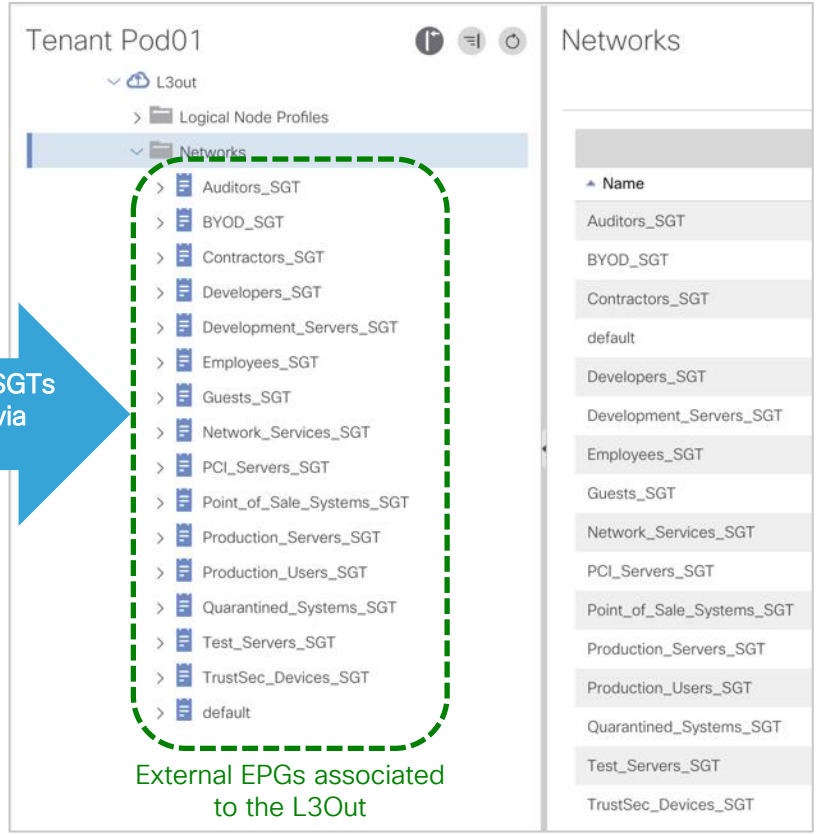
Where Should the Policy Be Applied?

Policy Enforcement on the Application Domain

Groups Provisioned from SDA to ACI (by ISE)



ISE provisions SGTs info to APIC via REST API*

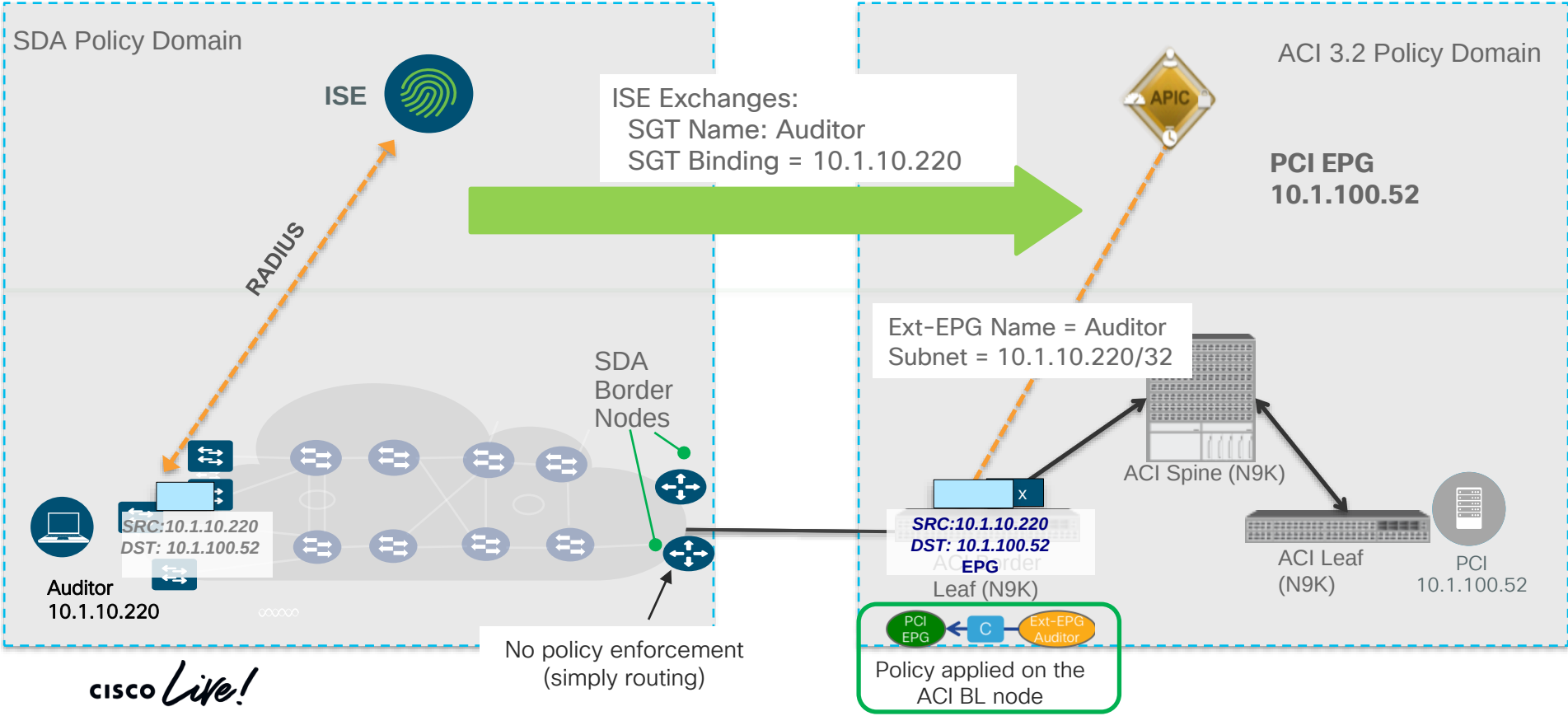


External EPGs associated to the L3Out

* Provisioning of SGTs to APIC can be controlled at the SGT level

Policy Enforcement on the Application Domain

Applying Policy on the ACI BL Nodes



Policy Enforcement on the Application Domain
Enforcement Scale in ACI

	EX	FX	FX2
SGT⇔External EPG	250	250	250
SGT IPv4 Mapping (/32)	64k	64k	64k
SGT IPv6 Mapping (/128)	24k	48k	24k
SGT IPv4+IPv6 Mapping (Dual-Stack)	24k + 24k	32k + 32k	24k + 24k
Policy	8k	128k	8k
Max. IPv4 Bindings per Ext-EPG	8k	8k	8k
Max. IPv6 Bindings per Ext-EPG	8k	8k	8k

Note: “Egress Policy Enforcement” on Campus Facing ACI Border Leaf

Policy Enforcement on the SDA Domain

SDA Learning Groups from ACI

Group-Based Access Control Policies | Scalable Groups

Name
AP_DB_EPG
AP_Development_Svr_EPG
AP_HIPAA_Svr_EPG
AP_Medical_Records_EPG
AP_PCI_Svr_EPG
AP_production_app_EPG
AP_production_web_EPG

SGTs created in DNAC

ISE retrieves Application EPGs from APIC via REST API*

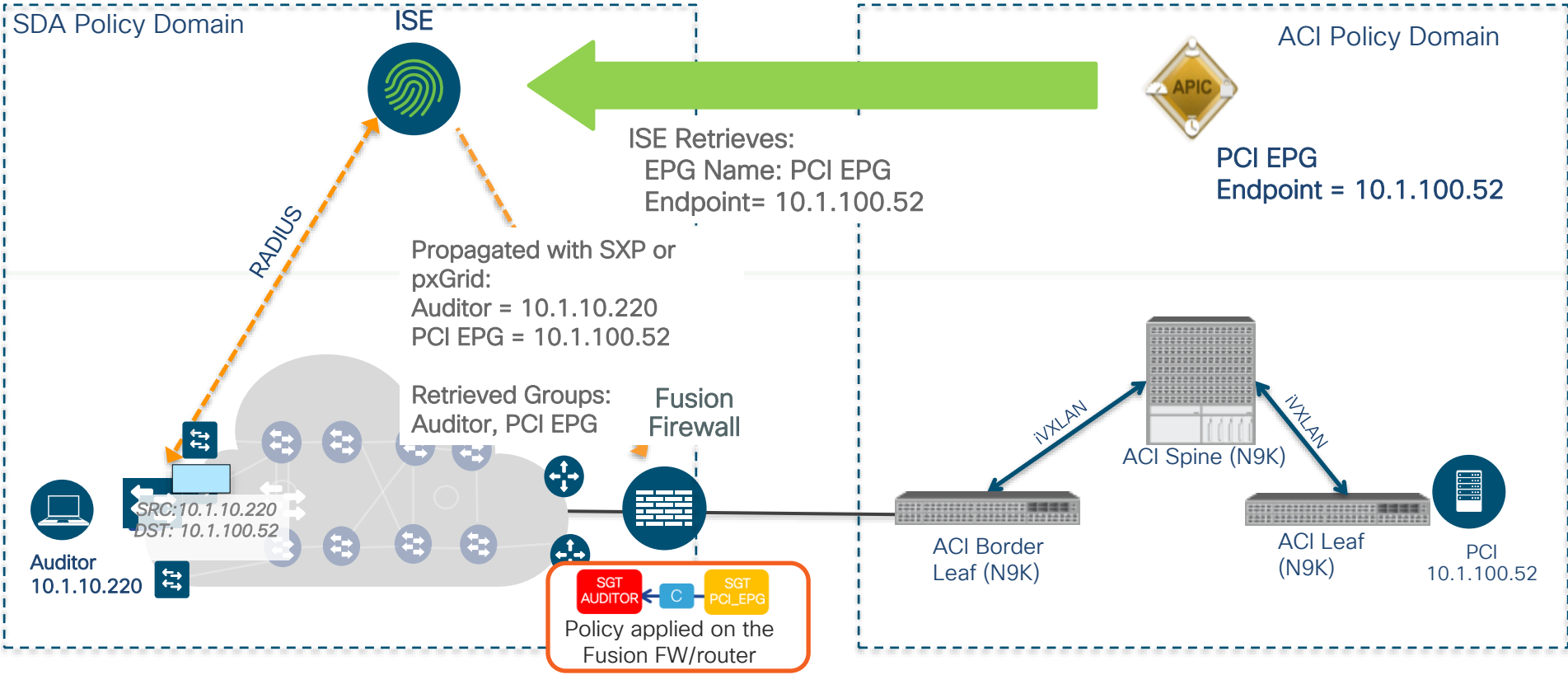
- Tenant Pod01
  - Application Profiles
    - AP
      - Application EPGs
        - DB
        - Development_Svr
        - HIPAA_Svr
        - Medical_Records
        - PCI_Svr
        - production_app
        - production_web

Application EPGs associated to the ACI Tenant integrated with ISE

* All the EPGs defined for the specific Tenant are retrieved

Policy Enforcement on the SDA Domain

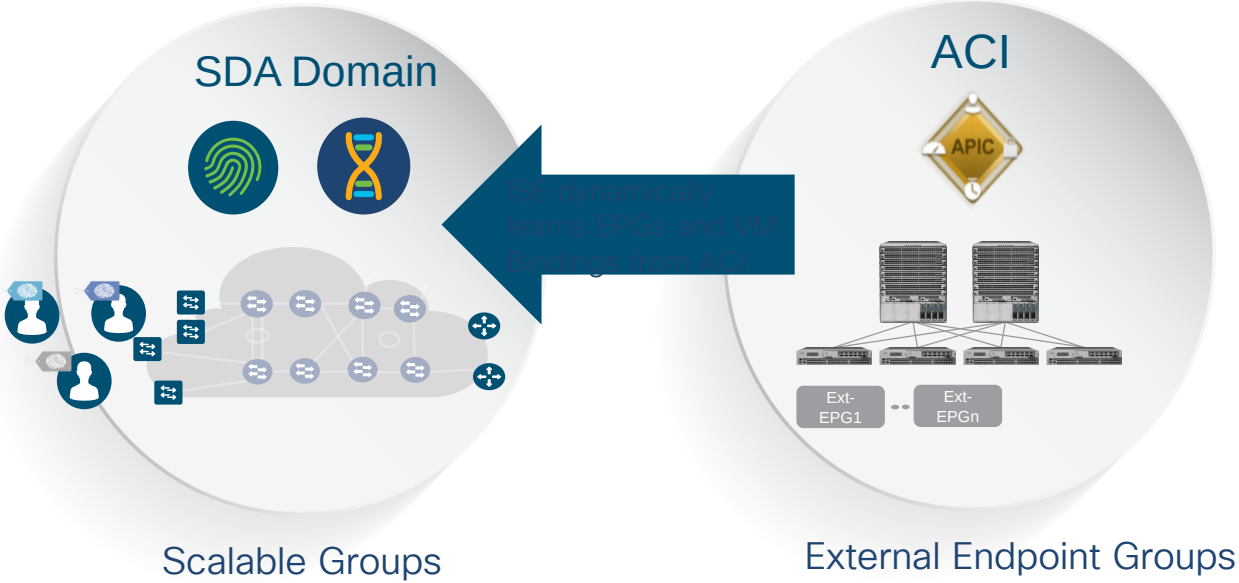
Apply Policy on the SDA Site (for Example on a Fusion Firewall)



Policy Enforcement on the SDA Domain
Scaling Enforcement in SDA Environment

ISE/SDA Scale	
Numbers of Groups	1000
Number of Mappings*	250k
SXP Peers*	200
pxGrid Peers**	200

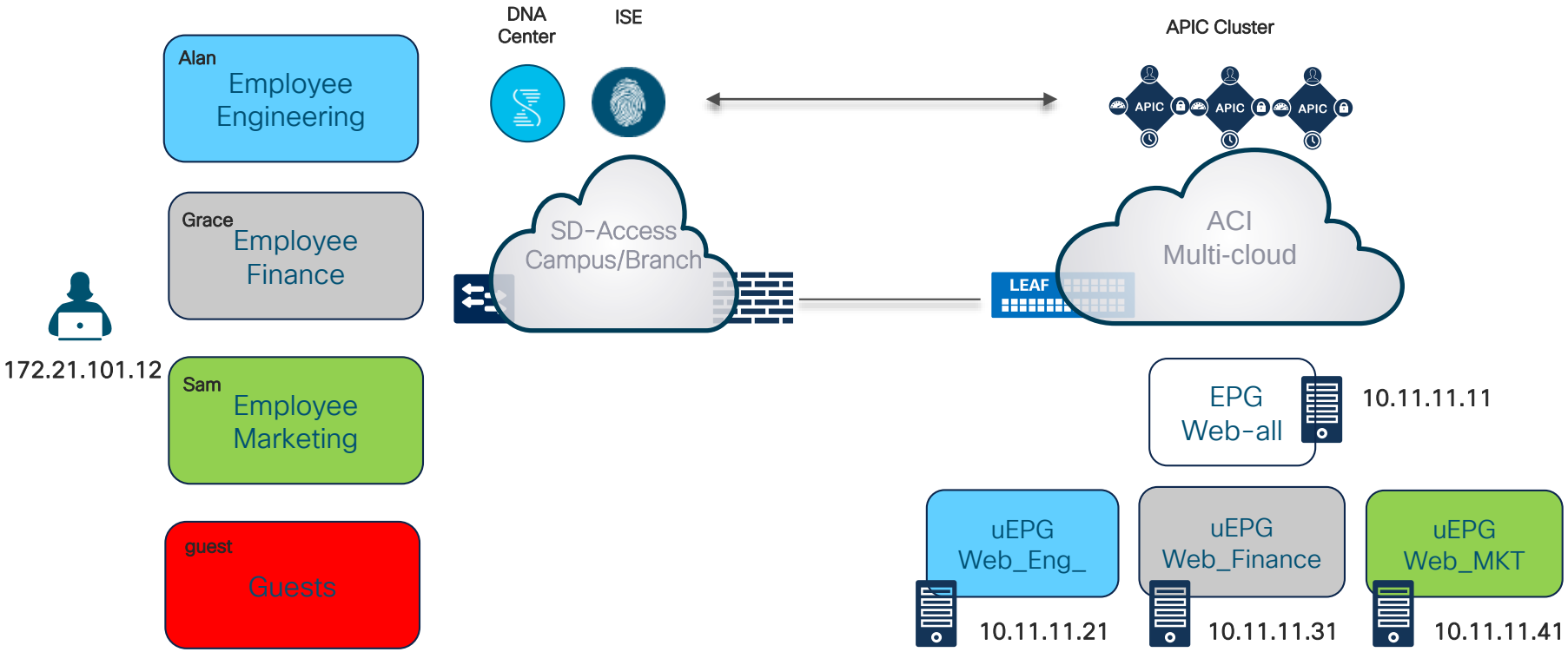
*Per pair of ISE SXP Nodes
** Per ISE pxGrid node



Fusion platforms capable of > 250k Mappings include:
ASR, ISR4k, C6800, ASA and FirePower Appliances

Demo 2: Multi-Domain Segmentation ACI and SDA

Topology ACI and SDA Integration - Phase 1



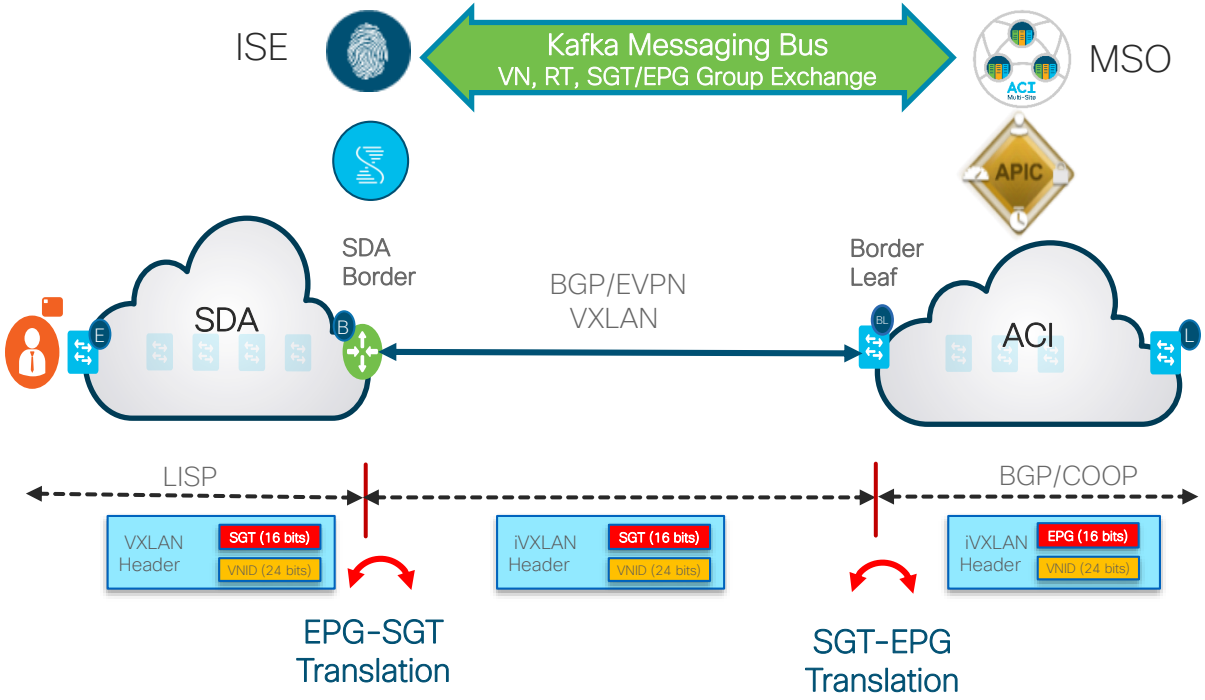
ACI and SDA Pairwise Integration – Phase 2

Disclaimer

- Phase 2 of ACI/SDA integration is currently planned for Q3CY20. As a consequence some of the specific implementation options described in the following slides may slightly change before FCS

Phase 2 SDA-ACI
Overall Architecture

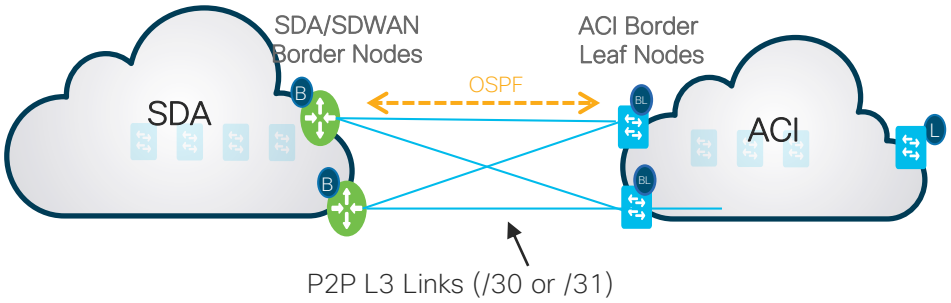
Q3CY20



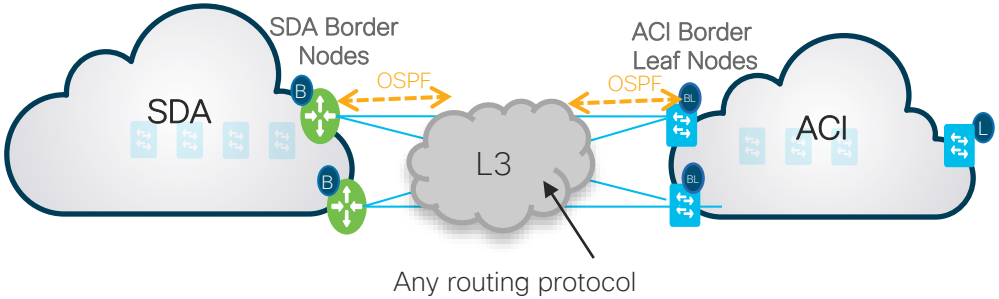
Area	Description
Policy Plane	• {VN, IP, group} sharing via Kafka Messaging Bus • Route Targets exchange between ACI and SDA (through ISE)
Control Plane	• BGP-EVPN with exchange of subnet information
Data Plane	• iVXLAN with group information • Endpoint data-plane learning on ACI BL nodes

Underlay Connectivity between SDA Border Nodes and ACI BL Nodes

Phase 2 SDA-ACI
Connecting Border Nodes



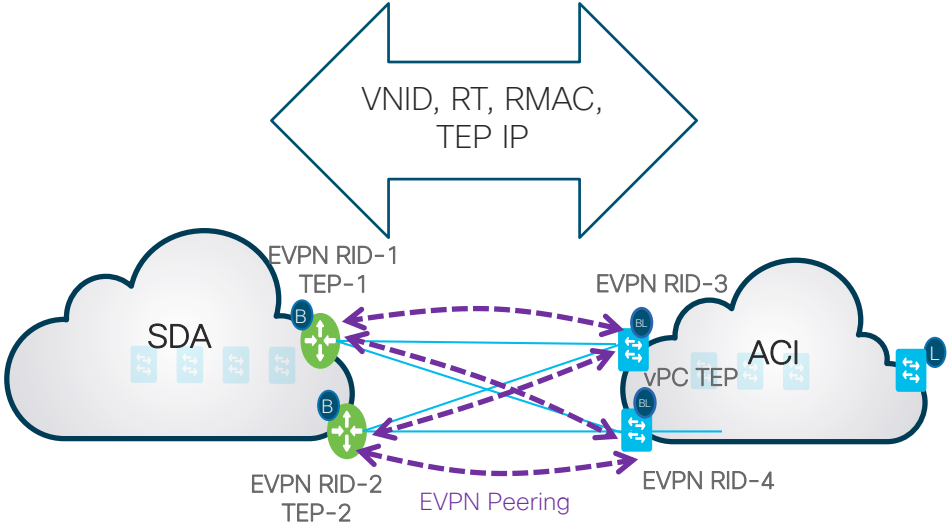
- Direct back-to-back connectivity between SDA Border Nodes and ACI Border Leaf Nodes
  - Square and full mesh topologies both supported
  - SDA Border Nodes could be co-located in the DC location
- Use of point-to-point Layer 3 interfaces (/30 or /31 mask)
- Point-to-point OSPF peerings to exchange TEP reachability information across domains
  - Independently configured in each domain



- SDA and ACI domains can also be connected through a generic Layer 3 infrastructure
  - Underlay configuration in the L3 network must be separately handled
  - Increased MTU support required in the L3 core
- OSPF still used between the border nodes and the first Layer 3 hop device in the L3 core
  - Possible to redistribute OSPF into a different protocol used in the core

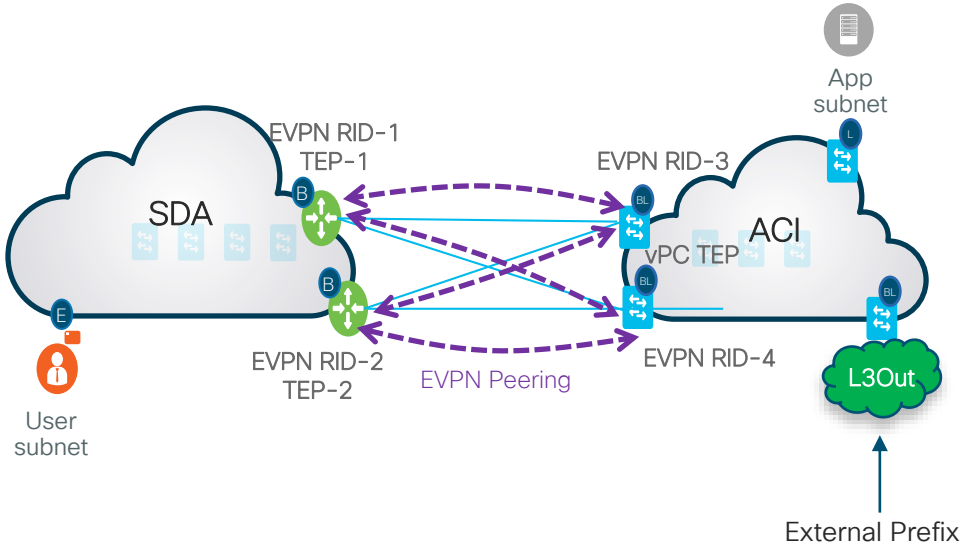
Control Plane Considerations

Control Plane Considerations
ACI-SDA BGP EVPN Peering



- BGP EVPN peerings to exchange prefix information for Campus and Application subnets
  - Full mesh BGP EVPN adjacencies between EVPN RIDs
- Information exchanged between border nodes:
  - VNIDs for the VRFs defined in each domain (downstream VNID assignment)
  - Route-Targets (RTs) value used to control import/export of prefixes into each VRF (Symmetric RT approach)
  - Router-MAC for the border node originating the prefix
  - TEP IP address to be used as next-hop
- VNID, RMAC, and TEP IP are used to construct VXLAN header for packets forwarded between SDA and ACI domains

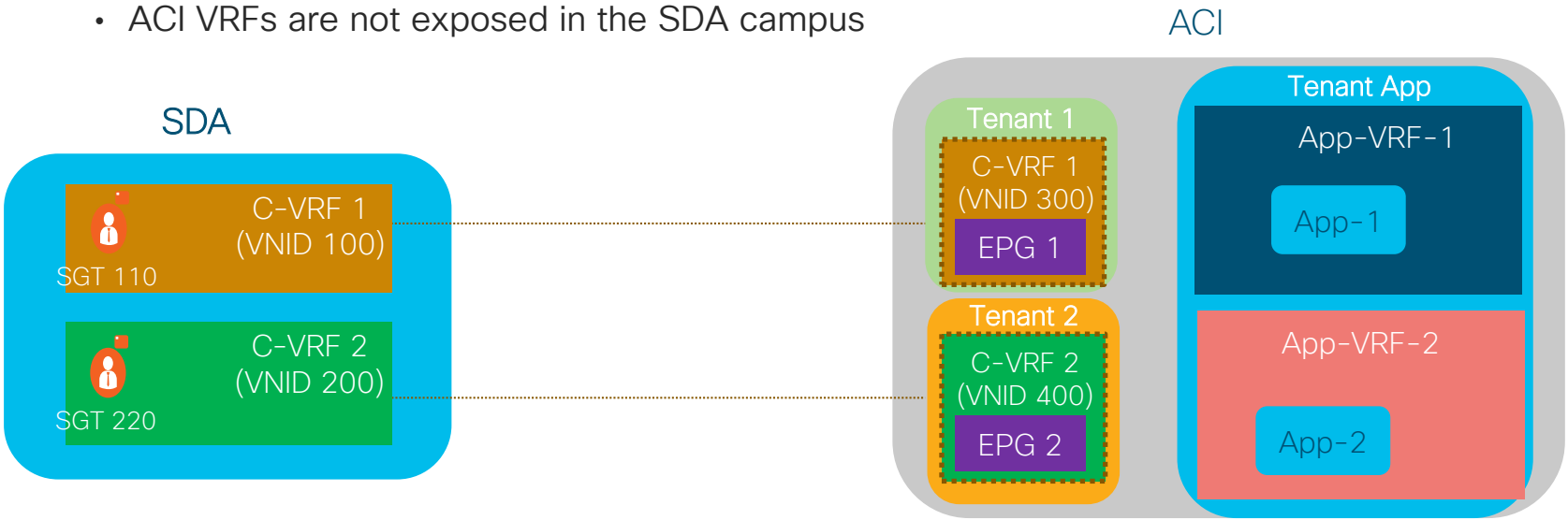
Control Plane Considerations
Route Exchange



- Routes are exchanged as EVPN type-5 routes with VNID of dedicated VRF
- Routes are imported within a domain based on the specific VRF RT import policies
- SDA border nodes push to ACI the subnets' routes for the users in the Campus requiring connectivity to the ACI services
- ACI BL nodes advertise the following routes to the SDA border nodes:
  - BD subnets for applications made available from ACI fabric
  - Prefix routes learnt in the ACI fabric from peers connected to other BL L3Outs
  - Specific /32 and /128 host routes for BDs that are stretched between ACI Pods/Sites (not at FCS)

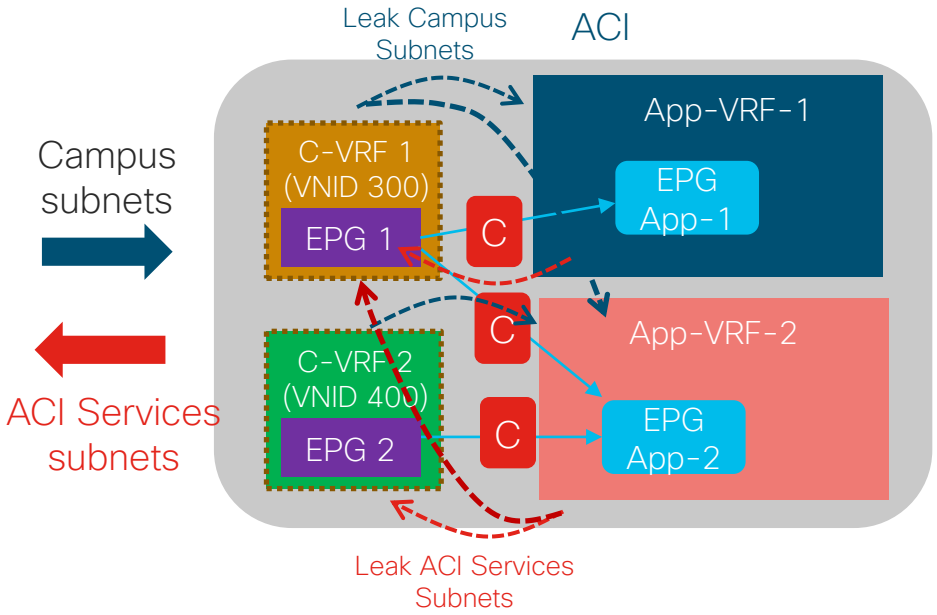
Campus VRF Extension into ACI

- Support multi-tenancy/multi-VRF design with minimal or no change to existing design on SDA and ACI side
- Allow campus to expose multiple VRFs to DC and ACI to expose apps from multiple VRFs to campus
  - SDA initiates a “Remote Tenant” setup in the ACI domain for each Campus VRF
  - For each defined Campus VRF (C-VRF) there is a corresponding C-VRF created on ACI
- ACI VRFs are not exposed in the SDA campus



Campus VRF Extension into ACI
Route-Leaking in ACI

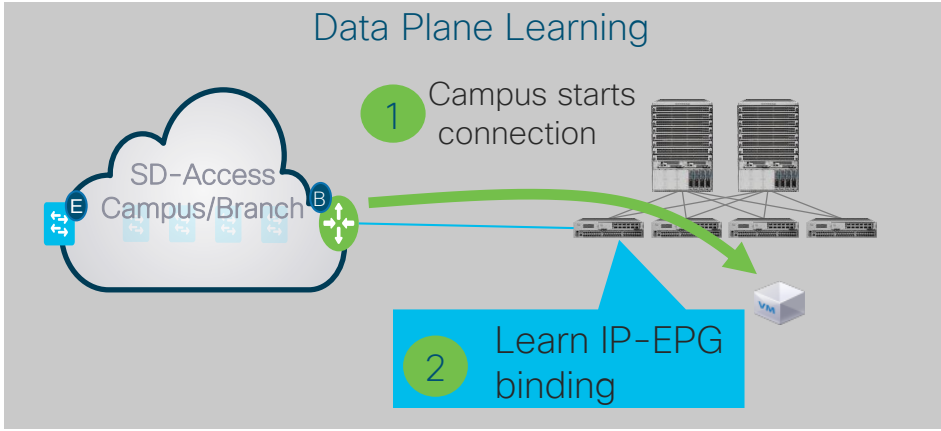
- **Campus SG consuming an ACI Service:** in ACI is represented as a “shared service” contract between C-VRF and the VRF(s) of the different Application EPGs representing the ACI services
- The subnets representing the ACI services will be leaked into C-VRF on the ACI Border Leaf nodes and advertised toward the Campus through BGP EVPN
- Similarly, the campus Subnets are advertised from the SDA border nodes into the C-VRF in ACI through BGP EVPN and leaked into one or more application VRFs

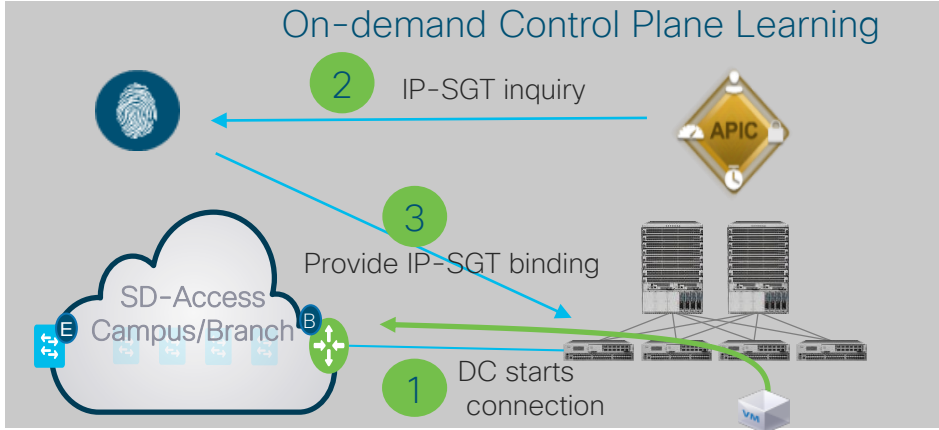


Data Plane Considerations

Phase 2 SDA-ACI
ISE-APIC/MSO Policy and Data Plane Learning

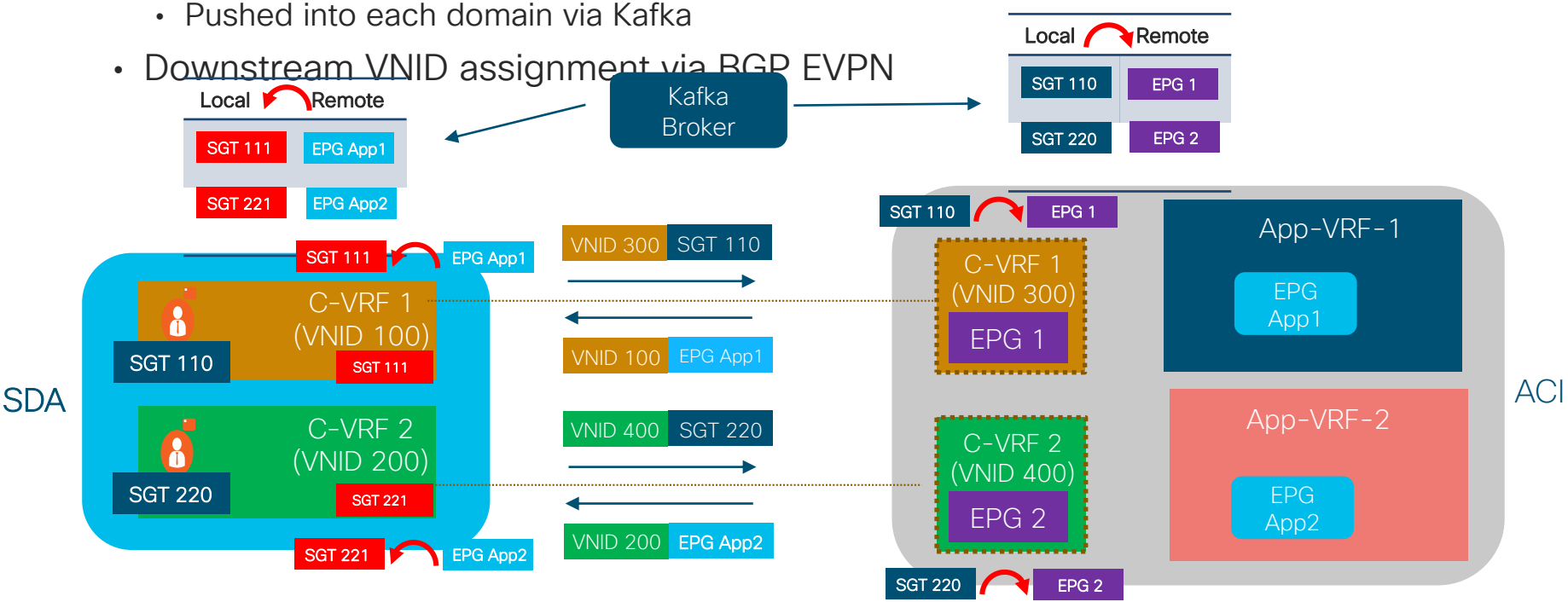
- Each campus SGT is represented as an External EPG in the ACI Border Leaf
- ACI BL learns the mapping of IP-SGT-EPG from data packet
  - Required because adjacent Campus IP addresses may be assigned to different SGTs
- Inquiry ISE for IP-SGT mapping when needed
  - E.g. when EP in ACI initiates the connection toward the campus
- Each domain can apply their policies independent of the other domain
  - On ACI the policy is always applied on the BL nodes (permit/deny/redirect to L4L7 graph, etc.)





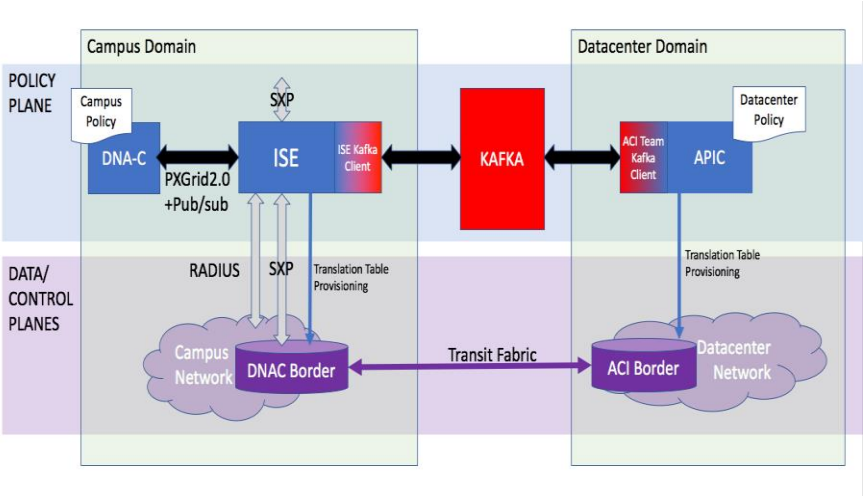
Campus VRF Extension into ACI
Class-ID Translation between Domains

- Class-ID translations to keep SDA and ACI separated domain for resource allocation
  - Pushed into each domain via Kafka
- Downstream VNID assignment via BGP EVPN



Cross-Domains Policies

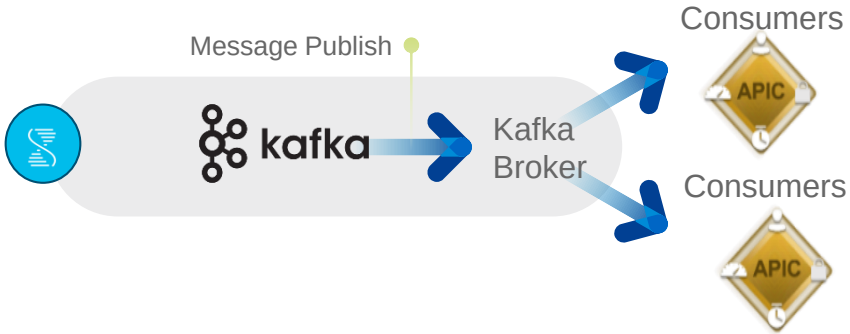
Cross-Domains Policies
Use of Kafka Communication Bus



- The Kafka broker does not need to run in a specific location
  - APIC controller cluster already has a Kafka cluster running
  - This Kafka cluster will be used for inter-domain communication at FCS
  - Post-FCS this functionality may be moved to MSO (or MDP)
- Kafka clients running on APIC and ISE
- Kafka Bus is used to join Domains: ACI is one domain (Datacenter), DNAC/ISE is another domain (Campus)
- The Kafka Interchange will allow the domains to peer, and to share data (Remote Tenant, Consumer/Service Gateway and Endpoint (IP-SG)) to allow the unification of Segmentation Policy

Phase 2 SDA-ACI
Cross-Domains Policy

- Use messaging bus (Kafka pub sub) model to publish/consume services between ACI and SDA domains
  - Flexible to construct cross domain policy without a 'uber controller'
  - BGP-EVPN and VXLAN config automation
  - VRF/Tenant instantiation
- Publish services provided by ACI
  - Domain, service name, EPGs, bindings, contracts, provider/consumer, protocol, ports
- Campus subscribe to DC service
  - ISE publishes Campus SGTs that represent the consumer and IP to SGT bindings



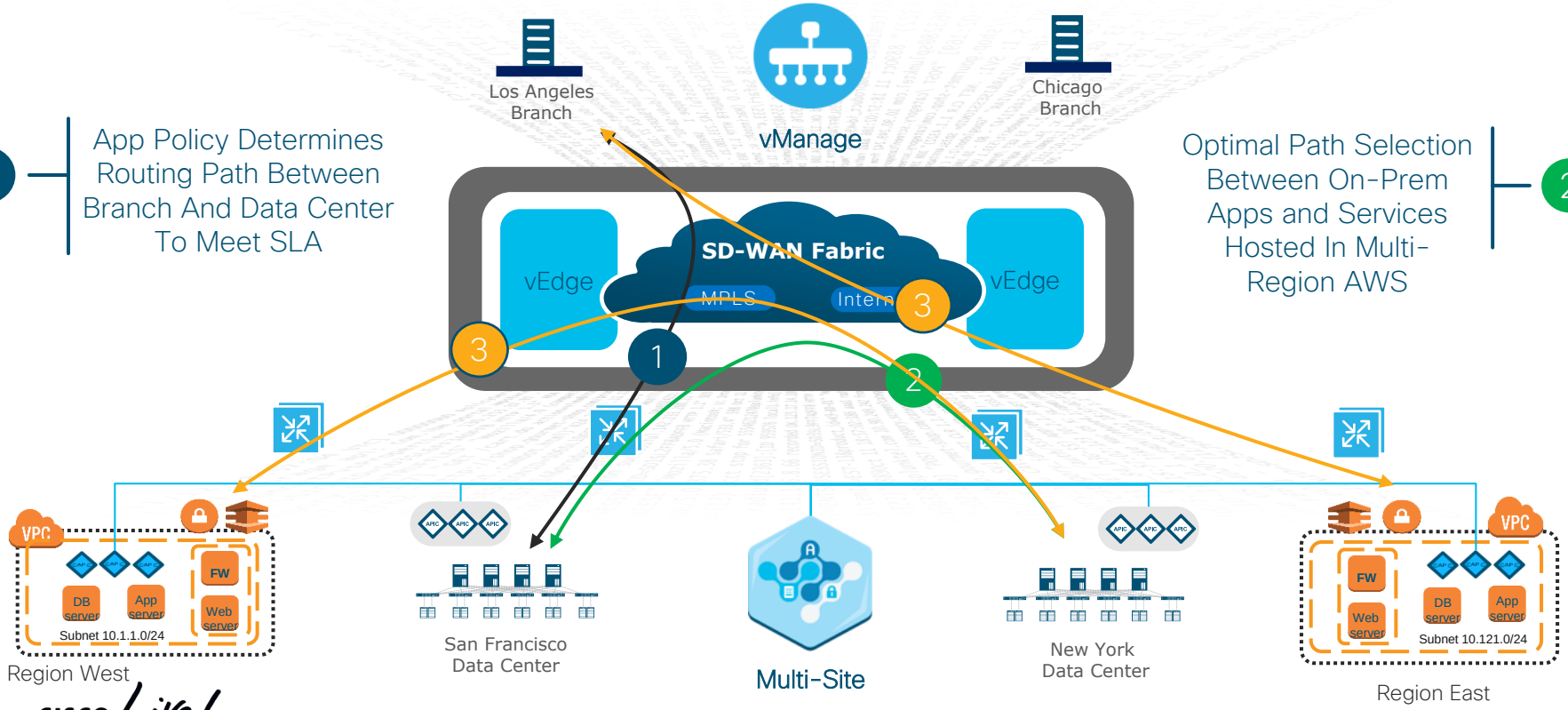
ACI and SDWAN Pairwise Integration

ACI and SD-WAN Pairwise Integration

Extend Operational Domain And Policy To Branch & Public Cloud

1 App Policy Determines Routing Path Between Branch And Data Center To Meet SLA

Optimal Path Selection Between On-Prem Apps and Services Hosted In Multi-Region AWS 2

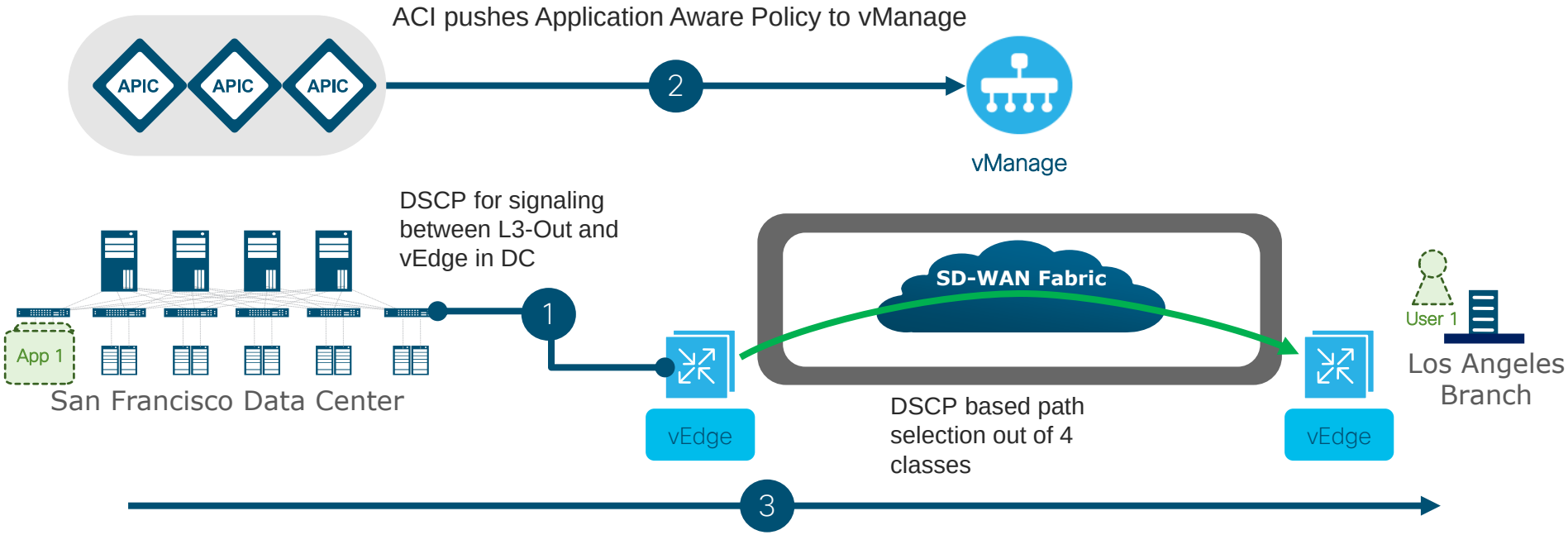


ACI and SD-WAN Pairwise Integration
Use Cases

1. Application Based Routing for North South Traffic
   a) ACI contract specifies outbound SD-WAN traffic engineering (tunnel selection, …) to vManage and marks outbound traffic accordingly
   b) ACI contract identifies IP/VIP to SD-WAN controller to mark inbound traffic
2. Remote Leaf, vPod and Multi-Site traffic optimization, transit of iVXLAN traffic across SDWAN
3. ACI Anywhere integration with SDWAN Cloud onRamp

ACI to SD-WAN (Viptela) Integration – Phase 1

ACI 4.1



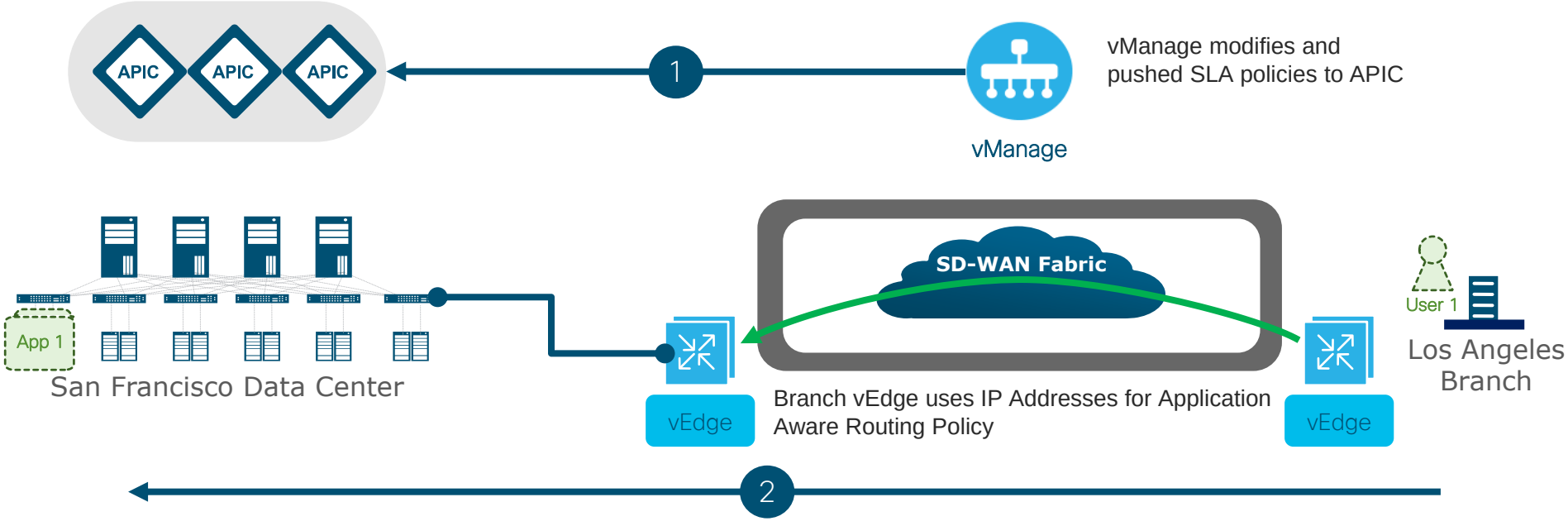
1) Physical Connectivity – L3Out per VPN

2) Application Policy – Export of Classification to vManage

3) Application Aware Routing – DC to Branch Ensured

ACI to SD-WAN (Viptela) Integration – Phase 1b

ACI 4.2



1) Application Policy – Export of Applications to vManage

2) Application Aware Routing – Branch to DC Ensured

ACI-SDWAN Release and Platform support

Release			Release
	Cisco ACI	4.2(X)/14.2(X)	
	vManage	19.2	
	IOS-XE SDWAN	16.12	

Platform Support			Platform Support
	ISR4K	ISR43xx, ISR-4431, ISR-4451	
	ASR1K	ASR1001-X, ASR1002-X, ASR 1001-HX, ASR 1002 -HX	
	ACI	EX, FX	

Configuration Details

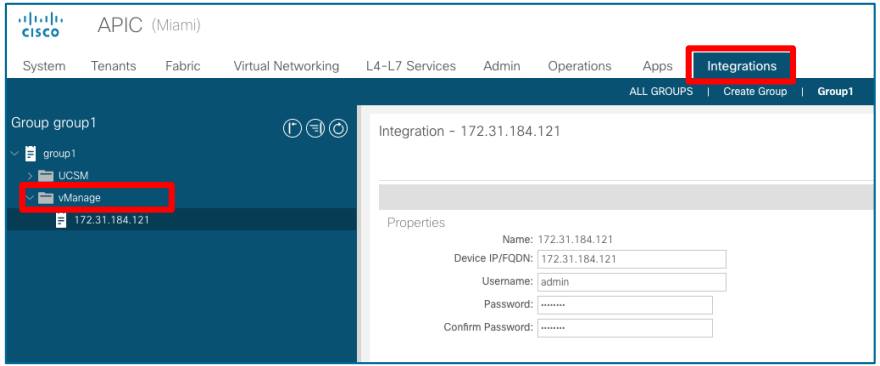
Configuration Overview

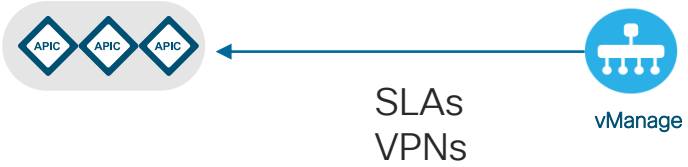
- ACI admin adds vManage integration to APIC
  - SD-WAN SLA policies are automatically pushed from vManage to APIC
  - SD-WAN VPN list is automatically pushed from vManage to APIC
- ACI admin assigns WAN SLA and VPN to the tenant VRF
- ACI admin assigns WAN SLA to L3Out contract

- SD-WAN admin provisions cEdge/vEdge devices and configures service VPNs towards ACI BL
- SDWAN admin adds a site list in vManage
- ACI partner registration appears in vManage
- SD-WAN admin attaches ACI controller to cEdge/vEdge devices
- SD-WAN admin configures central policy for SLA, VPN, and site list and pushes policy to vSmart
- vSmart pushes policy to cEdges/vEdges

Adding vManage Integration

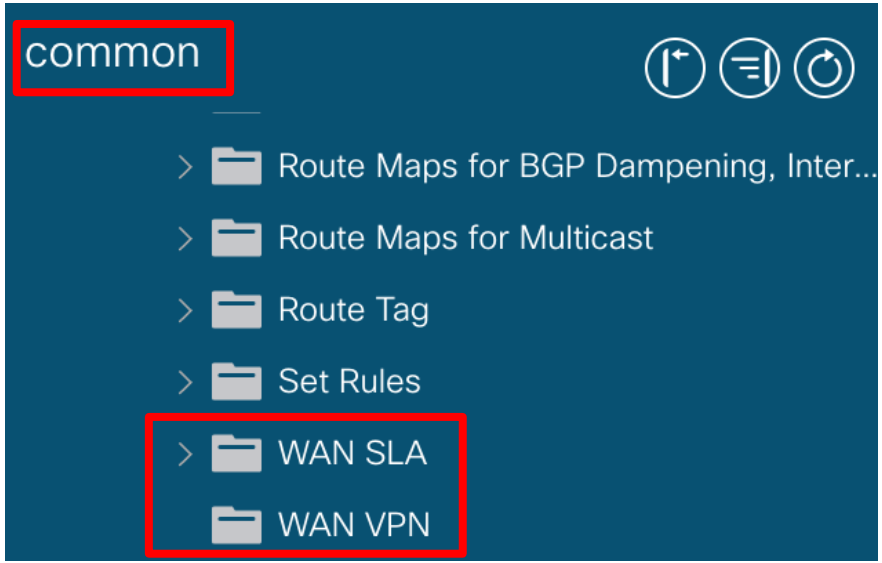
vManage is added to the APIC under the Integrations tab.





APIC pulls SLA policies and VPNs from vManage

ACI SD-WAN policies



From ACI 4.1 there are two new WAN folders under the common tenant protocol policies.

- **WAN SLA** displays SLA policies pulled from vManage
- **WAN VPN** displays VPNs pulled from vManage

ACI WAN SLA Policies on APIC

- The WAN SLA policies are defined as jitter, delay, and loss values of the WAN link
- SLA policy to DSCP value are randomly chosen
- When an SLA policy is added to an L3Out contract, the border leaf will mark traffic with the DSCP value mapped to that SLA class
- WAN cEdge/vEdge use the DSCP value to map to correct SLA class
- Modifying SLA policies is done on vManage
- Only 4 SLA classes are currently supported

Wan SLA Policies

Name	DSCP	Acceptable Jitter (ms)	Acceptable Delay (ms)	Acceptable Loss (%)
Bulk-Data	AF12 medium drop	100	300	10
Default	AF13 high drop	100	300	25
Transactional-Data	AF11 low drop	100	50	5
Voice-And-Video	AF21 low drop	100	45	2

ACI WAN SLA Policies on vManage

Cisco vManage

admin

CONFIGURATION | POLICIES Centralized Policy > Add Policy

Create Groups of Interest — Configure Topology and VPN Membership — Configure Traffic Rules — Apply Policies to Sites and VPNs

Select a list type on the left and start creating your groups of interest

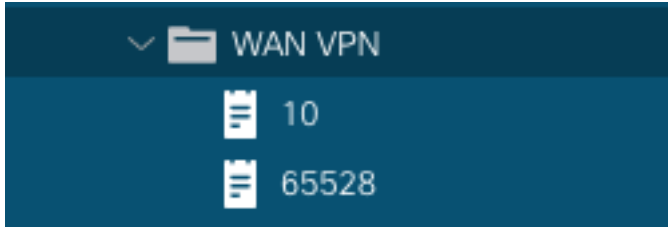
- Application
- Color
- Data Prefix
- Policer
- Prefix
- Site
- SLA Class
- TLOC
- VPN

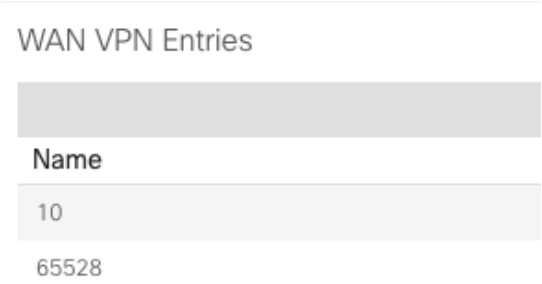
New SLA Class List

Name	Loss (%)	Latency (ms)	Jitter (ms)	Reference Count	Updated By	Last Updated	Action
Voice-And-Video	2	45	100	0	system	25 Sep 2019 1:27:43 PM	
Transactional-Data	5	50	100	0	system	25 Sep 2019 1:27:40 PM	
Default	25	300	100	0	system	25 Sep 2019 1:27:47 PM	
Bulk-Data	10	300	100	0	system	25 Sep 2019 1:27:49 PM	

SLA parameters modified on vManage are exposed to APIC

WAN VPNs





- Service VPN segments in vManage are pulled by APIC and appear under WAN VPN
- Each VPN segment represents a routing domain in the SD-WAN network and can be mapped to a VRF on the ACI domain

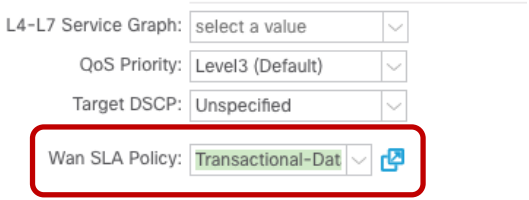
Tenant Configuration

Tenant admin configuration steps:

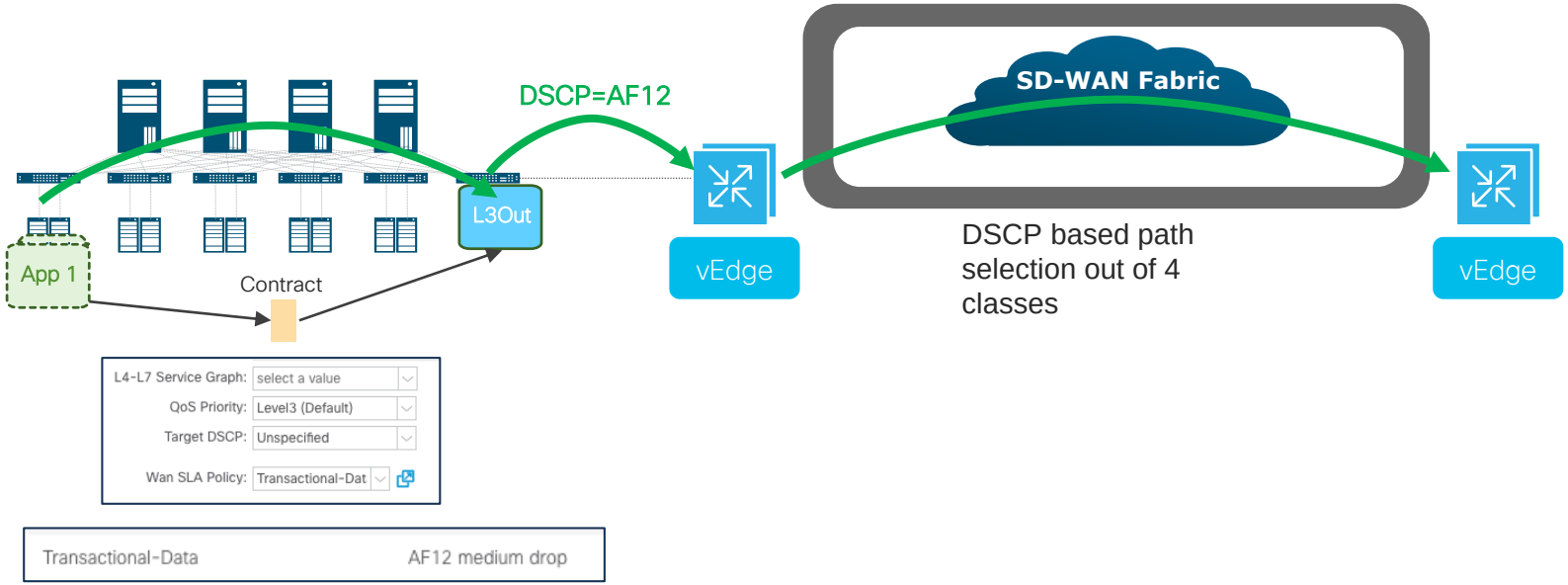
1. Assign VPN to VRF
2. Add SLA policy to contract subject

Under the VRF select a VPN. This is read from the common tenant WAN VPN list

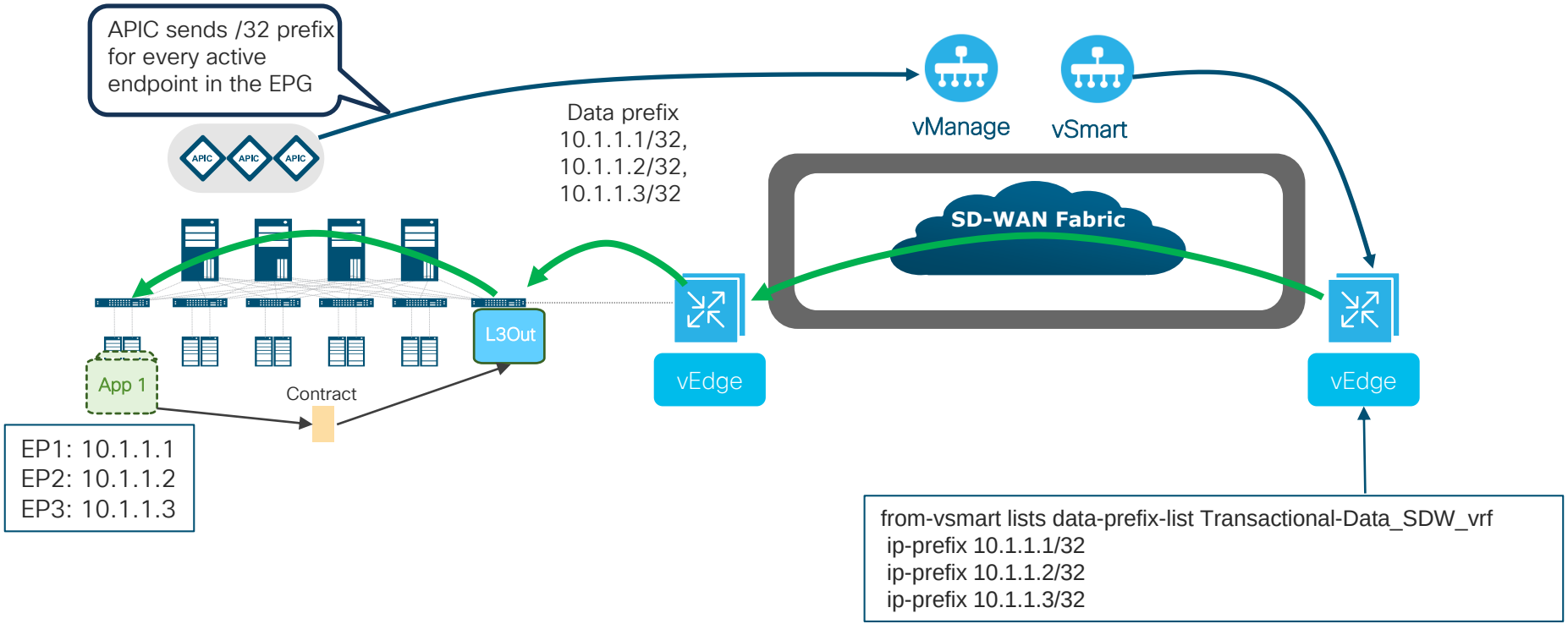
Under the contract subject select the WAN SLA policy and select a QoS Priority (cannot be undefined)



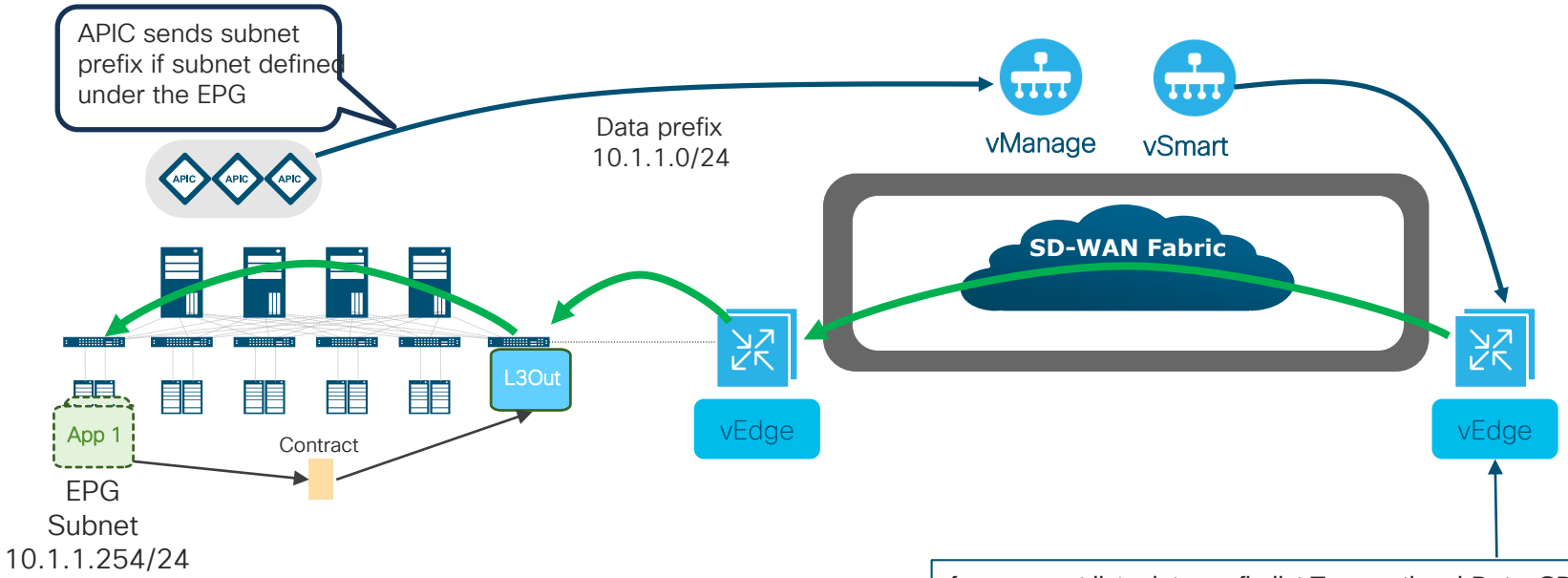
ACI Data Center to Branch



ACI Branch to Data Center



ACI Branch to Data Center



from-vsmart lists data-prefix-list Transactional-Data_SDW_vrf
ip-prefix 10.1.1.0/24

If the subnet is defined under the EPG, only the subnet prefix will be advertised to vManage (not the /32 prefixes)

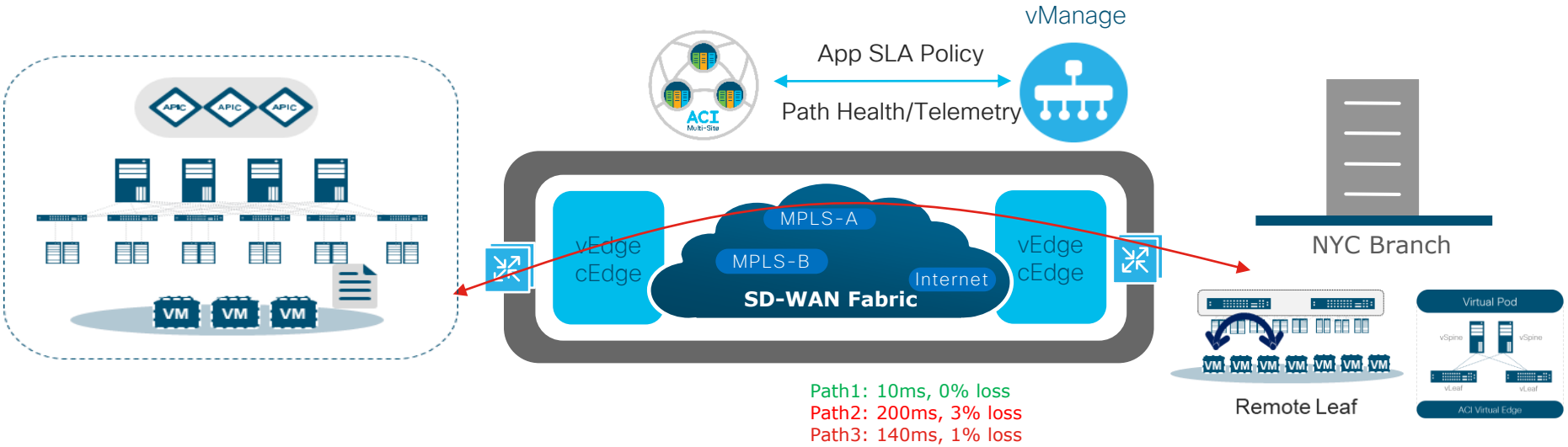
Use Case 2

Differentiate Traffic Flows between ACI Domains

Planned – 5.1 Release

ACI SD-WAN – Use Case 2

Branch with Remote Leaf or vPod Use Case



- App SLA policy determines routing path selection between ACI Fabric and RL or vPod branch location to meet SLA (E-W traffic)
- iVXLAN tunnel traffic is carried over SDWAN
- SLA policy is configured in the DSCP bits of the iVXLAN header

Planned – 5.1 Release

ACI SD-WAN – Use Case 2

ACI Multi-Site + SD-WAN Integration Use Case

1

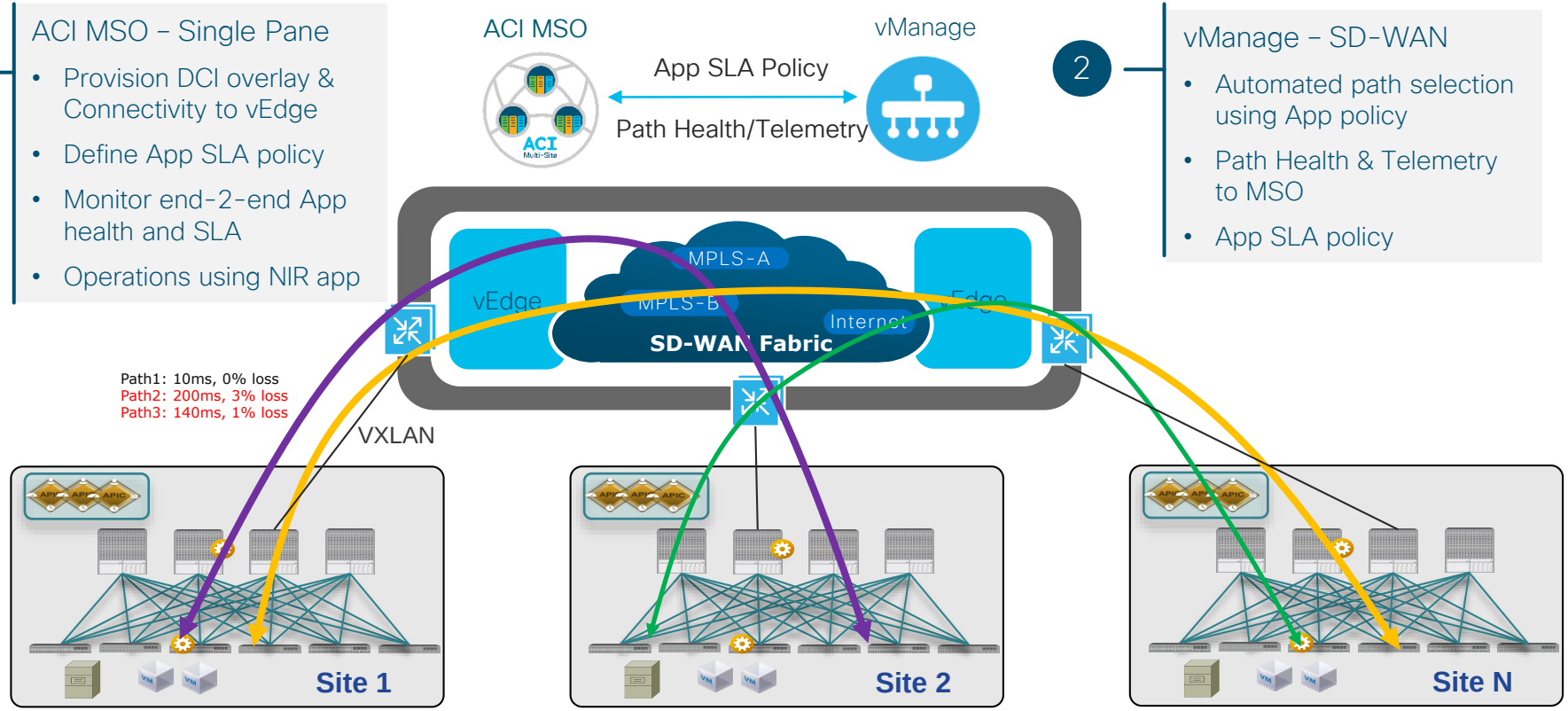
ACI MSO – Single Pane

- Provision DCI overlay & Connectivity to vEdge
- Define App SLA policy
- Monitor end-2-end App health and SLA
- Operations using NIR app

2

vManage – SD-WAN

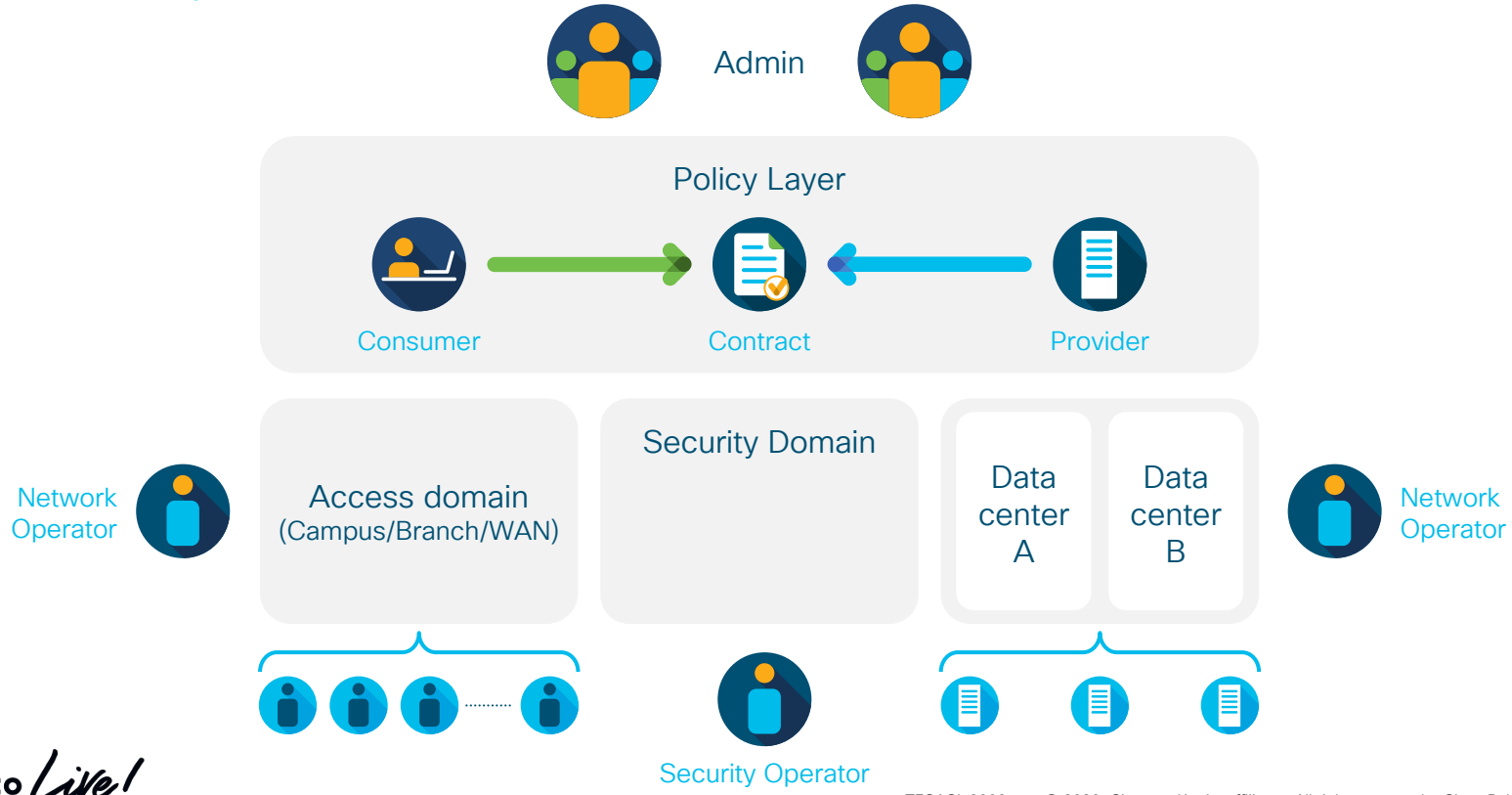
- Automated path selection using App policy
- Path Health & Telemetry to MSO
- App SLA policy



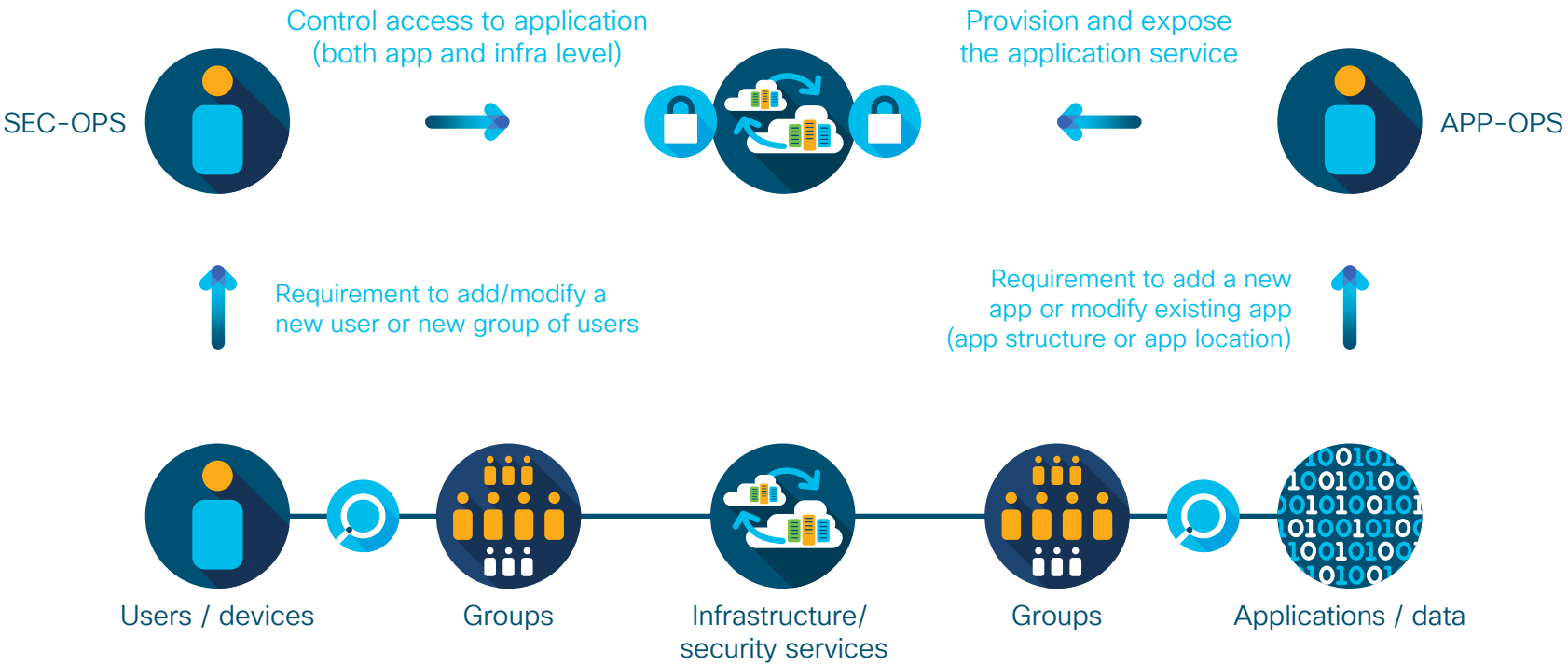
Next Steps

Unified policy language across domains

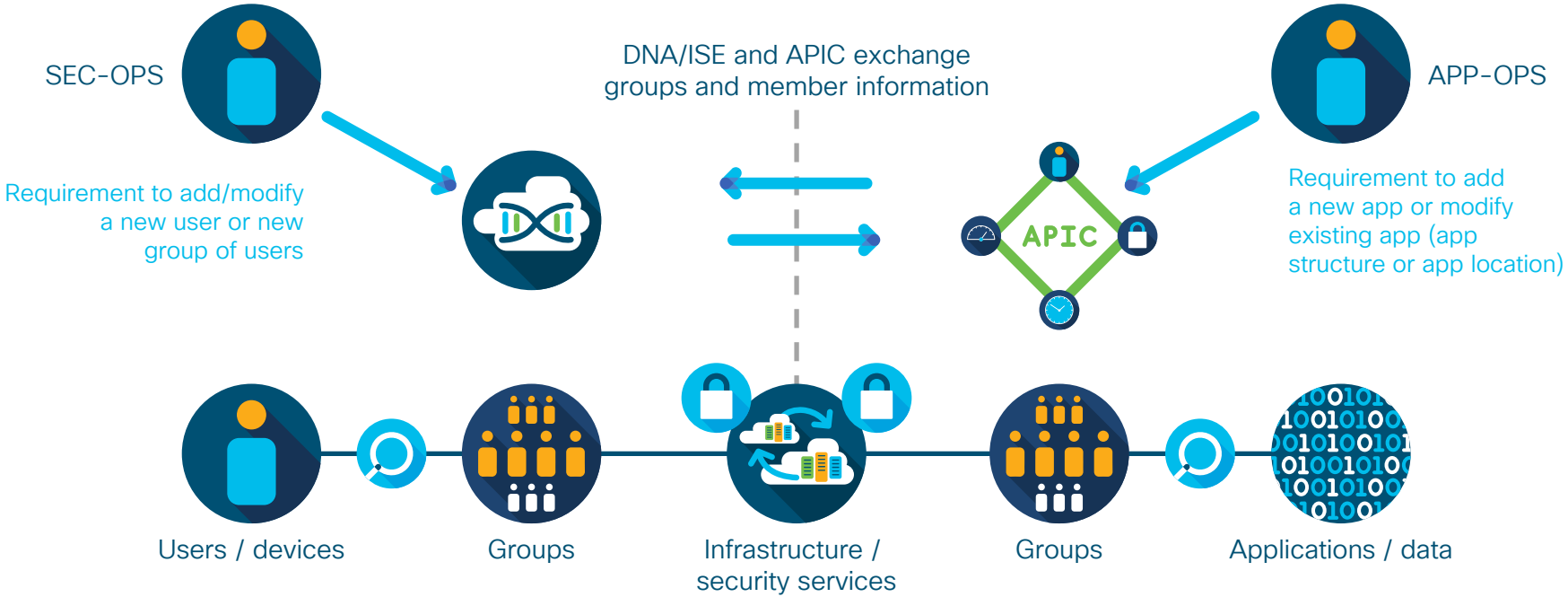
Consumer to provider



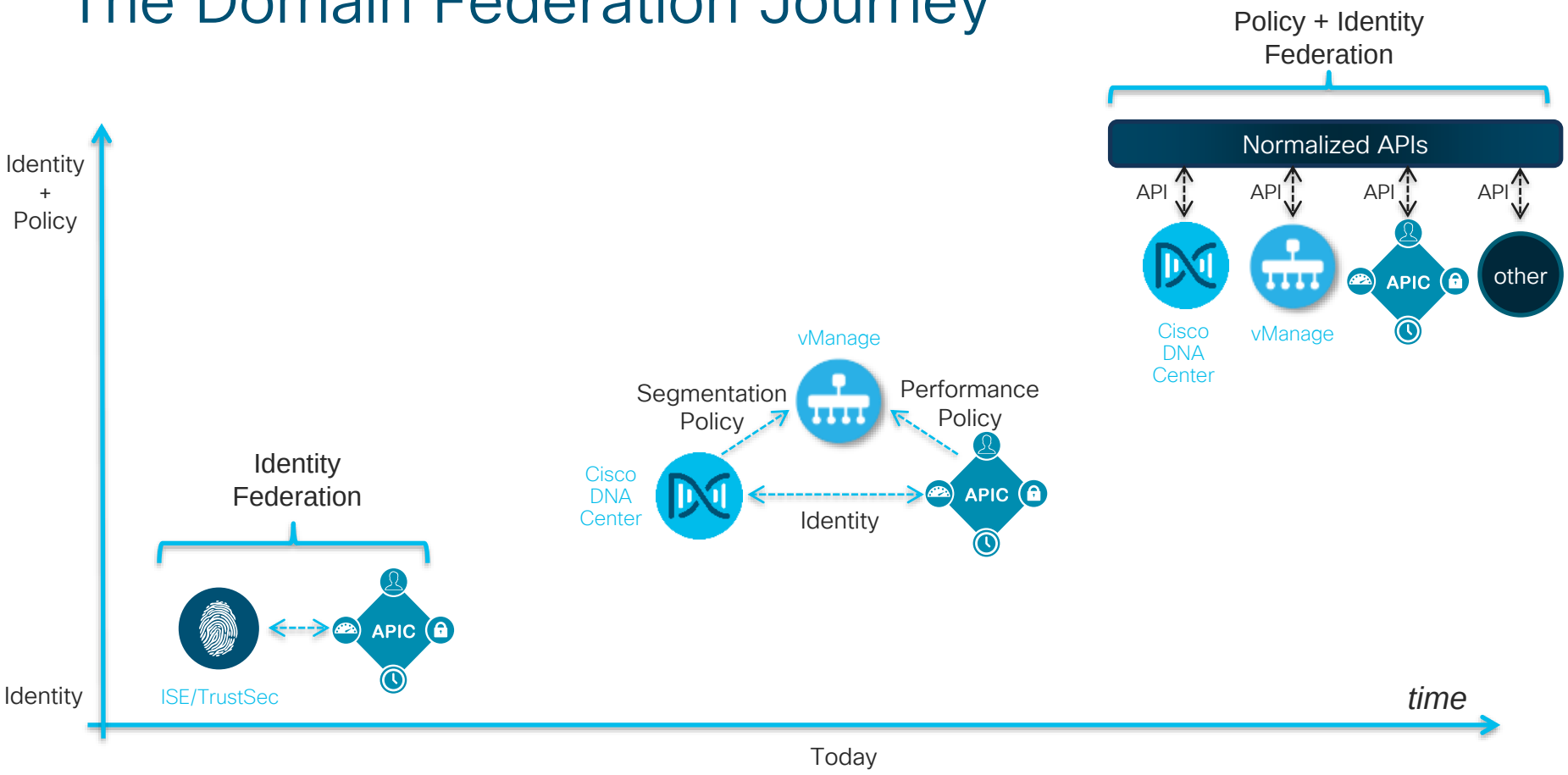
Business change view of policy



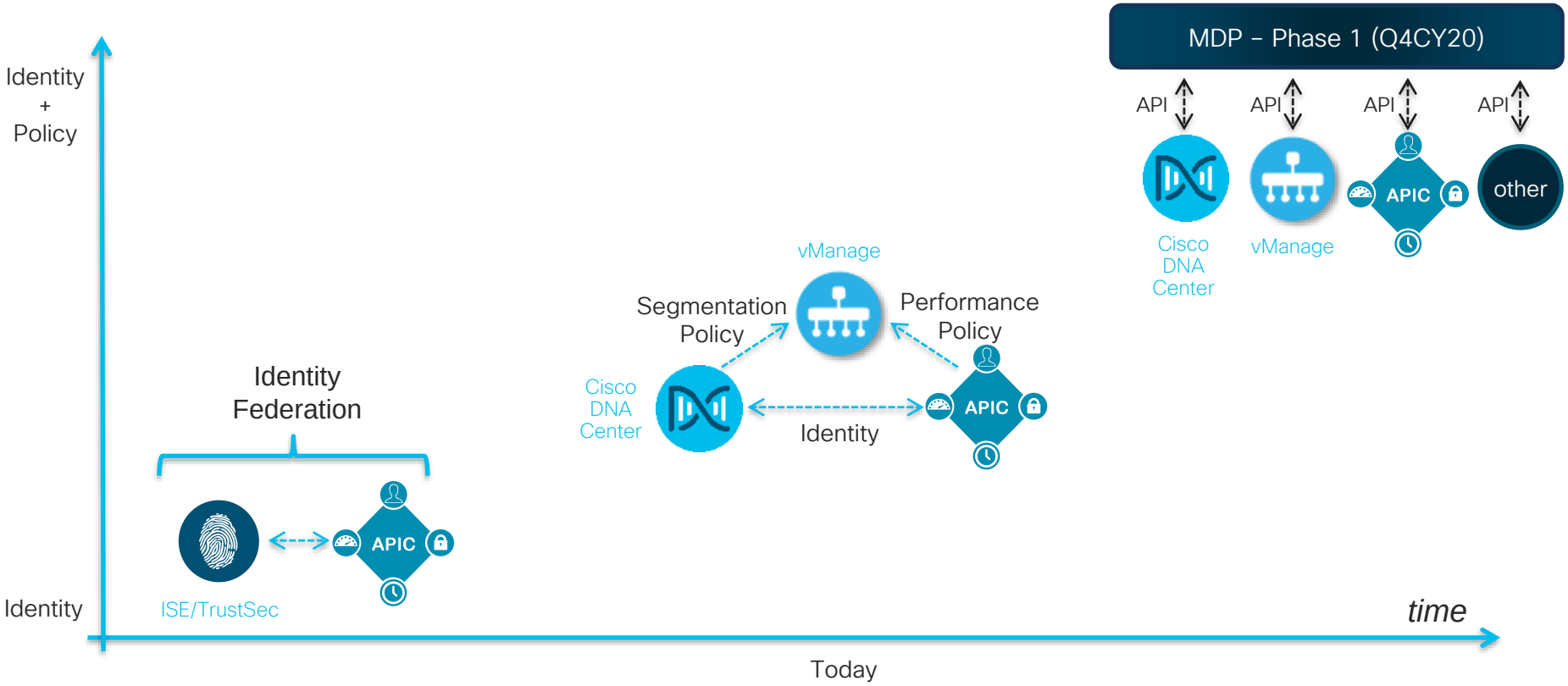
Federation of identity and policy change across domains



The Domain Federation Journey



The Domain Federation Journey



How to Operate ACI Fabric(s)

FCAPS and ITIL

FCAPS is a Network Management Framework (evolved from ISO Telecommunications Management Network)

ITIL v3 is a collection of (good) practices for IT as a whole. Focused on Lifecycle, Service Delivery, Support and Improvements

ITIL framework for service management and operations

IT Service Strategy - ITIL's Service Strategy provides a set of frameworks for determining what services are delivered, how their value is measured, how to measure cost and provide a measure of return on investment (ROI), and how to manage IT's relationship with its business partners.

IT Service Design - This area covers design of processes and how they relate to one another, Service-Level Agreements (SLA), capacity and availability management, business continuity management, security, and supplier management. IT Service Design also notes the need for a service catalog

IT Service Transition - Service Transition governs how services are delivered and deployed. Such areas as change management, release and deployment management, and service evaluation are typically part of the transition phase. When cloud is on scope, the actual tasks when deploying a service to the cloud change significantly. IT departments should set up a test cloud environment that mirrors the production environment in order to allow User Acceptance Testing (UAT).

IT Service Operation - Service Management covers the management and monitoring of services, and how issues are managed and resolved. Key to the Service Management component is the notion of a Service Desk

IT Continual Service Improvement - In Continual Service Improvement (CSI), IT personnel and business teams work together to ensure services can quickly meet new and emerging business requirements. CSI is heavily data-driven and relies upon operational statistics as well as business insights

ITIL framework for service management and operations

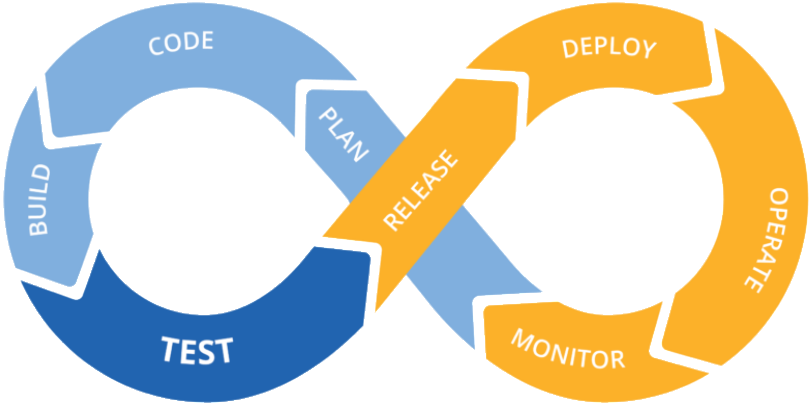
IT Service Strategy - ITIL's Service Strategy provides a set of frameworks for determining what services are delivered, how their value is measured, how to measure cost and provide a measure of return on investment (ROI), and how to manage IT's relationship with its business partners.

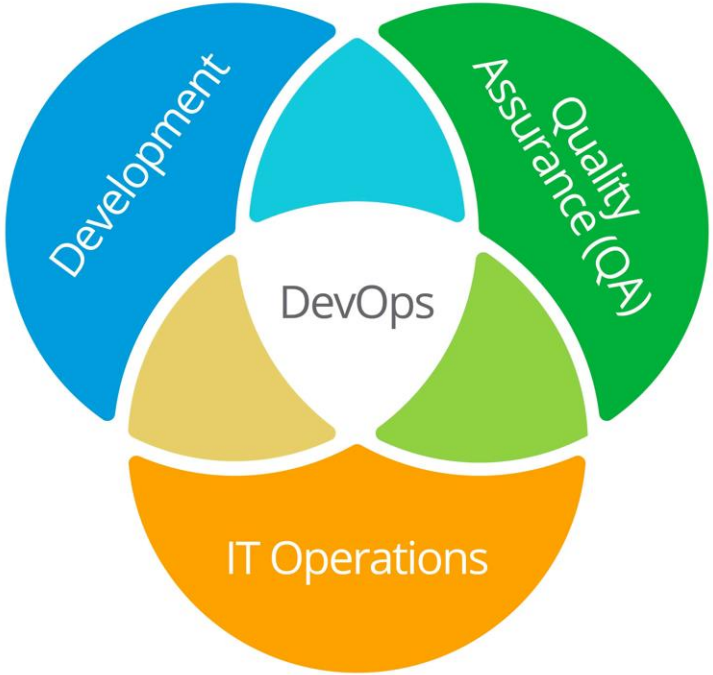
IT Service Design - This area covers design of processes and how they relate to one another, Service-Level Agreements (SLA), capacity and availability management, business continuity management, security, and supplier management. IT Service Design also notes the need for a service catalog

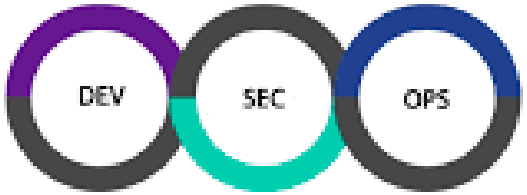
IT Service Transition - Service Transition governs how services are delivered and deployed. Such areas as change management, release and deployment management, and service evaluation are typically part of the transition phase. When cloud is on scope, the actual tasks when deploying a service to the cloud change significantly. IT departments should set up a test cloud environment that mirrors the production environment in order to allow User Acceptance Testing (UAT).

IT Service Operation - Service Management covers the management and monitoring of services, and how issues are managed and resolved. Key to the Service Management component is the notion of a Service Desk

IT Continual Service Improvement - In Continual Service Improvement (CSI), IT personnel and business teams work together to ensure services can quickly meet new and emerging business requirements. CSI is heavily data-driven and relies upon operational statistics as well as business insights

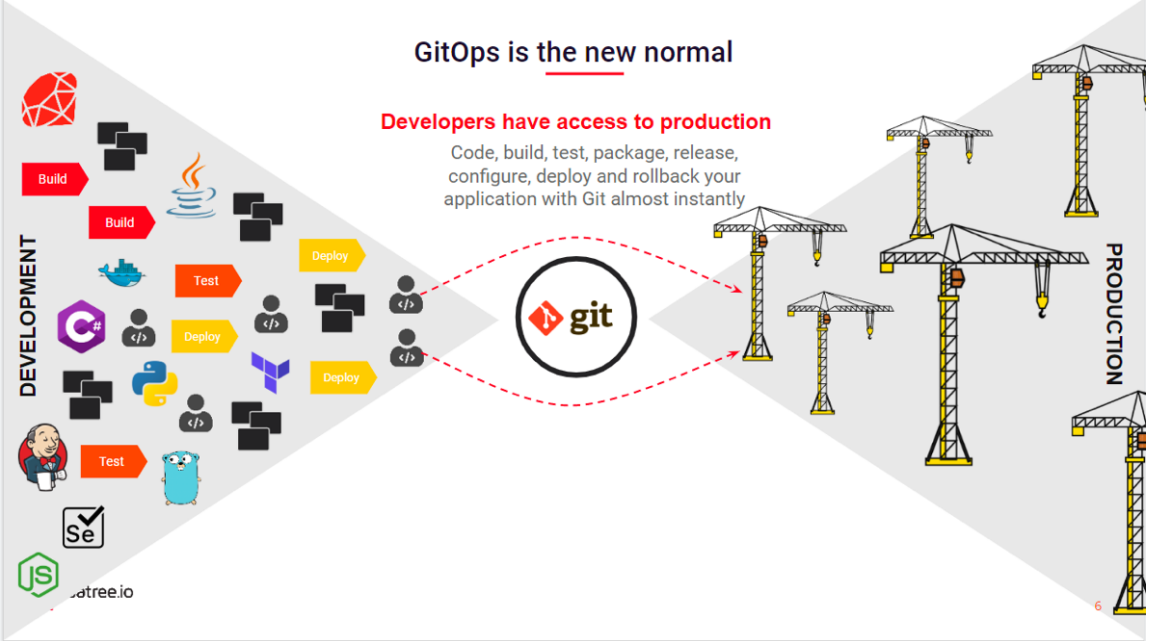


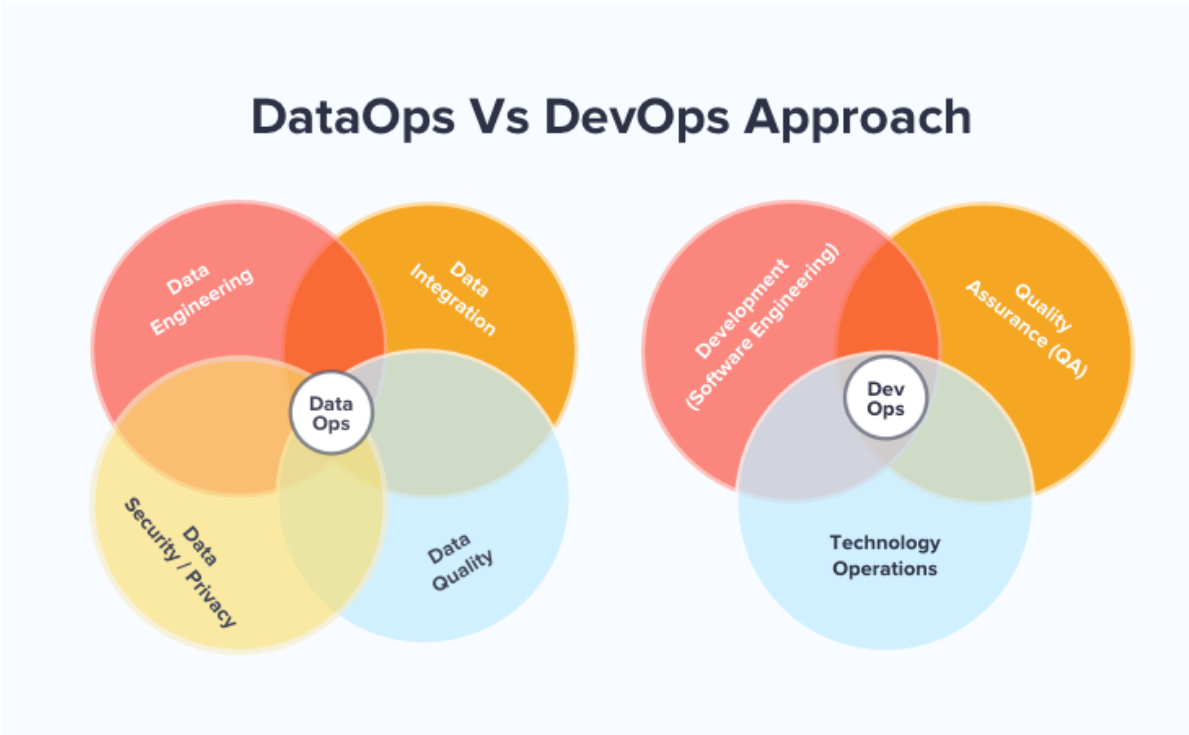




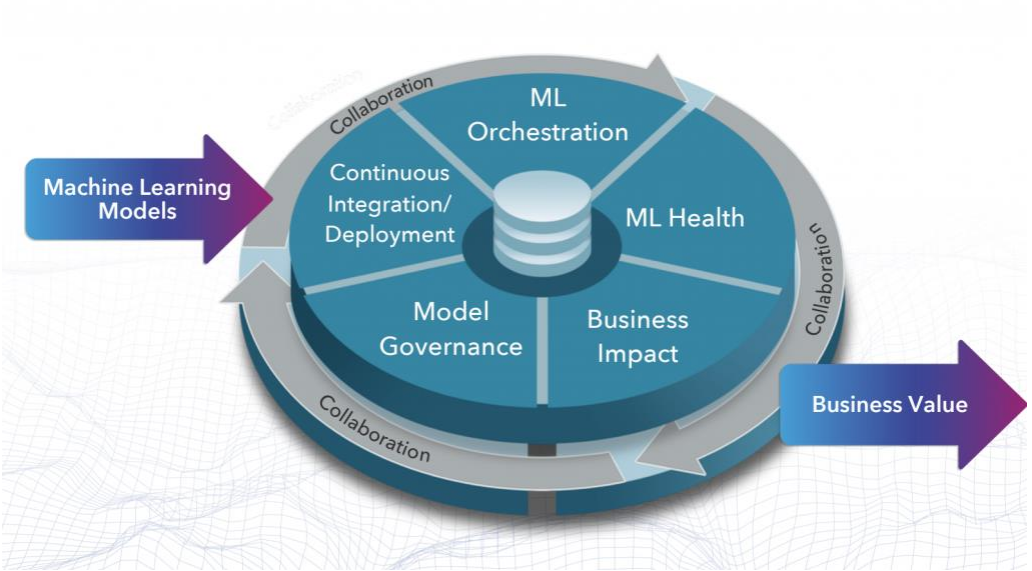
INTEGRATING SECURITY INTO PRODUCT LIFECYCLE: DEVSECOPS

gitops





MLOps



Infrastructure-as-Code (IaC) Goals

- **Unify** the view of resources
- **Support** the modern data center (IaaS, PaaS, SaaS)
- Expose a way for individuals and teams to **safely and predictably change** infrastructure
- Provide a workflow that is **technology agnostic**
- Manage anything with an **API**

Google has defined a model called **Site Reliability Engineering (SRE)** as a collection of best practices for any digital business

Example: Cloud Strategy Team at Microsoft IT

First and Best

Adoption Data, Lessons Learned, & Showcases

Enterprise Architecture

Cloud Patterns & Guidance Creation

Cloud Strategy Team

Service Capability & Improvement

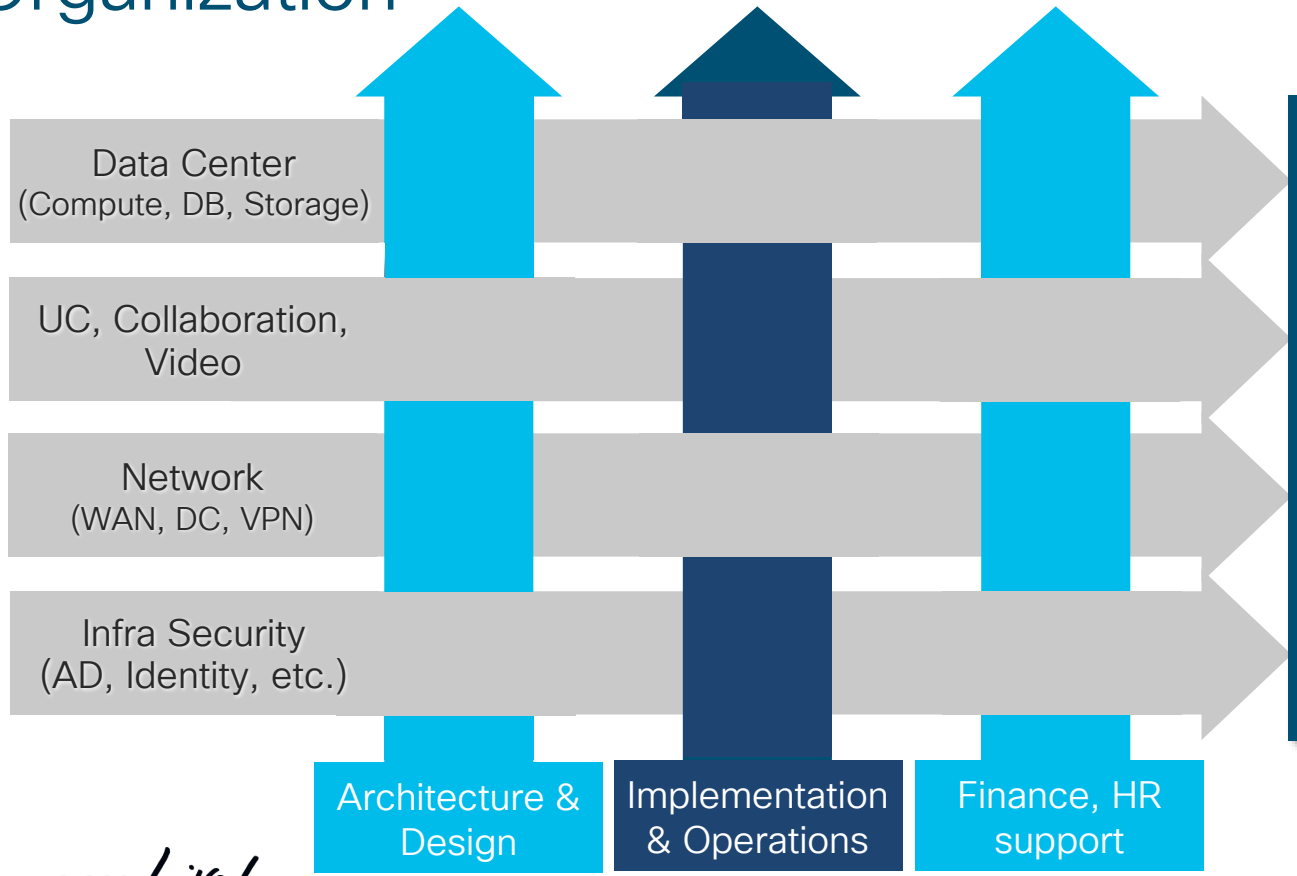
Cloud Providers

Cloud Analysis, Guidance Delivery, & Roadmaps

Corporate Functions IT	Devices & Services IT	Enterprise Commerce IT	Tech Support
Enterprise Infrastructure & Services IT	Information Security Risk Management	Marketing Products & Services IT	Strategic Enterprise Services IT

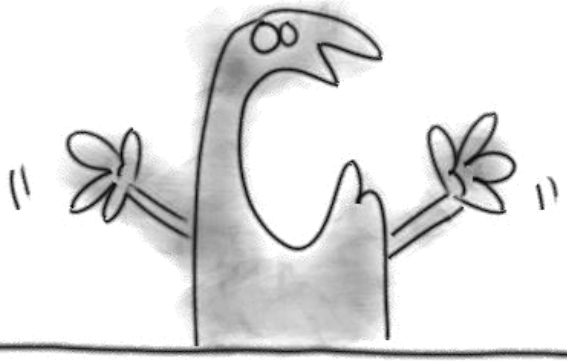
Enterprise Cloud Strategy - 2nd Edition – Microsoft Press

Example: Cisco IT Service-Based Organization



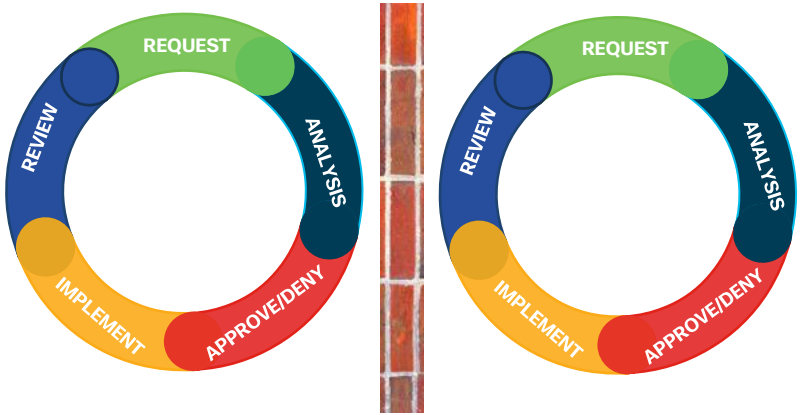
- Everything is a service
- Service owner is the GM
- Budget, roadmap, metrics, etc. are the responsibility of the service owner
- Interlaced with functions common across all services

Now What?!!



Current Culture and Mindsets within Enterprise

Traditional Mindset



Avoid failure

Change is Risky

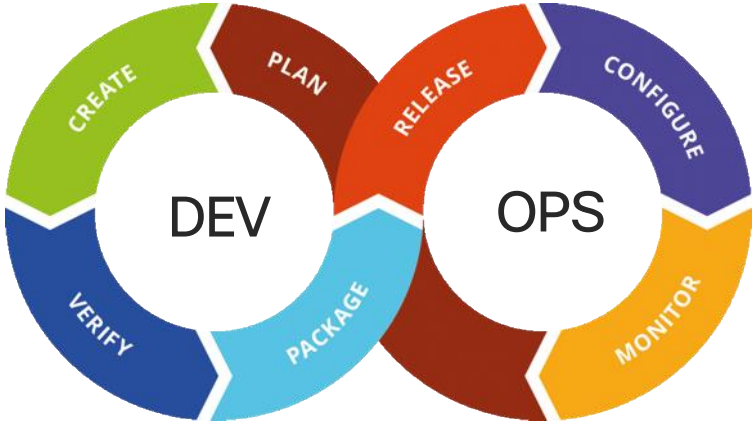
Change is Complex

Empowered accountability

Limited Feedback systems

Manual

DevOps Mindset



Embrace failure

Change is good

Active collaboration

Empowered accountability

Feedback systems

Automation

The Operations Facts

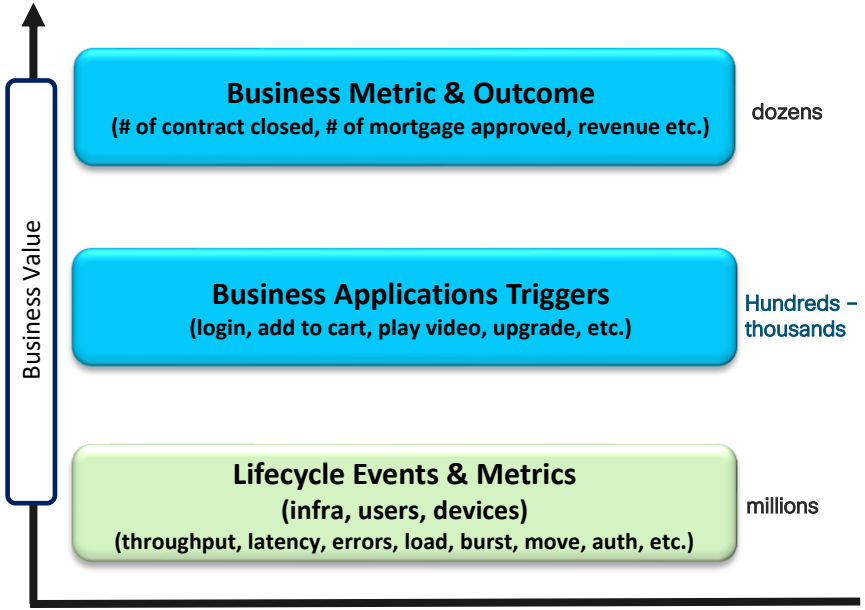
1) By Switching to a DevOps model of Continuous Integration and Continuous Deployment, enterprises often can realize much faster implementation of new features in their multicloud applications.

2) However, there are many applications for which change must be strictly controlled (e.g.: core Enterprise Resource Planning (ERP) system).

3) The need for agile as well as traditional frameworks (e.g.: ITIL) to control change - and the consequent risk - remains very valid.

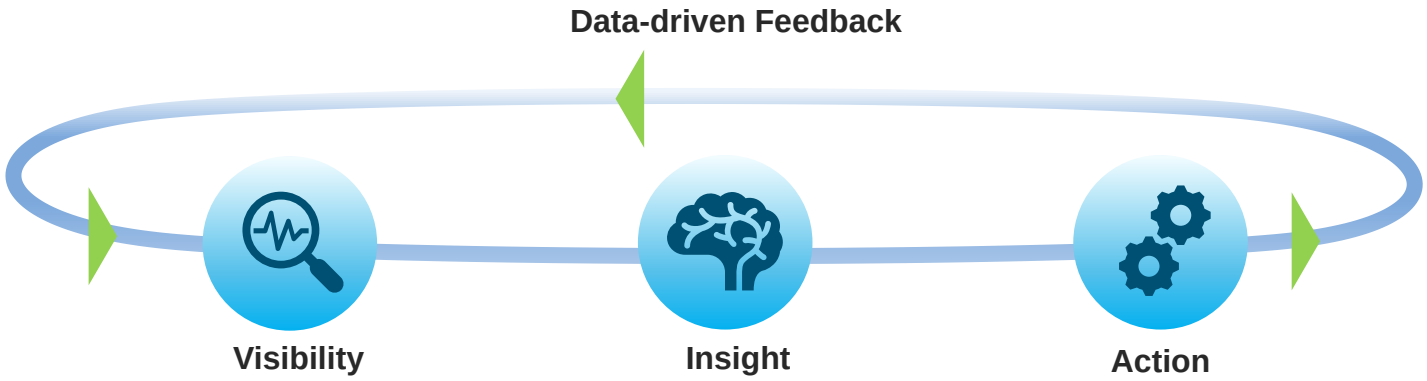
ACI Anywhere Operations address both ☺

Customers' business value chain

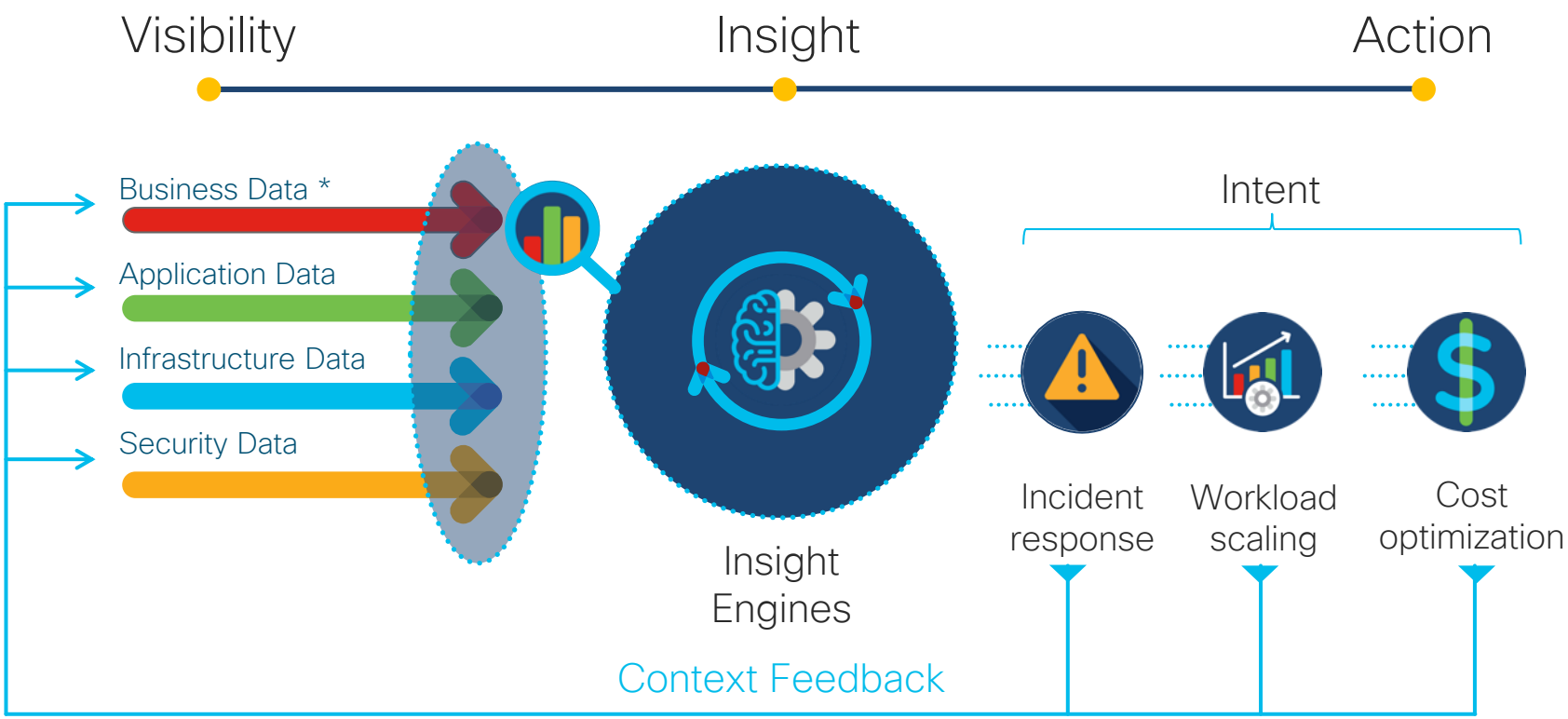


- Different parts of the organization (E.g. LOB vs IT) are focused on different metrics;
- Scale of metrics / events generated is far greater as we get closer to the infrastructure, users, devices;
- Increased evolution towards data augmented operating models (eg.: AIOps), however these must be driven by Real-Time data to be effective

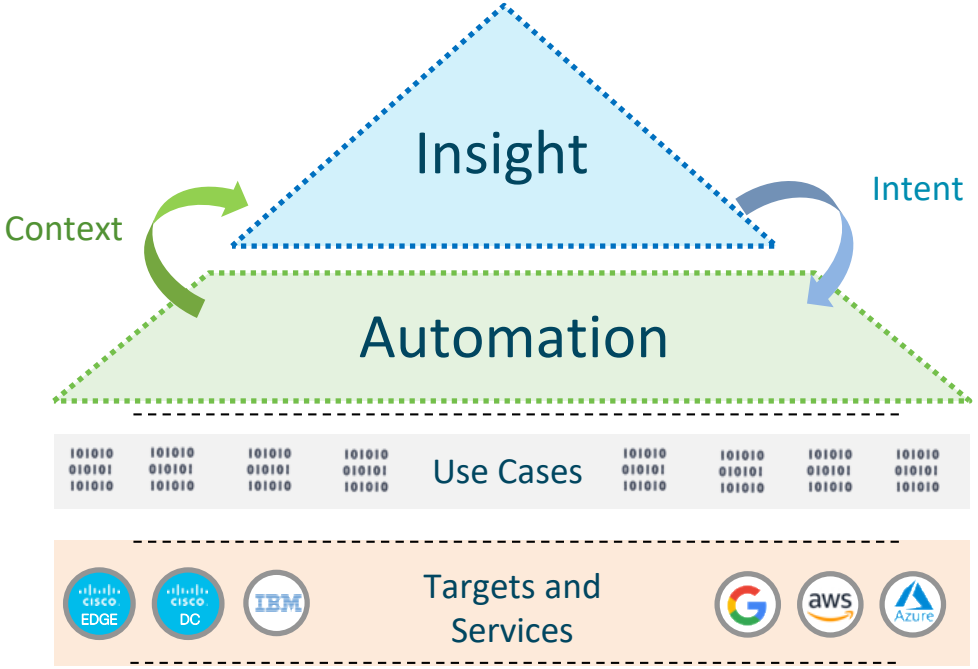
Cisco “Visibility-Insights-Actions” operations approach



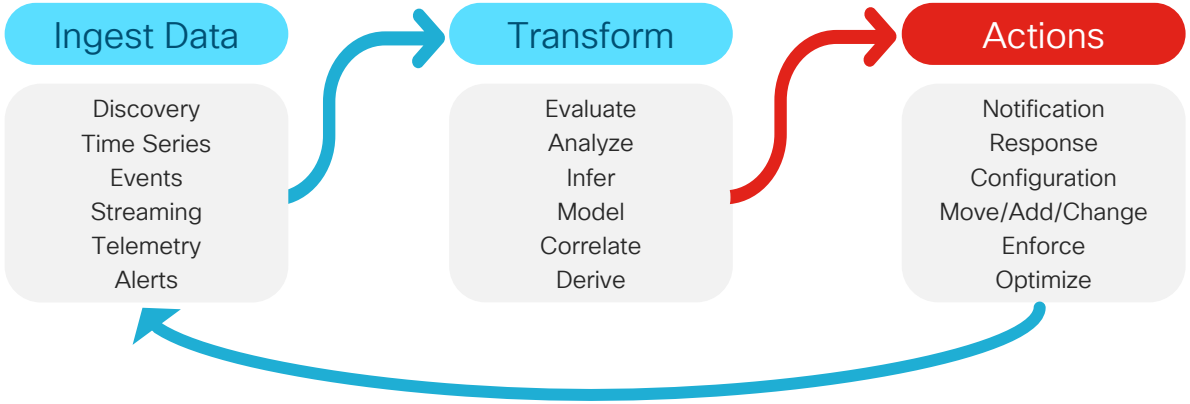
Cisco “Visibility-Insights-Actions” operations approach



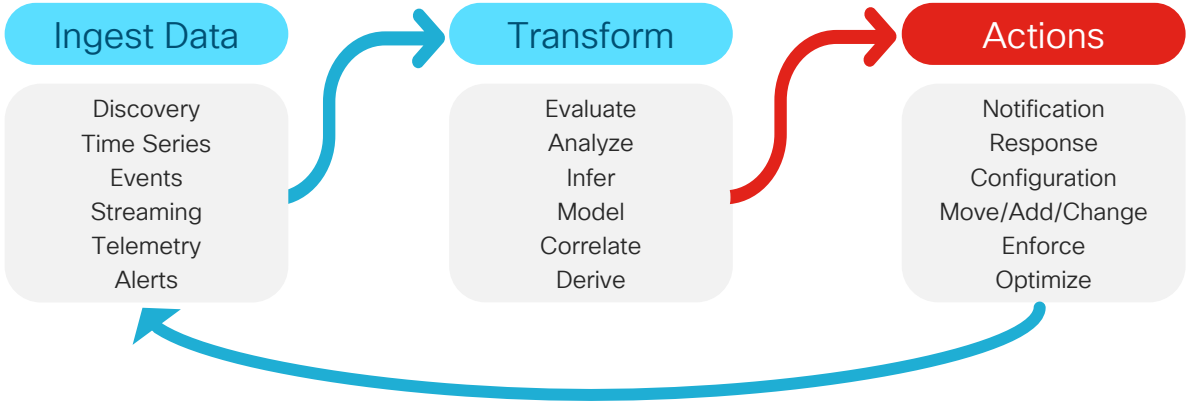
Cisco cross-architecture technology framework



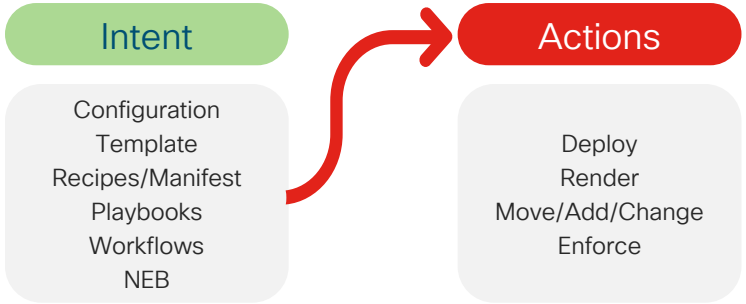
Data-driven INSIGHTS Engines



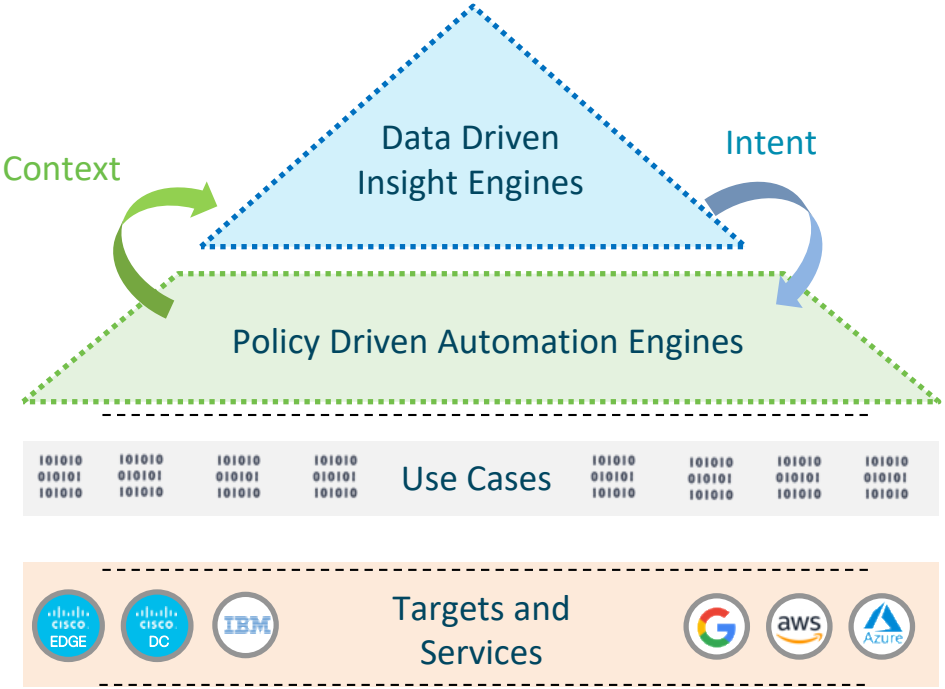
Data-driven INSIGHTS Engines

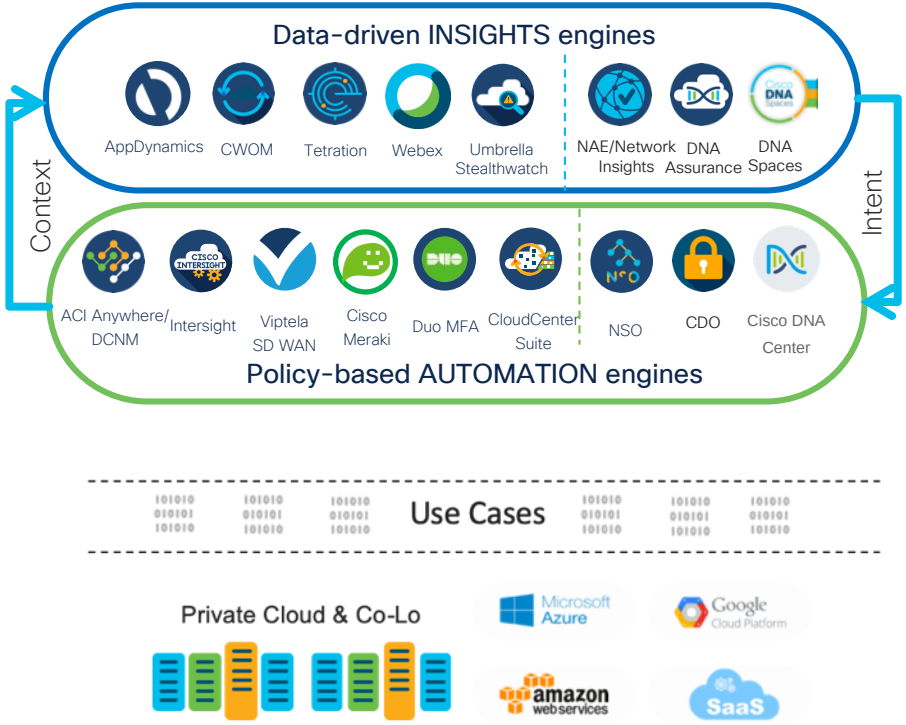


Policy-based AUTOMATION Engines

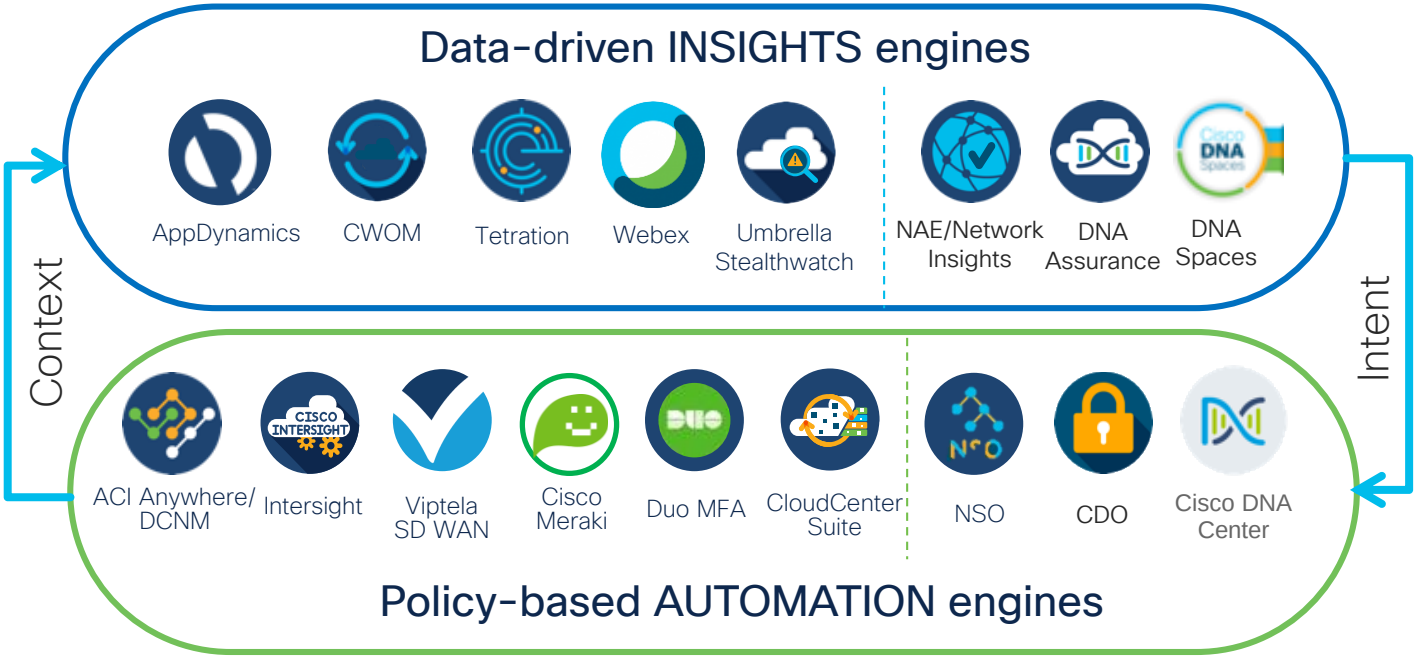


Technology framework detailed

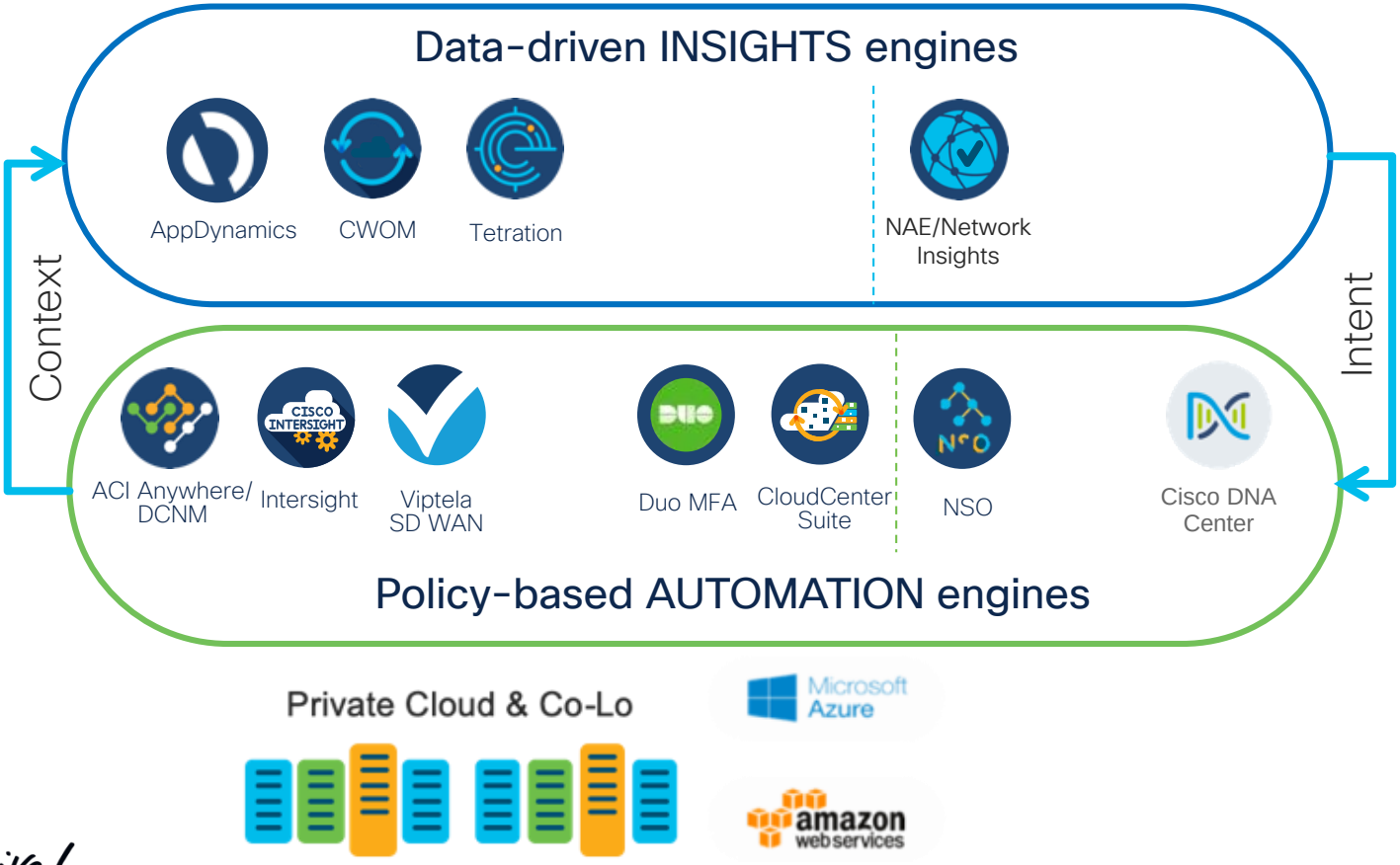




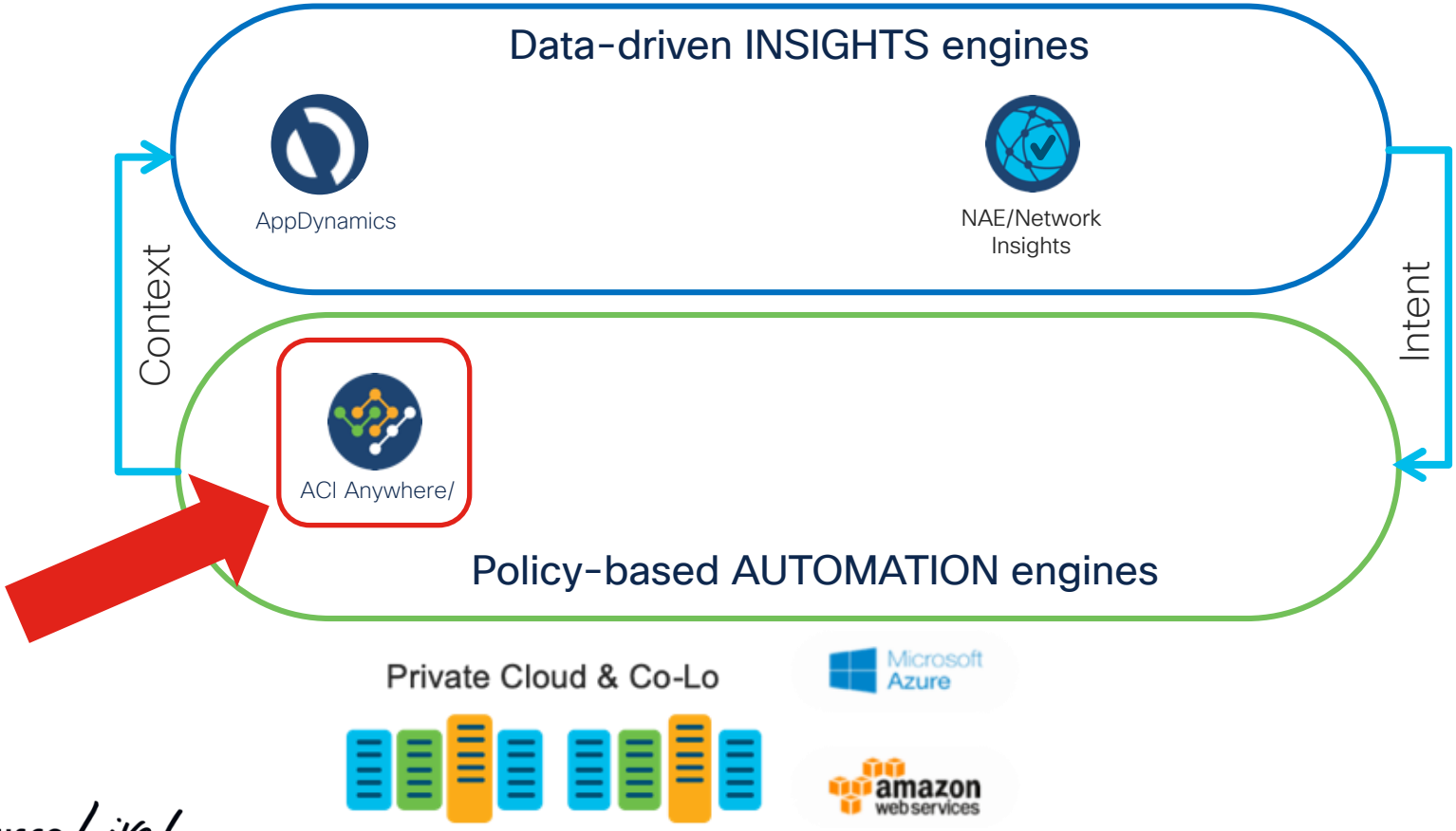
Technology framework detailed



ACI integrations & technology framework



Focus of this ACI Operations section

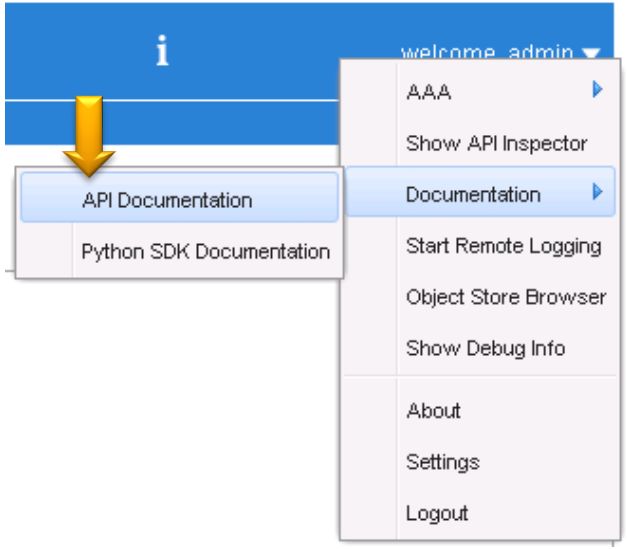


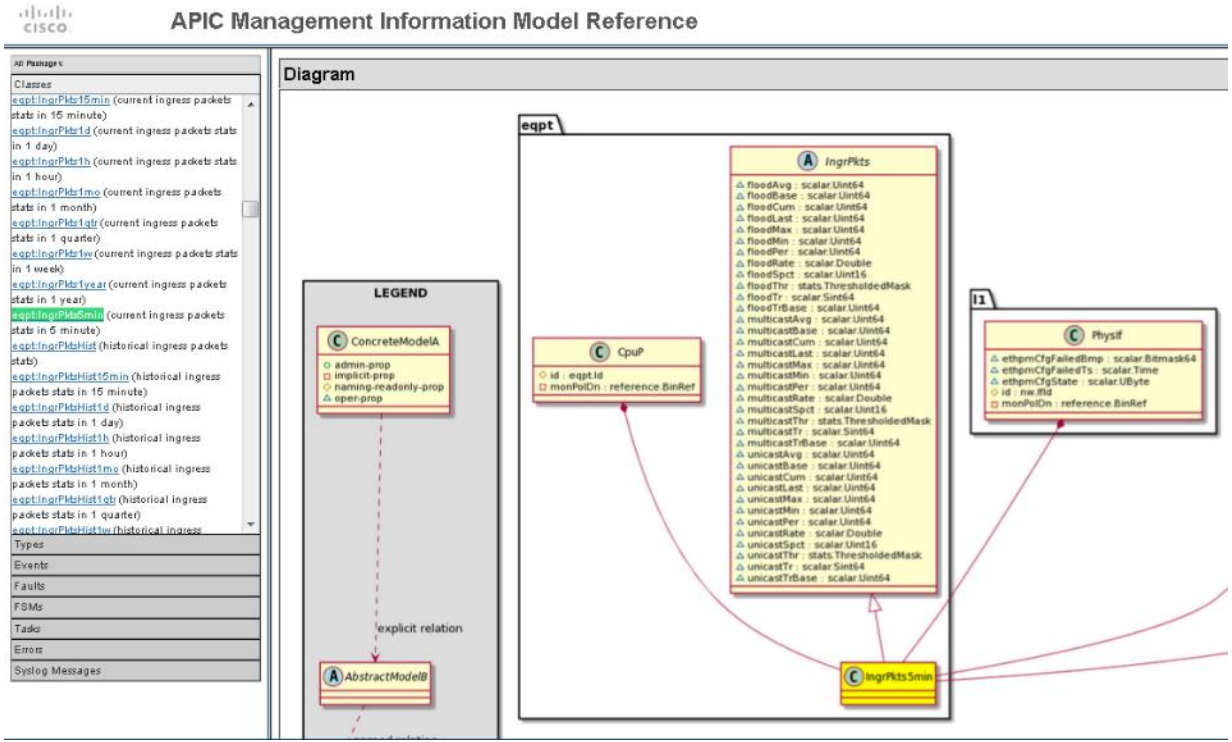
ACI Operations - Agenda

- Before getting started – setting the concepts stage

APIC Management Information Model Reference

From the WebUI





direct URL

https://<APIC IP address>/doc/html/

Cloud APIC Management Information Model Reference

ACI 4.1

Overview Naming Diagram Containers Contained Inheritance Stat Counters Stats Events Faults FSMs Properties Summary Properties Detail

Class cloud:EPg (CONCRETE)

https://<APIC IP address>/doc/html

Class ID:14644
Class Label: cloud EPg
Encrypted: false - Exportable: true - Persistent: true - Configurable: true - Subject to Quota: Disabled - Abstraction Layer: Logical Model - APIC NX Processing: Disabled
Write Access: [admin, tenant-epg, tenant-network-profile]
Read Access: [aaa, access-connectivity-l1, access-connectivity-l2, access-equipment, access-protocol-l1, admin, fabric-equipment, fabric-protocol-l2, fabric-protocol-mgmt, nw-svc-device, nw-svc-policy, tenant-connectivity-mgmt, tenant-connectivity-util, tenant-epg, tenant-network-profile, tenant-protocol-l2, tenant-protocol-l3, tenant-security, vmm-policy]
Creatable/Deletable: yes (see Container Mos for details)
Semantic Scope: EPG
Semantic Scope Evaluation Rule: Parent
Monitoring Policy Source: Parent
Monitoring Flags : [IsObservable: true, HasStats: true, HasFaults: true, HasHealth: true, HasEventRules: false]

Cloud EPg

Naming Rules

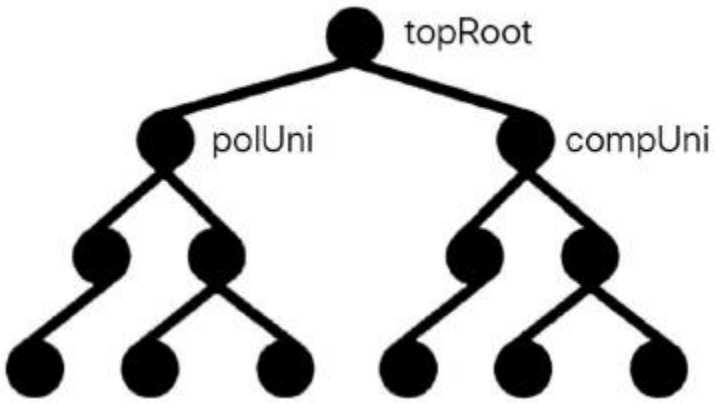
```
RN FORMAT: cloudepg-{name}

    [1] PREFIX=cloudepg- PROPERTY = name


DN FORMAT:

[1] uni/tn-{name}/cloudapp-{name}/cloudepg-{name}
```

Diagram

Super Mo: cloud:AAEPg,
Container Mos: cloud:App (deletable:yes),
Contained Mos: aaa:RbacAnnotation, cloud:AEPgSelector, cloud:EPSelector, cloud:RgInfoHolder, fv:CtrctCtxDefCont, fv:OrchsInfo, fv:RInfoHolder, fv:SharedService, fv:UpdateContract, health:NodeInst, orchs:LDevVipCfg, tag:AInst, tag:Annotation, tag:Tag, telemetry:MatchedSelector, vns:ACCfg, vns:SvcPol, vz:ConsCtrctLbl, vz:ConsLbl, vz:ConsSubjLbl, vz:ProvCtrctLbl, vz:ProvLbl, vz:ProvSubjLbl,
Relations From: cloud:Pool, file:ARemoteHost, vns:Chassis, span:ADest, vns:ALDev, vns:DevMgr, snmp:ClientGrpP, netflow:AExporterPol, extdev:MgrP, dbgac:FromEpgCmn, infra:AFunc, vns:LIfCtx, vmm:CtrlrP, datetime:NtpProv, poe:IfPol, dns:Profile, dhcp:RelayP, aaa:AProvider, span:ASrc, dnsepg:ASvrGrp, auth:Svr, vns:ATerm, dbgac:ToEpgCmn, dbgac:EpgToEpg, vns:VEpg, span:VSrcDef,
Relations To: fv:Ctx, vz:BrCP, vz:CPIf, qos:CustomPol, vz:Taboo, fv:EPg,
Relations: cloud:RsCloudEPgCtx, cloud:RtPoolToCloudEPg, fv:RsCons, fv:RsConsIf, fv:RsCustQosPol, fv:RsIntraEpg, fv:RsProtBy, fv:RsProv, fv:RsSecInherited, fv:RtARemoteHostToEpg, fv:RtChassisEpg, fv:RtDestEpg, fv:RtDevEpg, fv:RtDevMgrEpg, fv:RtEpg, fv:RtExporterToEPg, fv:RtExtdevMgrMgmtEPg, fv:RtFromAbsEpg, fv:RtFuncToEpg, fv:RtLIfCtxToInstP, fv:RtMgmtEPg, fv:RtNtpProvToEpg, fv:RtPoeEpg, fv:RtProfileToEpg, fv:RtProv, fv:RtSecInherited, fv:RtSecProvToEpg, fv:RtSrcToEpg, fv:RtSvcMgmtEpg, fv:RtSvrEpg, fv:RtSvrToMgmtEPg, fv:RtTermToEPg, fv:RtToAbsEpg, fv:RtToAbsEpgForEpgToEpg, fv:RtVConnToEpgEp, fv:RtVConnToEpgSubnet, fv:RtVsrcToEpg,

ACI Object Model Management Information Tree (MIT)

- Objects within APIC are structured in tree-based hierarchy
- Objects referred to as Managed Objects (MO)
- Every object has a parent, with exception of top:Root (top of tree)
- Relationships exist between objects



The REST API Exposes the Object Model

API Schema Follows Object Model Containment

http(s):// http or https protocol	host:port APIC host and port	/api API Operator	/{mo|class} Specify Managed Object or Class Operator	/{*dn*|*classname*} Distinguished name or Object Class	.{xml|json} Encoding for response	?[options] Specify filters, selectors or modifiers to query, joined using ampersand (&)

Read properties for an EPG by Distinguished Name:

```
http://apic/api/mo/uni/tn-Cisco/ap-Software/epg-Download.xml
```

Find all 10G ports on fabric:

```
http://apic/api/class/l1PhysIf.xml?query-target-filter=eq(l1PhysIf.speed,"10G")
```

Example: Invoking the API from Python

```python
import json
import requests

requests.packages.urllib3.disable_warnings()

base_url = 'https://apic-nicolas/api/'

# create credentials structure
name_pwd = {'aaaUser': {'attributes': {'name': 'admin', 'pwd': 'ins091209'}}}
json_credentials = json.dumps(name_pwd)

# log in to API
login_url = base_url + 'aaaLogin.json'
post_response = requests.post(login_url, data=json_credentials, verify=False)

# get token from login response structure
auth = json.loads(post_response.text)
login_attributes = auth['imdata'][0]['aaaLogin']['attributes']
auth_token = login_attributes['token']

# create cookie array from token
cookies = {}
cookies['APIC-Cookie'] = auth_token

# read the LED status from the chassis fan tray
sensor_url = base_url + 'mo/topology/pod-1/node-111/sys/ch/ftslot-1/ft/indled-1.json'
get_response = requests.get(sensor_url, cookies=cookies, verify=False)

# display sensor data structure
print()
print(json.dumps(get_response.json(), indent=4))
```

Object Store Browser (Visore)

- APIC has a built in object browser to navigate the object tree and inspect the state of objects.
- Point the web browser to Visore: http://<apic>/visore.html
- Search for a particular object or dn (fvTenant, topSystem, topology/pod-1/node-101)

topSystem	
address	10.0.82.223
childAction	
currentTime	2014-05-14T17:53:39.944+00:00
dn	topology/pod-1/node-101/sys
fabricId	1
fabricMAC	00:22:BD:F8:19:FF
id	101
inbMgmtAddr	0.0.0.0
lcOwn	local
modTs	2014-05-14T03:54:32.773+00:00
mode	unspecified
monPolDn	uni/fabric/monfab-default
name	services-leaf1
oobMgmtAddr	0.0.0.0
podId	1
role	leaf
serial	SAL1733B94B
state	in-service
status	
systemUpTime	00:14:19:03.000

Cloud APIC Visore – Web Base MO Query and Browser Tool

ACI 4.1

APIC Object Store https://<IP address>/visore.html

Class or DN or URL: fvTenant | Property | Operation: == | Value | Run Query

All MOs of class fvTenant — 6 objects found — Show URL and response of last query

fvTenant

dn	uni/tn-mgmt
annotation	
childAction	
descr	
extMngdBy	
lcOwn	local
modTs	2019-01-19T23:04:02.155+00:00
monPolDn	uni/tn-common/monepg-default
name	mgmt
nameAlias	
ownerKey	
ownerTag	
status	
uid	0

URL and Response of last query

Response Type: JSON XML

URL

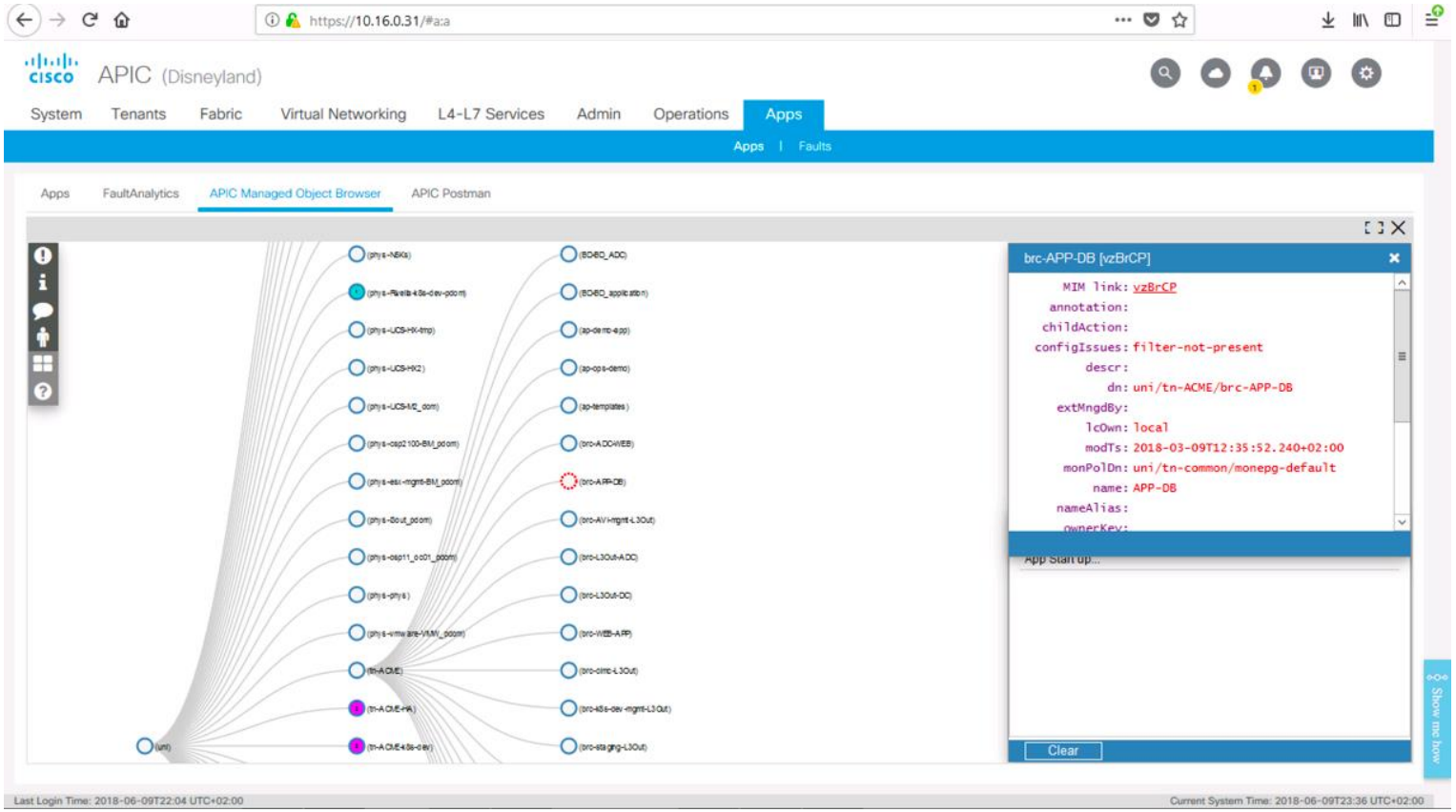
/api/node/class/fvTenant.json?

Response

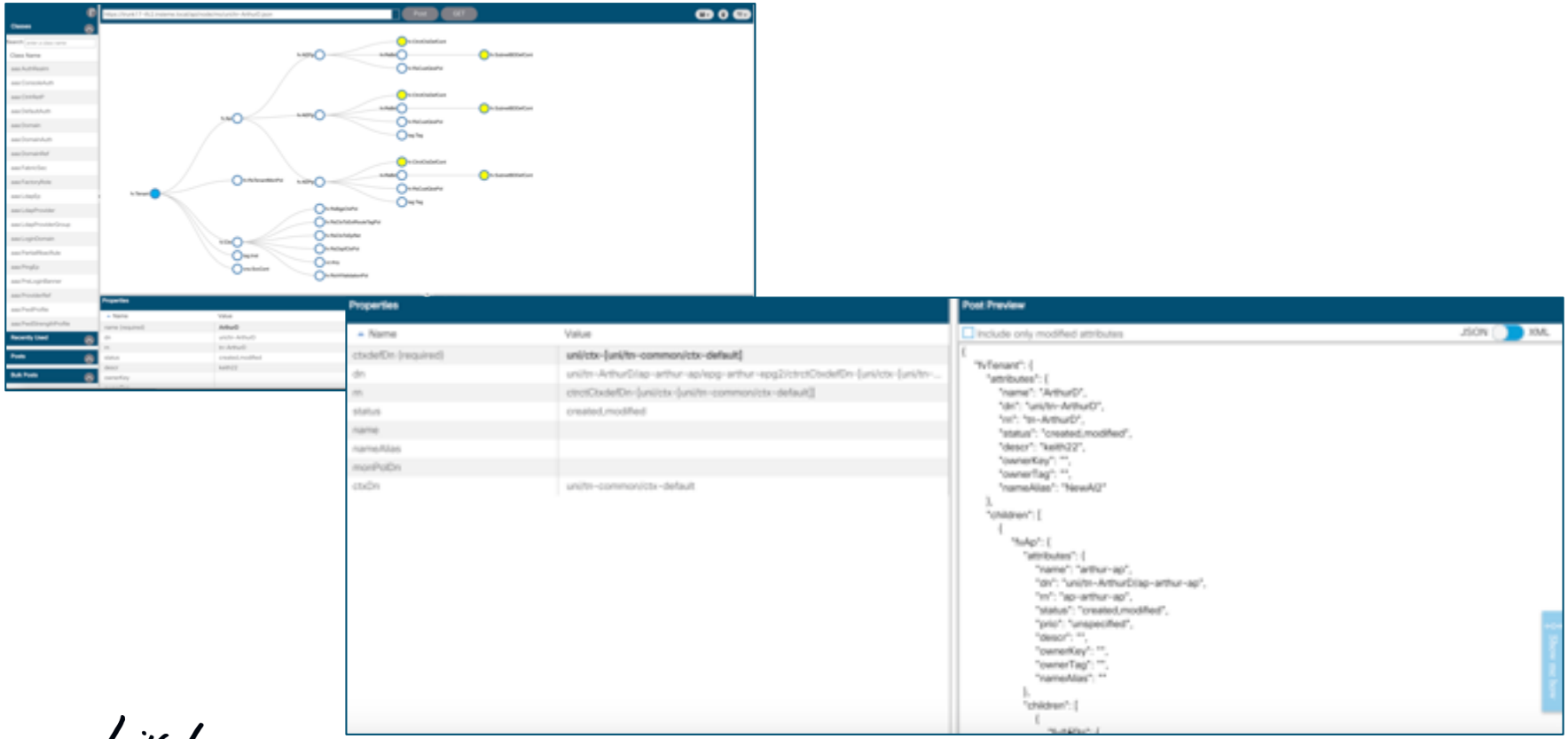

```
{
  "totalCount": "6",
  "imdata": [
    {
      "fvTenant": {
        "attributes": {
          "annotation": "",
          "childAction": "",
          "descr": "",
          "dn": "uni/tn-mgmt",
          "extMngdBy": "",
          "lcOwn": "local",
          "modTs": "2019-01-19T23:04:02.155+00:00",
          "monPolDn": "uni/tn-common/monepg-default",
          "name": "mgmt",
```


Close

APIC Managed Object Tree Browser App



APIC Postman App



Easier way to access the Object Store Browser

Right click (since ACI 3.2)

System | Tenants | Fabric | Virtual Networking | L4-L7 Services | Admin | Operations | Apps

ALL TENANTS | Add Tenant | Tenant Search: name or descr | common | Baekerei | Candid_Tenant_One | Candid_Tenant_Two | dngo-T1

Tenant Baekerei

- Quick Start
- Tenant Baekerei
  - Application Profiles
    - MNLB-test
    - VRF-sharing-test
    - client-to-server
      - Application EPGs
        - inside1
        - inside2
        - inside3
        - private-vlan-EP
      - uSeg EPGs
      - L4-L7 Service Para
  - Networking
  - Contracts
  - Policies

Create EPG Subnet
Add VMM Domain Association
Add Physical Domain Association
Add L2 External Domain Association
Add Fibre Channel Domain Association
Deploy Static EPG on PC, VPC, or Interface
Add Taboo Contract
Add Provided Contract
Add Consumed Contract
Add Consumed Contract Interface
Add Intra-EPG Contract
Create L4-L7 IP Address Pool
Delete
Save as ...
Post ...
Share
Open In Object Store Browser

EPG - inside1

Summary | Policy | Operational | Stats | Health | Faults | History

Static EPG Members: 1 Total — 0

Dynamic EPG Members: 4 Total — 0

Domains (VM and Bare Metals)

Contracts

Show me how

Last Login Time: 2018-06-09T22:42 UTC+02:

Current System Time: 2018-06-09T22:57 UTC+02:00

ACI Operations – Agenda

- Before getting started – setting the concepts stage

Visibility	Insights	Actions
Faults, events, stats, health, logs, trails	Application dependency	Incident Troubleshooting

Sounds familiar ?

If someone else hit this issue and its fixed, why not notify that It could impact me!

Why is support asking for access to my setup three times for this issue?

Is my configuration consistent, at the right scale and following best practices?

How many times do I need to gather these, always rolling over logs to resolve this case?

Are my network security patches up to date?

Is the code I'm running since 2016 still recommended?

Biggest Consumer of IT Time: 43% Troubleshooting

4x

Network operators spend more time collecting data than analyzing while troubleshooting

Replication challenge

Troubleshooting an issue can be impossible if IT can't replicate the issue or see the issue as is it happening real time

Slow resolution

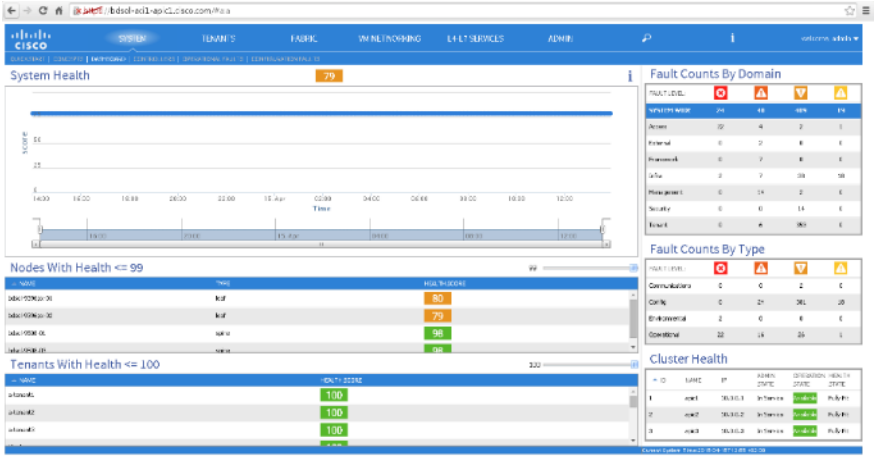
Downtime is expense; unplanned downtime cost Fortune 1000 $1.25-2.5B annually

. [1] McKinsey Study of Network Operations for Cisco – 2016

How do we want to troubleshoot the network?

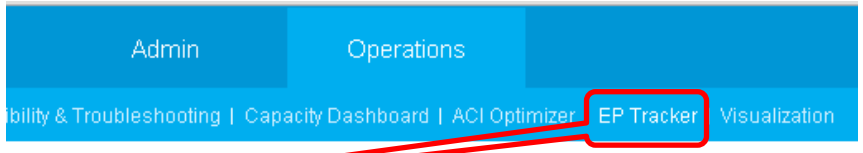
	Switch 1	Switch 2	Switch 3	…
Hardware	✓	✓	✓	
Cabling	✓	✓	✓	
Software	✓	✓	✓	
Configuration	✓	✓	✓	
Operations	✓	✓	✓	OR
Switching	✓	✓	✓	
Routing	✓	✓	✓	
…	✓	✓	✓	

The ACI way:
One view for the whole Fabric!



The way we're used to troubleshoot legacy …

End Point Tracker /Search



We can search End Point by IPv4, IPv6 or MAC address

172.16.1.0/24

SEARCH

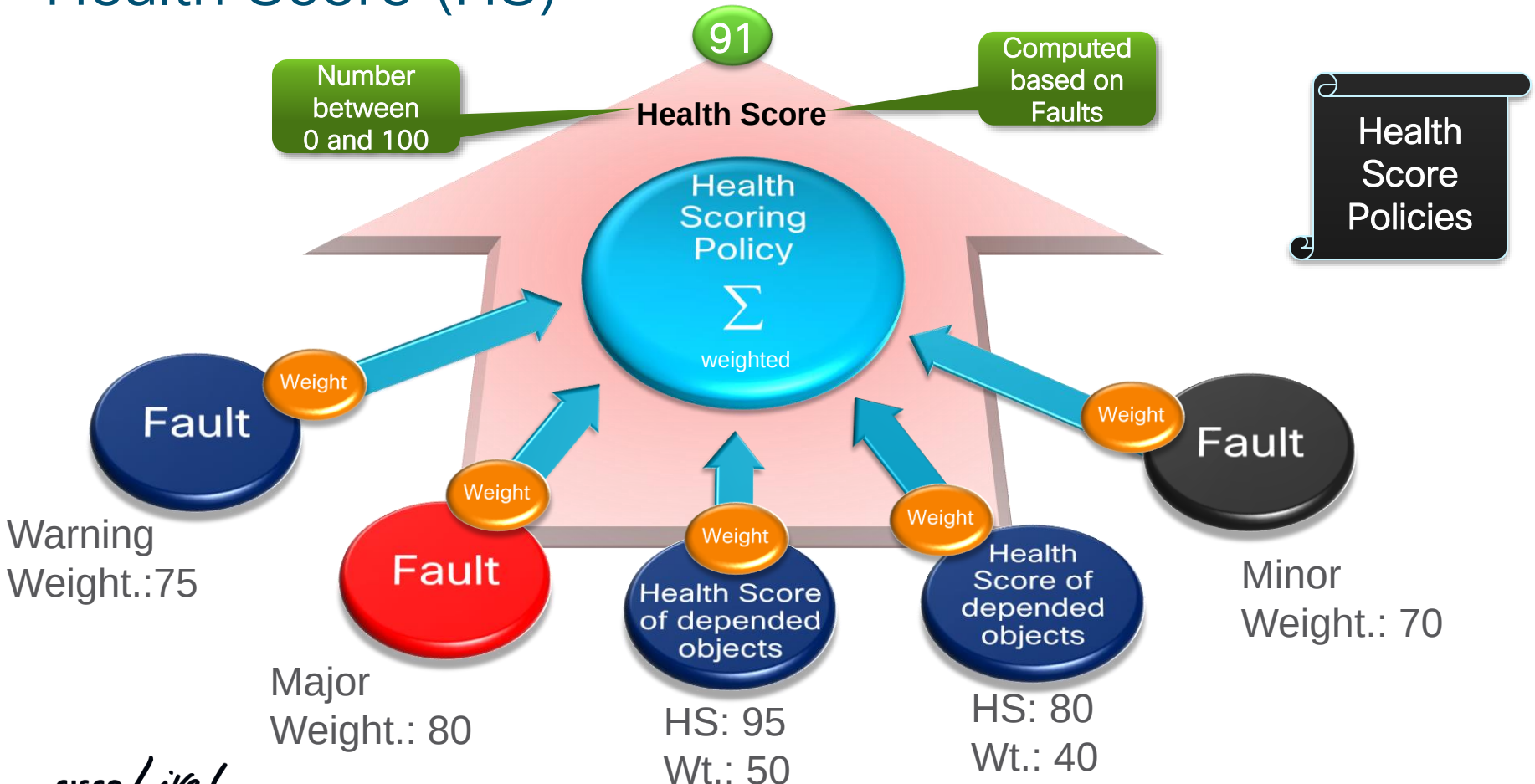
Learned At	Tenant	Application	EPG	IP
Leaf:101, Port:eth1/33	mio	mioAP1	mioEPG2	172.16.1.0/24

State Transitions

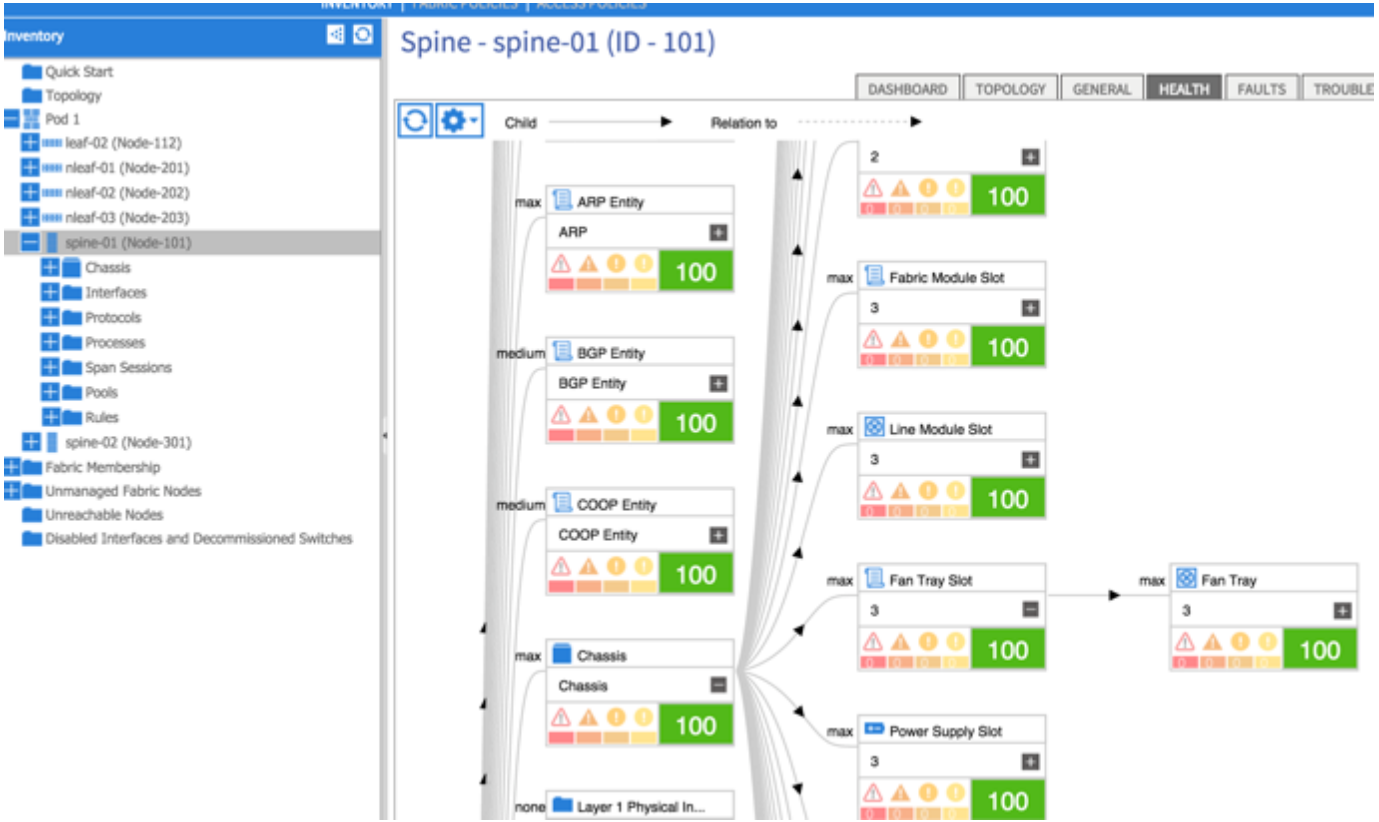
Date	IP	MAC	EPG	Action	Node	Interface	Encap
2015/12/31 10:42...	172.16.1.212	00:50:56:03:02:02	mio/mioAP1/mioEPG2	detached	Node-102	eth1/33	vlan-3398
2015/11/05 15:54...	172.16.1.212	00:50:56:03:02:02	mio/mioAP1/mioEPG2	detached	Node-101	eth1/33	vlan-3398
2015/11/05 15:53...	172.16.1.212	00:50:56:03:02:02	mio/mioAP1/mioEPG2	attached	Node-102	eth1/33	vlan-3398
2015/10/01 16:38...	172.16.1.212	00:50:56:03:02:02	mio/mioAP1/mioEPG2	attached	Node-101	eth1/33	vlan-3398
2015/10/01 16:38...	172.16.1.212	00:50:56:03:02:02	mio/mioAP1/mioEPG2	detached	Node-102	eth1/33	vlan-3398
2015/10/01 16:30...	172.16.1.212	00:50:56:03:02:02	mio/mioAP1/mioEPG2	attached	Node-102	eth1/33	vlan-3398
2015/10/01 16:06...	172.16.1.212	00:50:56:03:02:02	mio/mioAP1/mioEPG2	detached	Node-101	eth1/33	vlan-3398
2015/10/01 15:46...	172.16.1.212	00:50:56:03:02:02	mio/mioAP1/mioEPG2	detached	Node-102	eth1/33	vlan-3398
2015/10/01 15:40...	172.16.1.212	00:50:56:03:02:02	mio/mioAP1/mioEPG2	attached	Node-101	eth1/33	vlan-3398
2015/10/01 15:26...	172.16.1.212	00:50:56:03:02:02	mio/mioAP1/mioEPG2	attached	Node-102	eth1/33	vlan-3398
2015/10/01 15:25...	172.16.1.212	00:50:56:03:02:02	mio/mioAP1/mioEPG2	detached	Node-101	eth1/33	vlan-3398
2015/10/01 15:20...	0.0.0.0	00:50:56:03:02:02	mio/mioAP1/mioEPG2	attached	Node-101	eth1/33	vlan-3398

Page 2 Of 2 Objects Per Page: 15

Health Score (HS)



Drill down the reason for the lowered spine Health Score



Visibility and Troubleshooting

Operations

Visibility & Troubleshooting

0 Session Name: SR6123CLEUR

Description: TSW Demo for #CLEUR

Source

External IP

1 10.10.101.1

SEARCH

Learned At	Tenant	Application	EPG	IP
101-102, vPC: ESX01-vpc	Services			10.10.10...

Destination

External IP

10.201.1.11 2

SEARCH

Learned At	Tenant	Application	EPG	IP
Leaf:103, Port:eth1...	DC	App	EPG2	10.201.1...

GENERATE REPORT | SHOW GENERATED REPORTS | START 3

0 define session name

1 select end point 1

2 select end point 2

3 start

Q: Endpoints unable to communicate to each other? We're unsure where the impacted hosts and what's the datapath between them?

A: We select End Points we'd like to troubleshoot visually. The rest is done by Visibility and Troubleshooting tool

Home

05/19/18 10:05 PM - 05/19/18 11:05 PM

Demo - Healthy

Overview | Unified Monitoring | Getting Started

Recently Visited

- EComm-NPM-Demo - Dashboard
  Applications > EComm-NPM-Demo - Dashboard
- EComm-NPM-Demo - Network Dashboard
  EComm-NPM-Demo - Network Dashboard
- EComm-NPM-Demo
  Applications > EComm-NPM-Demo
- Order-Tier - Network Dashboard
  EComm-NPM-Demo > Tiers & Nodes > Order-Tier - Network Dashboard
- Order-Tier - Dashboard
  Applications > EComm-NPM-Demo > Tiers & Nodes > Order-Tier - Das...
- EComm-NPM-Demo
  Applications > EComm-NPM-Demo
- EComm-NPM-Demo
  Applications > EComm-NPM-Demo
- Baselines

Applications 7

0 critical, 0 warning, 7 normal

- EComm-NPM-Demo
- NPMApplication
- NPMApplication-1
- NPMApplication-2
- NPMApplication-3
- NPMApplication-4
- NPMApplication-5

User Experience

Browser Apps 0

No Browser Apps

Get Started

Mobile Apps 0

No Mobile Apps

Get Started

Databases 0

No Databases

Get Started

Servers 0

No Servers

Get Started

Analytics

0 Transactions

0 Logs

0 Browser Requests

ACI Contract Logging – Denied Packets

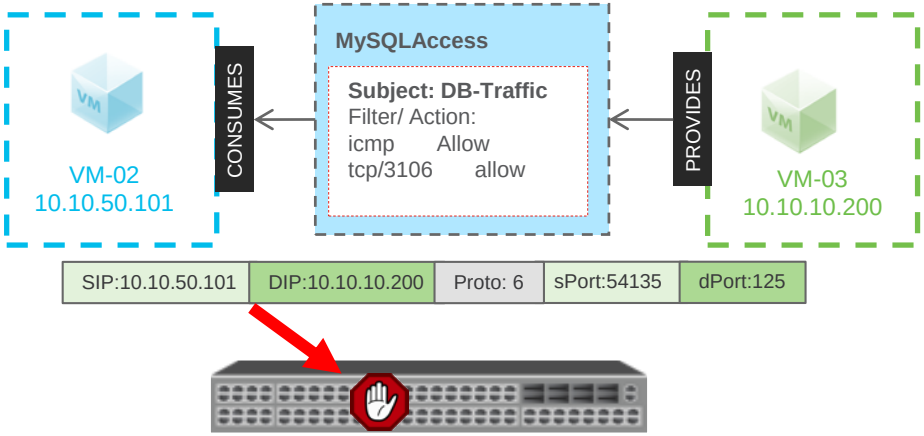
Logging Deny

- ACI can **log implicit deny hits**
  - For Bare Metal, VMware VDS and MSFT Domains logs generated by Leaf
  - For AVS/AVE logs may be generated on Leaf or vLeaf
  - For OpenStack ML2 mode, logs configured external to the fabric at the host
- Syslog is exported according to monitoring policies and configured External Data Collectors
- Logs **include Tenant/VRF, EPG VLAN encap, ingress interfaces** and offending packet details

- **Software Dependency: supported on all software releases**
- **Hardware Dependency: supported on all hardware models**

ACL deny not logged by default:

Fabric -> Fabric Policies -> Monitoring Policies -> Common Policy -> Syslog Message Policies -> Policy for system syslog messages -> Change 'default' to 'info'



Feb 04 10:26:54 troy-leaf1 %LOG_LOCAL7-6-SYSTEM_MSG [E4204936][transition][info][sys] %ACLLOG-5-ACLLOG_PKTLOG_DENY: CName: Test-Tenant:Test-Tenant-VRF(VXLAN: 2162689), VlanType: FD_VLAN, Vlan-Id: 21, SMac: 0x00505690b43a, DMac:0x0022bdf819ff, SIP: 10.10.50.101, DIP: 10.10.10.200, SPort: 54135, DPort: 125, Src Intf: port-channel2, Proto: 6, PktLen: 74

ACI Contract Logging – Permitted Packets

Logging Permit

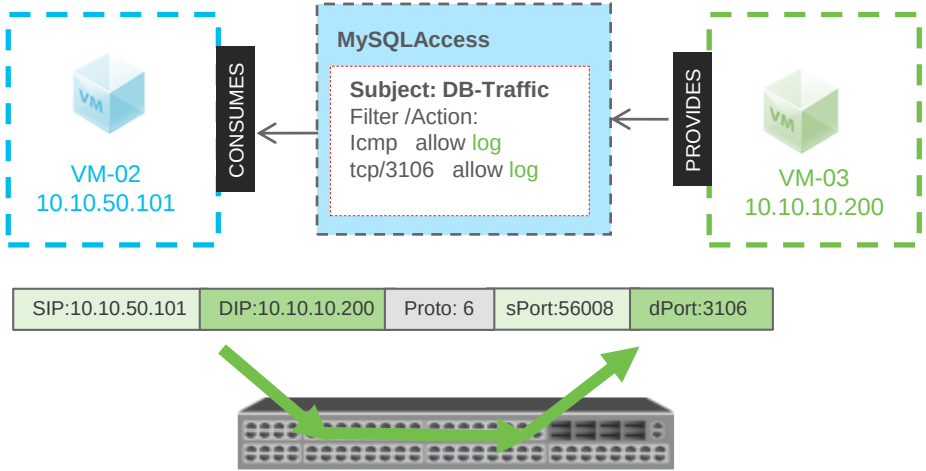
- Permit logging is configured per Filter
  - For Bare Metal, VDS and MSFT Domains logs generated by Leaf
  - For AVS/AVE logs may be generated on Leaf or vLeaf
  - For OpenStack ML2 mode, logs configured external to the fabric at the host
- Syslog is exported according to monitoring policies and configured External Data Collectors
- Logs include Tenant/VRF, EPG VLAN encap, ingress interfaces and offending packet details

- **Software Dependency: 2.2(1n) or higher**
- **Hardware Dependency: requires EX models or newer**

Permit log configured at the subject on a per filter basis.

Filters:

Name	Tenant	Directives
icmp-traffic	Test-Tenant	log
mysql-traffic	Test-Tenant	log



Feb 04 10:14:44 troy-leaf1 %LOG_LOCAL7-6-SYSTEM_MSG [E4204936][transition][info][sys] %ACLLOG-5-ACLLOG_PKTLOG_PERMIT: CName: Test-Tenant:Test-Tenant-VRF(VXLAN: 2162689), VlanType: FD_VLAN, Vlan-Id: 21, SMac: 0x00505690b43a, DMac:0x0022bdf819ff, SIP: 10.10.50.101, DIP: 10.10.10.200, SPort: 56008, DPort: 3106, Src Intf: port-channel2, Proto: 6, PktLen: 98

Audit Logs

- Any configuration changes on the system is recorded as audit logs.
- User action which could be create, modify and delete
- Objects affected by the policy that got implemented

TIME STAMP	USER	ACTION	AFFECTED OBJECT	DESCRIPTION
2015-03-04T01:03:50.288+00:00	admin	deletion	uni/tn-Customer/acEpToEp-yong_67_dst_src	EpToEp yong_67_dst_src deleted
2015-03-04T01:03:50.288+00:00	admin	deletion	uni/tn-Customer/acEpToEp-yong_67_dst_src/rstoEpForEpToEp-[uni/tn-Customer/ap-DB/epg-SQLServerPool-1/cep-00:00:1F:FB:01:11]	RsToEpForEpToEp uni/tn-Customer/ap-DB/epg-SQLServerPool-1/cep-00:00:1F:FB:01:11 deleted
2015-03-04T01:03:50.288+00:00	admin	deletion	uni/tn-Customer/acEpToEp-yong_67_dst_src/rsfromEp-[uni/tn-Customer/ap-Apps/epg-AppServerPool-1/cep-00:00:20:03:9D:4F]	RsFromEp uni/tn-Customer/ap-Apps/epg-AppServerPool-1/cep-00:00:20:03:9D:4F deleted
2015-03-04T01:03:50.228+00:00	admin	deletion	uni/tn-Customer/acEpToEp-yong_67_src_dst	EpToEp yong_67_src_dst deleted
2015-03-04T01:03:50.228+00:00	admin	deletion	uni/tn-Customer/acEpToEp-yong_67_src_dst/rstoEpForEpToEp-[uni/tn-Customer/ap-Apps/epg-AppServerPool-1/cep-00:00:20:03:9D:4F]	RsToEpForEpToEp uni/tn-Customer/ap-Apps/epg-AppServerPool-1/cep-00:00:20:03:9D:4F deleted
2015-03-04T01:03:50.228+00:00	admin	deletion	uni/tn-Customer/acEpToEp-yong_67_src_dst/rsfromEp-[uni/tn-Customer/ap-DB/epg-SQLServerPool-1/cep-00:00:1F:FB:01:11]	RsFromEp uni/tn-Customer/ap-DB/epg-SQLServerPool-1/cep-00:00:1F:FB:01:11 deleted
2015-03-04T01:01:27.905+00:00	admin	creation	uni/tn-Customer/acEpToEp-yong_67_dst_src	EpToEp yong_67_dst_src created
2015-03-04T01:01:27.905+00:00	admin	creation	uni/tn-Customer/acEpToEp-yong_67_dst_src/rstoEpForEpToEp-[uni/tn-Customer/ap-DB/epg-SQLServerPool-1/cep-00:00:1F:FB:01:11]	RsToEpForEpToEp uni/tn-Customer/ap-DB/epg-SQLServerPool-1/cep-00:00:1F:FB:01:11 created
2015-03-04T01:01:27.905+00:00	admin	creation	uni/tn-Customer/acEpToEp-yong_67_dst_src/rsfromEp-[uni/tn-Customer/ap-Apps/epg-AppServerPool-1/cep-00:00:20:03:9D:4F]	RsFromEp uni/tn-Customer/ap-Apps/epg-AppServerPool-1/cep-00:00:20:03:9D:4F created
2015-03-04T01:01:27.834+00:00	admin	creation	uni/tn-Customer/acEpToEp-yong_67_src_dst/rstoEpForEpToEp-[uni/tn-Customer/ap-Apps/epg-AppServerPool-1/cep-00:00:20:03:9D:4F]	RsToEpForEpToEp uni/tn-Customer/ap-Apps/epg-AppServerPool-1/cep-00:00:20:03:9D:4F created
2015-03-04T01:01:27.834+00:00	admin	creation	uni/tn-Customer/acEpToEp-yong_67_src_dst/rsfromEp-[uni/tn-Customer/ap-DB/epg-SQLServerPool-1/cep-00:00:1F:FB:01:11]	RsFromEp uni/tn-Customer/ap-DB/epg-SQLServerPool-1/cep-00:00:1F:FB:01:11 created
2015-03-04T01:01:27.833+00:00	admin	creation	uni/tn-Customer/acEpToEp-yong_67_src_dst	EpToEp yong_67_src_dst created
2015-03-04T01:01:20.705+00:00	admin	deletion	uni/tn-Customer/trEp-yong_67_dst_src	TrEp yong_67_dst_src deleted

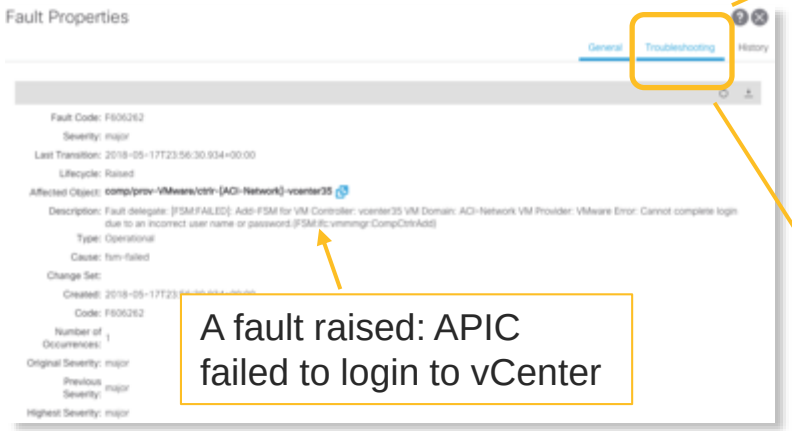
Faults and Audit Logs/Events Association

Since APIC 3.2, the UI introduces an enhancement for easy association of faults and audit logs/ events. For a given fault, the APIC UI can display the audit logs and events right before the fault is raised.

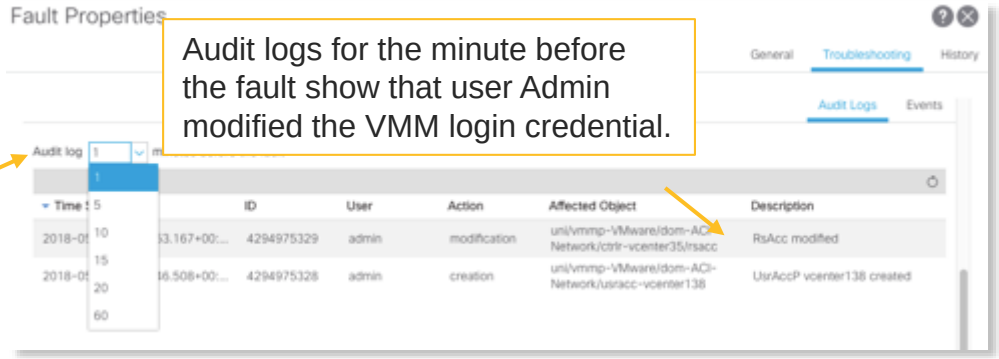
To use the association for fault investigation, open up the specific fault, click on “Troubleshooting”, then view the audit logs and event logs in the pop-up window.

User can choose the time range of the correlation, such as 1, 5, 10, 15, 20 or 60 minutes prior to the fault.

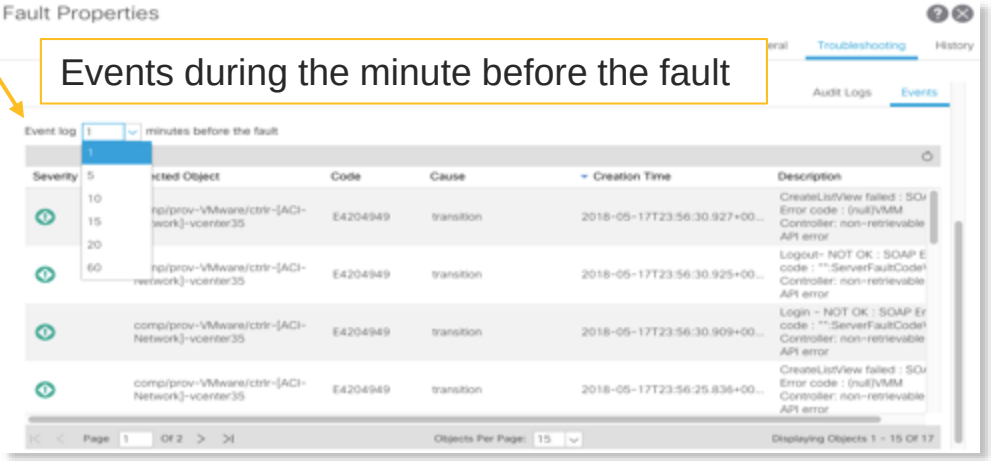
Events/Audit Log Fault Association



A fault raised: APIC failed to login to vCenter



Audit logs for the minute before the fault show that user Admin modified the VMM login credential.



Events during the minute before the fault

Cloud ACI – Event Analytics

ACI 4.1

Cloud APIC

- Dashboard
- Topology
- Application Management
- Cloud Resources
- Operations
  - Active Sessions
  - Event Analytics
  - Schedulers
- Infrastructure
- Admin

Cloud APIC

Event Analytics

Faults | Events | Audit Logs

Filter by attributes or search by keyword (Coming soon!)

Actions

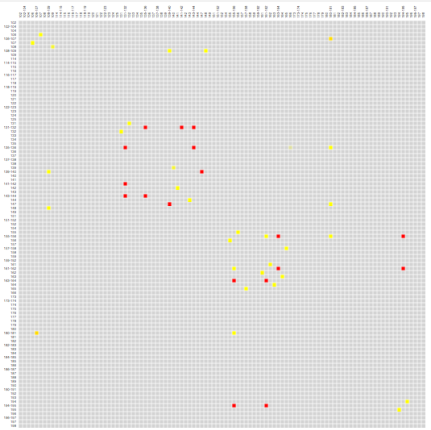
Severity	Code	Affected object	Description	Creation Time
Critical	F0323	topology/pod-1/node-1/lon/svc-ifc_edmgr	Lost connectivity to leader for some data subset(s) of Access Service ifc_edmgr on node 1	Jan 22 2019 11:14:58am
Critical	F0325	topology/pod-1/node-1/lon	Connectivity has been lost to the leader for some data subset(s) of a service on node 1, the service may have unexpectedly restarted or failed.	Jan 22 2019 11:14:58am
Critical	F0323	topology/pod-1/node-1/lon/svc-ifc_licensemgr	Lost connectivity to leader for some data subset(s) of Access Service ifc_licensemgr on node 1	Jan 22 2019 11:14:58am
Critical	F0323	topology/pod-1/node-1/lon/svc-ifc_domainmgr	Lost connectivity to leader for some data subset(s) of Access Service ifc_domainmgr on node 1	Jan 22 2019 11:14:58am
Critical	F0323	topology/pod-1/node-1/lon/svc-ifc_vmmmgr	Lost connectivity to leader for some data subset(s) of Access Service ifc_vmmmgr on node 1	Jan 22 2019 11:14:58am
Critical	F0323	topology/pod-1/node-1/lon/svc-ifc_topomgr	Lost connectivity to leader for some data subset(s) of Access Service ifc_topomgr on node 1	Jan 22 2019 11:14:58am
Minor	F3361	acct-[infra]/region-[us-west-1]/context-[overlay-1]/csr-[ct_routerp_us-west-1_1:0]/physical-1/oper	Operational State of Physical Interface is down due to \u0002	Jan 22 2019 11:20:15am
Major	F607450	topology/pod-1/node-1/sys	[FSM:FAILED]: Task for updating topSystem changes for chassis on Node 1 to Topology Manager(TASK:ifc:ae:TopSystemControllerChassis)	Jan 22 2019 11:28:49am
Major	F606274	topology/pod-1/node-1/sys	[FSM:FAILED]: Task for updating topSystem changes for Node 1 to Topology Manager(TASK:ifc:ae:TopSystemSendTopSystem)	Jan 22 2019 11:28:49am

ACI Traffic Map

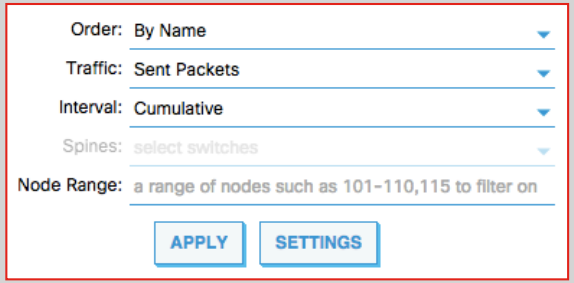
Help visualize and quickly spot high traffic density and underutilized nodes in the Cisco ACI™ fabric.

A grid is presented with a list of node IDs or vPC pairs on each axis. Traffic flow between a given pair of nodes or between a vPC pair is presented using color-coded cells on the heat map.

Traffic density is presented in a range of colors, from lightest (yellow), to shades of orange, to red (highest). Traffic statistics are collected using atomic counters.

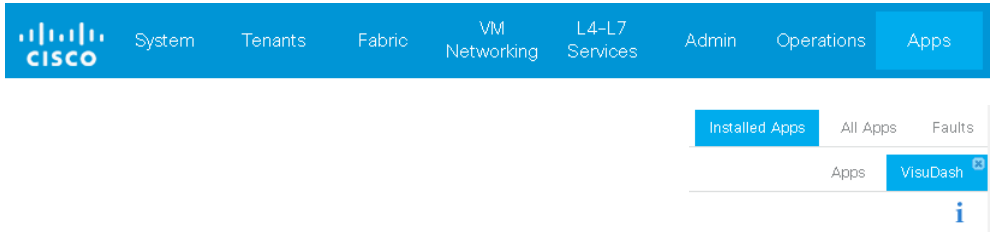


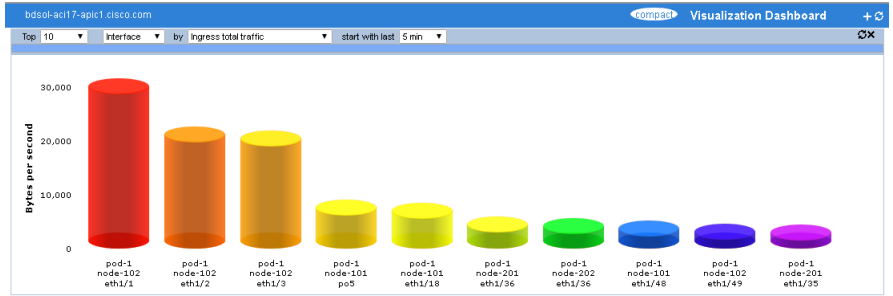
- You can order by name or by traffic.
- Traffic can be seen by:
  - Sent packets
  - Received packets
  - Dropped packets
  - Excess packets



Another way to check traffic on Fabric level

- Visualize utilization on Fabric level using APIC Apps
- We can monitor different parameters at Fabric Level
- VisuDash App:
  - Top 10 Tenants ranked by number of End-points
  - Top 20 interface by utilization



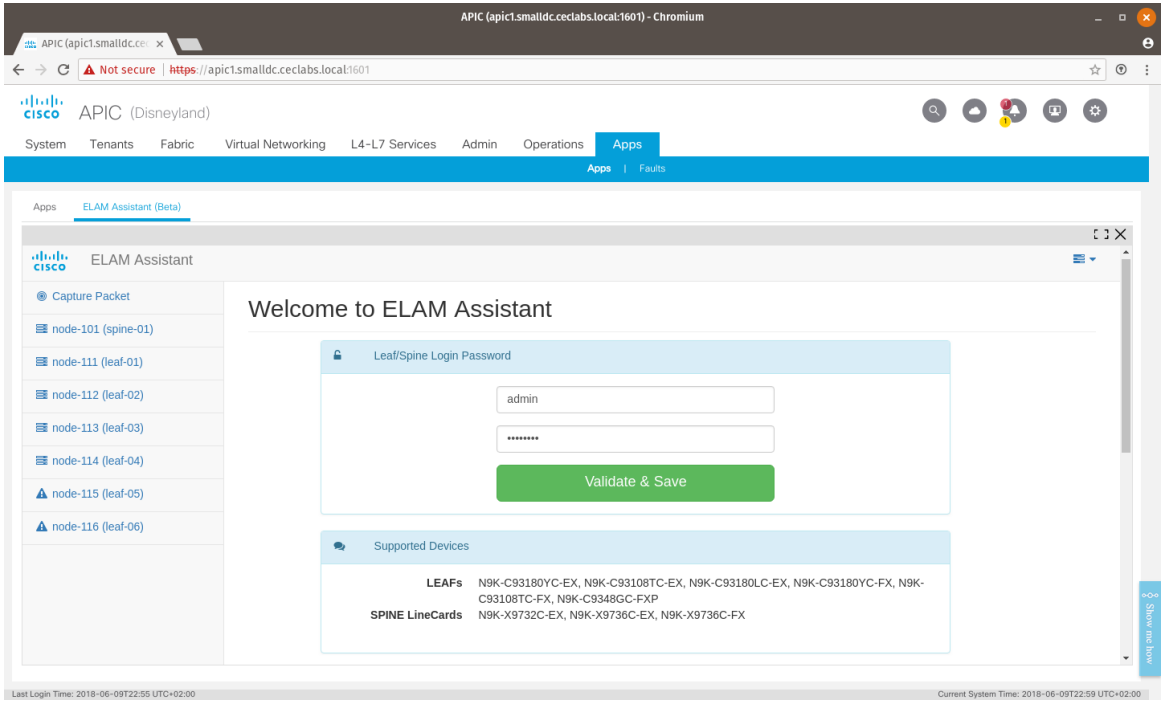


Going deeper – fTriage

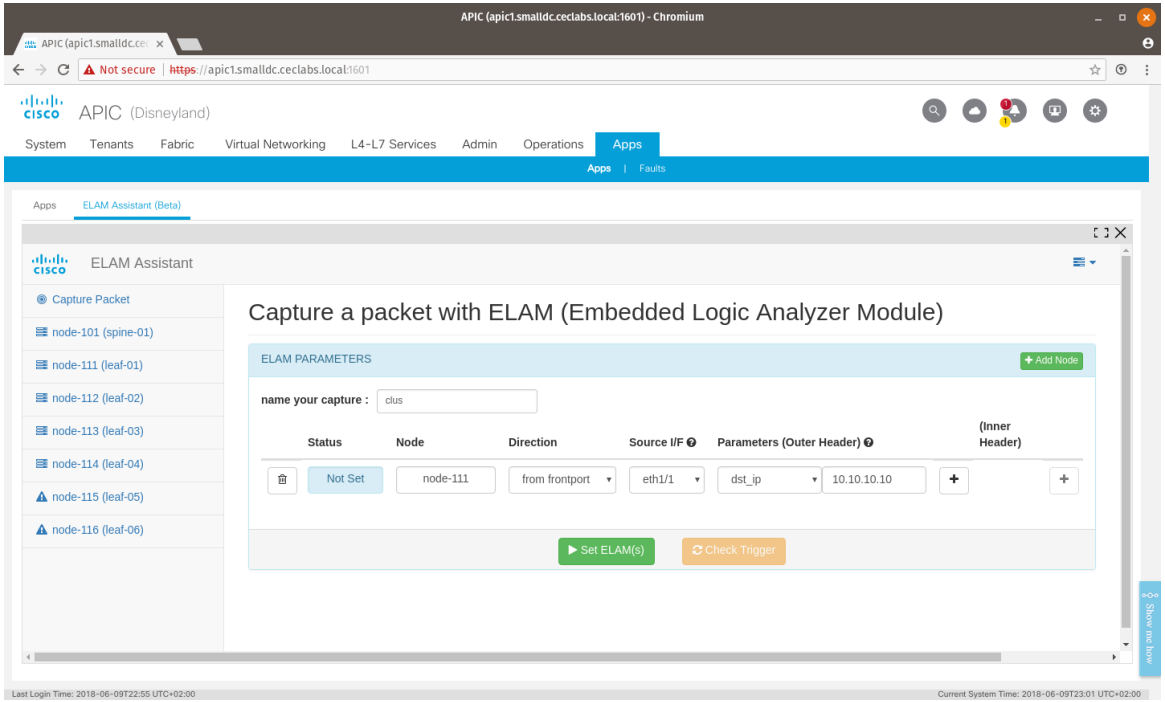
- Fabric triaging tool (used by TAC)
- Python utility, runs on APIC in admin mode.
- Logs into switch nodes to capture requested data with commands/query/ELAM.
- Driven by specific user inputs which are validated first.
- Runs ELAM to determine packet data path
- Runs show commands/moquery on APIC/switch to determine control plane
- Traces a packet hop by hop until the point where it exits the fabric or gets dropped.
- Provides detailed info on interfaces + nodes where packet capture is attempted
- Drop reason is provided along with node, interface info.
- Requires consistent traffic stream flowing, as it relies on ELAM, it must have data packet flowing.
- Not a tool to detect transient and partial drops

ELAM = Embedded Logic Analyzer Module

ELAM Assistant (graphical interface of fTriage) – 1/2



ELAM Assistant (graphical interface of fTriage) – 2/2



If you REALLY want to get deeper and love CLI

fTriage Route

```
sd-tb-99-ifcl# ftriage route -h
usage: ftriage route [-h] -ii intf [-ie encap] [-iie encap] -ei intf
                     [-ee encap] [-eie encap] -dip addr [-sip addr]

optional arguments:
  -h, --help  show this help message and exit
  -ii intf    ingress if (siteid::(node:if|VPC:<vpc-name>|PC:<pc-
              name>)[+vtep])
  -ie encap   ingress encap (VLAN/VNID)
  -iie encap  ingress encap (VLAN/VNID)
  -ei intf    egress if (siteid::(node:if|VPC:<vpc-name>|PC:<pc-name>)[+vtep])
  -ee encap   egress encap (VLAN/VNID)
  -eie encap  egress encap (VLAN/VNID)
  -dip addr   destination IP
  -sip addr   source IP
```

If you REALLY want to get deeper and love CLI

fTriage Bridge

```
sd-tb-99-ifc1# ftriage bridge -h
usage: ftriage bridge [-h] -ii intf [-ie encap] [-iie encap] -ei intf
                      [-ee encap] [-eie encap] -dmac addr [-smac addr]

optional arguments:
  -h, --help  show this help message and exit
  -ii intf    ingress if (siteid::(node:if|VPC:<vpc-name>|PC:<pc-
              name>)[+vtep])
  -ie encap   ingress encap (VLAN/VNID)
  -iie encap  ingress encap (VLAN/VNID)
  -ei intf    egress if (siteid::(node:if|VPC:<vpc-name>|PC:<pc-name>)[+vtep])
  -ee encap   egress encap (VLAN/VNID)
  -eie encap  egress encap (VLAN/VNID)
  -dmac addr  destination MAC
  -smac addr  source MAC
```

If you REALLY want to get deeper and love CLI

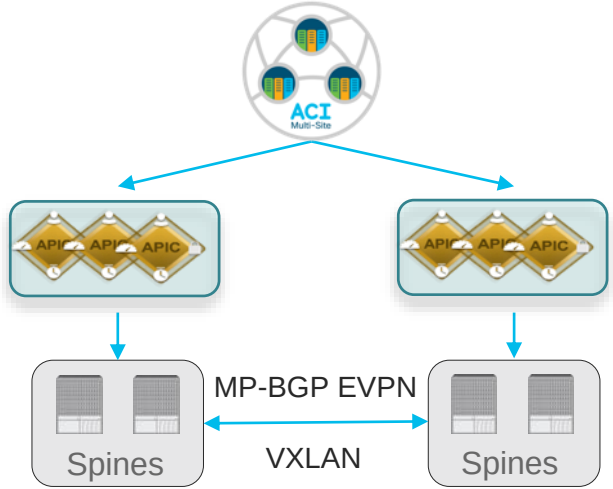
fTriage Logs

```
sd-tb-99-ifc1# ftriage route -ii sd-tb-99-leaf1:Eth1/47 -ie 101 -ei sd-tb-99-leaf2:Eth1/47 -ee 203 -sip 101.1.1.10 -dip 101.1.2.10
ftriage: info : 2018-03-13 09:06:33.286: Building egress BD(s), Ctx
ftriage: info : 2018-03-13 09:06:34.675: Egress BD(s) {sd-tb-99-leaf2: 'bd-[vxlan-16154554]'}
ftriage: info : 2018-03-13 09:06:34.675: Egress Ctx ctx-[vxlan-2916352]
ftriage: info : 2018-03-13 09:06:34.675: Building ingress BD(s), Ctx
ftriage: info : 2018-03-13 09:06:36.297: Ingress BD(s) {sd-tb-99-leaf1: 'bd-[vxlan-16613251]'}
ftriage: info : 2018-03-13 09:06:36.297: Ingress Ctx ctx-[vxlan-2916352]
ftriage: info : 2018-03-13 09:06:36.297: Capturing L3 packet on [sd-tb-99-leaf1]
ftriage: info : 2018-03-13 09:07:09.560: L3 packet seen on sd-tb-99-leaf1:Eth1/47
ftriage: info : 2018-03-13 09:07:16.149: SIP 101.1.1.10 DIP 101.1.2.10
ftriage: info : 2018-03-13 09:07:16.149: sd-tb-99-leaf1: Ingress function
ftriage: info : 2018-03-13 09:07:27.851: sd-tb-99-leaf1: Dst EP is remote
ftriage: info : 2018-03-13 09:07:33.464: sd-tb-99-leaf1: DMAC(00:22:BD:F8:19:FF) same as RMAC(00:22:BD:F8:19:FF)
ftriage: info : 2018-03-13 09:07:33.464: sd-tb-99-leaf1: L3 packet getting routed in SUG
ftriage: info : 2018-03-13 09:07:37.586: sd-tb-99-leaf1: Dst IP is present in SUG L3 tbl
ftriage: info : 2018-03-13 09:07:42.223: sd-tb-99-leaf1: RwDMAC DIPo(10.0.232.66) is one of dst TEPs ['10.0.232.66']
ftriage: info : 2018-03-13 09:07:46.732: Computing next set of nodes
ftriage: info : 2018-03-13 09:07:56.872: Capturing L3 packet on [sd-tb-99-spine1]
ftriage: info : 2018-03-13 09:08:31.547: L3 packet seen on sd-tb-99-spine1:Eth1/1
ftriage: info : 2018-03-13 09:08:40.749: sd-tb-99-spine1: Transit function
ftriage: info : 2018-03-13 09:08:40.749: sd-tb-99-spine1: Capturing L3 packet on egress LC
ftriage: info : 2018-03-13 09:08:44.656: sd-tb-99-spine1: Infra route 10.0.232.66 present in RIB
ftriage: info : 2018-03-13 09:09:21.082: sd-tb-99-spine1: L3 packet seen on egress LC1
ftriage: info : 2018-03-13 09:09:24.989: Computing next set of nodes
ftriage: info : 2018-03-13 09:09:36.779: Capturing L3 packet on [sd-tb-99-leaf2]
ftriage: info : 2018-03-13 09:10:09.364: L3 packet seen on sd-tb-99-leaf2:Eth1/53
ftriage: info : 2018-03-13 09:10:15.207: sd-tb-99-leaf2: Egress function
ftriage: info : 2018-03-13 09:10:23.674: sd-tb-99-leaf2: Dst EP is local
ftriage: info : 2018-03-13 09:10:23.674: sd-tb-99-leaf2: EP if(Eth1/47) same as egr if(Eth1/47)
ftriage: info : 2018-03-13 09:10:27.359: sd-tb-99-leaf2: Dst IP is present in HOM L3 tbl
ftriage: info : 2018-03-13 09:10:29.223: sd-tb-99-leaf2: RW seg_id in HOM same as EP fd seg_id
ftriage: hunch: None
```

ACI Multi-Site

Day-2 Operations: Full-Stack Consistency Checker

Since ACI 3.2

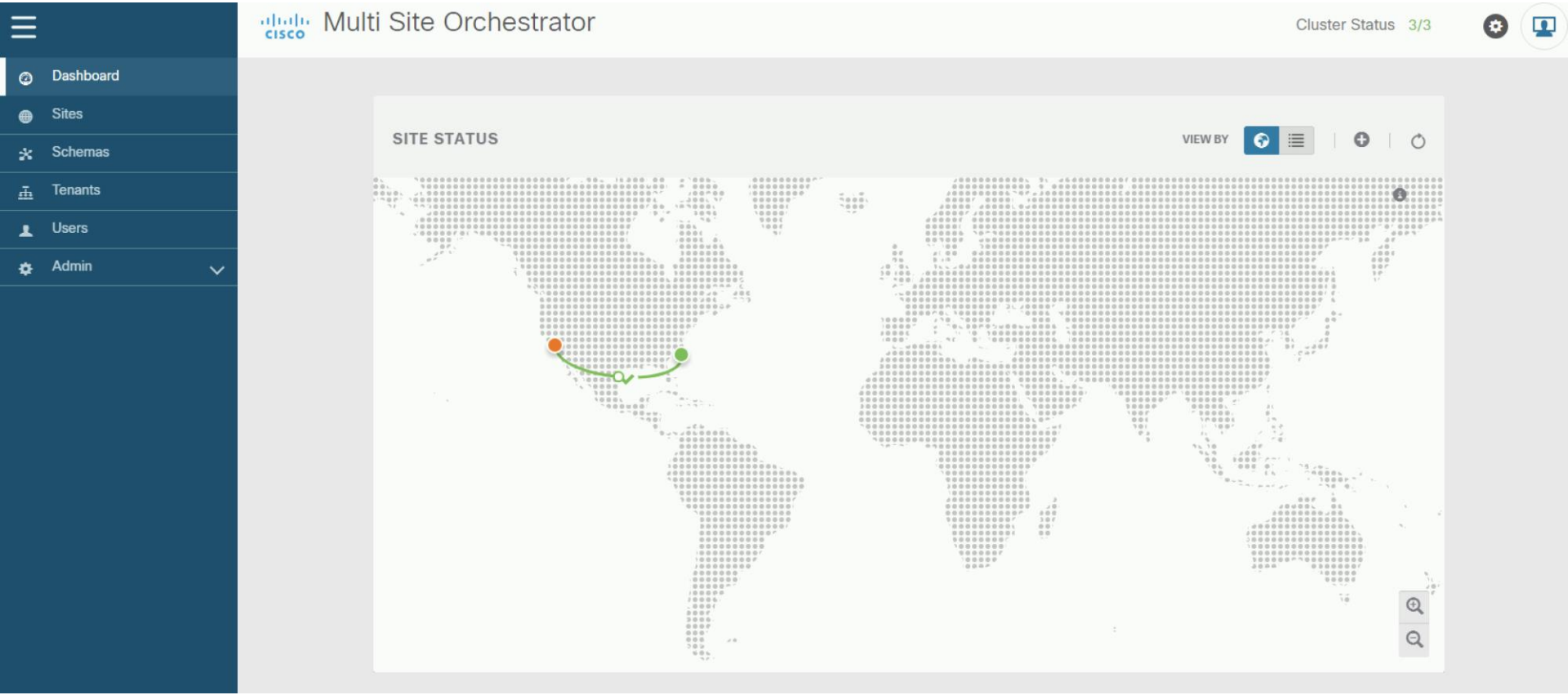


- Multi-Site Infra: Unicast, Multicast, BGP TEPs and Tunnel state
- Multi-Site Tenant and EPG granularity:
  - Inspect and validate full-stack programming: Multi-Site Controller (MSC), APICs and Spine translations
  - Validate the consistency of local and remote inter-site EPGs, BD, VRF, External EPG, policies, etc.
  - Root cause configuration programming issues without calling TAC*
- GUI and APIs are both supported

** Restrictions apply. View NIR/NIA*

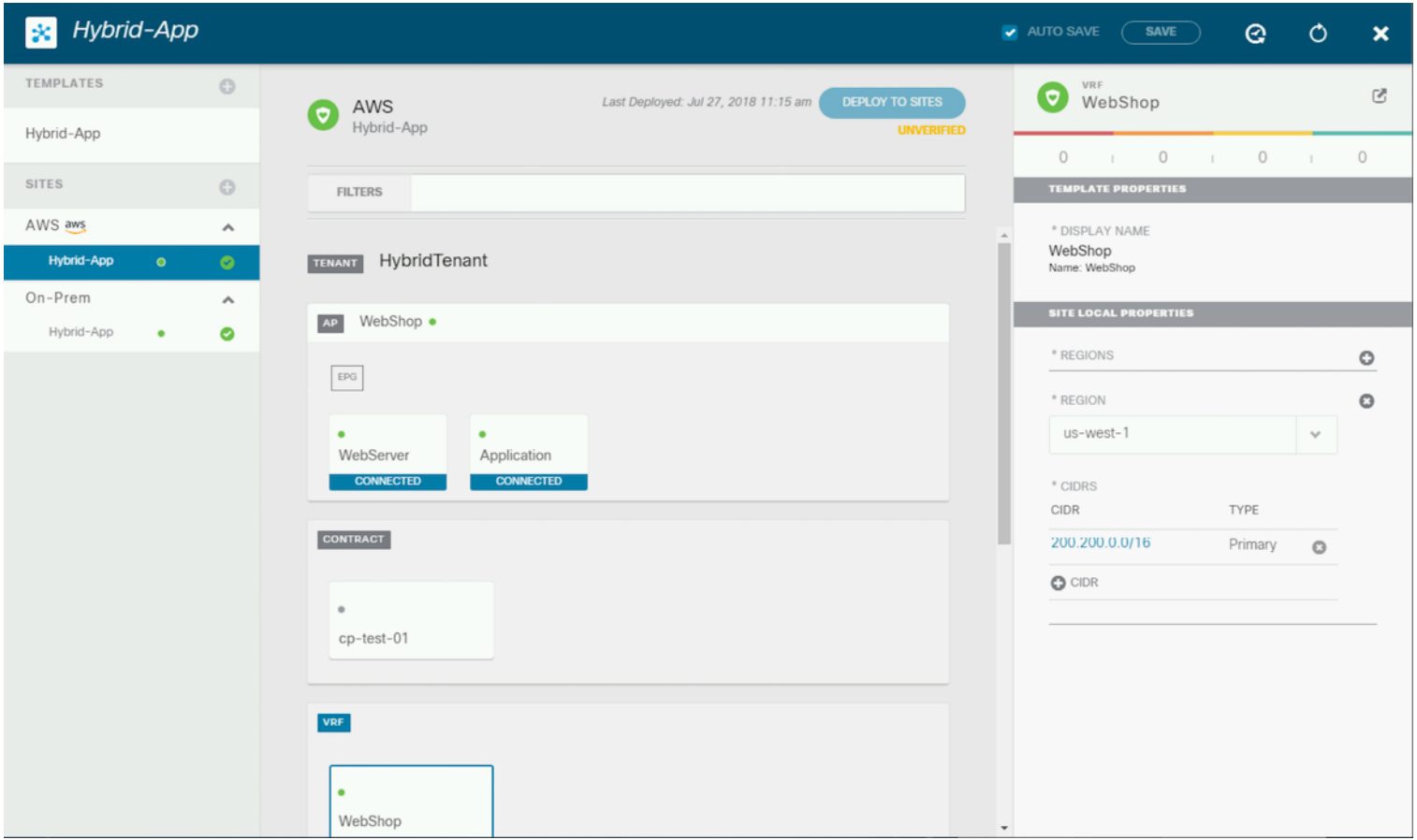
Eg.: Multi-Site Orchestrator - Dashboard

ACI 4.1



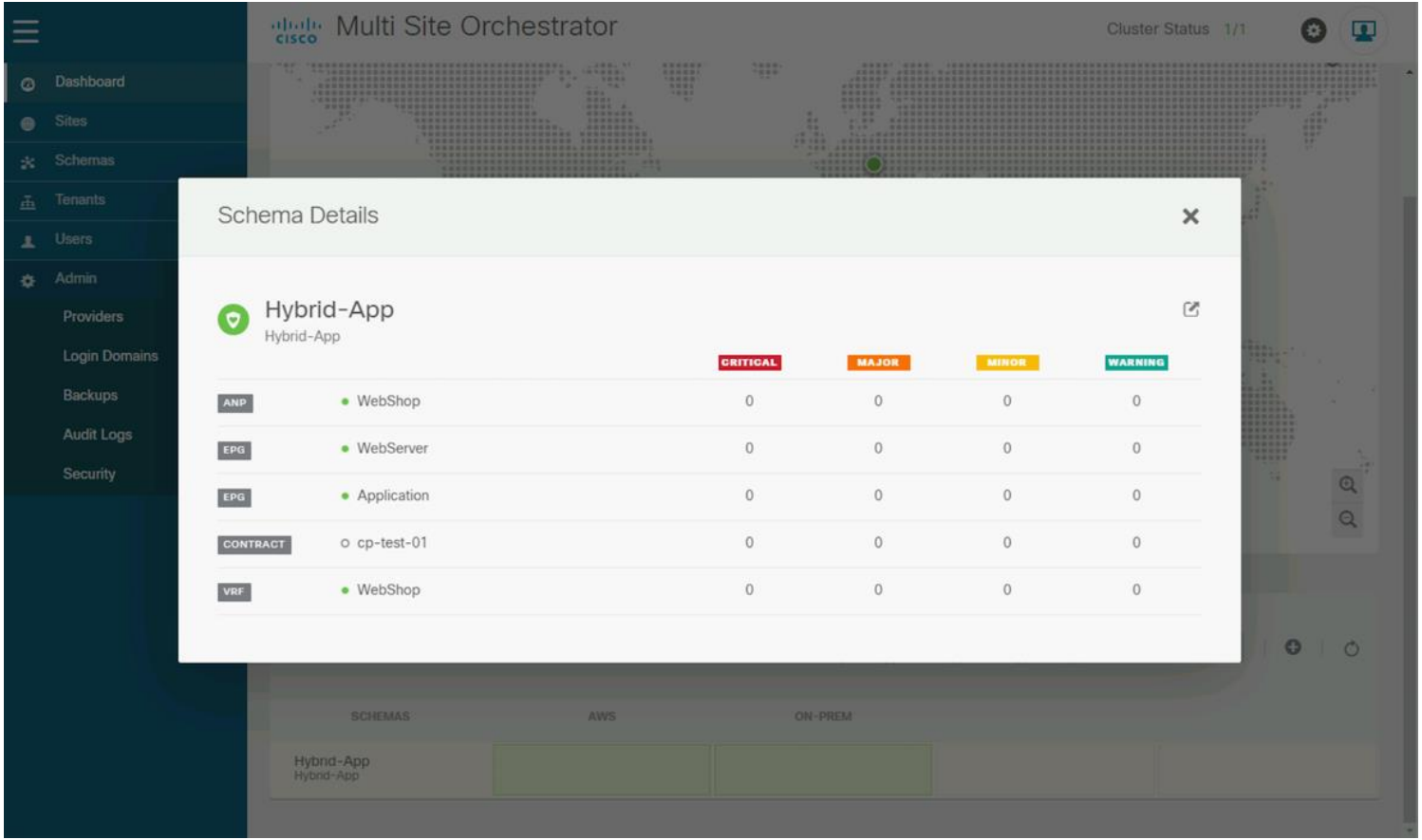
Eg.: Multi-Site Orchestrator – Schema details

ACI 4.1



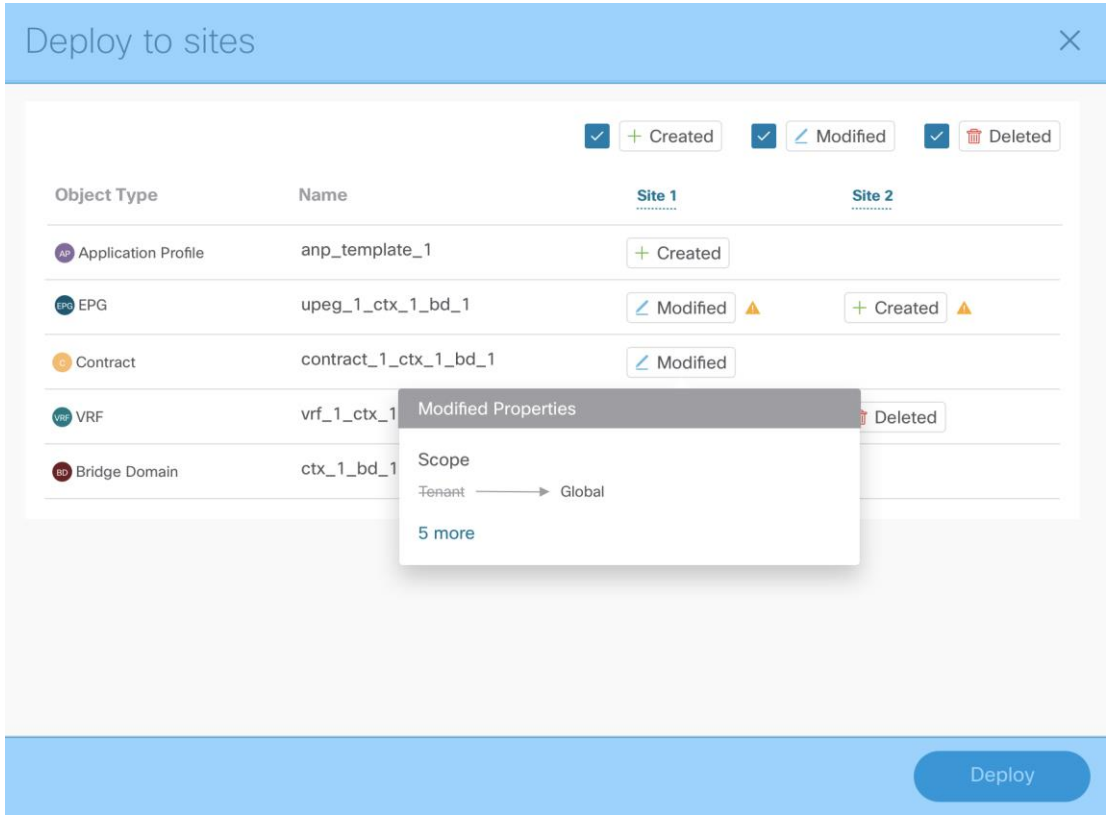
Eg.: Multi-Site Orchestrator – Schema deployed

ACI 4.1



Eg.: Multi-Site Orchestrator 2.2.3 - diffs

ACI 4.2



Eg.: Multi-Site Orchestrator Audit Logs

Multi Site Orchestrator

Cluster Status 1/1

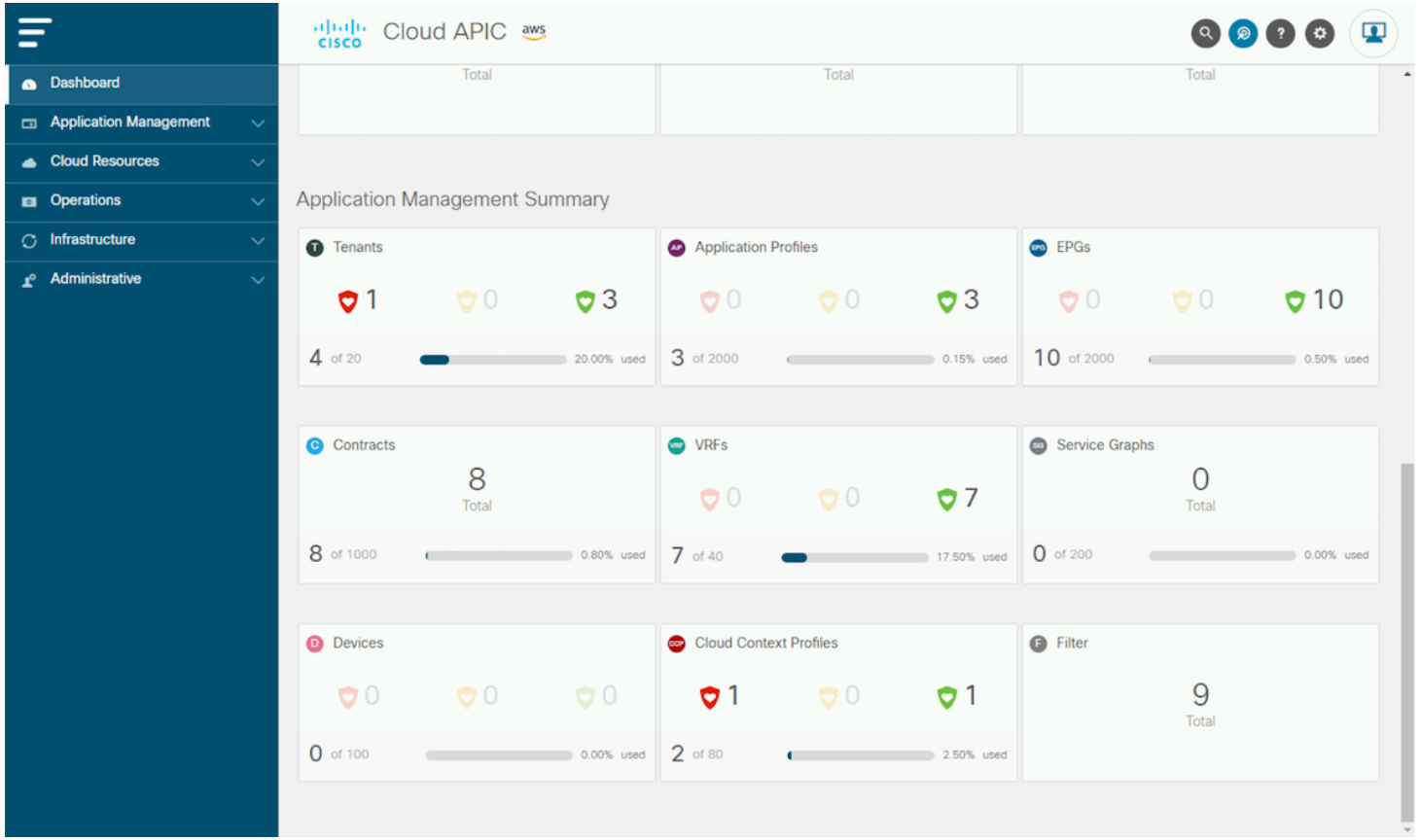
- Dashboard
- Sites
- Schemas
- Tenants
- Users
- Admin
  - Providers
  - Login Domains
  - Backups
  - Audit Logs
  - Security

Audit Logs (592 Logs)

Most Recent

DATE	ACTION	TYPE	DETAILS	USER
Aug 2, 2018 7:39:19 AM	Logged In	Authentication	User admin has successfully logged in	admin (Admin User) Local
Aug 1, 2018 11:12:01 PM	Logged In	Authentication	User admin has successfully logged in	admin (Admin User) Local
Aug 1, 2018 10:43:24 PM	Logged In	Authentication	User admin has successfully logged in	admin (Admin User) Local
Aug 1, 2018 10:11:41 PM	Logged In	Authentication	User admin has successfully logged in	admin (Admin User) Local
Aug 1, 2018 10:11:25 PM	Login Failed	Authentication	Login failed for admin	-
Jul 29, 2018 8:44:45 PM	Logged In	Authentication	User admin has successfully logged in	admin (Admin User) Local
Jul 27, 2018 11:15:33 AM	Deployed	Schema Site	Template Hybrid-App of Hybrid-App was deployed to On-Prem	admin (Admin User) Local
Jul 27, 2018 11:15:33 AM	Deployed	Schema Site	Template Hybrid-App of Hybrid-App was deployed to AWS	admin (Admin User) Local
Jul 27, 2018 11:15:31 AM	Updated	Schema	Schema Hybrid-App was updated	admin (Admin User) Local
Jul 27, 2018 11:15:31 AM	Updated	Template	Template Hybrid-App on Schema Hybrid-App was updated	admin (Admin User) Local
Jul 27, 2018 11:15:31 AM	Updated	EPG	EPG Application on Template Hybrid-App was updated	admin (Admin User) Local

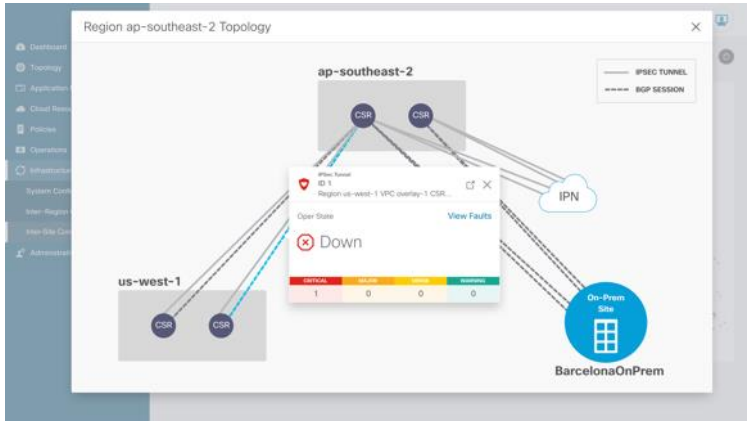
Cloud ACI – Dashboard (AWS example)

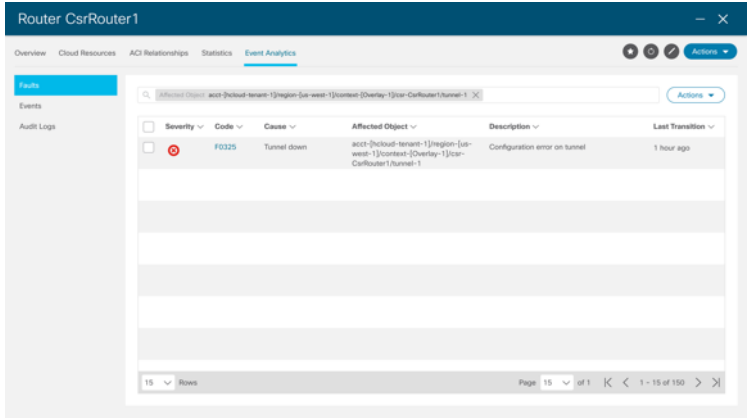


Cloud Native Events, Faults, Stats and Audit Trails

- Cloud events and stats are captured through native services like:
  - Flow Logs – VPC, NIC Flow Stats.
  - Cloud Trail – Audit History of changes to various cloud resources.
  - AWS CloudWatch – Health Monitoring metrics and Log monitoring for events.
  - AWS Config service – Configuration changes to the resources.
  - Additional statistics and events are collected from CSR1000V, like tunnel health, BGP session health etc.
- The cloud events and stats are converted into the ACI Object Model to provide consistent events, faults, stats, audit like on-premises ACI deployment.
- APIC manages cloud inventory objects per account. Aggregation and reporting can be handled at logical objects like CloudEP, CloudEPG , VRF (or) at cloud resource objects such as NIC, host instance, VPC.

Topology Health

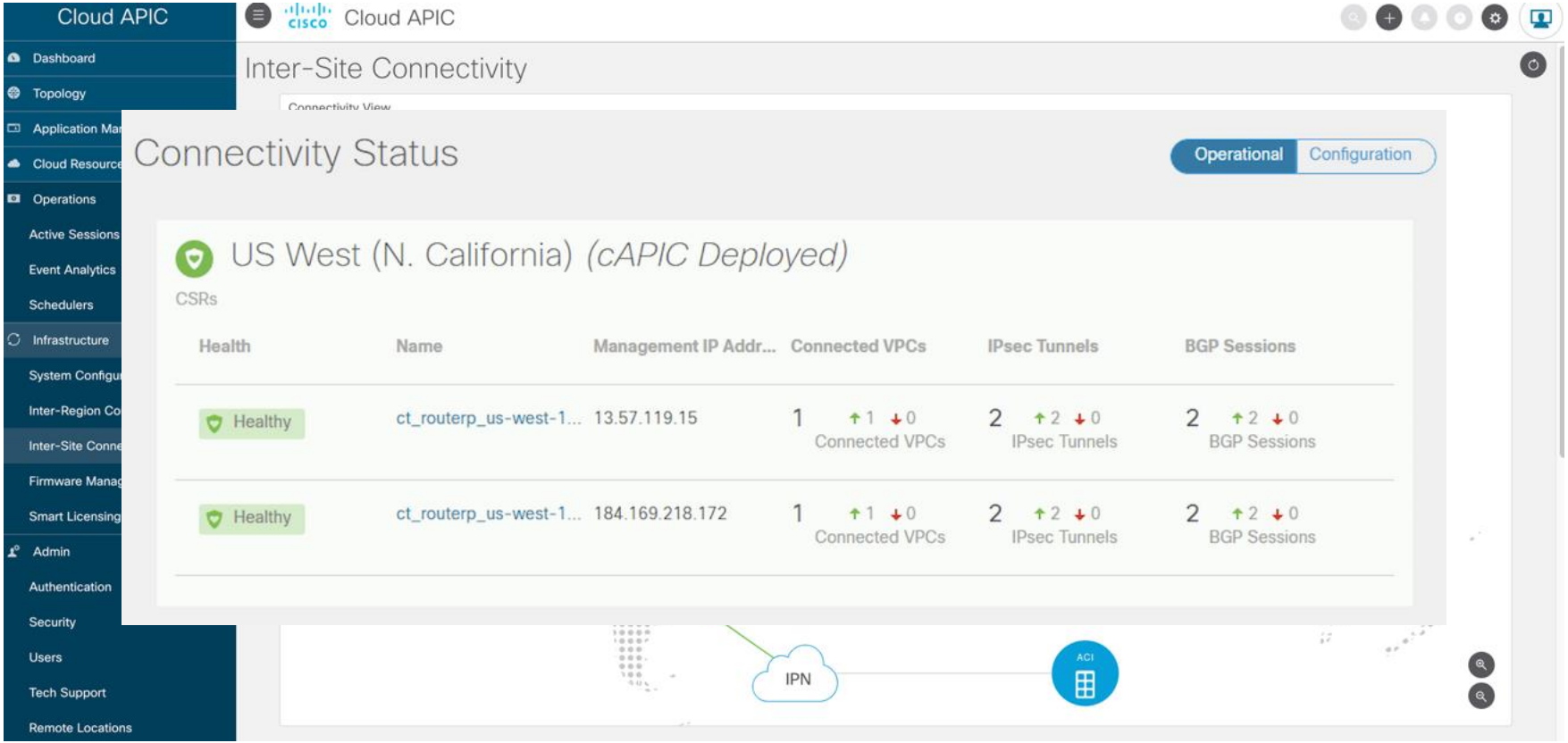




- Network connectivity and Health

Cloud ACI – Inter-Site Connectivity

ACI 4.1



Cloud ACI – Events analytics

ACI 4.1

Tenant HybridTenant

Overview | Cloud Resources | Application Management | Statistics | Event Analytics

Actions

Faults

Events

Audit Logs

Description contains **clus**

ID
Cause
Affected object
Description
Action
Local User

ID	C	cted object	Description	Action	Local User	Creation Time
4294969932	t	-[uni/tn-HybridTenant/cloudapp-Shop/cloudepg-S01/rscons---msc--WebShop-S01]	RsCons --msc--WebShop-CLUS01 created	creation	admin	Jun 08 2019 07:47:36pm
4294969929	transition	-[uni/tn-HybridTenant/cloudapp-Shop/cloudepg-CLUS01/rsprov---msc--WebShop-CLUS01]	RsProv --msc--WebShop-CLUS01 created	creation	admin	Jun 08 2019 07:47:36pm
4294969935	transition	subj-[uni/tn-HybridTenant/cloudapp-WebShop/cloudepg-CLUS01]	EPg CLUS01 created	creation	admin	Jun 08 2019 07:47:36pm
4294969923	transition	subj-[uni/tn-HybridTenant/cloudapp-WebShop/cloudextepg---msc--CLUS01/rscons---msc--WebShop-CLUS01]	RsCons --msc--WebShop-CLUS01 created	creation	admin	Jun 08 2019 07:47:36pm
4294969922	transition	subj-[uni/tn-HybridTenant/cloudapp-WebShop/cloudextepg---msc--CLUS01/rsprov---msc--WebShop-CLUS01]	RsProv --msc--WebShop-CLUS01 created	creation	admin	Jun 08 2019 07:47:36pm

Cloud ACI – Infra CSR1000v Routes

ACI 4.1

context overlay-1

Overview | Cloud Resources | ACI Relationships | Event Analytics | Actions

General

Account
infra

Region
us-west-1

Tags

No tags set

Cloud Resources

2	2
Availability Zones	Routers

Configuration Relationships

3	1
EPGs	Cloud Context Profiles

Settings

Encap
16777199

Encap Type
Vxlan

Cloud Context Profile
ct_ctxprofile_us-west-1
Tenant infra

CIDRs

Address	Subnets	Primary
10.0.1.0/25	10.0.1.0/28 7 more	yes

Route Tables

destinationAddress	Subnets
0.0.0.0/0 1 more	10.0.1.96/28
0.0.0.0/0 1 more	10.0.1.32/28
0.0.0.0/0 1 more	10.0.1.0/28
0.0.0.0/0 1 more	10.0.1.64/28
52.53.52.97 1 more	10.0.1.48/28
52.53.52.97	10.0.1.16/28

Health

N/A

Faults

CRITICAL	MAJOR	MINOR	WARNING
0	0	1	0

Audit Logs

0	0	0
Deletion	Creation	Modification

Events

0	0	0	0
Critical	Major	Minor	Other

Cloud ACI – Statistics example

ACI 4.1

Cloud ACI – Events Analytics

ACI 4.1

Cloud APIC aws

Dashboard
Application Management
Cloud Resources

Event Analytics

Faults Events Audit Logs

Fault F609093

General

Actions

Details

Severity
Major

Code
F609093

Cause
fsm-failed

Affected object
uni/controller/setuppol/setupp-1

Description
[FSM:FAILED]: Remove fabricPod after deleting the pod policy. (TASK:ifc:policymgr:FabricSetupPPodCleanup)

Last Transition
May 01 2019 12:59:34pm

View More

Change Set

No	Major	F606666	uni/fabric/macprotp-default/macexpg-default	[FSM:FAILED]: Assign Virtual IP address fo Protection Group Name(TASK:ifc:policymgr:FabricAProtGEp	May 01 2019 12:59:40pm

ACI Operations – Agenda

- Before getting started – setting the concepts stage

Visibility	Insights	Actions
Faults, events, stats, health, logs, trails	Application dependency	Incident Troubleshooting
Configuration	Containers Integration	Change Management

ACI Configuration Rollback

APIC (ACI-SJC-1)

What are you looking for?

admin

System | Tenants | Fabric | Virtual Networking | L4-L7 Services | Admin | Operations | Apps | Integrations

AAA | Schedulers | Historical Record Policies | Firmware | External Data Collectors | Config Rollbacks | Import/Export | Downloads

Config Rollbacks for: Tenant

Snapshots	File Name	Description	File Size (bytes)
2018-03-08 15:25:1...	ce2_defaultOneTime-2018-03-08T13-25-13.ta...	before tenant Rivella ...	54667
2018-06-10 22:50:2...	ce2_defaultOneTime-2018-06-10T22-50-13.ta...	CLUS18-SNAP1	580416
2018-06-10 22:52:1...	ce2_defaultOneTime-2018-06-10T22-52-15.ta...	CLUS18-SNAP2	580698
2018-06-10 22:59:3...	ce2_defaultOneTime-2018-06-10T22-59-29.ta...	CLUS18-SNAP3	578894
2018-12-12 18:33:4...	ce2_defaultOneTime-2018-12-12T17-33-42.ta...	pre-CCP-2.1	823781
2018-12-14 17:05:4...	ce2_defaultOneTime-2018-12-14T16-05-39.ta...	Pre-HX2	817899
2018-12-21 12:27:4...	ce2_defaultOneTime-2018-12-21T11-27-38.ta...	chrischtchindli	818092
2019-05-16 10:53:4...	ce2_config_backup-2019-05-16T10-53-38.tar....	Backups taken before...	850052
2019-05-16 10:57:5...	ce2_config_backup-2019-05-16T10-57-44.tar....	Backups taken before...	848866
2019-05-16 11:00:5...	ce2_config_backup-2019-05-16T11-00-49.tar....	Backups taken before...	848304
2019-05-16 11:03:5...	ce2_config_backup-2019-05-16T11-03-50.tar....	Backups taken before...	850067
2019-05-16 11:05:5...	ce2_config_backup-2019-05-16T11-05-49.tar....	Backups taken before...	850481
2019-05-16 11:10:5...	ce2_config_backup-2019-05-16T11-10-49.tar....	Backups taken before...	849511
2019-05-16 11:16:0...	ce2_config_backup-2019-05-16T11-15-55.tar....	Backups taken before...	850002
2019-05-16 11:17:4...	ce2_config_backup-2019-05-16T11-17-39.tar....	Backups taken before...	850175
2019-05-16 12:42:2...	ce2_config_backup-2019-05-16T12-42-20.tar....	Backups taken before...	850673

Actions

Rollback
Select any one snapshot on left to start.

Take a snapshot
Location: APIC
Description: optional
Create a snapsho

Import export file to snapshot
Click icon on top

Modify import/export security settings
Click icon on top

Create recurring snapshots
Click icon on top

Automatically create snapshot — Enable

Snapshots taken every

Mon Tue Wed Thu Fri Sat Sun

at Hour: 0 Minute: 0

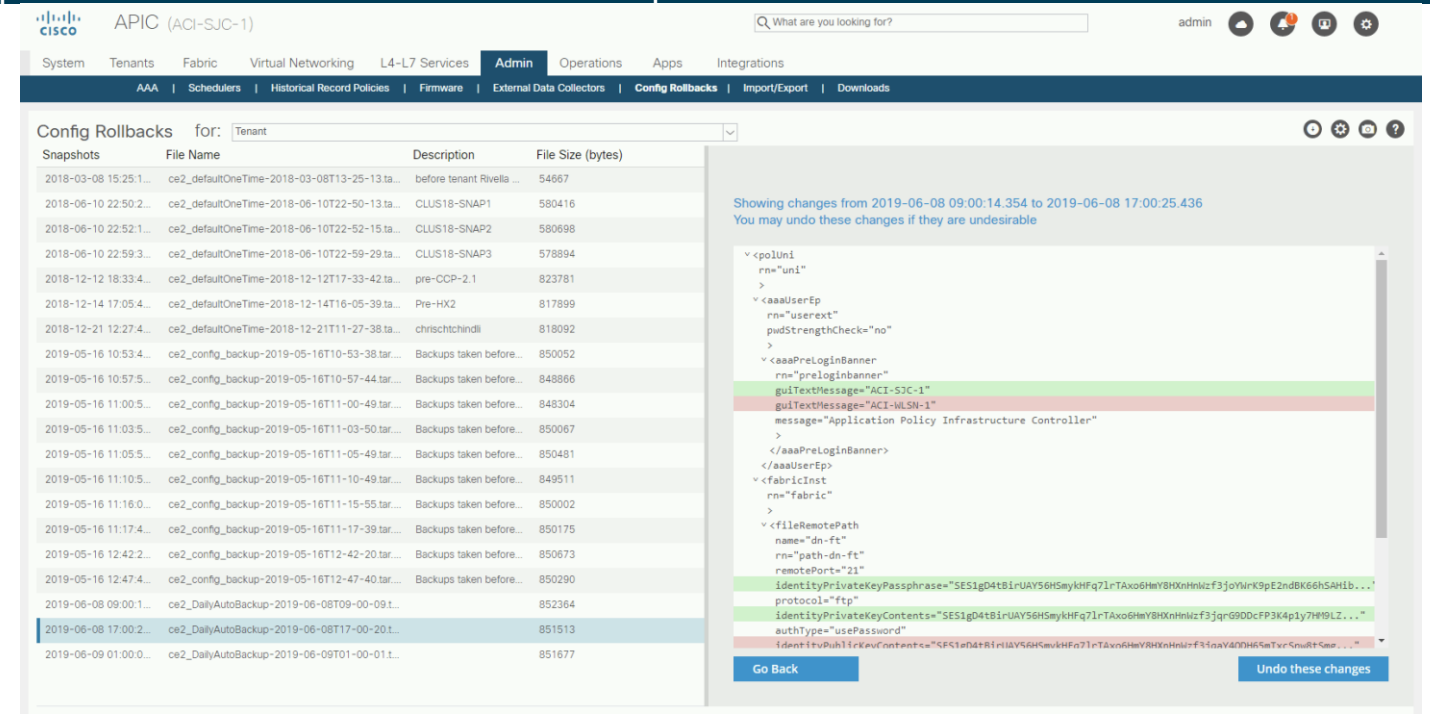
Save to Remote Location instead: select a location

Create a remote location:

Cancel Submit

ACI Configuration Rollback

- Snapshot for whole fabric / per tenant basis. You can create snapshots manually or periodic
- Example shows the difference between 2 snapshots at fabric level



Duplicate IP

APIC (apic1.smalldc.ceclabs.local:1601) - Chromium

Not secure | https://apic1.smalldc.ceclabs.local:1601/#a:a

APIC (Disneyland)

System | Tenants | Fabric | Virtual Networking | L4-L7 Services | Admin | Operations | Apps

QuickStart | Dashboard | Controllers | System Settings | Smart Licensing | Faults | Config Zones | Events | Audit Log | Active Sessions

System Health 92

Zoom 1H 1D All

score 50 0

22:20 22:30

9. Jun 06:00

Alert List

To disable blinking of the alert icon's badge, remove all Critical alerts. A disabled close button on a Critical alert indicates that you must resolve the underlying issue.

Same IP Address used for multiple MAC's
One or more duplicate IP addresses are being used by multiple MAC addresses. Please resolve this to avoid conflicts.
View all Duplicate IP Usage

Smart Licensing is not configured.
The evaluation period has 71 day(s) remaining. There will be no impact on the functionality of the ACI fabric at the end of evaluation period.
Go to Smart Licensing

Critical: 1 | Major: 0 | Minor: 1 | Info: 0

Close

Fault Counts By Domain

Hide Acked Faults | Hide Delegated Faults

2	106	52	74
2	0	0	8
0	24	5	1
0	10	0	0
0	70	7	61
0	0	0	0
0	0	0	0
0	2	40	4

Nodes With Health ≤ 99

Name	Pod ID		
leaf-01	1	leaf	78
leaf-02	1	leaf	98
leaf-03	1	leaf	98
leaf-05	1	leaf	98
leaf-06	1	leaf	98

s By Type

Hide Delegated Faults

Communications	2	18	0	8
Config	0	56	43	64
Environmental	0	0	7	0
Operational	0	32	2	2

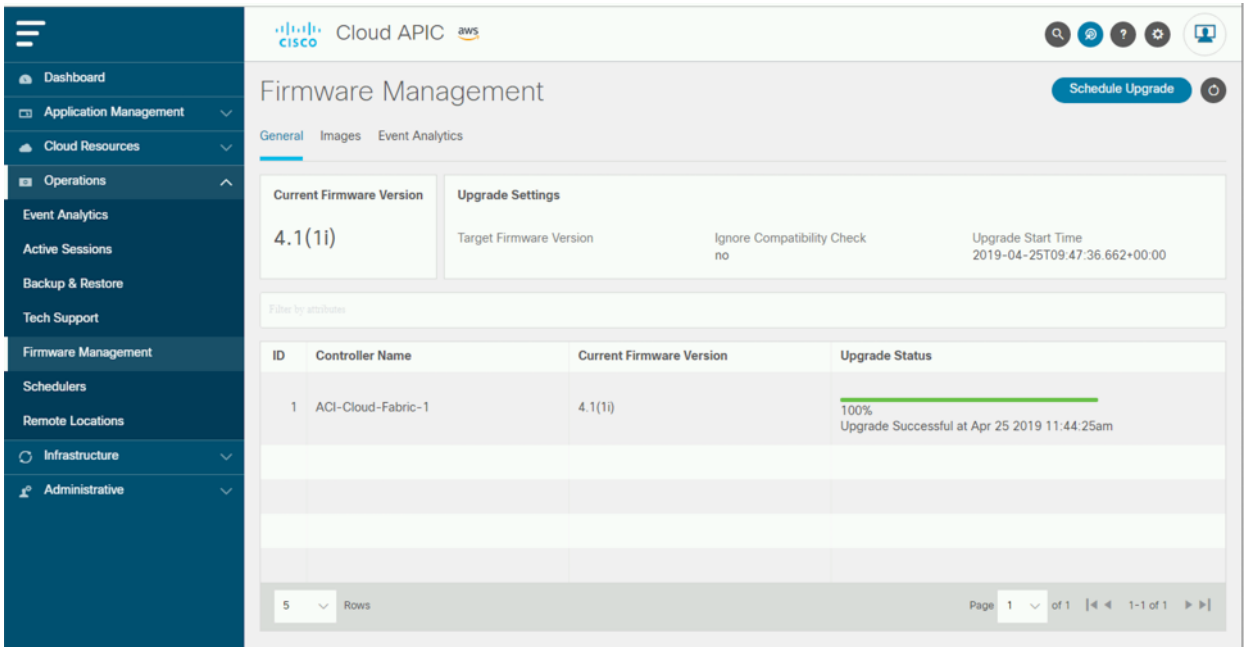
Show me how

Last Login Time: 2018-06-09T22:55 UTC+02:00

Current System Time: 2018-06-09T23:16 UTC+02:00

Cloud APIC – Upgrade

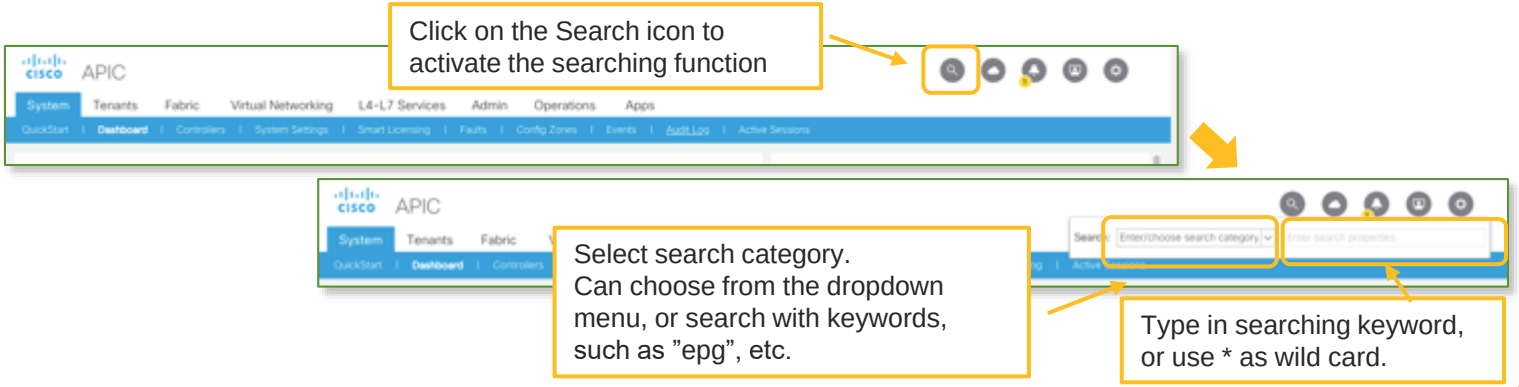
ACI 4.1



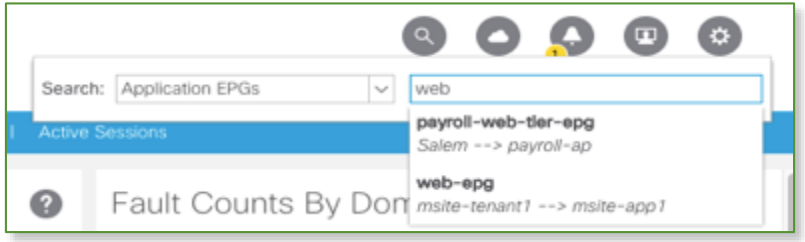
- Similar steps as APIC
- Under Firmware Management select image location
- Schedule a time to upgrade
- Once done, it will show upgrade got completed

Improved Native Searching Function

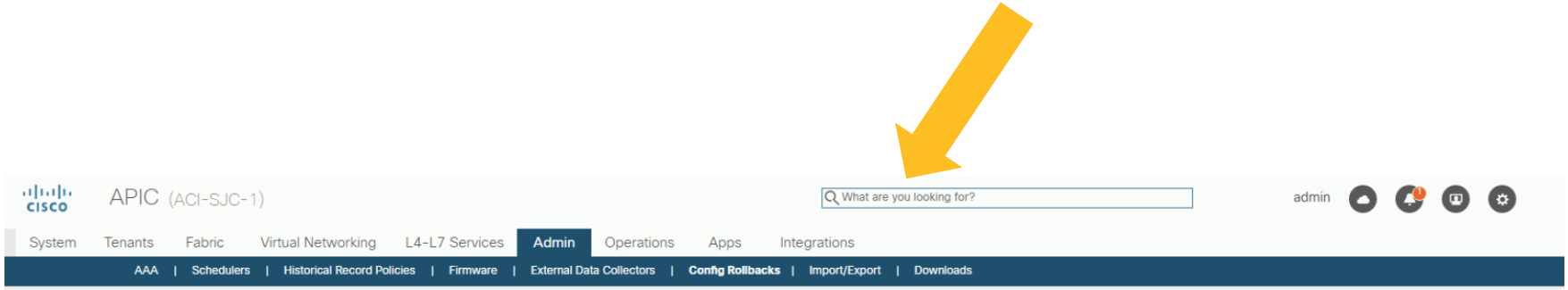
Since the APIC 3.2, the UI provides a more flexible and easy-to-use native searching function with the support of search keywords and wild card.



Search Output Example:



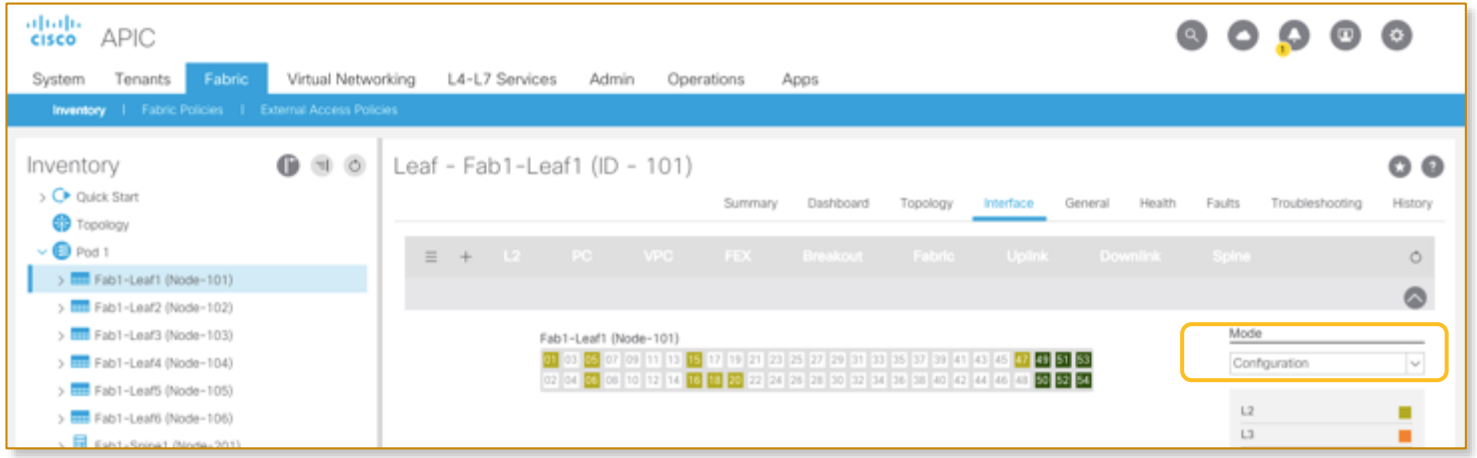
Since ACI 4.1 – search to rescue …



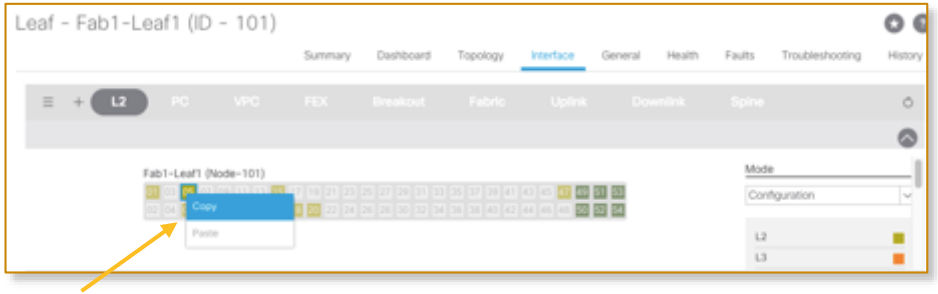
Copy/Paste L2 Port Configuration

Since the APIC 3.2, the UI adds the copy/paste function to replicate L2 port configuration on leaf switches. This is an alternative method to quickly configure L2 ports individually. The system will automatically create the associated objects behind the scenes.

Navigate to the function via Fabric → PoD → Leaf Switch → Interface, then toggle the mode to Configuration



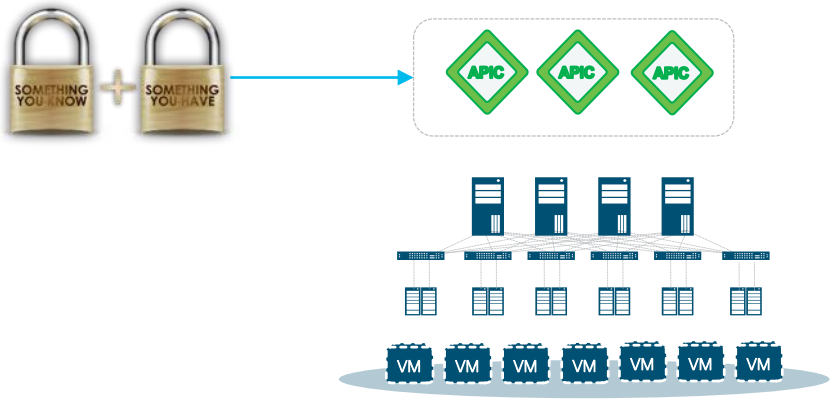
Copy/Paste L2 Port Configuration



Select the copy-from port

Select and confirm the copy-to port

ACI 2-Factor Authentication Options



ACI 3.0

External Authentication via SAML and IDPs supported Okta & MSFT ADFS

ACI 3.0

Local Authentication TOTP using Google Authenticator for 2nd factor pin/barcode

ACI 3.1

RSA SecureID

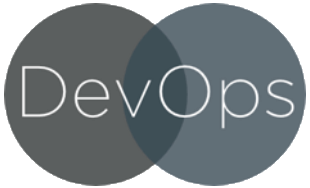
ACI 3.2

PingFederate SSO PingID 2-FA

ACI 4.0

Federal Common Access Card (CAC)

Culture change @ network operations …

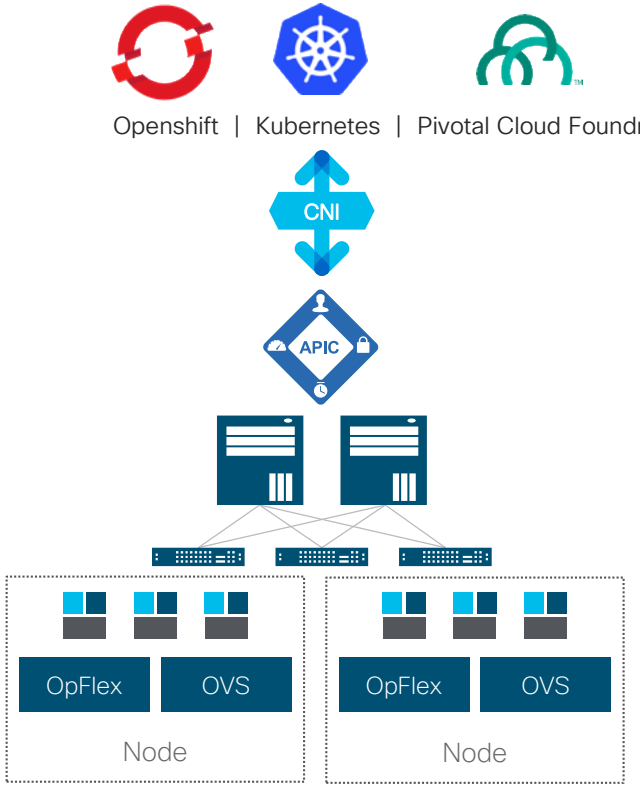


"ACI is my SDN solution and I only need access to its APIs as it must be part of my CI/CD pipeline for upgrades so I can have the latest integrations with the tools I need."

"ACI is the production network of my key business processes, so any upgrade goes through change management and as/if needed, must be done at a maintenance window"

Since ACI 3.0

ACI Integration with Container Application Platforms

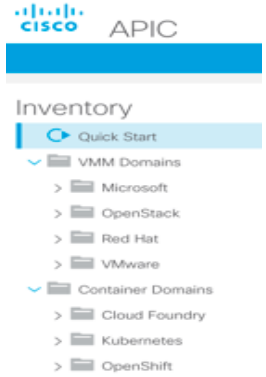


ACI and Containers

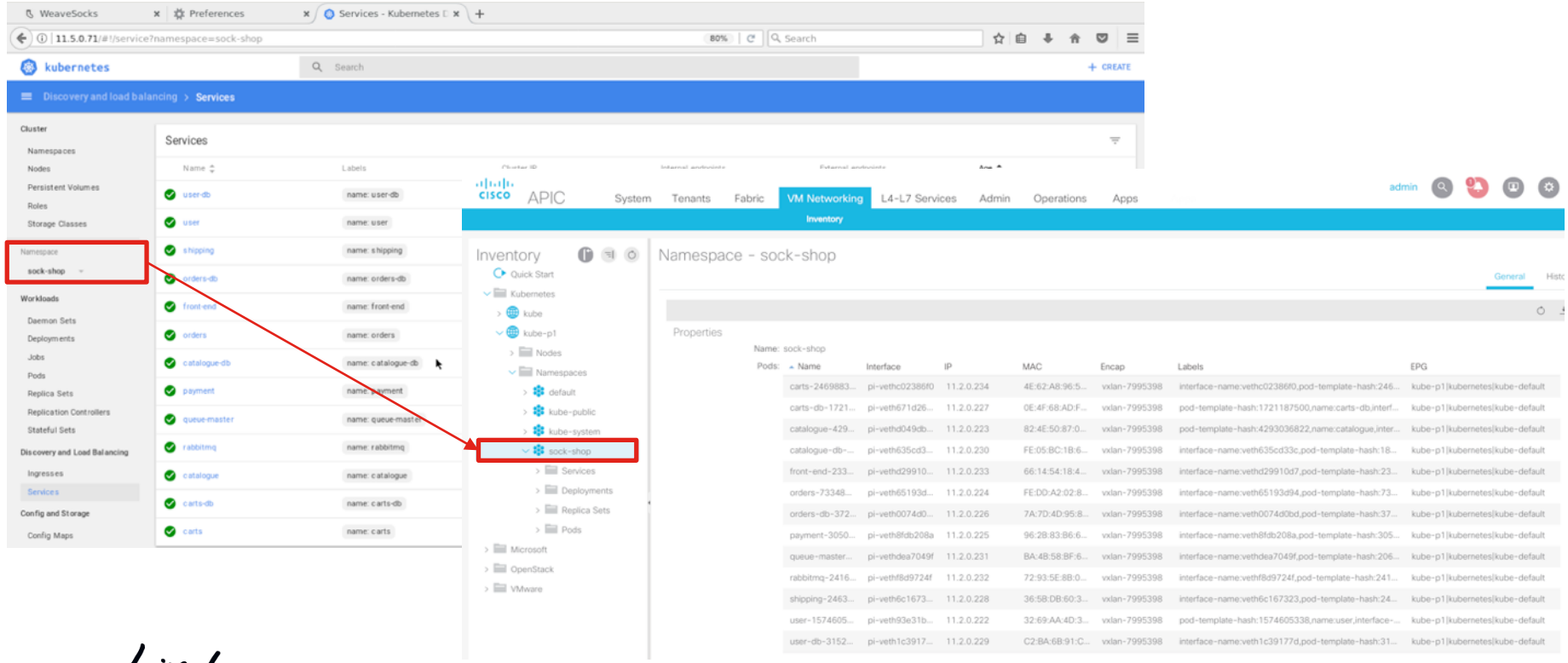
Unified networking:
Containers, VMs, and bare-metal

Integration of containers network policies and ACI policies

Visibility: Live statistics in APIC per container and health metrics



VMM domain helps to bridge the gap between Kubernetes admin and network operations



Fabric Administrator has inventory of Kubernetes objects – simplify operations

Fabric admin can search APIC for k8s nodes, masters, pods, services …

APIC System Tenants Fabric VM Networking L4-L7 Services Admin Operations Apps

admin

Inventory

Inventory
- kubernetes
  - Controllers
    - kubernetes
      - Nodes
        - fab10-compute-1
        - fab10-compute-2
        - fab10-compute-3
        - fab10-compute-4
        - fab10-compute-5
      - Namespaces
        - default
          - Services
          - Deployments
          - Replica Sets
          - Pods
        - kube-system
          - Services
          - Deployments
          - Replica Sets
          - Pods
- Microsoft
- OpenStack
- VMware

Hypervisor - fab10-compute-2

Topology General Health Faults History

100

Properties

Name: fab10-compute-2
Hostname: fab10-compute-2
IP Address: 10.0.48.65
Status: Connected
Pods:

Name	Interface	IP	MAC	Encap	Namespace	Labels	EPG
kubedn...	pi-veth6ca705f9	10.1.2.2	B2:5C:...	vxlan-7634944	kube-system	kubernetes.io/cluster-service:...	k8s|k8s-app|kube...
kubedn...	pi-vethf97eca53	10.1.2.4	0A:95:...	vxlan-7634944	kube-system	interface-name:vethf97eca53...	k8s|k8s-app|kube...

APIC keeps inventory of pods and their metadata (labels, annotations), deployments, replicasets, etc.

View pods per node, map to encapsulation, physical point in the fabric.

Supported Container Application Platforms

ACI 4.2

	Baremetal	ESXi	KVM (Openstack)
Open source Kubernetes 1.6 – 1.15	✓	✓	Future
Cisco Container Platform	Future	✓	✓
Openshift 3.6, 3.9, 3.11	✓	✓	✓
Pivotal Cloud Foundry (PCF) 2.1.1*	✓	✓	Future
Docker EE 2.1 (only with Kubernetes and/or Openshift)	✓	✓	Future
Mesosphere	Not currently planned		

ACI virtualization support matrix – https://www.cisco.com/c/dam/en/us/td/docs/Website/datacenter/aci/virtualization/matrix/virtmatrix.html

Demo 3: ACI Operations

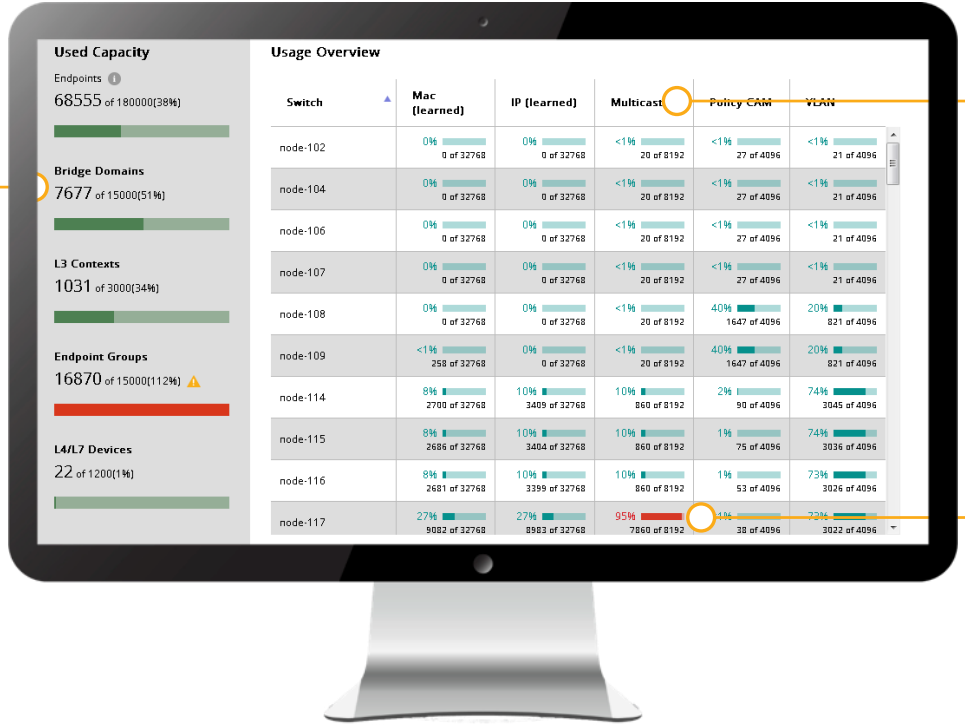
ACI Operations – Agenda

- Before getting started – setting the concepts stage

Visibility	Insights	Actions
Faults, events, stats, health, logs, trails	Application dependency	Incident Troubleshooting
Configuration	Containers Integration	Change Management
Capacity, Fabric metrics (utilization, flows, states, environmental, etc.), Telemetry	Anomalies detection (via SW & HW correlation) Trends	Increase Performance, Availability & Reliability Prevent Outages

ACI Capacity Dashboard

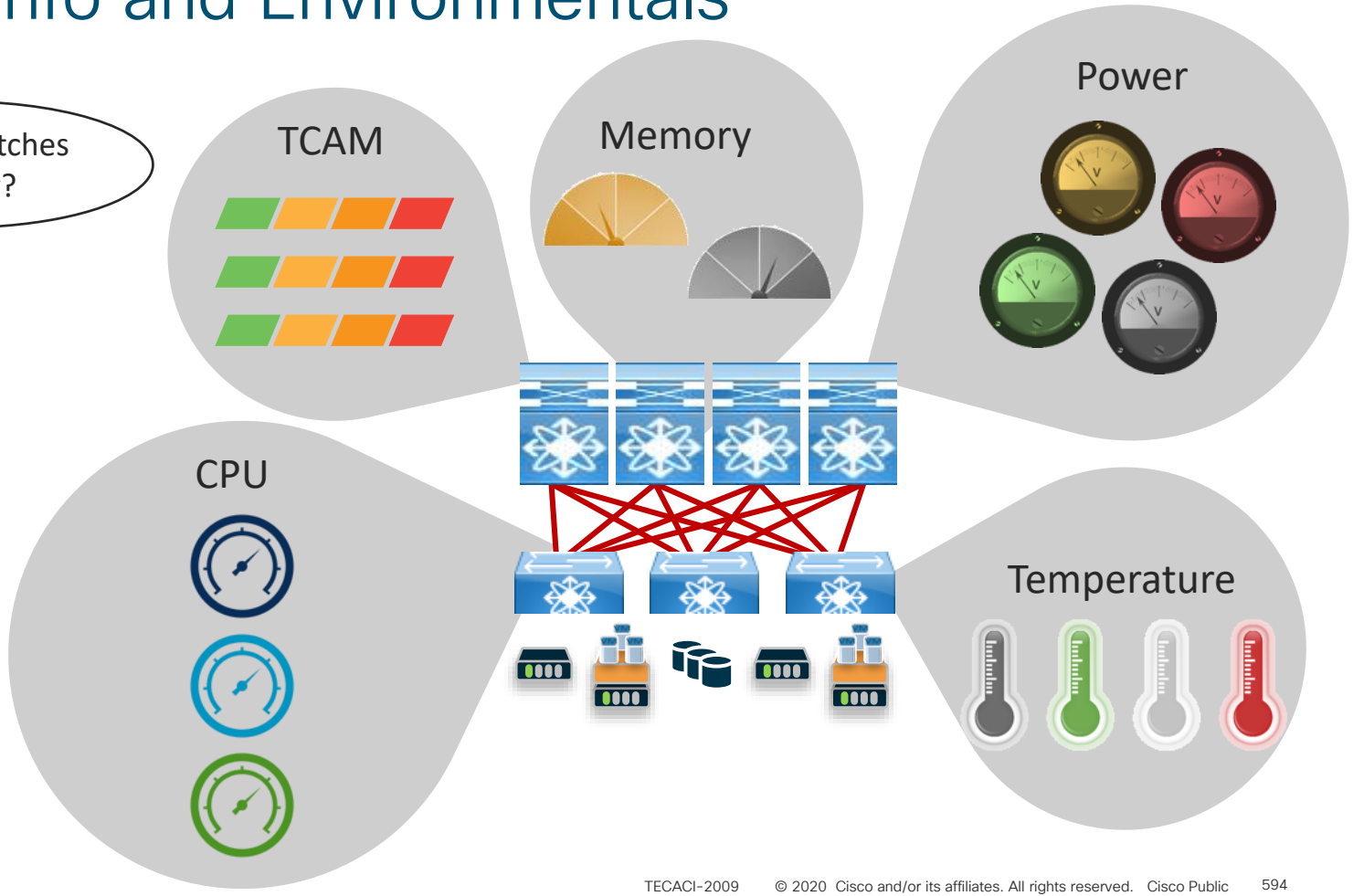
Comprehensive Management and Automation



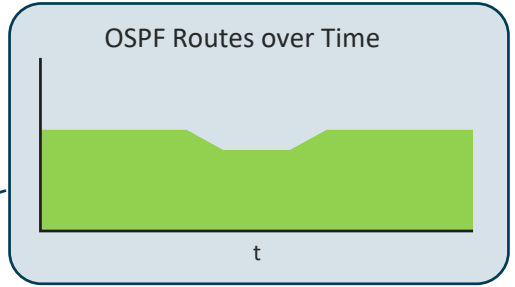
Understand easily ALL major aspects of capacity (MAC tables, Policy CAM etc.)

Instantly identify capacity issues for the ENTIRE fabric with clear visuals

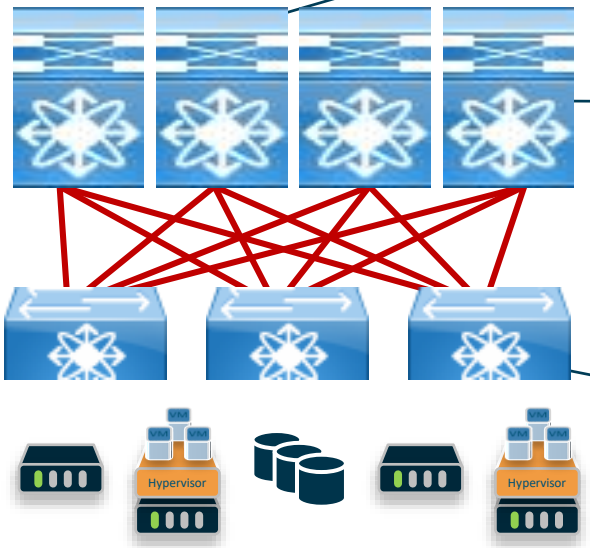
System Info and Environmentals



Protocol State and Events



Is routing working as expected?



Alert:

! Neighbor Lost!

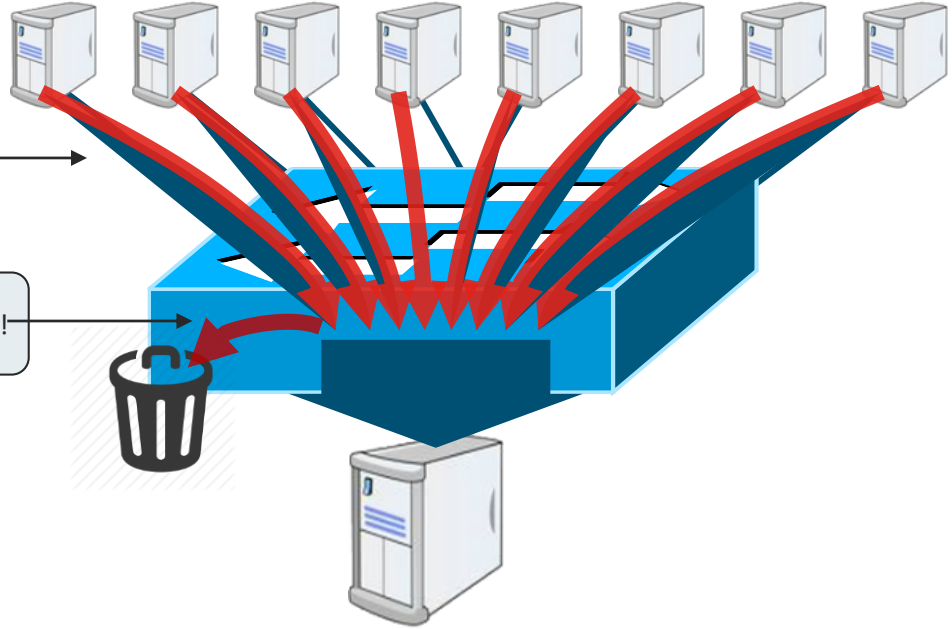
OSPF Process State	
Process ID	10
Router ID	10.1.1.1
Area	0.0.0.0
OSPF Interfaces	105
LSAs	7

Monitoring Buffer Utilization and Drops

I see queue drops – but who's affected?!

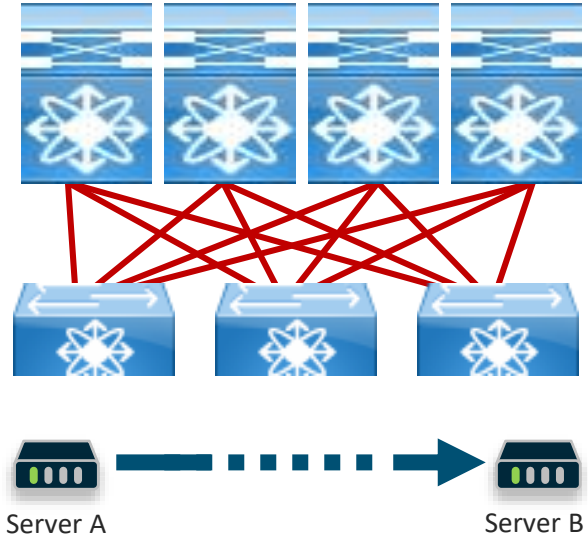
Incast or other oversubscription

Packet drops!

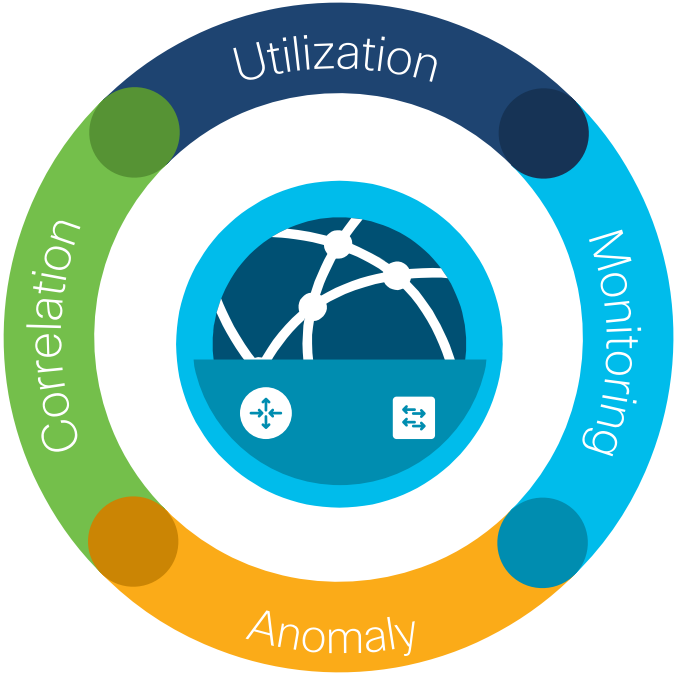


Network Path and Latency Measurement

Application performance is slow between Server A & Server B!

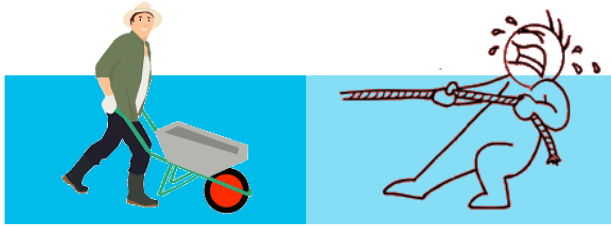


ACI Telemetry to rescue



Telemetry is the only true way of seeing data that represents what the network is experiencing at any point in time…

Key Telemetry Characteristics



Push not Pull

Efficient Delivery

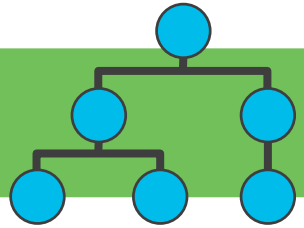
{JSON} JavaScript Object Notation

protobuf Protocol Buffers

GRPC http:// UDP

Analytics-ready Data

Tool-Chain consumption and Integration



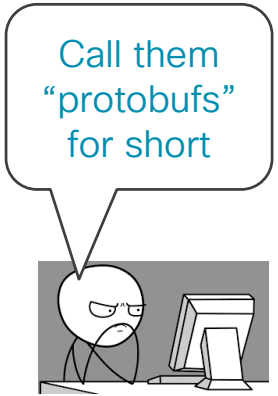
Data-model Driven Consistent format

Structure and Automation

Encoding Options

Google Protocol Buffers (GPB)

- Designed for simplicity, performance
- Intended for machine consumption, not human consumption



protobuf
Protocol Buffers

JavaScript Object Notation (JSON)

- Human-readable, self-describing, text-based encoding format
- Open-standard
- Not designed with performance or extensibility in mind

{JSON}
JavaScript Object Notation

Transport Options

gRPC

- Modern, open source RPC framework
- Low latency, scalable, distributed
- Enables extension such as authentication, load balancing, logging and monitoring etc.

HTTP

- Ubiquitous transport option
- Many available open source stacks on multiple operating systems
- Well understood in industry

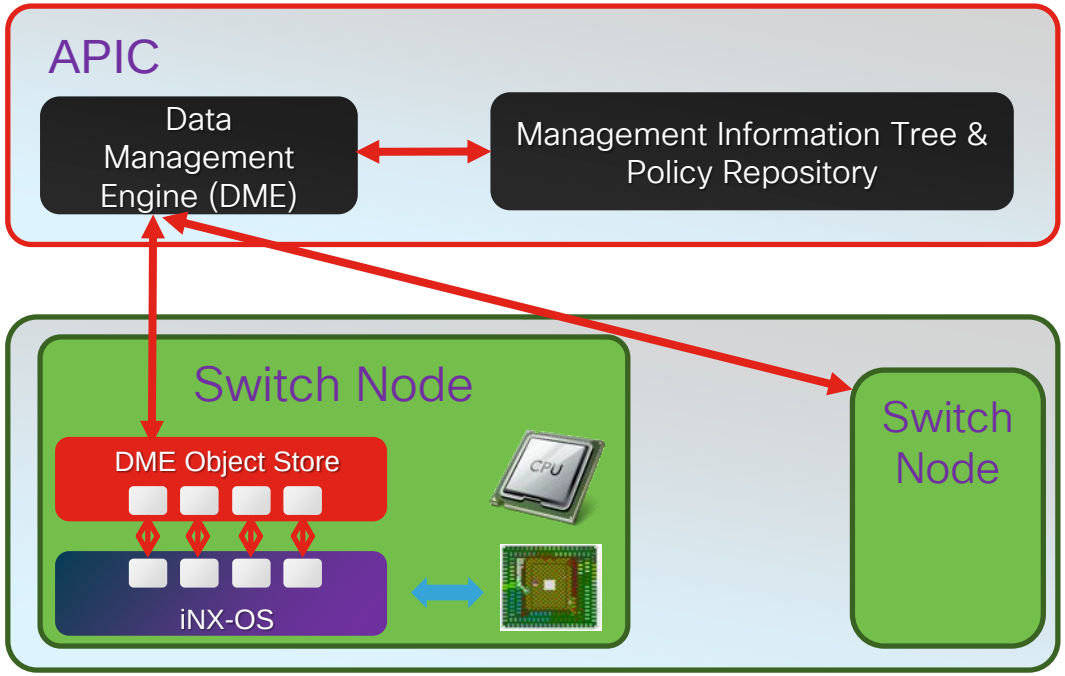
UDP

- Connectionless transport
- No upper-layer overhead

ACI Model Based Telemetry

Embedded in the Architecture

- ACI leverages a distributed DME (distributed database)
- Provisioning State is durably stored in fault-resistant manner by keeping multiple in-synch replicas
- Active Hardware and Software state (counters, faults, logs, …) is distributed via DME as well
- Centralized Reporting and Correlation
- Application-level visibility



Cloud Scale Streaming Hardware Telemetry

Flow Table (FT)

Captures full data-plane packet flow information, plus metadata

Flow Table Events (FTE)

Triggers notifications based on thresholds / criteria met by data-plane packet flows

Data-Plane Flow Data

Streaming Statistics Export (SSX)

Streams ASIC statistics based on user configuration

ASIC State

Available only on Cisco Cloud Scale platforms!

Telemetry Data Source

Flow Table (FT)

- Captures full data-plane packet flow information, plus metadata
  - 5-tuple flow info
  - Interface/queue info
  - Flow start/stop time
  - Flow latency

Direct hardware export with low flush times (100 milliseconds)

Streaming Statistics Export (SSX)

- Streams statistics and other ASIC-level state data based on user config
  - Interface statistics (packets/bytes/drops)
  - Buffer depth
  - Queue-level microburst stats

Direct hardware export with very low collection intervals (10's of microseconds)

Flow Table Events (FTE)

- Triggers notifications based on thresholds / criteria met by data-plane packet flows
  - 5-tuple flow info
  - Interface/queue info
  - Forwarding drop indication
  - Buffer drop indication
  - Latency/burst threshold indication

Direct hardware export with flow-level and global throttling

9300-FX / 9300-FX2 platforms support triggered flow table events

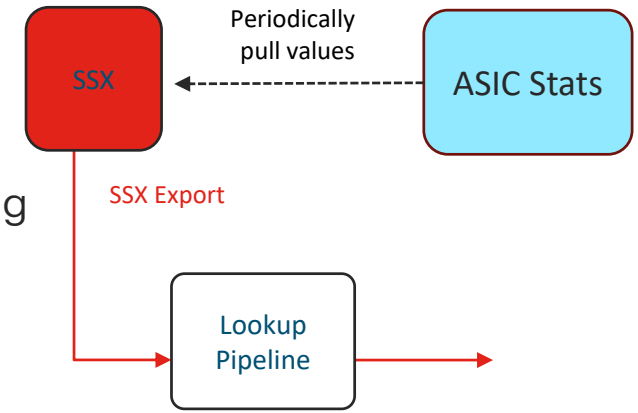
Flow Table

- Collects full flow information plus metadata
  - 5-tuple flow info
  - Interface/queue info
  - Flow start/stop time
  - Flow latency
- 32K flow table entries per ASIC slice
- Direct hardware export
- EX / FX / FX2 based 9k platforms support hardware flow table

Flow Table Events

- Triggers notifications based on criteria / thresholds met by data-plane packet flows
- Collects full flow information plus metadata
  - 5-tuple flow info with timestamp
  - Interface/queue info
  - Buffer drop indication
  - Forwarding drop, ACL drop, policer drop indication
  - Latency/burst threshold exceeded indication
- Direct hardware export, with flow-level and global throttling
- FX / FX2 based 9k platforms support triggered flow table events

Streaming Statistics Export (SSX)



- Streams ASIC statistics at rapid cadence based on user config
  - Interface counters (packets/bytes/drops)
  - Ingress/Egress queue depth
  - Ingress/Egress queue drops
  - Egress queue microbursts
  - Buffer depth
- User defines streaming parameters – which statistics, how often, and to which collector
- Direct export from ASIC to front-panel port – no switch CPU involvement
- Hardware support in 9364C / 9300-FX2/GX / 9500-FX2/GX

Hardware Telemetry Platform Support

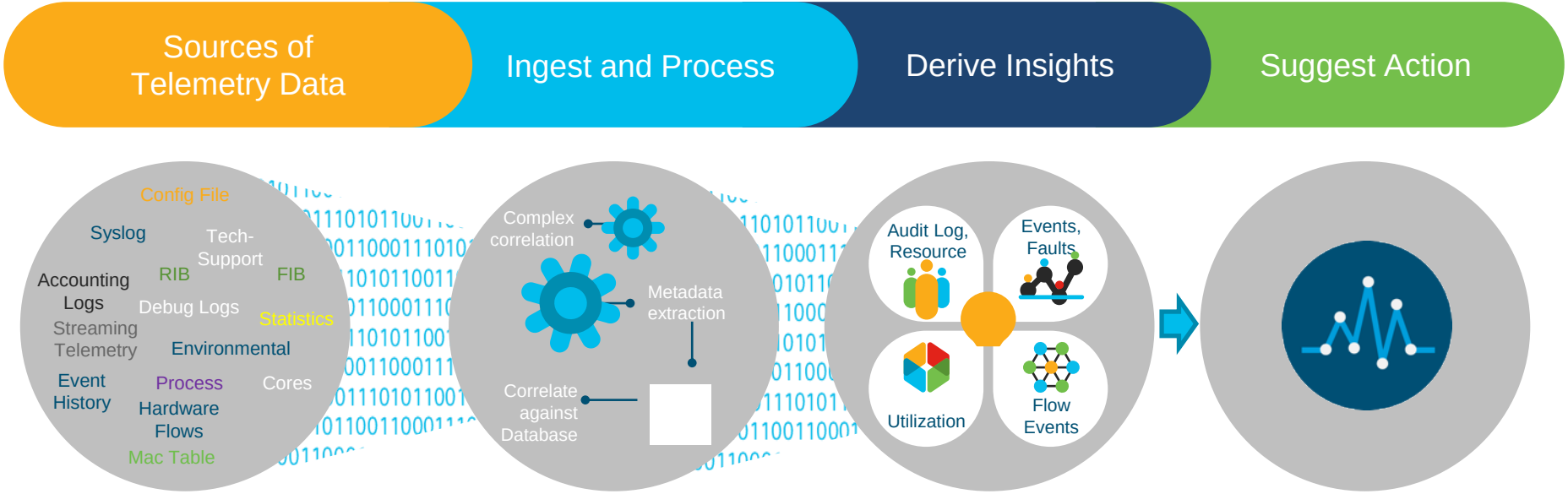
Platform	FT	FTE (roadmap)	SSX (roadmap)
9300/9500-EX	✓	X	X
9300/9500-FX	✓	✓	X
9364C	X	X	✓
9300-FX2	✓	✓	✓
9k GX platforms	✓	✓	✓

Next step: Build your own Telemetry Platform ?

- Too many open-source options to pick from
- Complex interactions between services
- Auto scaling of the application
- Building telemetry dashboards
- Long term support

Investing in a software development team

Network Insights Framework – Enabling proactive action



Increase Availability, Performance and Visibility

Leverage Knowledge Base

ACI | NX-OS

Network Insight Telemetry Applications

Providing Network Health Visibility & Enabling Proactive Insights

ACI App

Network Availability

Advisor

Proactive Software Recommendations/Notifications
Issue Vulnerability Detection & Remediation

Network Health

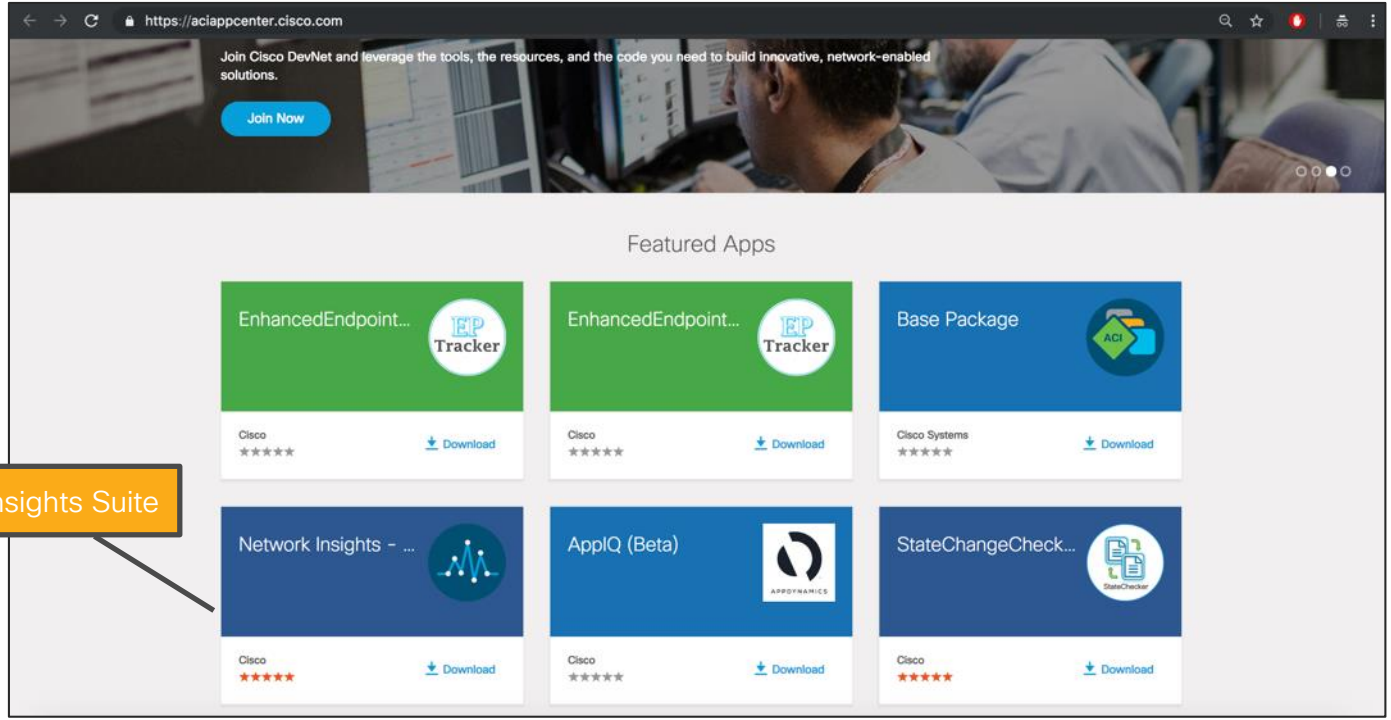
Resource Analysis

Physical/Logical Network Capacity & Utilization
Data & Control Plane & Environmental Health

Enhance Availability, Uptime & Network Wide Visibility

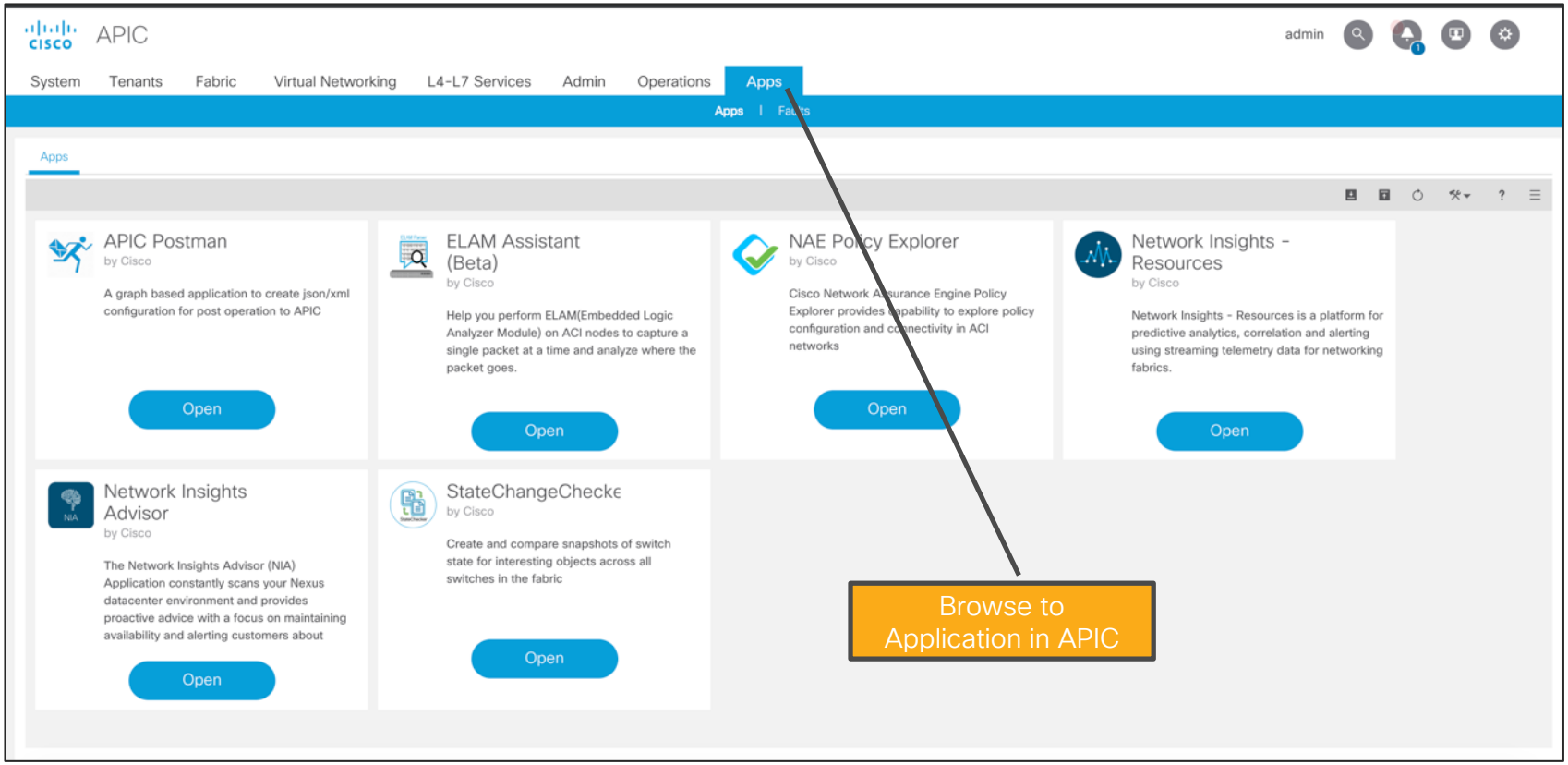
Download application from the App Store

Common app store for ACI and NXOS – https://aciappcenter.cisco.com/

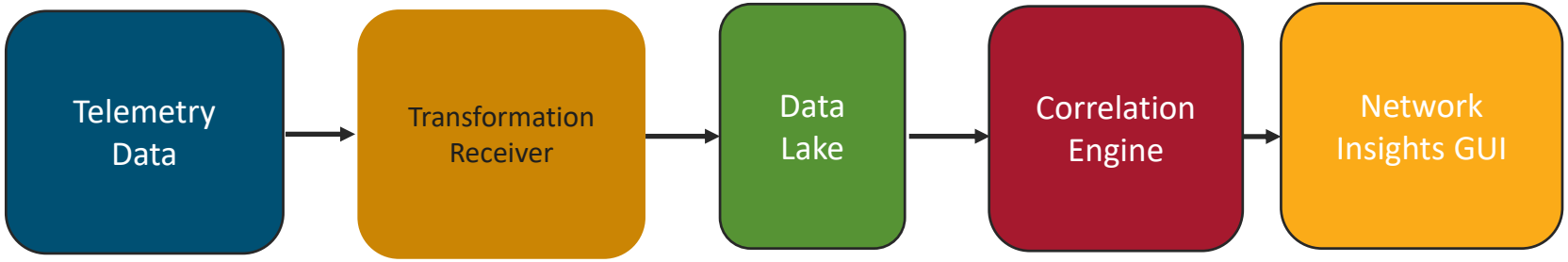


Network Insights Suite

Accessing the application from APIC



Network Insights App Architecture



- GPB over Kafka
- Flow tables over UDP
- Buffer and queue stats over UDP

Collect and Normalize Telemetry data

Store correlated data for use by GUI

Analyze

Visualize

cisco

elasticsearch

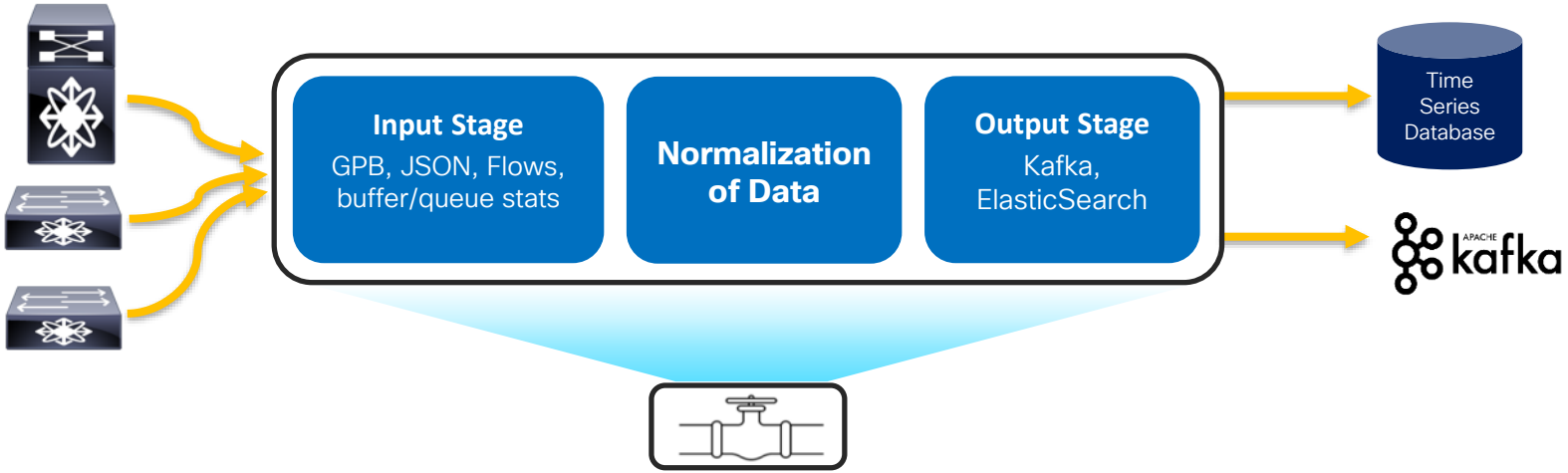
React

Microservices

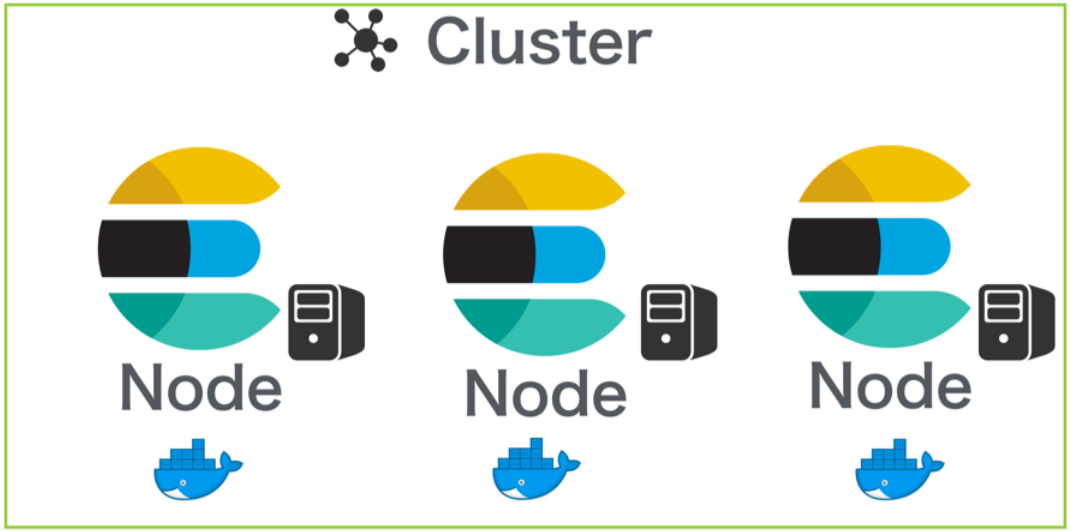
docker

Transformation Receiver

- Telemetry Gateway Platform
- Pipeline normalizes disparate telemetry inputs to a common output
- Scalable architecture leveraging container-based scale-out model



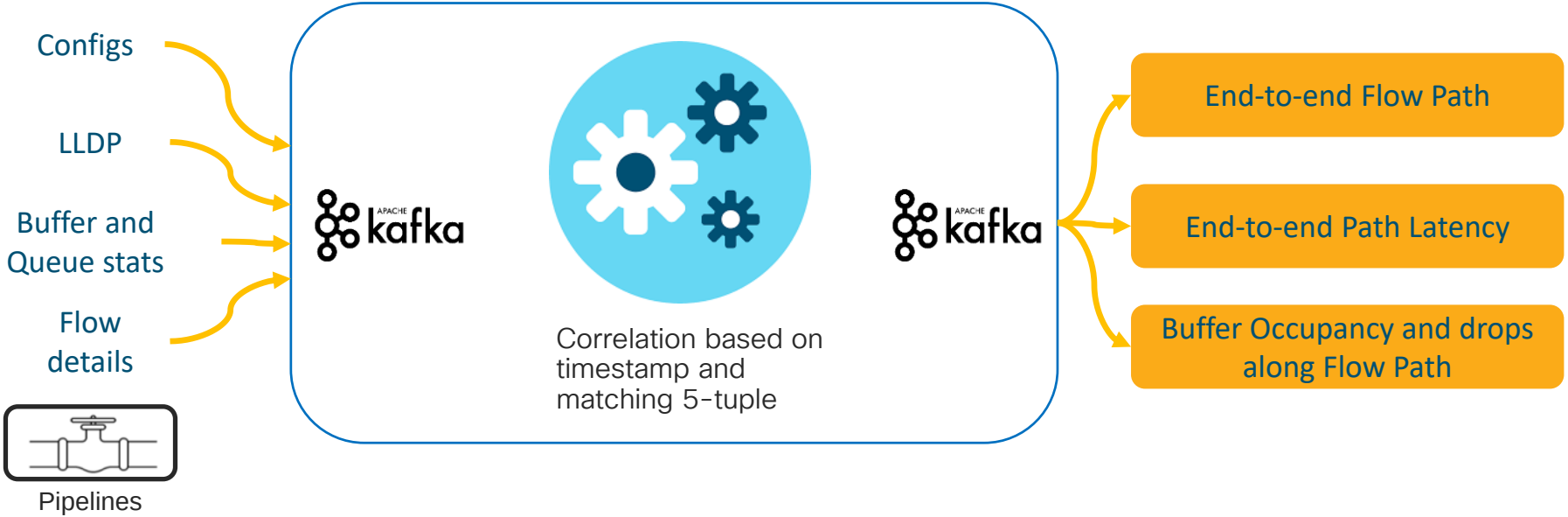
Data Lake



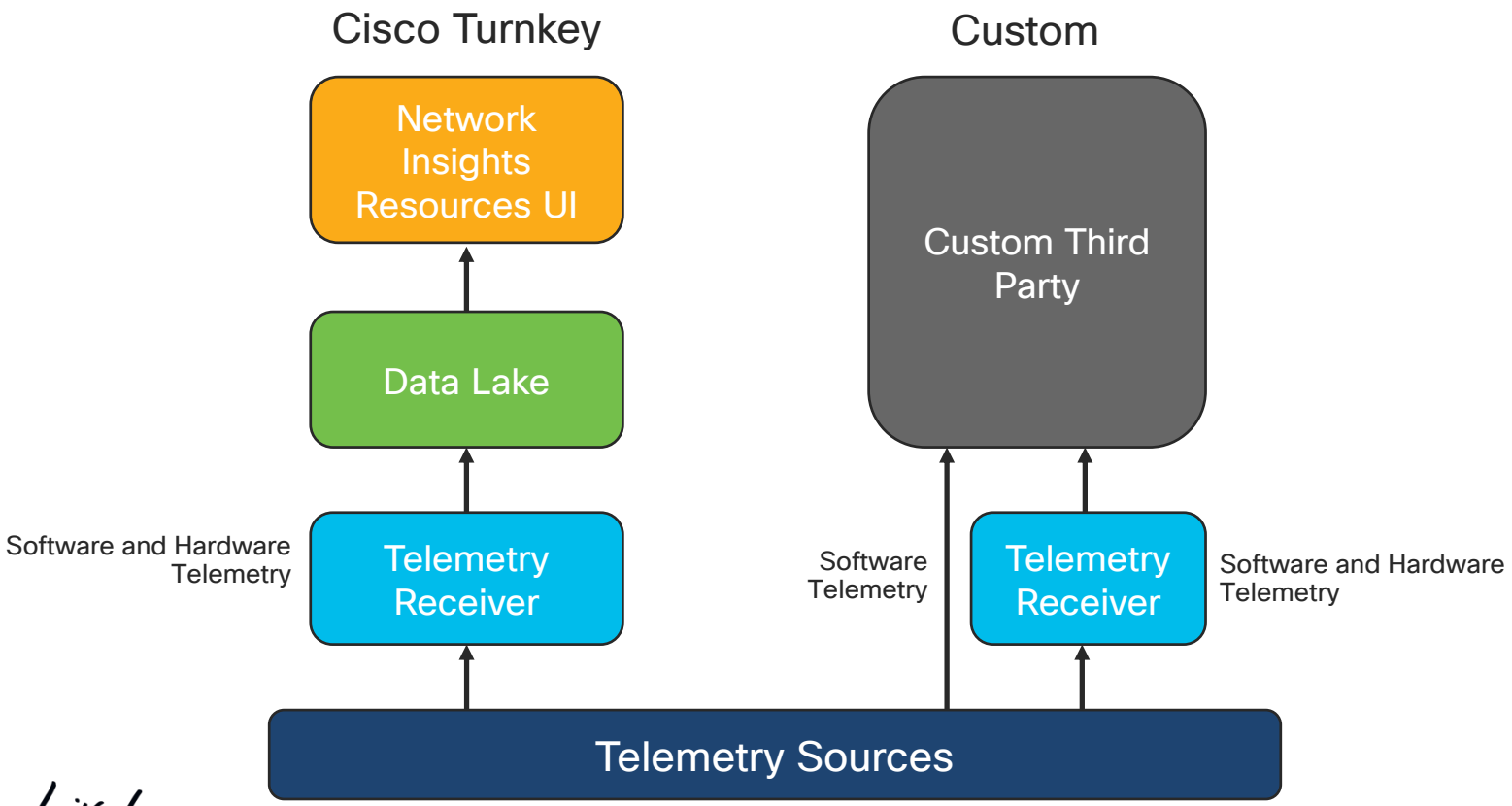
- Horizontally scalable Elasticsearch Cluster used for storing correlated data
- Up to 2 TB of storage to store 3-5 days of flow telemetry (30 days for s/w telemetry)

Correlation Engine

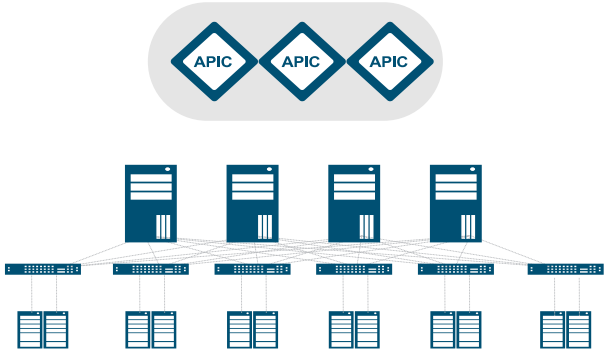
Correlate normalized telemetry data streams from Transformation Receiver



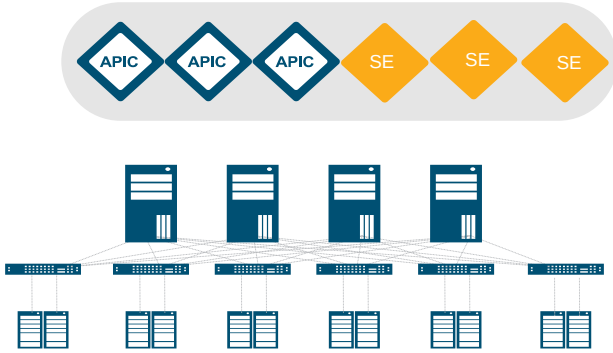
Telemetry and Analytics Deployment Models



Telemetry and Analytics Deployment Options



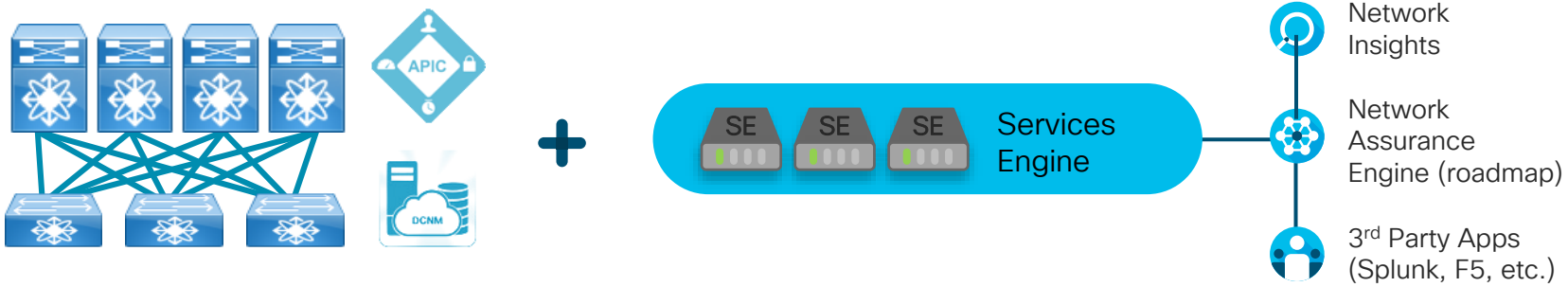
1 Use current APIC and store whatever is available



2 Use additional 3 or more Application servers (aka. Service Engines) to host dedicated Telemetry data

Introducing the Services Engine

Application Hosting Platform for APIC / DCNM



Dual Boot Option | Cluster for Redundancy

What Is Services Engine and When Is It Needed?

- Provides additional compute & storage to NIA/NIR for both APIC and DCNM
- Required to support large-scale fabrics, and for flow data collection due to increased ingest rate, correlation workload, and data consumption on disk
- Operates as a cluster of 3 compute nodes
- Can boot in “ACI mode” or “DCNM mode”

2 x 10-core 2.2GHz CPUs

256GB memory

9.6TB storage

Network Insights Resources

- Analysis and correlation of software and hardware telemetry data with focus on **network operations use-cases**
- Integrates directly into APIC with common visualizations
- Focus on **anomalies** and **quick drill-down** to specific issues

Pre-requisites for using Network Insights on APIC

- NTP needs to be configured and working for all nodes to enable software telemetry
- Inband management needs to be setup on both APIC and switch nodes
- PTP needs to be enabled in the fabric. PTP is used to sync the Leaves and Spines for Flow Analytics. GrandMaster clock is not required
- Flow Analytics must be enabled and flow rules must be configured

Forwarding of Anomalies

- Every anomaly is treated as a fault
- Every fault is written to Kafka topic
- 3rd party applications (like ServiceNow) can subscribe to these topic to retrieve the faults and process/analyze them further

Integration with External System

NIR has REST-APIs exposed to pull the data available within the UI to interface with 3rd party tools

REST-APIs are available to interact with:

- Anomalies
- Resources
- Events
- Nodes

REST-API examples

Network Insights - Resources

Network Insights - Resources

Time Range: Apr 10th 2019, 4:50 PM - Apr 10th 2019, 5:50 PM

Dashboard

System

Resource Utilization

Environmental

Operations

Dashboard

Total Nodes

0 Spines
0 Leaves

Back Forward Print Search Go

Contents

Network Insight - Resources Application

Troubleshooting NIR

Upgrading or Downgrading Cisco NIR

Network Insight - Resources Application

Network Insight - Resources Application

This chapter contains the following sections:

- About Network Insight - Resources Application
- NIR Guidelines and Limitations
- Using NIR
- NIR REST API Examples

Back Forward Print Search Go

Contents

Network Insight - Resources Application

About Network Insight - Resources Application

NIR Guidelines and Limitations

Using NIR

NIR REST API Examples

Troubleshooting NIR

Upgrading or Downgrading Cisco NIR

Network Insight - Resources Application | NIR REST API Examples

NIR REST API Examples

The following is the list of REST API examples for the Network Insight - Resources application:

- all_resources()
- anomalies_details()
- anomalies_summary()
- anomalies_top_flows()
- anomalies_top_nodes()
- anomalies_top_processes()
- events_buckets()
- events_details()
- events_summary()
- fabric_summary()
- get_nodes_list()
- get_processes_list()
- get_stats_details()
- get_stats_nodes()
- health_diagnostics()
- service_health()
- set_logging()
- sysmetrics_detail_stats()
- sysmetrics_get_top_nodes()
- sysmetrics_get_top_processes( )
- sysmetrics_summary()
- utilization_node_details()
- utilization_top_nodes()

Expand All | Collapse All

Back Forward Print Search Go

Contents

Network Insight - Resources Application

About Network Insight - Resources Application

NIR Guidelines and Limitations

Using NIR

NIR REST API Examples

all_resources()

anomalies_details()

anomalies_summary()

anomalies_top_flows()

anomalies_top_nodes()

anomalies_top_processes()

events_buckets()

events_details()

events_summary()

fabric_summary()

get_nodes_list()

get_processes_list()

get_stats_details()

get_stats_nodes()

health_diagnostics()

service_health()

set_logging()

sysmetrics_detail_stats()

sysmetrics_get_top_nodes()

sysmetrics_get_top_processes(

sysmetrics_summary()

utilization_node_details()

utilization_top_nodes()

Troubleshooting NIR

Upgrading or Downgrading Cisco NIR

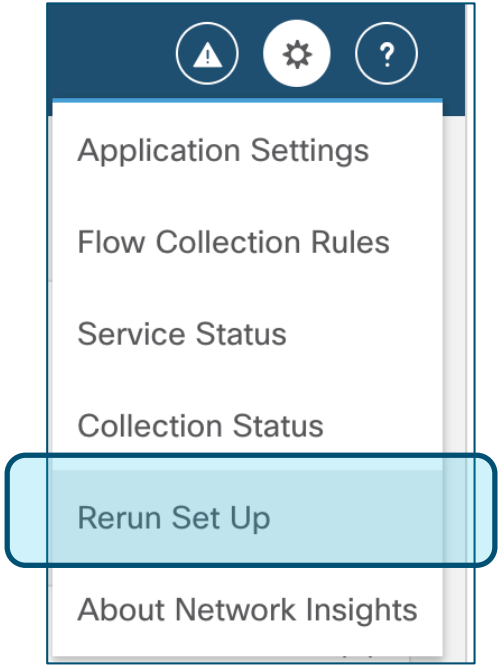
Network Insight - Resources Application | NIR REST API Examples | all_resources()

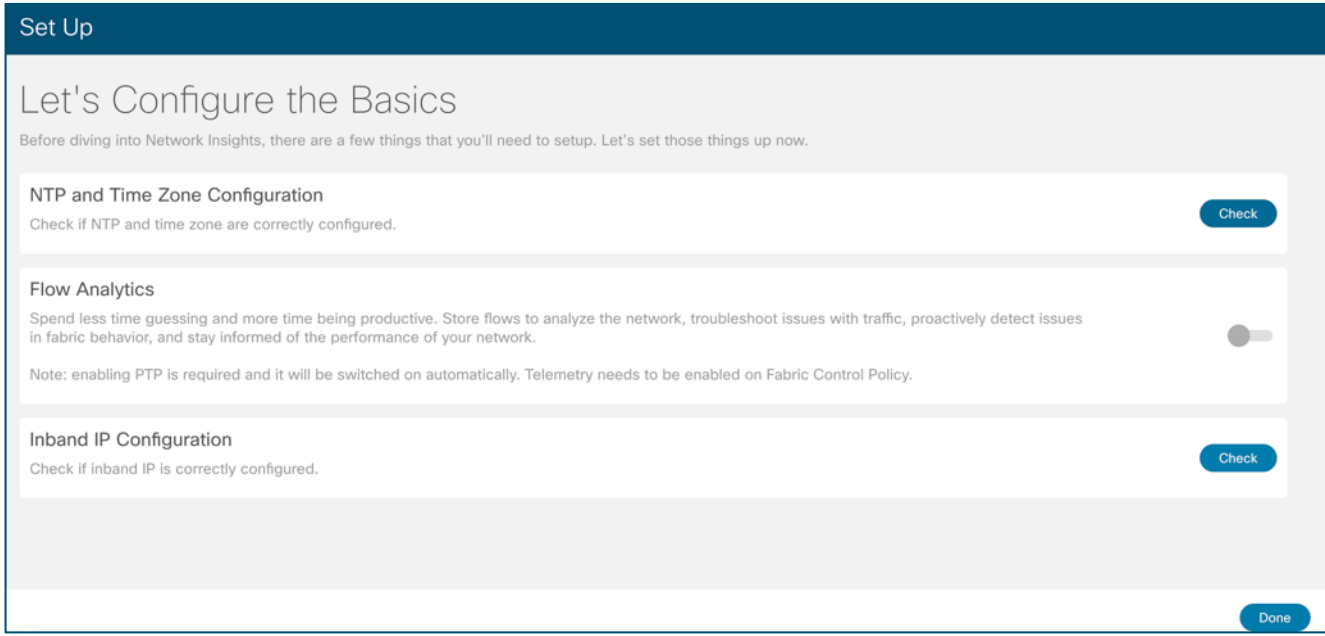
all_resources()

```
Get all resources  .
REST URL   :
   GET /api/telemetry/utilization/resources.json
Parameters :
   None
Example    :
   curl -k -i -XGET 'https://ip:port/api/telemetry/utilization/resources.json'
Response   :
   {
       "totalResultsCount": 5,
       "totalItemsCount":5,
       "entries": [
           {
               "categoryName": "",
               "resourceName": "EndPoints",
           }
           <-- SNIP LIST OF ALL OTHER RESOURCES -->
           {
           }
       ]
   }
```

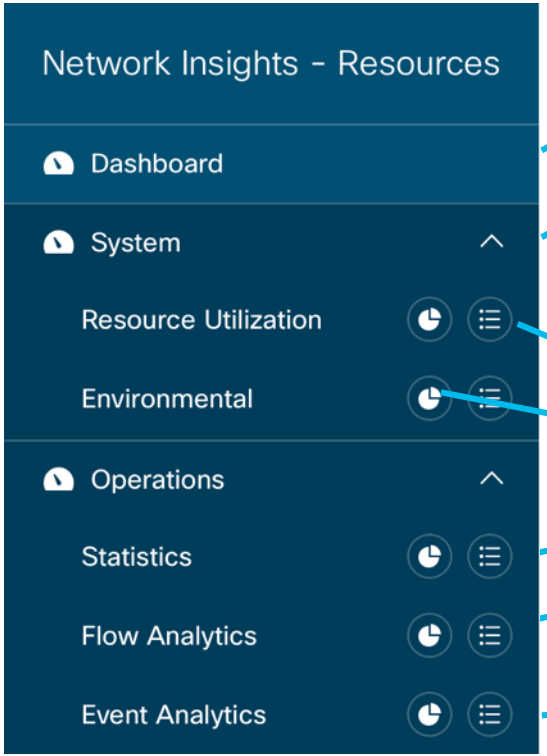
APIC has a wizard to help you get started!

Follow all the steps till the setup is finished





Network Insights Resources – Tabs



Dashboard : Summary of anomalies or unusual behavior

System : Software telemetry data

- Utilization, trends and anomalies around operational, config and hardware resources
- Utilization, trends and anomalies around environmental

Browse view for details

Dashboard view for quick view

Statistics : Interface counters, LLDP, CDP and ISIS errors

Flow analytics : End to end flow details, drop reasons – directed flow monitoring of 10k flows/sec

Event analytics : Audit logs, Events, Faults

Network Insight Resources – What you get

System (shipping)

- Resource Utilization [Fabric Wide]
  - Trend Monitoring (rising/falling), anomaly detection, prediction and alerting
  - Fabric Capacity
    - Endpoints, Bridge Domains, VNIs, Virtualization ratio.
    - Per device stats.
  - Environmental
    - CPU Power, Storage, Fan, Memory.
    - Side-by-side analysis

Operations

- Statistics **(shipping)**
  - Protocol Stats (Errors, ingress/egress..)
  - Interface Stats (Utilization, CRC, FW..)
  - Error Trends, **anomaly detection, prediction and alerting**
  - Bandwidth
- Flow Analytics (shipping)
  - **Anomaly detection**
  - **Path Tracing & Latency**
  - **MAC moves**
- Flow Analytics **(Phase 2)**
  - Microburst Detection
- Event Analytics **(shipping)**
  - Audit and Correlation
  - Hot-Spot / Congestion Monitoring

Resource Utilization Dashboard

Time Range: Mar 13th 2019, 6:13 PM - Mar 13th 2019, 7:13 PM

Resource Utilization Dashboard

APIC Capacity

ACI

Fabric wide utilization – current versus maximum allowed

Top nodes by resources and which resource within the node is contributing to the node being a top node

Bridge Domains 28 of 15000	1% used	
Contracts 13 of 10000	1%	
Endpoint Groups	1%	
Hypervisors		
IP Endpoints 28 of 225000	1% used	
L3 Contexts 16 of 3000		
Proxy Database Entries 55 of 450000	1% used	
Tenants 7 of 3000		

Top Nodes by Capacity

Leaf106	
EPG	3751 of 3960
BD	2901 of 3500
LPM	30%
IPV4 (learned)	7 of 24K
Port Usage	16 of 54 ports

Leaf107	
IPV4 (learned)	15 of 24K
LPM	29%
IPV6 Host Routes	3 of 12K
Port Usage	15 of 54 ports
Policy TCAM	17 of 64K

Leaf105	
Multicast Routes	6 of 8K
IPV4 (learned)	6 of 24K
Port Bandwidth	192 of 720 Gbps
IPV6 (learned)	4 of 12K
VRF	223 of 800

Leaf108	
LPM	30%
Policy TCAM	19 of 64K
VLAN	1148 of 3960
IPV6 Host Routes	3 of 12K
Multicast Routes	2 of 8K

Spine1	
Port Usage	6 of 66 ports

Spine2	
Port Usage	5 of 66 ports

Resource Utilization Browse View

Anomaly score accounts for anomalies as seen in dashboard

Sort nodes by operational, config and hardware resources

Utilization and trend of operational resources

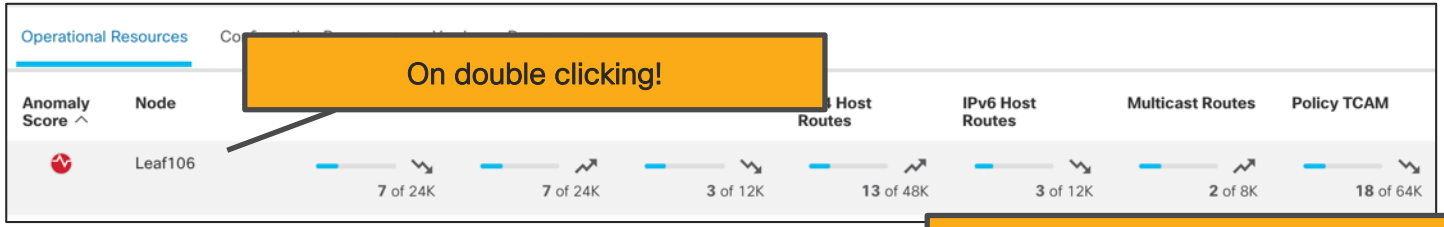
Utilization and trend of configuration resources

Utilization and trend of hardware resources – how many ports are admin up and how much bandwidth is being used

cisco Live!

On double clicking a line item..

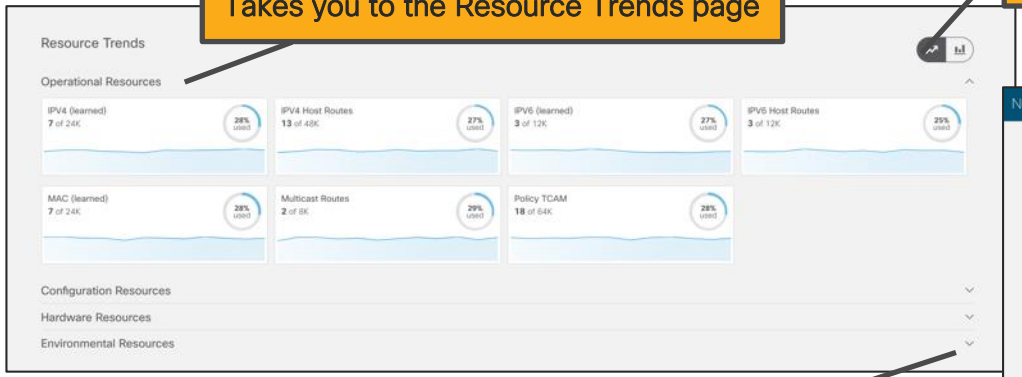
Applicable to all the resources



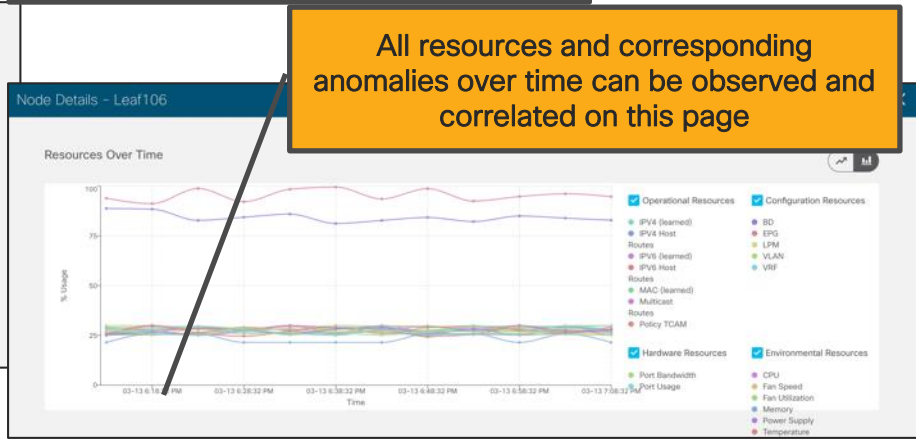
On double clicking!

The icon represents a time graph

Takes you to the Resource Trends page



All resources and corresponding anomalies over time can be observed and correlated on this page



Trends across resources can be seen here

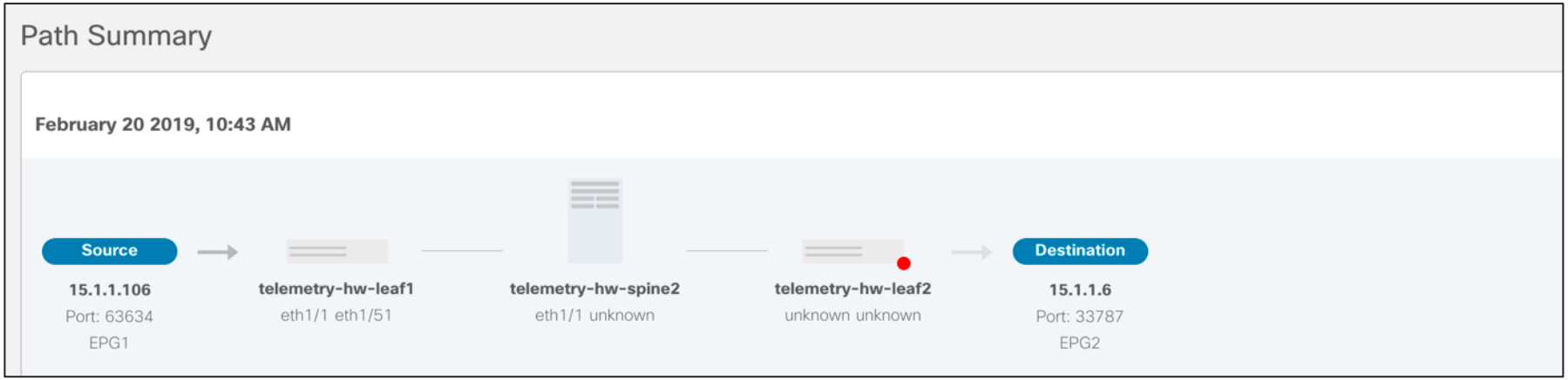
Resource Analysis – Flow Analytics

ACI

Proactive Anomaly Detection for ACI Deployments

Targeted Flow Monitoring Use Cases –

- Application Performance Issues:
  - Forwarding/policy Drops indicating congestion
  - High end to end application latency
- Application Downtime Event –
  - Policy misconfiguration due to ACL's



The intent – Network Insights and Cloud ACI

Network Insights - Resources

Flow Details from 50.4.0.3 to 50.10.0.2

Dec 1, 2019, 1:00PM - Dec 5, 2019, 1:00PM

Page 1 of 1 1-1 of 1

Path Summary

Site: us-east-1 aws

Site: San Francisco On-Prem ACI

Source
50.4.0.3
Port: 5444
EPG: epg1

FW
ASA-Firewall
Service Graph: Service-Graph

LB
F5 Load Balancer
Service Graph: Service-Graph

VGW
vgw1
VPC: vpc-1
Region: us-east-1

CSR
csr1
VPC: vpc-1
Region: us-east-1

AWS Direct Connect

SF-leaf2
eth1/35 - eth1/2

SF-spine1
eth5/30 - eth1/11

SF-leaf1
eth1/21 - eth1/1
View Events

Destination
50.10.0.2
Port: 5543
VLAN: -

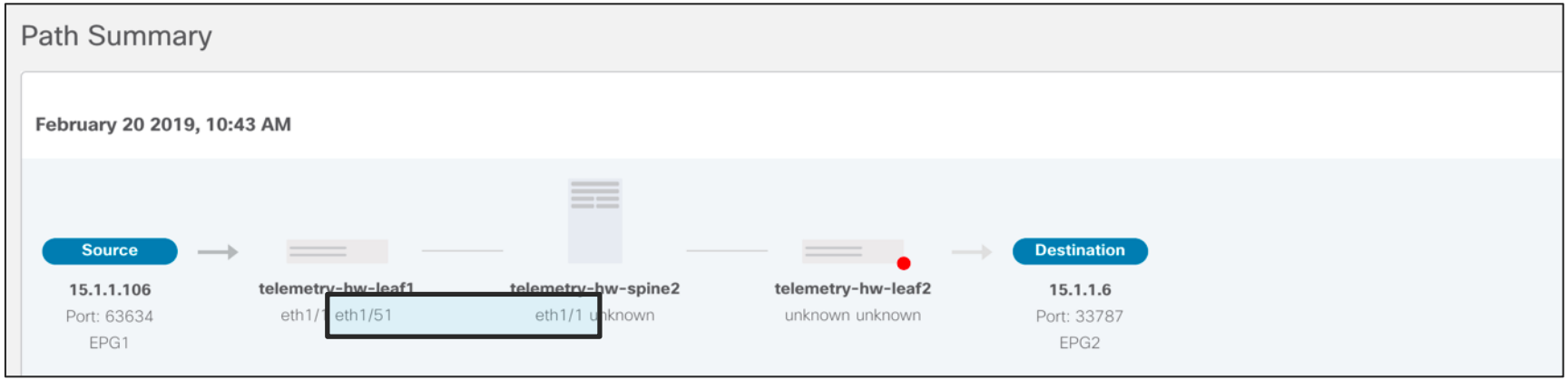
View reverse path

Average Latency By : Trending

Burst By: Trending

When Spine doesn’t support FT

Even if Spine doesn’t support FT (for example Nexus 9364 as Spine), Ingress Interface is gathered via LLDP between Ingress Leaf and Spine. The flow will look like the below when exported to NIR -

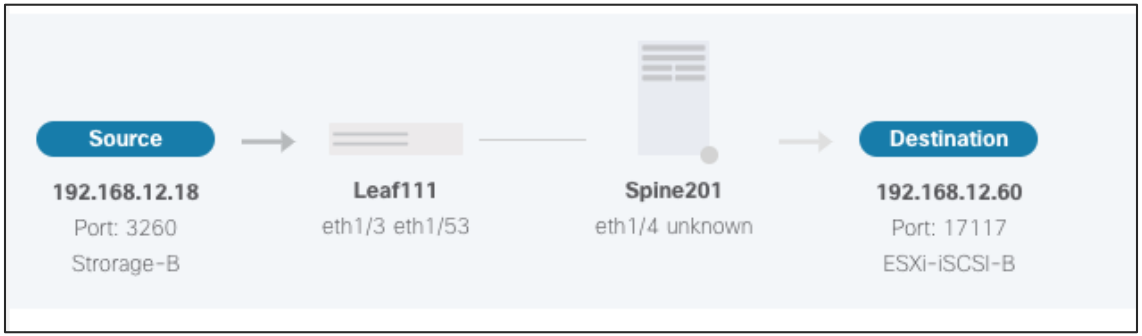


Limitation of FX/FX2

FX/FX2 can not determine the ingress port, so it is expected to see ‘unknown’ ingress port on egress leaf. The flow will look like the below when exported to NIR -

Mixed Fabric with FT and non-FT capable Devices

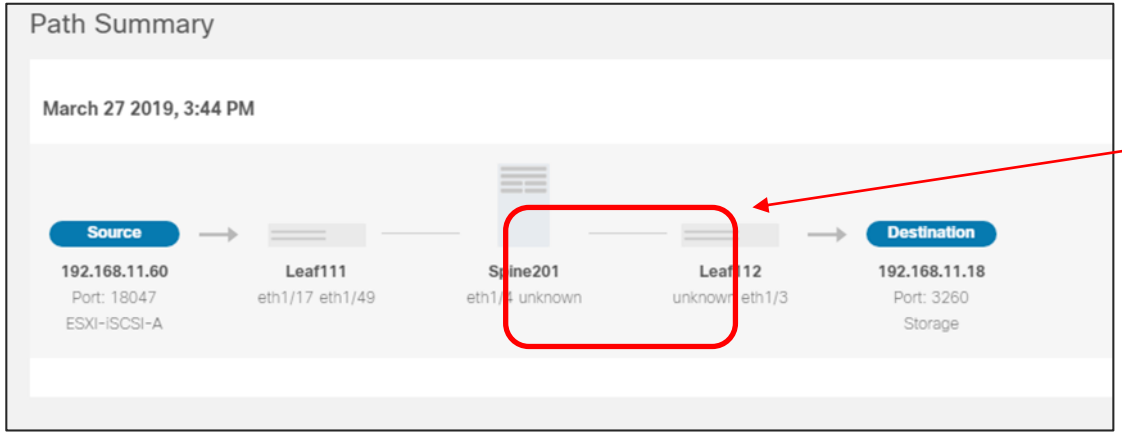
EP 192.168.12.60 is attached to non-FT capable leaf. The flow will look like the below when exported to NIR -



Multipod Support

Multipod is supported. Following limitations and guidelines apply:

- IPN is not visible in path trace
- Spine egress interface to IPN is only visible if spine is FT capable
- Spine in the other POD is only visible in pathtrace if it is FT capable



Both spines are not FT capable in this MPOD scenario. So no egress interface information nor spine in other POD is visible

Directed flow monitoring – Supported in Nexus 9K-FX/EX and FX2 in NIR 2.1

- Phase1 will support 10k flows/sec across the fabric
- User must define flow rules in APIC after which only those flows will be pushed to NIR for flow analysis

User defined flow rules

ACI

Settings -> Flow Collection Rules

Time Range: Mar 10th 2019, 12:40 PM - Mar 17th 2019, 12:40 PM

Flow Analytics Dashboard (Limited Availability)

Application Settings
Flow Collection Rules
Service Status

Top 10 Nodes by Average Latency

Latency (µs)

VRF Based Rules

Add a rule

Flow Collection Rules

Define parameters – Name, Tenant, VRF, Subnets

VRF Based Rules

Name	Tenant	VRF
Demo		
NewTestRule	mgmt	inb
vEPC	vEPC	VRF

Demo

Rule will be applied to all relevant switches.

Subnets

Subnet: 10.10.10.0/24 Add Subnet

Cancel Save

Flow Analytics Dashboard

ACI

Each chart is clickable to drill down on details

Flow Analytics Dashboard (Limited Availability)

Top 10 nodes by Average Latency of flows – useful to detect unusual latency

Top 10 Nodes by Average Latency

Average Latency (µs)

8
6
4
2
0

03-08 4:11:00 PM
03-09 7:03:00 PM
03-10 9:55:00 PM
03-12 12:47:00 AM
03-13 3:39:00 AM
03-14 6:31:00 AM
03-15 9:23:00 AM

Spine1
Leaf101
Leaf104
Spine2

Top 10 nodes by packet drops – Quick drill down on packet drops per node

Top 10 nodes by End Point Moves – track EP moves

Top 10 Nodes by EP Move Indicator

EP Move Indicator

320
240
160
80
0

03-11 6:04:00 PM
03-15 9:23:00 AM

Time (403m interval)

Leaf101
Leaf104
Spine1
Spine2

Top 10 Nodes by Packet Drop Indicator

Packet Drop Indicator

1.2e+6
900000
600000
300000
0

03-11 6:04:00 PM
03-15 9:23:00 AM

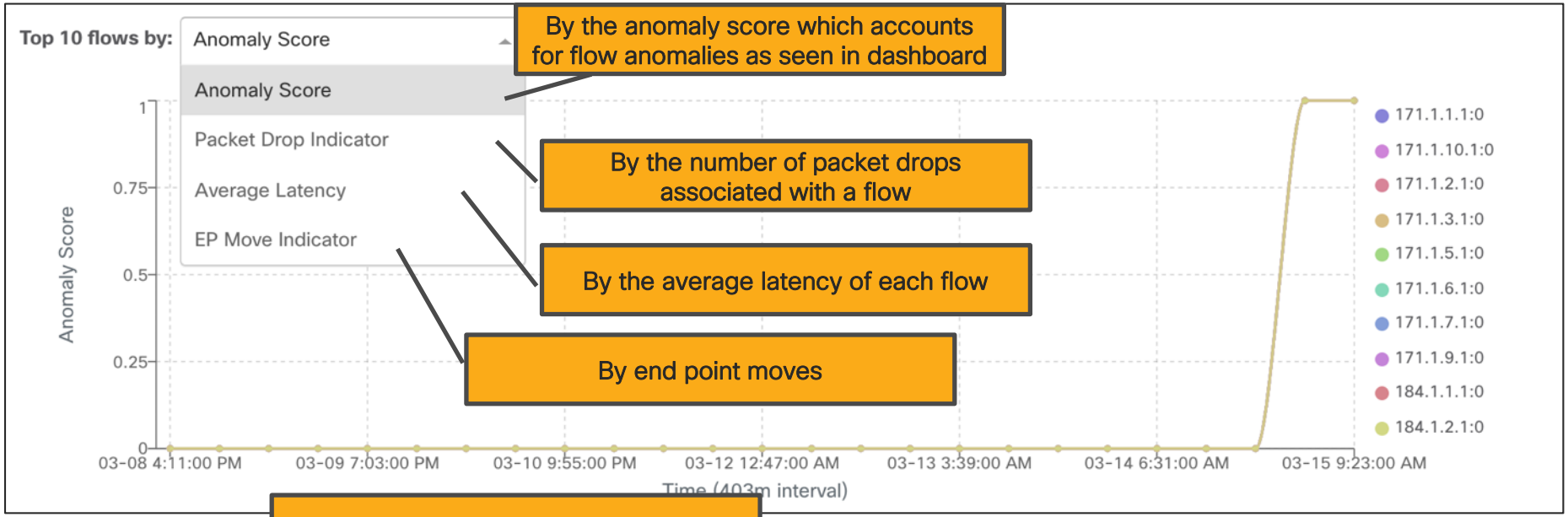
Time (403m interval)

Leaf101
Leaf104
Spine1
Spine2

Flow Analytics Browse View

Sort top 10 flows over a selected period of time

ACI

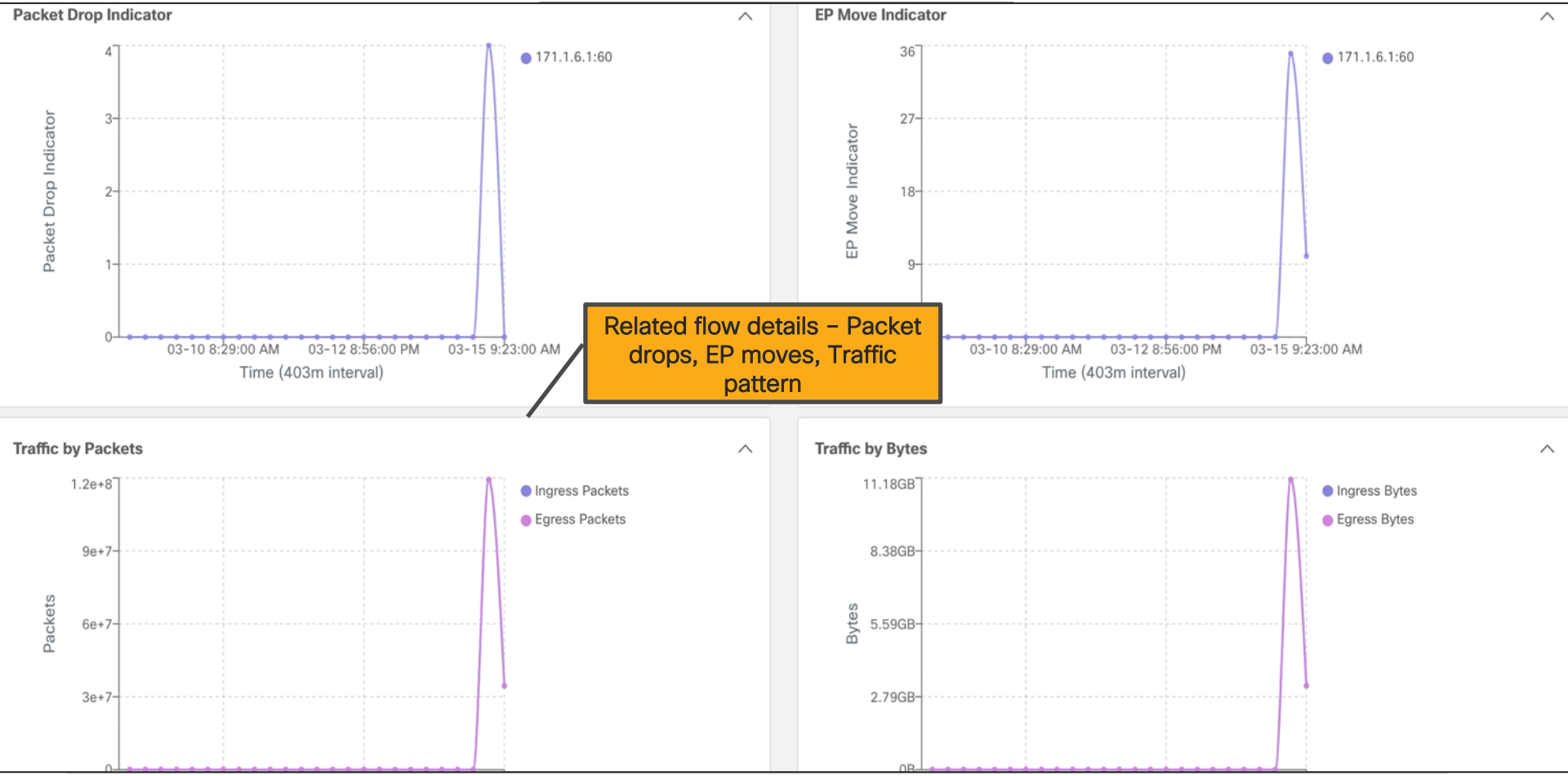


Populates all the flows seen by the app

							Source			Destination	
Anomaly Score	Origination Timestamp	Nodes	Ingress Nodes	Egress Nodes	Ingress Tenant	Egress Tenant	EPG	Address	Port	EPG	Address
	Mar 15 2019 11:33:31am	Leaf101, Leaf104, Spine1	Leaf101	Leaf104	vEPC	vEPC	PGW	171.1.1.1	60	Ext-EPG	174.1.1.

On double clicking a flow in the browse view

ACI



Related flow details – Packet drops, EP moves, Traffic pattern

Network Insights Advisor (NIA)

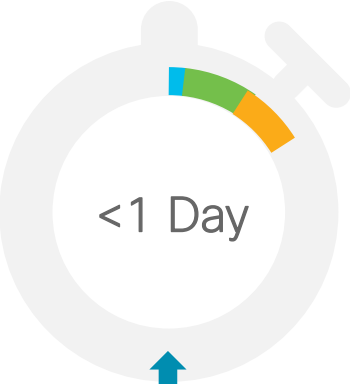
Reduce Downtime/Outages

Before Network Insights Advisor

After Network Insights Advisor

Time to resolve issue in Hours

	Notice/Anomaly Detection	4
360	Case Accepted, Outputs Attached	2
240	Tech Support Analysis	3
120	Back and Forth Communication	
120	Remediation, Close case	1



Downtime/Outages to the Network cost **Millions**

Success metrics

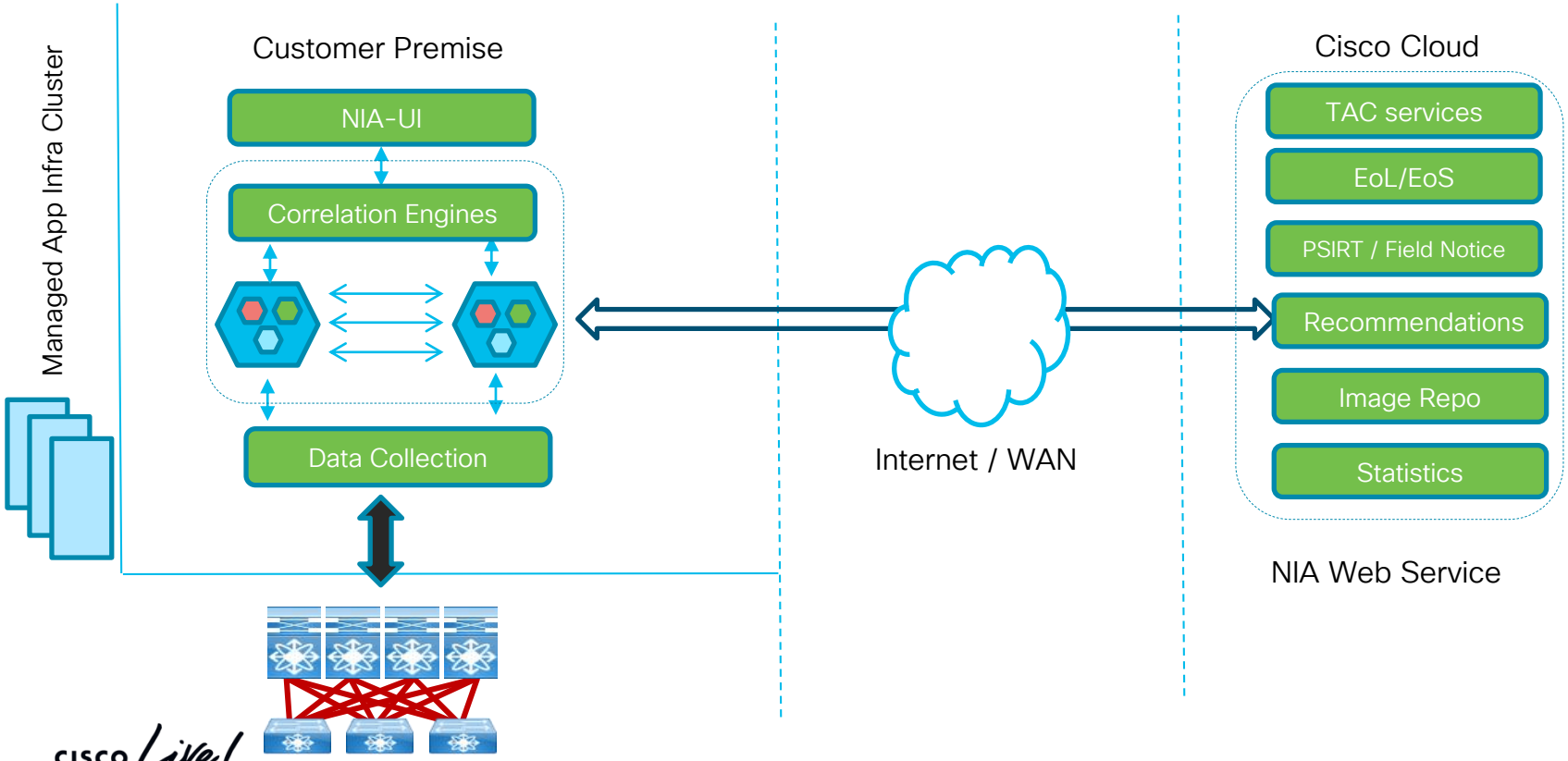
- NIA immediately flags anomalies and optimizes your network
- NIA helps prevent downtimes/outages
- Significant OPEX, CAPEX and time savings

Network Insights Advisor – Customer Benefits

Network Insights Advisor

Advisories	Software/Hardware Recommendations Workarounds	Avoid multiple TAC calls
Notices	EOL/EOS Field Notices SMUs	Keep Network up to date Adhere to Cisco policies Recommendations
Anomalies	Known Bugs/PSIRTs Unknown runtime Config anomalies	Remove Complexity Avoid Outages Faster Deployment times
Compliance	Version Scale Limits/Hardening Check Configuration	Significant CAPEX And OPEX Savings
Diagnostic	Forwarding State Check Network Anomaly (Loops) Detection	Prevent traffic black holing Avoid downtimes

NIA High Level Architecture

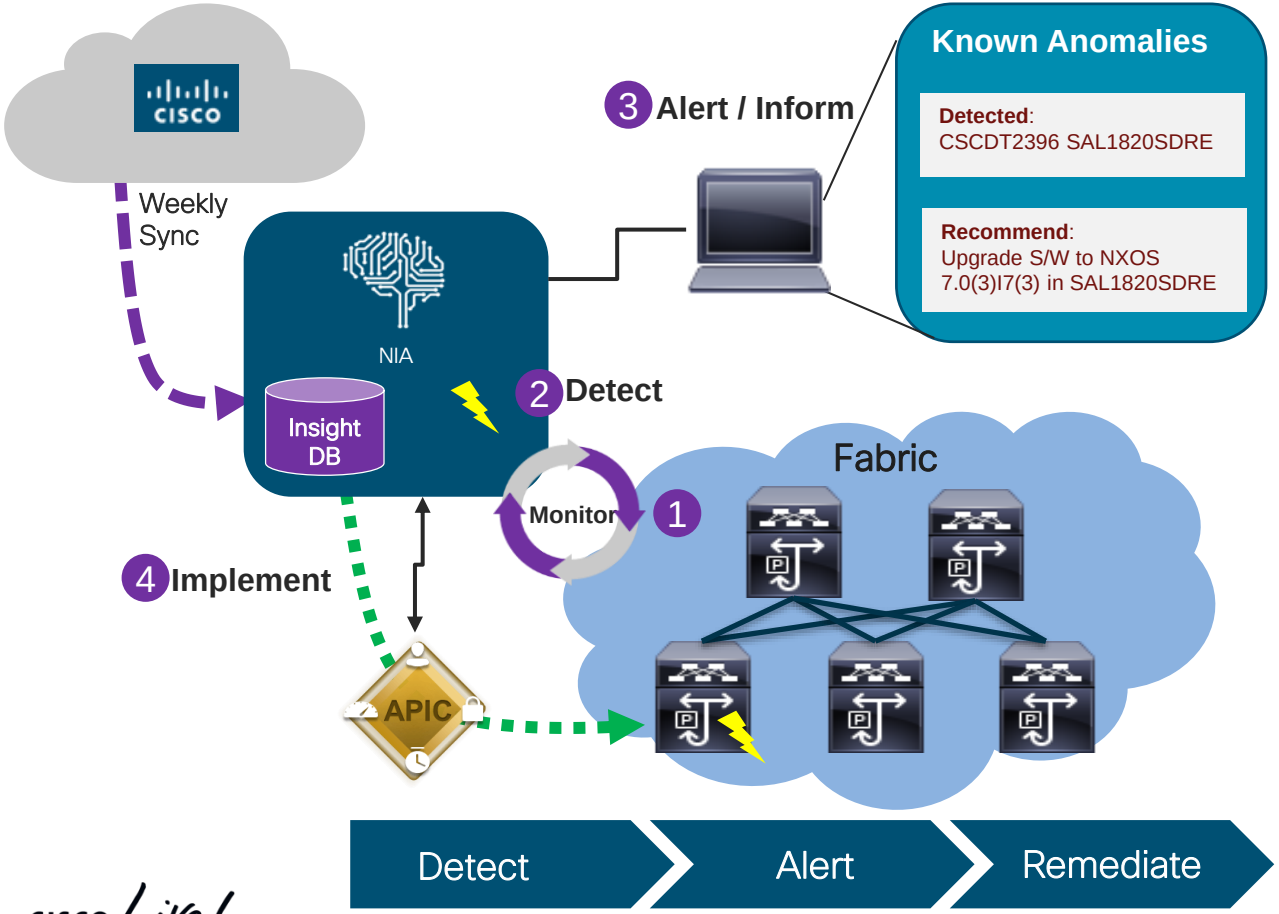


Network Insights Advisor Architecture

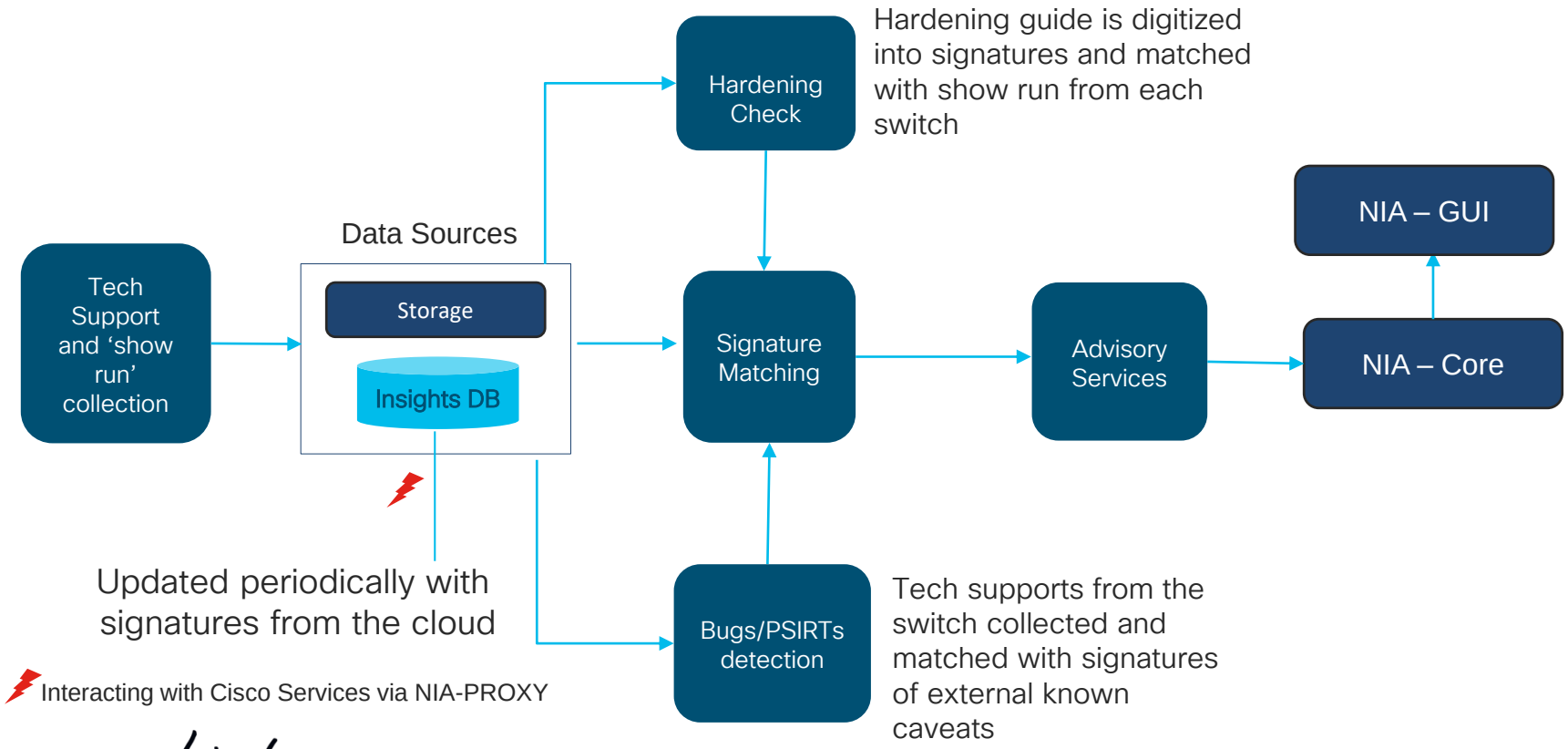
- NIA application resides on-premise within the controller and is based on a micro-services architecture. It collects information relevant information from the switches (tech-support/running config) and processes it for further analysis
- NIA also communicates with Cisco Cloud services periodically to get the latest signatures related to known bugs, field notices, EOL/EOS information on SW and HW
- NIA requires secure internet access from APIC to be able to interface with the cloud services
- From NIA TAC assist, users can collect logs in a timely manner which can be attached to an SR
- All the updates to the app will be available from within the app store and don't require controller or switch upgrade

Note: No user data will be sent to Cloud

Use Case – Notify About Anomalies

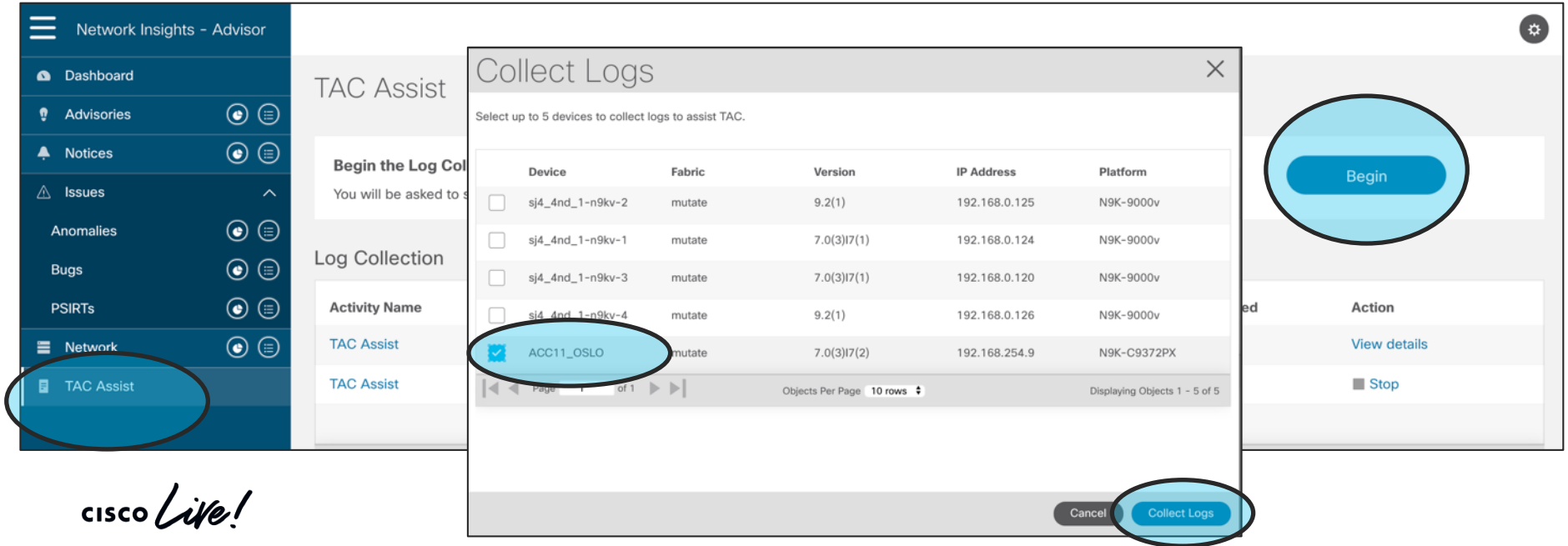


Network Insights issue detection

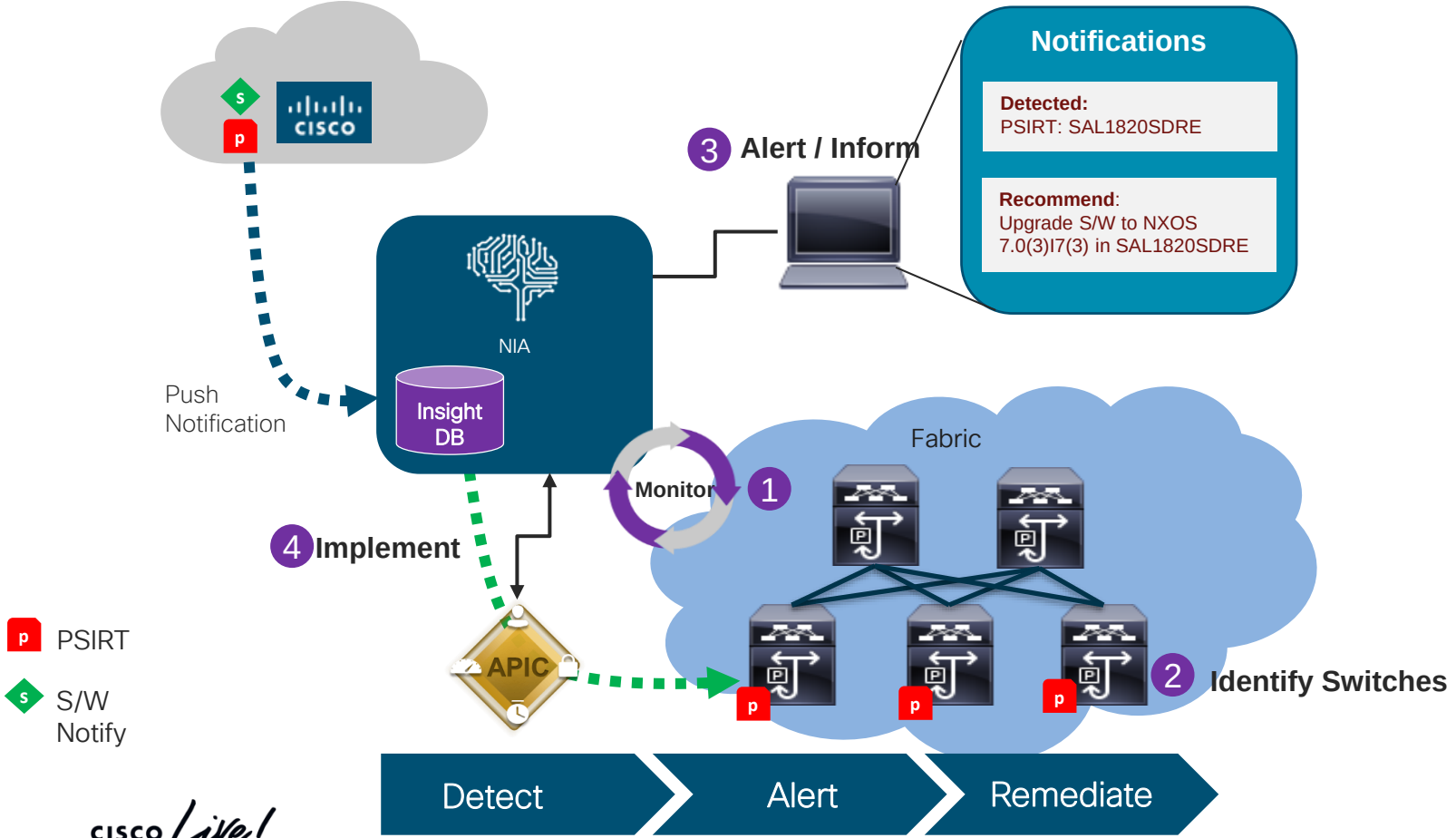


TAC ASSIST

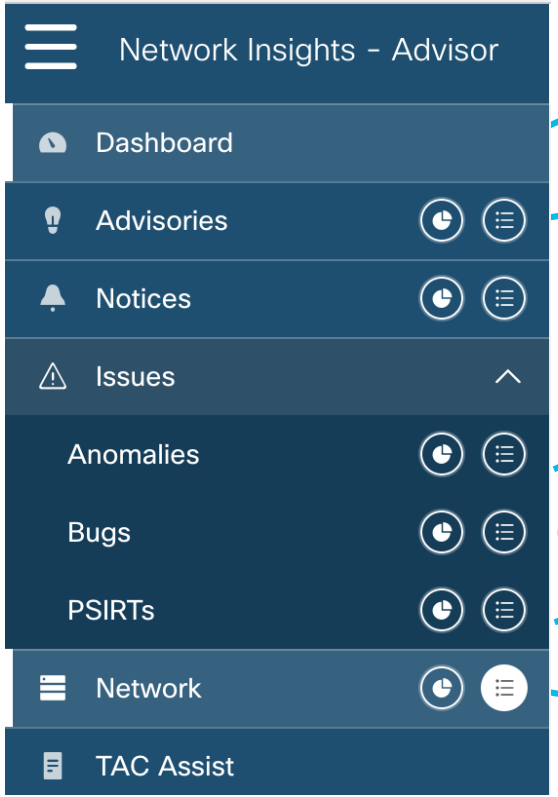
Helps with collecting logs when a user notices abnormal behavior in the fabric. These logs can then be attached to an SR



Use Case – Notify Me About New Releases



Network Insights Advisor Targeted Use Cases
Proactive supportability insights



Dashboard ”Give me a summary of issues”

Advisories
Provides advisories based on anomalies, bugs, PSIRTs and field notices. Measure upgrade impact

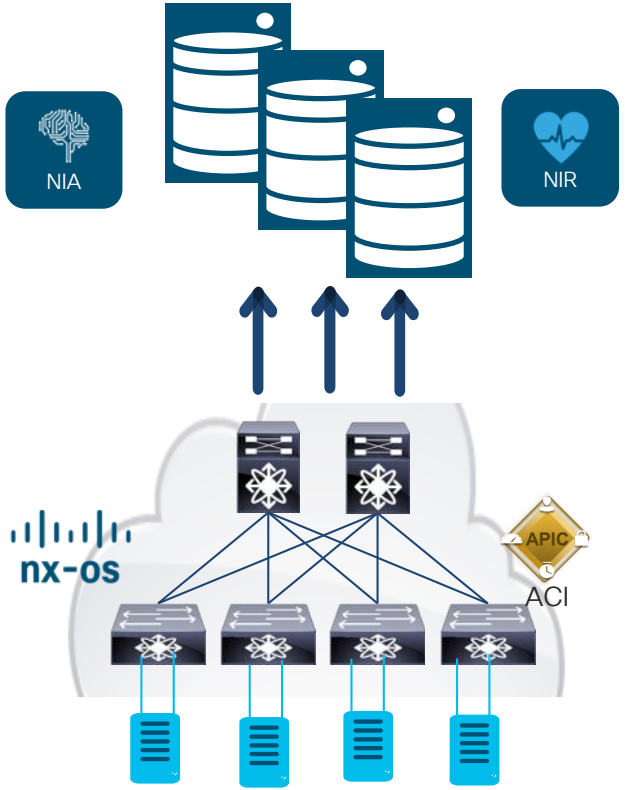
Anomalies
hardening checks, scale checks

Bugs and PSIRTs
Known bugs and vulnerabilities in the system

Fabric wide analysis

Network Insights Solution Dependencies

Targeted Scale, ACI / DCNM / NX-OS Versions



Compute Footprint*

- Two Options - Node Server Appliance or OVA Cluster
- 3 Compute nodes each with **Memory**:192GB, 2.4 TB Hard Drive, 400GB SSD, 32 vCPUs

Fabric Scale & Platforms*

- Fabric Scale up to 300 switches target at FCS
- Multi fabric support, up to 300 switches
- Directed Flow Monitoring up to 10k Flows/s
- Streaming intervals: 30sec S/S, 1sec H/W

APIC / DCNM & NX-OS Target Versions*

- APIC / ACI Minimum Release 4.1 onwards
- NX-OS 7.0(3)i7(6) onwards
- DCNM 11.3 (Aragon MR2) onwards

*under evaluation / pre engineering commit

Demo 4: ACI Operations

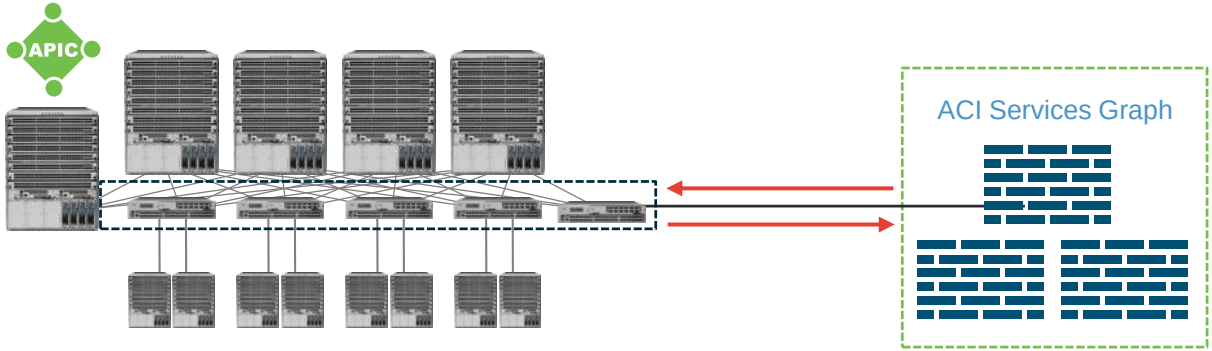
ACI Operations – Agenda

- Before getting started – setting the concepts stage

Visibility	Insights	Actions
Faults, events, stats, health, logs, trails	Application dependency	Incident Troubleshooting
Configuration	Containers Integration	Change Management
Capacity, Fabric metrics (utilization, flows, states, environmental, etc.), Telemetry	Anomalies detection (via SW & HW correlation) Trends	Increase Performance, Availability & Reliability Prevent Outages
Security Intent, Policy	Network Isolation Network Modeling	Segmentation Assurance

ACI Security

Automated Security With Built In Multi-Tenancy



APIC Hardening - Cent OS 7.2

Line Rate Security Enforcement

Open: Integrate Any Security Device

PCI, FIPS, CC, UC-APL, USG-v6

Embedded Security | Micro-Segmentation | Security Automation | Encryption | Analytics

ACI Security Certifications – by Jun/2018

Certification	ACI
PCI Security Standards Council	Certified
UC APL	Certified
FIPS 140-2 FIPS CRYPTOGRAPHY	Certified
Common Criteria	Certified
Vulnerability Scanners • Nessus, Fuzzing, etc … • Port Scan, AppScan	Certified (Ran every release)

ACI Hardening- Every Software Release

Flooding Attacks
SYN-FLOOD: Remain stable during SYN flooding attack
EST-FLOOD: Remain stable during ESTABLISHED flooding attack
LASTACK-FLOOD: Remain stable during LASTACK flooding attack
FINWAIT-FLOOD: Remain stable during FINWAIT flooding attack
CLOSING-FLOOD: Remain stable during CLOSING flooding attack
Port and Service Scans
DEF-CRED: No default authentication credentials
RECON-PORT-TCP: Remain stable during TCP port scan
RECON-PORT-UDP: Remain stable during UDP port scan
RECON-OSID: Remain stable during OS Fingerprinting
RECON-IP-PROT: Remain stable during IP protocol scan
NESSUS-SCAN: Known vulnerability scanner- Nessus
WEB-DEFECT: Known webserver and application defects
WEB-ID: Remain stable during web fingerprinting
Fuzzing
ESIC: UUT must endure malformed Ethernet packets
ICMPSIC: UUT must endure malformed ICMP packets
ISIC: UUT must endure malformed IPv4 packets
TCPSIC: UUT must endure malformed TCP packets
UDPSIC: UUT must endure malformed UDP packets
ICMPSIC6: UUT must endure malformed ICMPv6 packets
ISIC6: UUT must endure malformed IPv6 packets
TCPSIC6: UUT must endure malformed TCP over IPv6 packets
UDPSIC6: UUT must endure malformed UDP over IPv6 packets
Web Scan
Nexpose
IBM AppScan
OpenVas

Platform	Hardening
APIC+N9k	✓
ACI Multisite	✓
AVE	✓
Telemetry	✓

ACI 1st Hop Security Enhancements

Since ACI 3.2

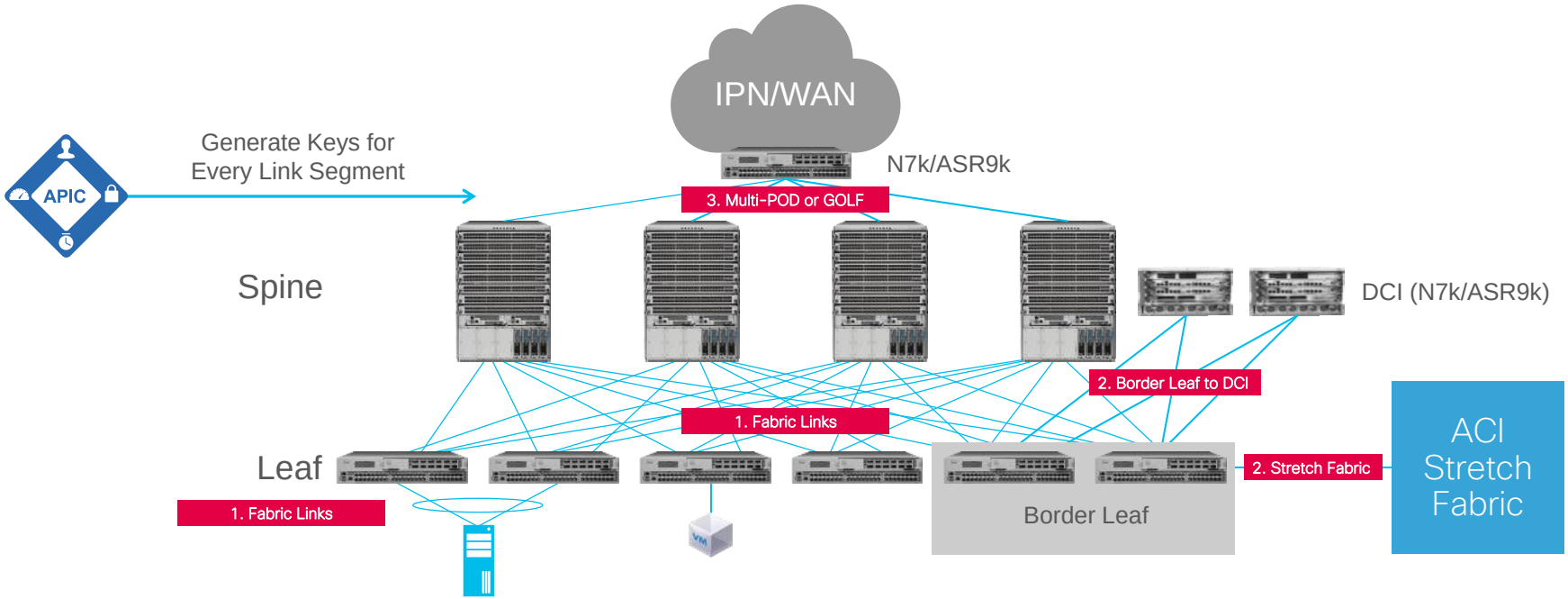
- Theft detection for VXLAN based endpoints
- Coexistence with multiple VPC interfaces
- IP Inspect feature support for IPv4 Local Endpoints
- IP Source Guard feature support for IPv4 Local Endpoints
- IPv4 Static endpoint configuration push to Cisco AVS

Caveats

- All FHS are always enabled for v4 and v6 together, HW cannot support each address family v4 or v6 separately
- FHS disables HW learning of endpoints in all ASIC, hence both v4 and v6 must be enabled together
- IP-Source Guard is enabled for v4 and v6 together (EX/FX leaf switches)
- In future ASIC, we will support FHS enablement for v4 or v6 separately

ACI MACSec

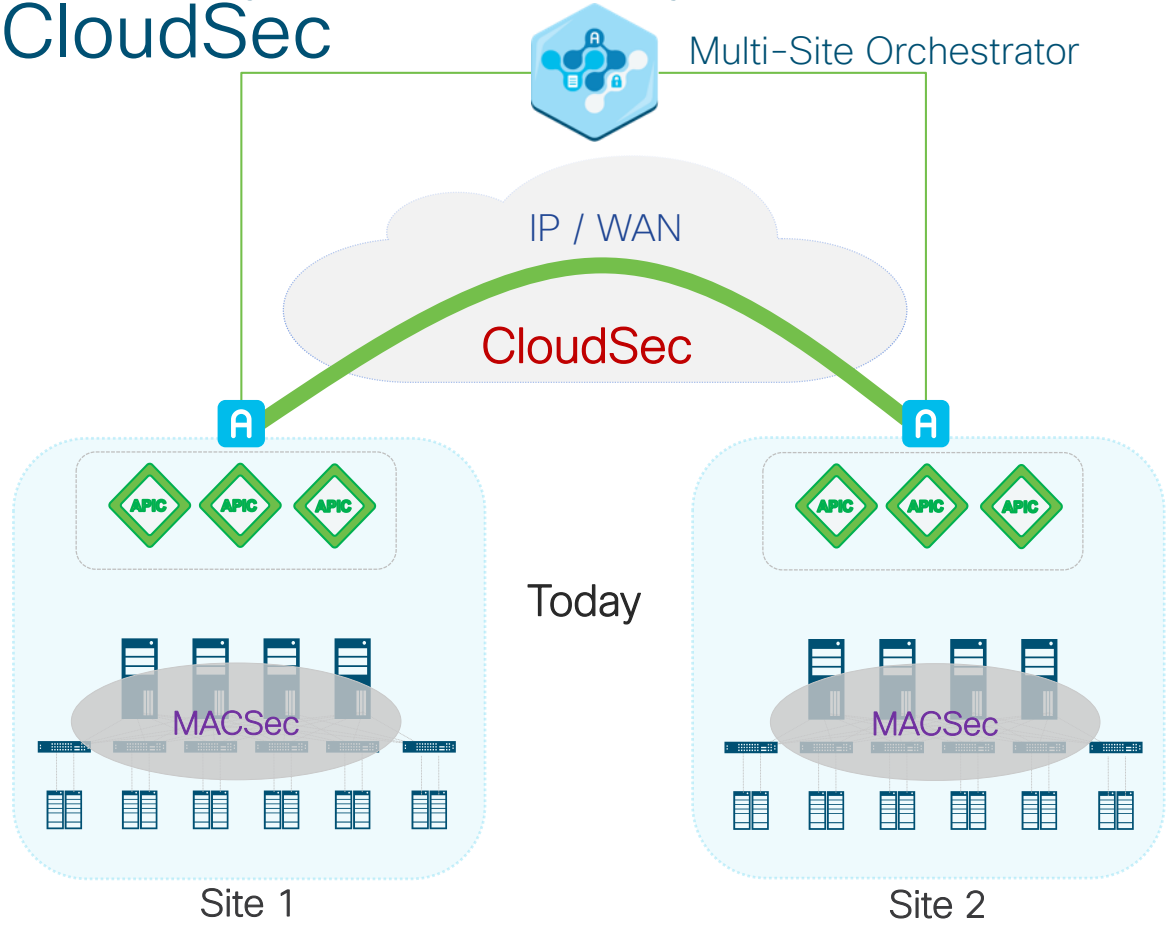
Since ACI 3.1



MACSEC Link Encryption
MKA Key Exchange

APIC Centralized Key
Management

ACI Anywhere Encrypted DCI Connectivity – CloudSec



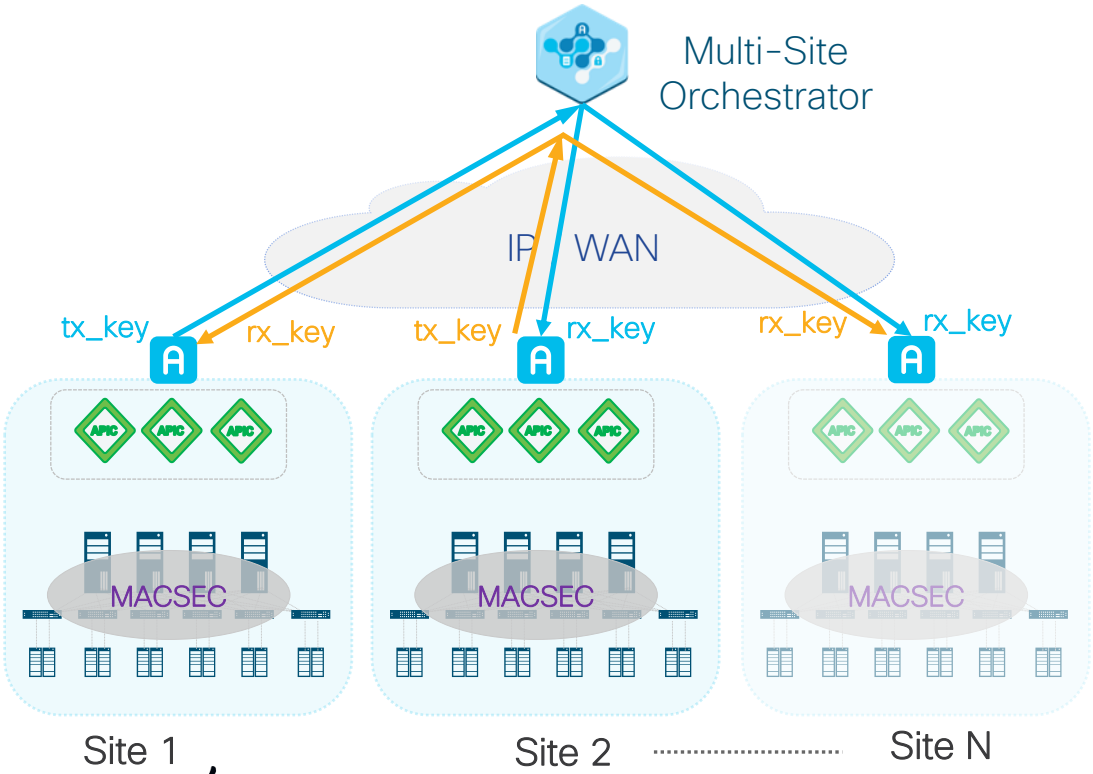
Encrypted VXLAN Overlay for Inter-Site Traffic

MKA Key Exchange over BGP-EVPN Protocol

Supported Hardware:
X9736C-FX LC
Nexus 9364C & 9332C
Nexus 9332C

ACI CloudSec Keys operations

Automated Key Distribution & Re-Key

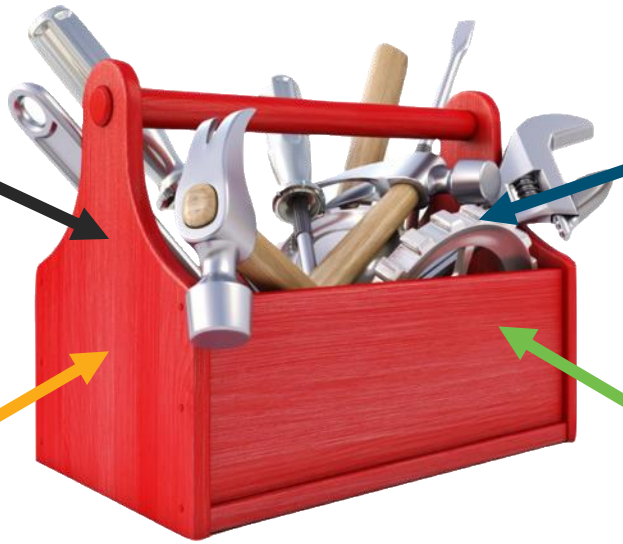


- Multi-site Orchestrator driven
  - No protocol dependency
- Reliable and secure key transport
  - rx_key installed before tx_key
- Non disruptive re-key
  - Hitless make before break
- Always encrypted
  - In case of programming errors, stay encrypted with previous key

The ACI Segmentation and Isolation Toolbox

EPGs & Contracts
ACI Policy Model

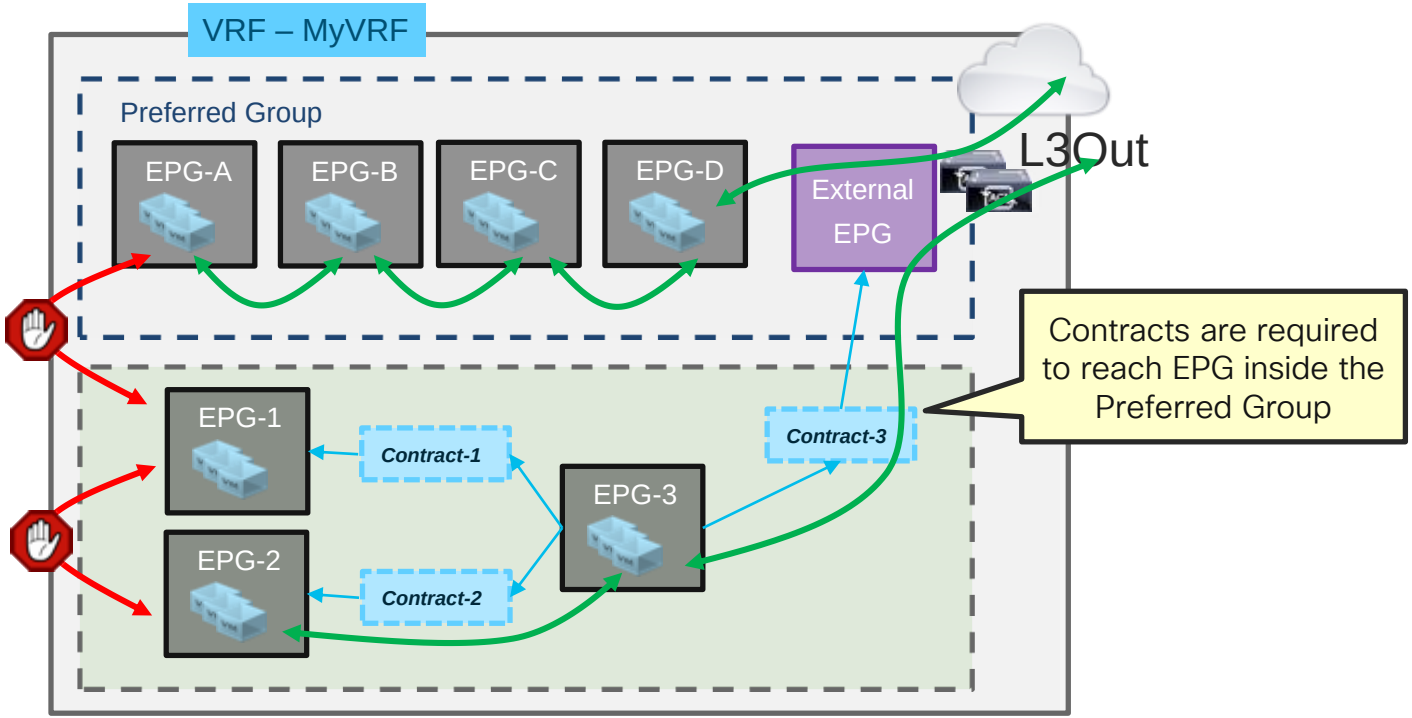
Micro-segmented EPGs
with attributes



Intra EPG isolation

Integration with L4/L7
Services ecosystem

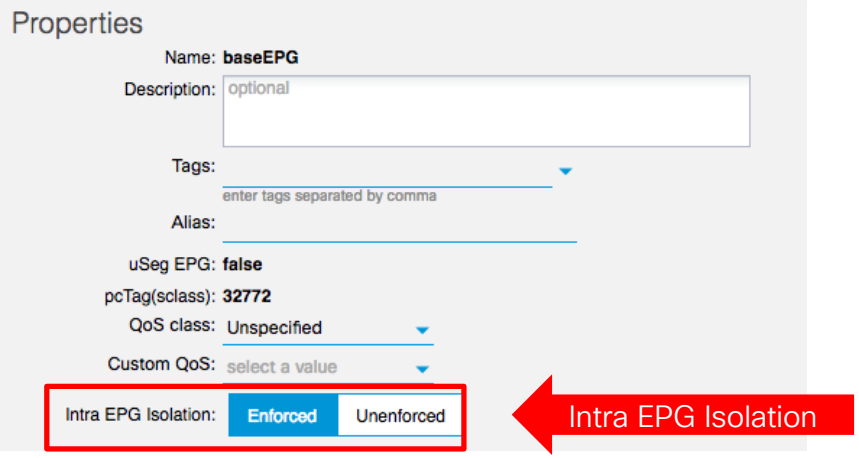
Isolation: EPG Contract Preferred Groups



Intra EPG Isolation

Intra EPG Isolation

- Intra EPG Isolation **blocks communication between all endpoints** inside the group
- Supports mixing of Physical and Virtual endpoints in same EPG
- Can be configured on **all type of EPG**

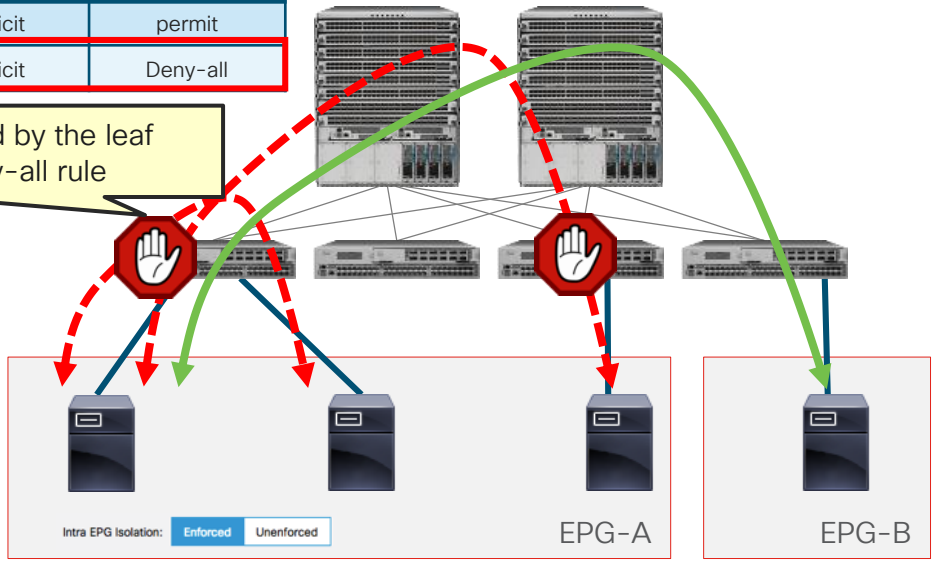



```
<fvTenant name="Tenant1">
  <fvAp name="ap1">
    <fvAEPg isAttrBasedEPg="no" matchT="AtleastOne" name="baseEPG" pcEnfPref="enforced" prefGrMemb="exclude" prio="unspecified">
      <fvRsBd tnFvBDName="bd"/>
    </fvAEPg>
  </fvAp>
</fvTenant>
```

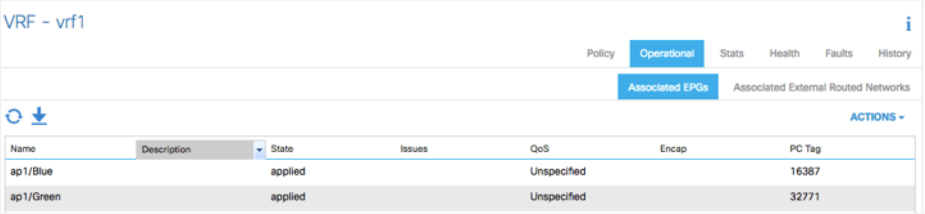
Intra EPG Isolation – Zoning Rules

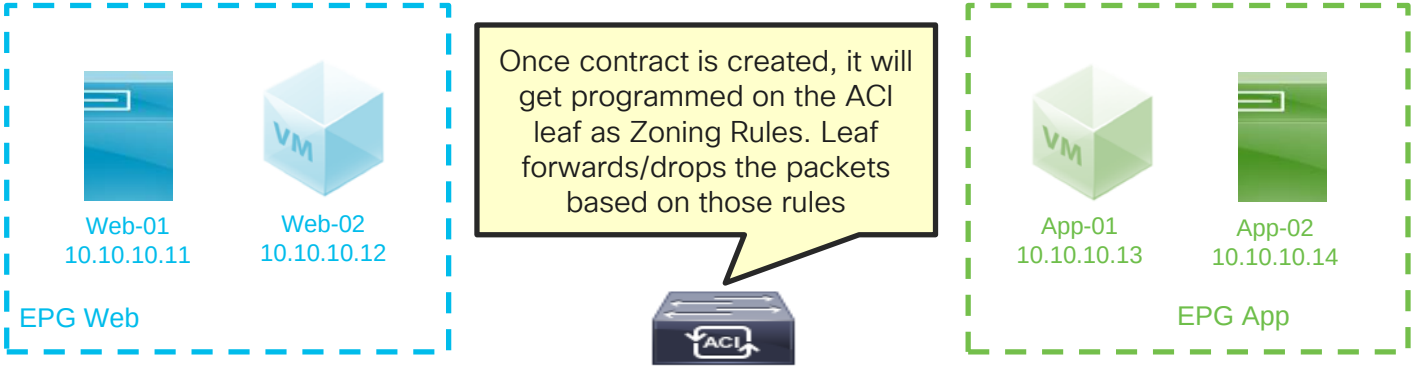
Source	Destination	Filter	Action
EPG-A	EPG-B	implicit	permit
EPG-A	EPG-A	implicit	Deny-all

Intra EPG traffic will be dropped by the leaf because of the implicit deny-all rule



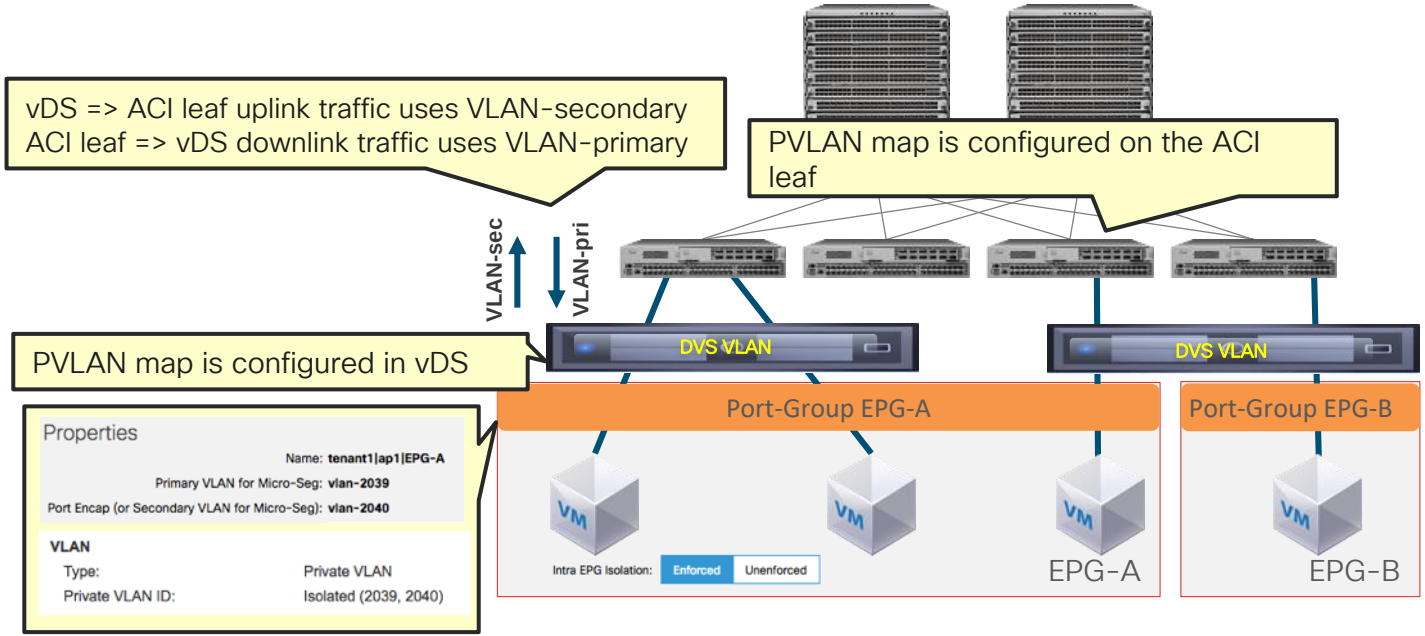
ACI Leaf Uses Zoning Rules to Forward or Drop the Traffic






```
leaf1# show zoning-rule scope 2162697 | egrep -E "Scope|32771|16387"
Rule ID  SrcEPG   DstEPG   FilterID operSt   Scope    Action   Priority
4616     16387    32771    5        enabled  2162697  permit   src_dst_any(8)
4617     32771    16387    5        enabled  2162697  permit   src_dst_any(8)
```

VMWare DVS Intra EPG Isolation



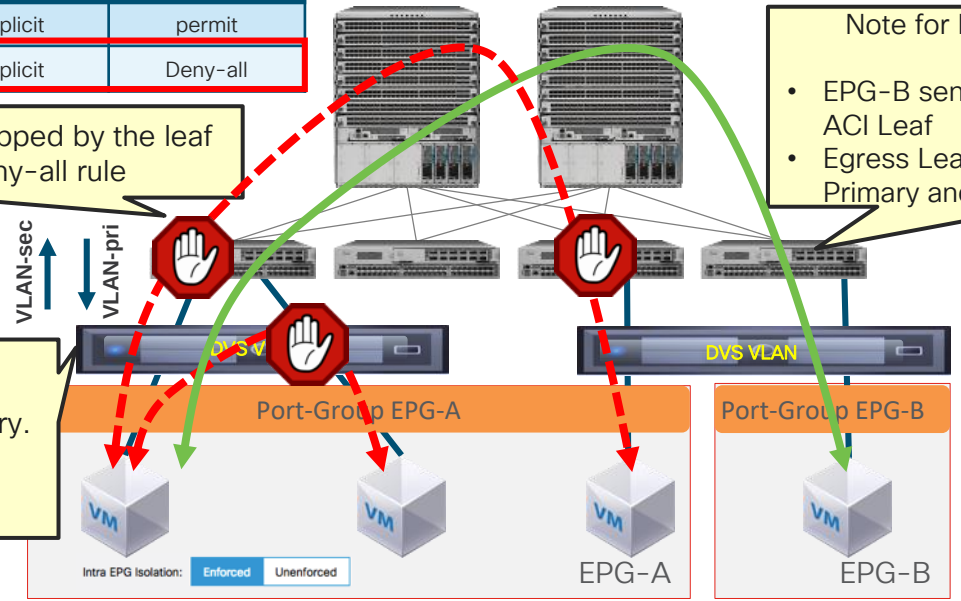
VMWare DVS Intra EPG Isolation

Source	Destination	Filter	Action
EPG-A	EPG-B	implicit	permit
EPG-A	EPG-A	implicit	Deny-all

Inter-ESXi host traffic will be dropped by the leaf because of the implicit deny-all rule

Note for Inter-EPG Traffic with Isolation Enabled:

- EPG-B sends traffic over regular VLAN to ACI Leaf
- Egress Leaf will encapsulate traffic in VLAN-Primary and send towards EPG-A VMs

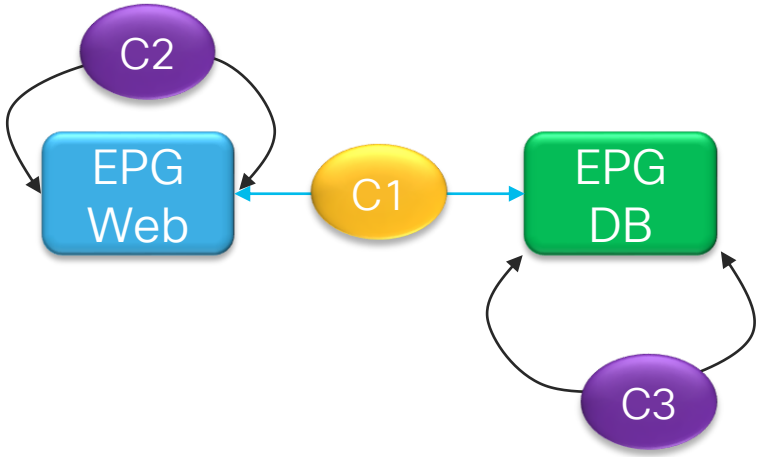


- Intra-ESXi host traffic is encapsulated in VLAN-secondary.
- vDS denies local intra-EPG VM traffic via PVLAN

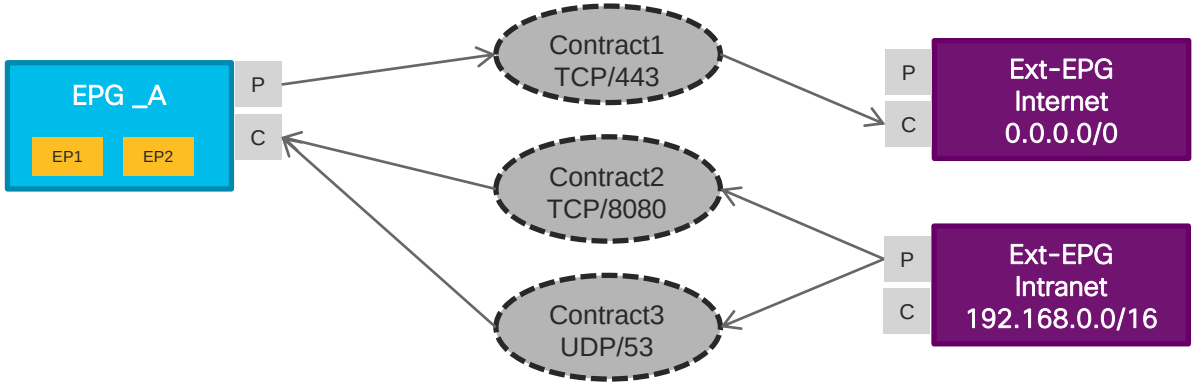
Intra-EPG Contracts

- Since release 3.0, ACI supports “Intra-EPG Contracts”
  - Allows whitelist policy enforcement of Intra-EPG traffic
  - Can co-exist with Inter-EPG contracts
  - Eliminates the need to create uSeg EPGs or deploy external FW for Intra-EPG segmentation
  - Enforcement is on Leaf switch (ie. Nexus 9000-EX models or above)
  - Same as regular contract scale

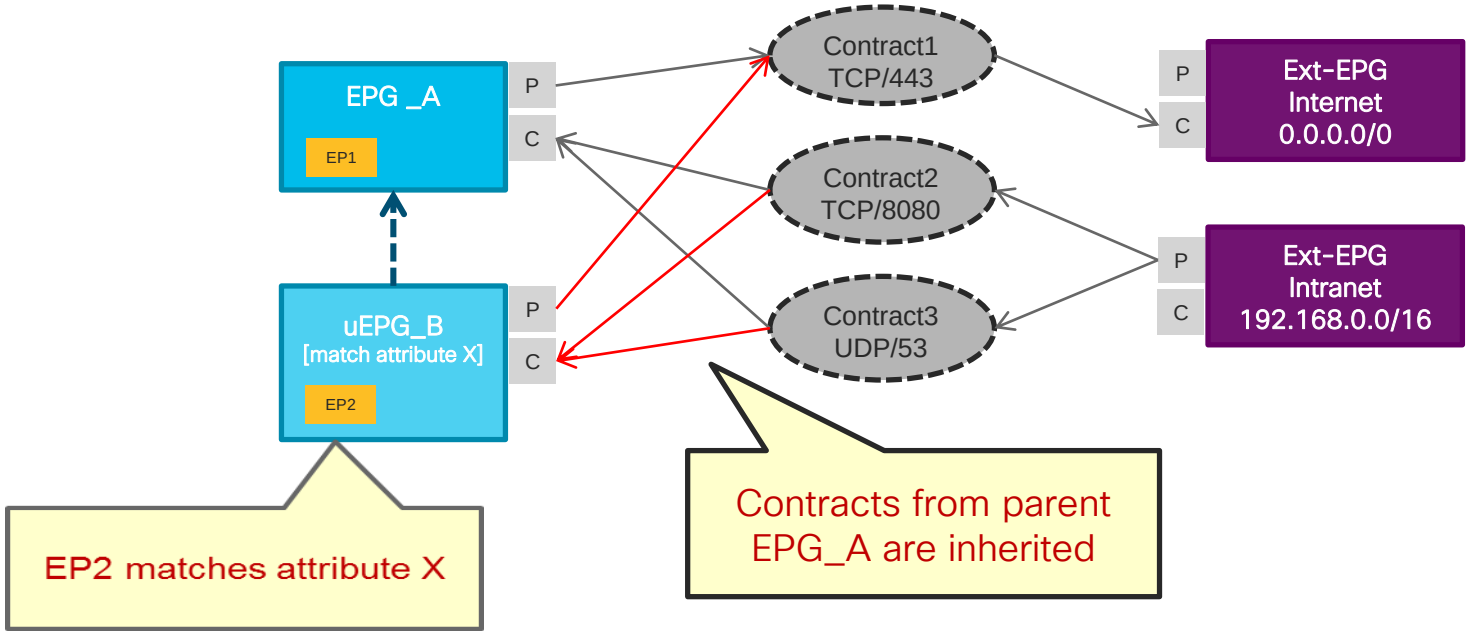
Intra-EPG Contract:
- Src-Class = Dest-Class
- Src-Class, Src-Class, Contract



Contract Inheritance



Contract Inheritance



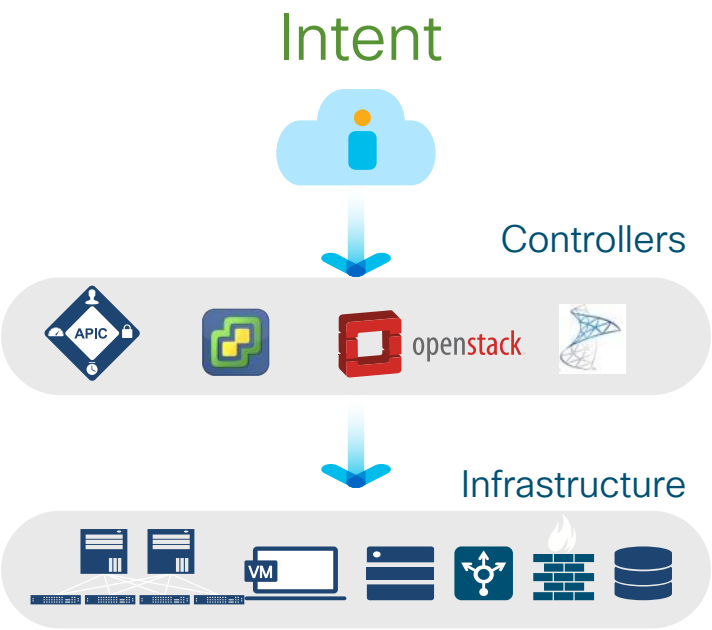
What is Network Assurance?

"**Assurance** is the **Guarantee** that the Network is doing what you **Intended** it to do"

Intent Encompasses Data Center Operations

Bridging, Routing, Security, QOS, Service Chaining, VM placement … Compliance, Audits

Assurance Gap in Today's Networks



1. How do I have confidence that I don't have errors due to my changes?
2. How do I easily understand the state of my entire infrastructure?
3. How do I rapidly analyze the network to identify issues?

Key Insight: Networks are Deterministic

If we **understand** the entire state of the network, we can accurately **predict** it's behavior

… without a single packet flowing through it

Cisco Day 2 Operations Stack (aka “Opstack”)

Network Insights & Assurance

+

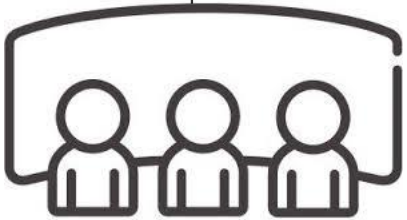
Insights: Health and Availability

Assurance: Moving from Reactive to Proactive

ACI & NX-OS Fabrics

ACI Fabrics

Opstack benefits

	Architecture and Planning teams	Network Operations	Network Administration and Maintenance
			
NAE Policy Explorer	Verification of design/connectivity mandates Ad-hoc policy exploration Ad-hoc Connectivity and Segmentation Analysis	Accelerate ACI on-ramp	Light weight book-keeping procedures
NAE	Capacity Planning Design and verify compliance mandate and posture Design and verify security mandate and posture	Proactive Assurance and Compliance Faster incident and problem management Shrink change management windows Accelerate ACI on-ramp	Execute high confidence production maintenance and upgrades
NIR/NIA	Trend based capacity planning Design trend-based environmental site operating procedures	Fabric wide anomaly detection based on - Resource monitoring - Flow monitoring Faster low-level troubleshooting and diagnostics	Proactively reduce vulnerability exposure Improve Site reliability

Opstack- Available via Premier License

Network Insight & Network Assurance

NAE Policy Explorer *

- Network Policy exploration
- Ad-hoc connectivity and segmentation discovery

Network Assurance Engine

- Policy/ Control/Data plane Assurance
- Incident and Problem Management
- Compliance and Audit

Network Insights Resources *

- Fabric wide resource utilization & trends
- Anomaly detection – environmental, config & operational resources, interface errors
- End-to-end flow path, latency and drop reason

Network Insights Advisor **

- Notifications of EOS/EOL of H/W & S/W
- Security Advisory Notification Updates (PSIRTs)
- Recommended S/W Release Updates and upgrade impact analysis
- TAC assist

Architecture and Planning

Network Administration and Maintenance

Network Operations

* Available as App on APIC

Cisco Network Assurance Engine

Based on mathematical models of the network

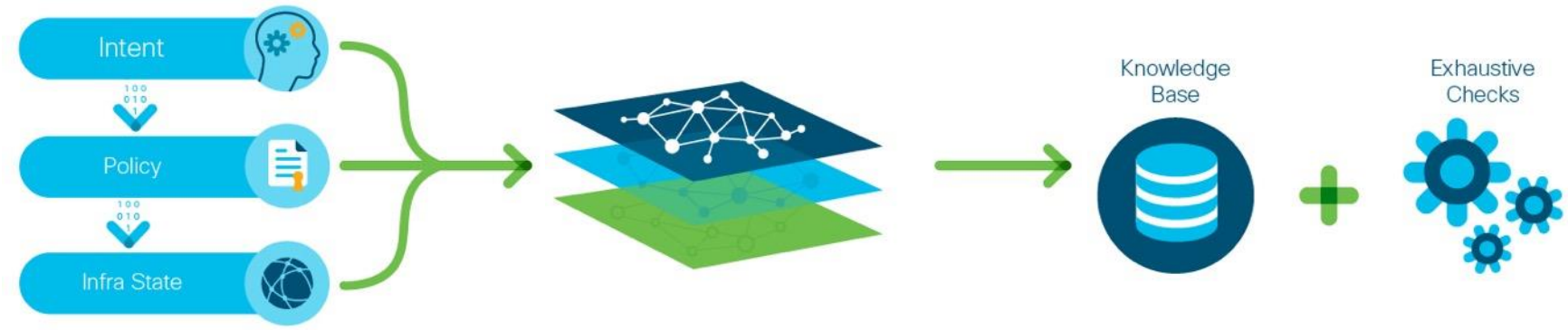
Continuously verifies and validates the entire ACI network

Proactively delivers the confidence that the ACI network is operating correctly

Always-On
Network Assurance

Comprehensive, Intelligent, Continuous

Cisco Network Assurance Engine: How It Works



Data Collection

Captures all non-packet data: intent, policy, state across data center network

Comprehensive Network Modeling

Mathematically accurate models spanning underlay, overlay and virtualization layers

Intelligent Analysis

5000+ domain knowledge-based error scenarios built-in, codified remediation steps

Use Cases & Benefits

Making Operations Fundamentally Proactive

PREDICT THE IMPACT OF CHANGES

- Drive change agility
- Minimize human errors and eliminate configuration drift
- Accelerate migrations

PROACTIVELY VERIFY NETWORK-WIDE BEHAVIOR

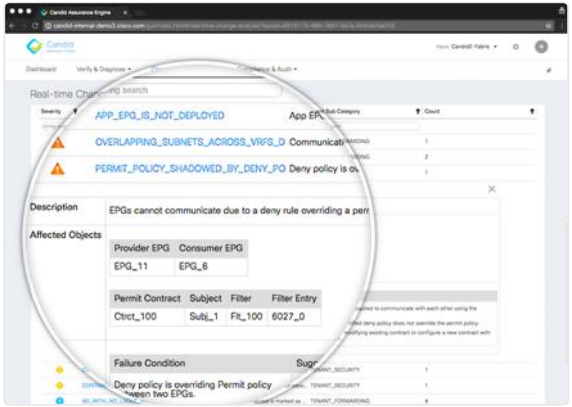
- Ensure connectivity
- Proactively eliminate potential network outages or vulnerabilities
- Enhance SLAs

ASSURE NETWORK SECURITY POLICY AND COMPLIANCE

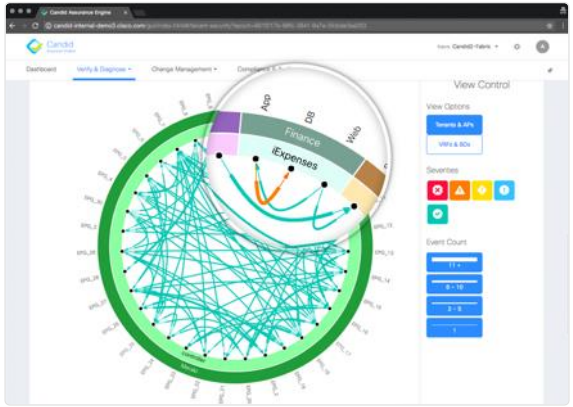
- Reduce security risk
- Achieve provable compliance by design, continuously

User Interface: Centered Around “Smart Events”

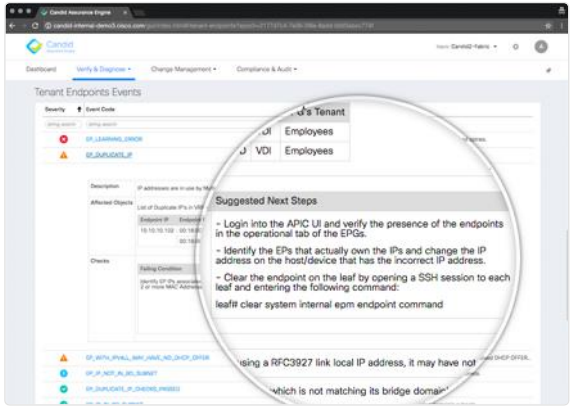
Change Management



Search and Visualization



Incidence and Problem Management



Smart Events: What, Where, Why, and How

Ex.: NAE UI for Compliance Analysis

Cisco Network Assurance Engine is currently running with a TRIAL license. View License Details

Cisco Network Assurance Engine

Assurance Group Can5_BM28_mod28.tar.gz-20181108-1739

Dashboard | Change Management | Verify & Diagnose | Epoch Analysis | Optimize | Compliance | Smart Events

Continuous Compliance

Compliance Analysis

Manage Compliance Requirements

Generate Assurance Report

view compliance analysis results

Configure compliance requirements

01:01 PM 07/20/2018 | 01:15 | 01:30 | 01:45 | 02 PM | 02:15 | 02:30 | 02:45 | 03 PM

Live Update | Zoom Level | All | 1 m | 1 w | 1 d | 12 h | 6 h | 1 h | Custom

Dashboard

Epoch | Trend

Smart Events

Critical	Major	Minor	Warning	Info	Total
325	965	204	294	3,240	5,028

Real-time Change Analysis

Top Tenants by Policy Issue Severity

Tenant			
abc02	0	48	0

Top VRFs by Policy Issue Severity

VRF	Issues	
abc02_ctx02		24

Top Tenants by Policy Issue Subcategory

Tenant	Forwar...	Security	Fabric	Others
abc02	48	0	0	0

Agenda

Data Center / Virtualization Cisco education offerings

Course	Description	Cisco Certification
Introducing Cisco Data Center Networking (DCICN) Introducing Cisco Data Center Technologies (DCICT)	Get job-ready foundational-level certification and skills in installing, configuring, and maintaining next generation data centers.	CCNA® Data Center
Implementing Cisco Data Center Unified Computing (DCUCI) Implementing Cisco Data Center Infrastructure (DCII) Implementing Cisco Data Center Virtualization and Automation (DCVAI) Designing Cisco Data Center Infrastructure (DCID) Troubleshooting Cisco Data Center Infrastructure (DCIT)	Obtain professional level skills to design, configure, implement, troubleshoot next generation data center infrastructure.	CCNP® Data Center
Product Training Portfolio:DCAC9K, DCINX9K, DCMDS, DCUCS, DCNX1K, DCNX5K, DCNX7K, CACND, DSACI, HFLEX UCSDF, UCSDACI, DCUCCEN	Gain hands-on skills using Cisco solutions to configure, deploy, manage and troubleshoot unified computing, policy-driven and virtualized data center infrastructure.	
Designing the FlexPod® Solution (FPDESIGN) Implementing and Administering the FlexPod® Solution (FPIMPADM)	Learn how to design, implement and administer FlexPod® solutions	Cisco and NetApp Certified FlexPod® Specialist
Designing the VersaStack Solution (VSDESIGN) Implementing and Administering the VersaStack Solution (VSIMP)	Learn how to design, implement and administer VersaStack solutions	

For more details, please visit: http://learningnetwork.cisco.com
Questions? Visit the Learning@Cisco Booth

Network Programmability Cisco education offerings

Course	Description	Cisco Certification
Developing with Cisco Network Programmability (NPDEV)	Provides Application Developers with comprehensive curriculum to develop infrastructure programming skills; Addresses needs of software engineers who automate network infrastructure and/or utilize APIs and toolkits to interface with SDN controllers and individual devices	Cisco Network Programmability Developer (NPDEV) Specialist Certification
Designing and Implementing Cisco Network Programmability (NPDESI)	Provides network engineers with comprehensive soup-to-nuts curriculum to develop and validate automation and programming skills; Directly addresses the evolving role of network engineers towards more programmability, automation and orchestration	Cisco Network Programmability Design and Implementation (NPDESI) Specialist Certification
Programming for Network Engineers (PRNE)	Learn the fundamentals of Python programming – within the context of performing functions relevant to network engineers. Use Network Programming to simplify or automate tasks	Recommended pre-requisite for NPDESI and NPDEV Specialist Certifications
Cisco Digital Network Architecture Implementation Essentials (DNAIE)	This training provides students with the guiding principles and core elements of Cisco's Digital Network Architecture (DNA) architecture and its solution components including; APIC-EM, NFV, Analytics, Security and Fabric.	

For more details, please visit: http://learningnetwork.cisco.com
Questions? Visit the Learning@Cisco Booth

Complete your online session survey

- Please complete your session survey after each session. Your feedback is very important.
- Complete a minimum of 4 session surveys and the Overall Conference survey (starting on Thursday) to receive your Cisco Live t-shirt.
- All surveys can be taken in the Cisco Events Mobile App or by logging in to the Content Catalog on ciscolive.com/emea.

Cisco Live sessions will be available for viewing on demand after the event at ciscolive.com.

Continue your education

Demos in the Cisco Showcase

Walk-In Labs

Meet the Engineer 1:1 meetings

Related sessions

Thank you

CISCO

You make **possible**