cisco

Cisco Connect

19 - 21 March, 2018

Rovinj, Croatia

ACI Anywhere: Extending the ACI Fabric

Max Ardica

Principal Engineer – Cisco INSBU

Agenda

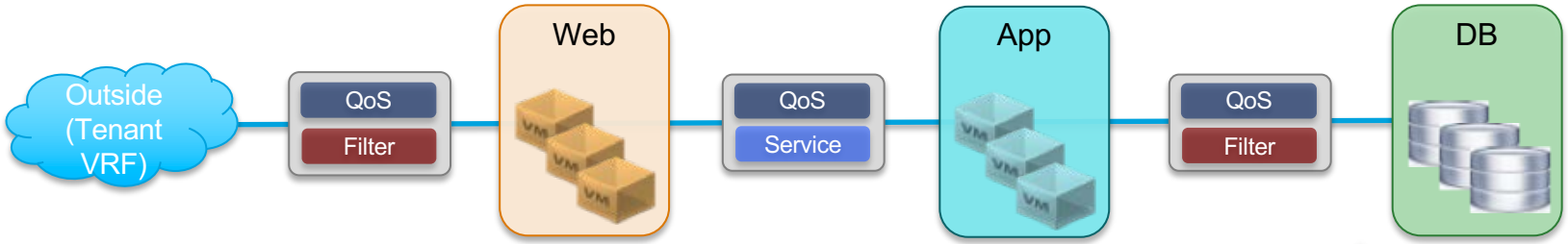
Cisco Connect

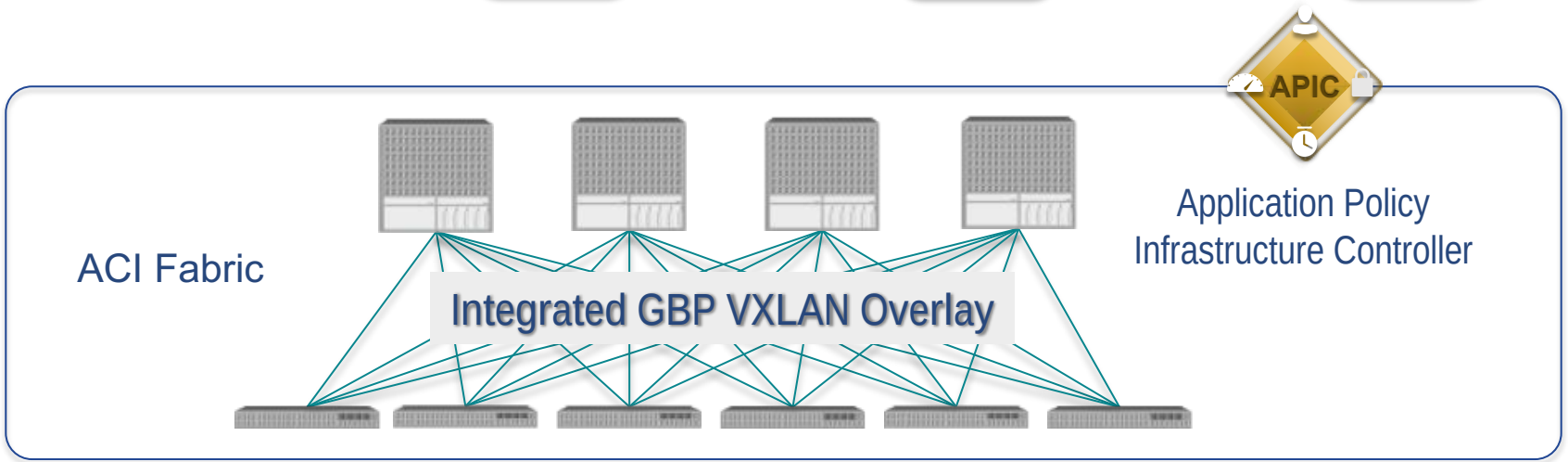
- ACI Network and Policy Domain Evolution
- ACI Multi-Pod
- ACI Multi-Site
- ACI Remote Physical Leaf
- ACI Remote Virtual Leaf (vPod)
- ACI Extensions to Multi-Cloud
- Conclusions and Q&A

Cisco
Connect

ACI Network and Policy Domain Evolution

Introducing: Application Centric Infrastructure (ACI)





Cisco ACI: Industry Leader

4,800+
ACI Customers

46+%
ACI Attach Rate

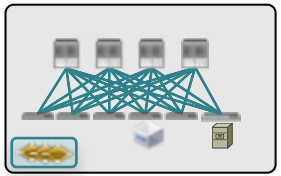
65+
Ecosystem Partners

Ecosystem Partners

Cisco ACI
Fabric and Policy Domain Evolution

Cisco Connect

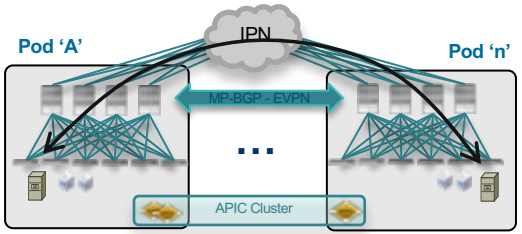
ACI Single Pod Fabric



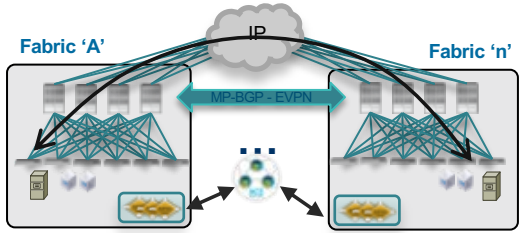
ACI 1.0 - Leaf/Spine Single Pod Fabric

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

ACI Multi-Pod Fabric



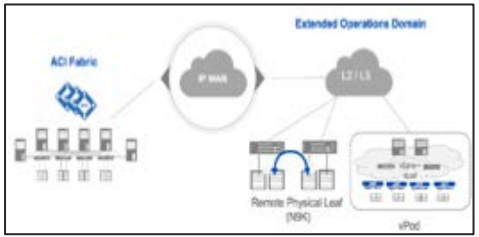
ACI Multi-Site



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

ACI 3.1/3.2 - Remote Leaf and vPod extends an Availability Zone (Fabric) to remote locations

ACI Remote Leaf



ACI Multi-Cloud

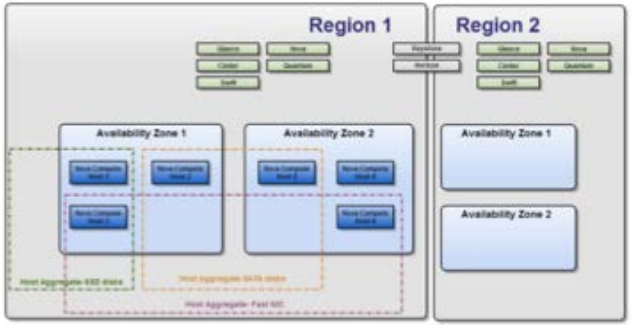
Future – ACI Extensions to Multi-Cloud

Regions and Availability Zones
OpenStack and AWS Definitions

Cisco Connect

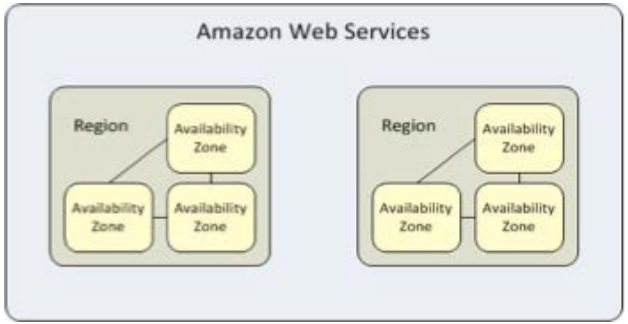
- Regions - Each Region has its own full OpenStack deployment, including its own API endpoints, networks and compute resources
- Availability Zones - Inside a Region, compute nodes can be logically grouped into Availability Zones, when launching new VM instance, we can specify AZ or even a specific node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

OpenStack

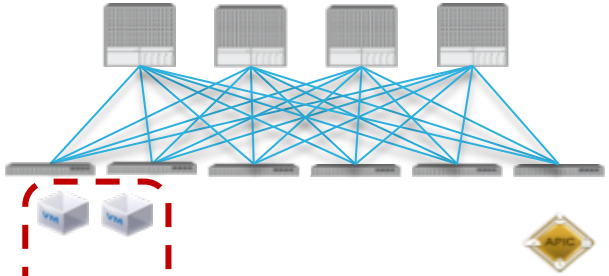


- Regions – Separate large geographical areas, each composed of multiple, isolated locations known as Availability Zones
- Availability Zones - Distinct locations within a region that are engineered to be isolated from failures in other Availability Zones and provide inexpensive, low latency network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

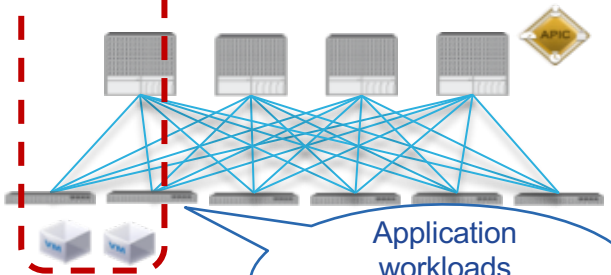
Amazon Web Services



Typical Requirement
Creation of Two Independent Fabrics/AZs



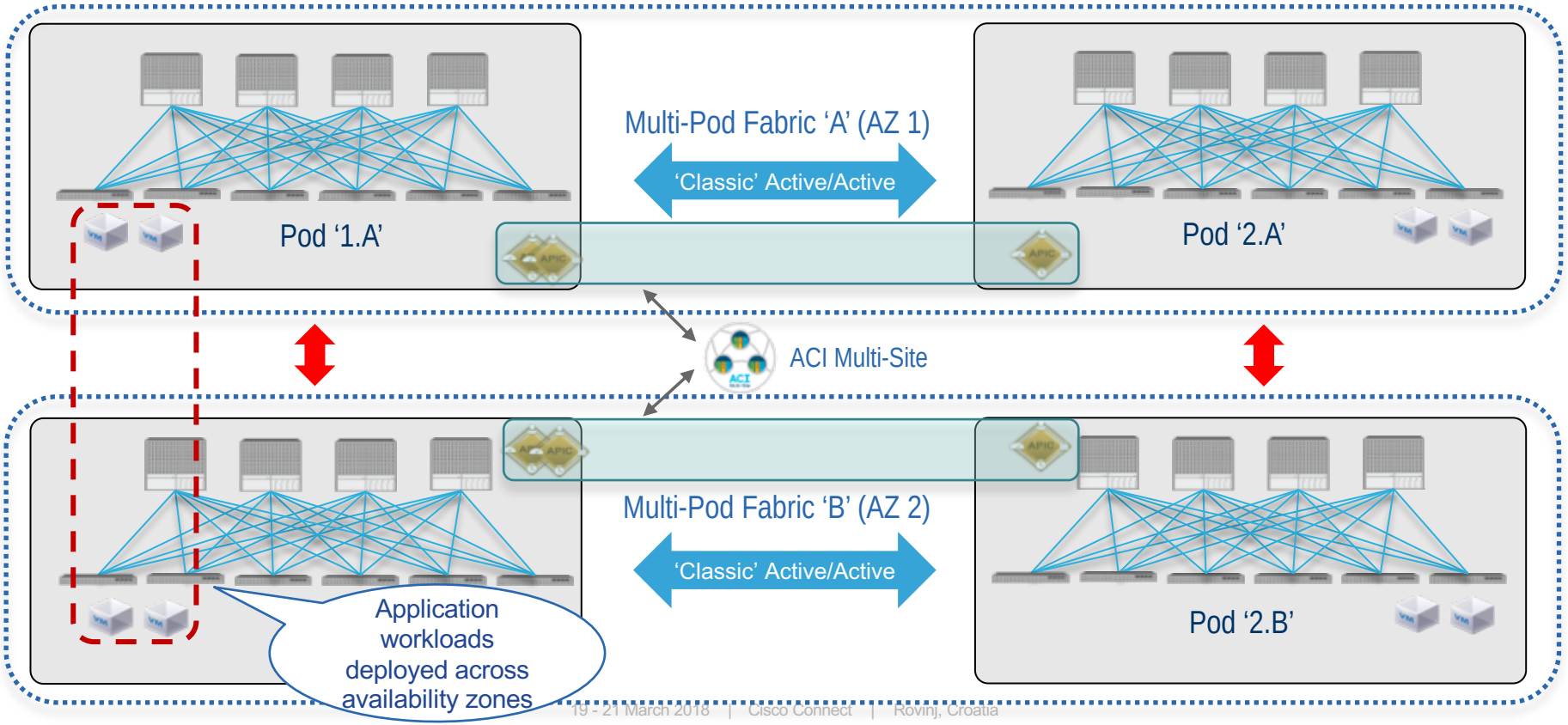
Fabric ‘A’ (AZ 1)



Fabric ‘B’ (AZ 2)

Application workloads deployed across availability zones

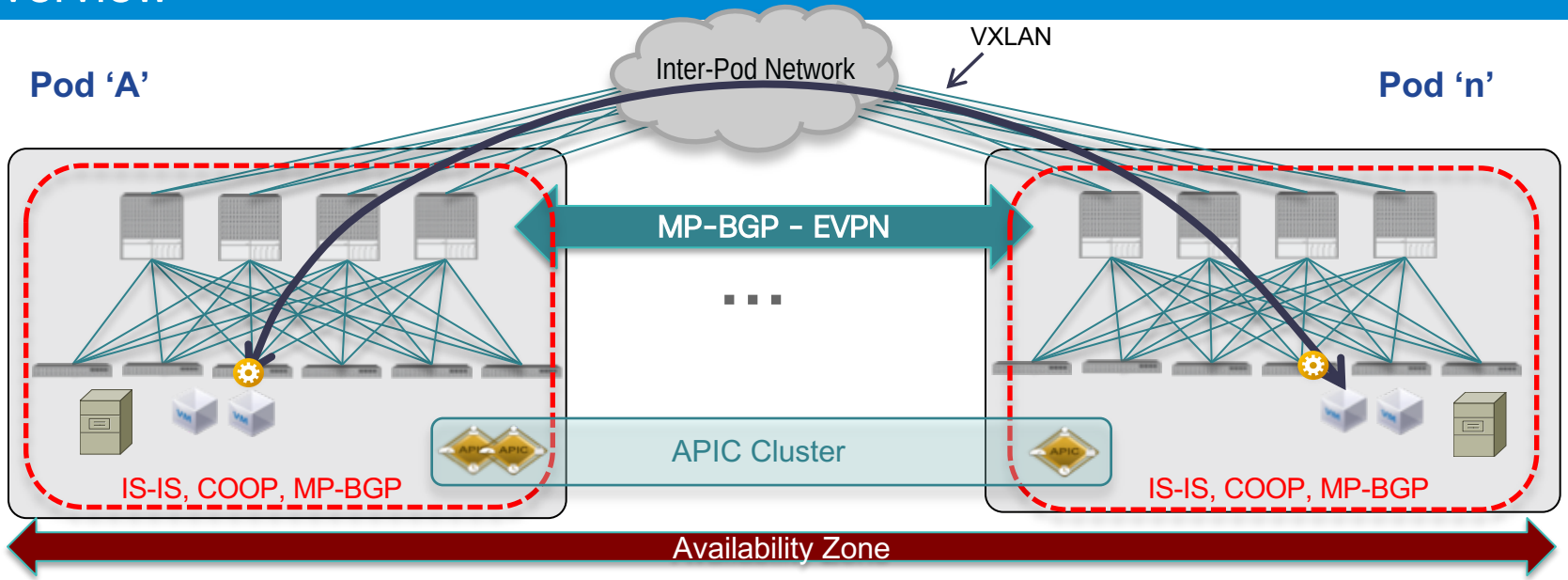
Typical Requirement
Creation of Two Independent Fabrics/AZs



Cisco
Connect

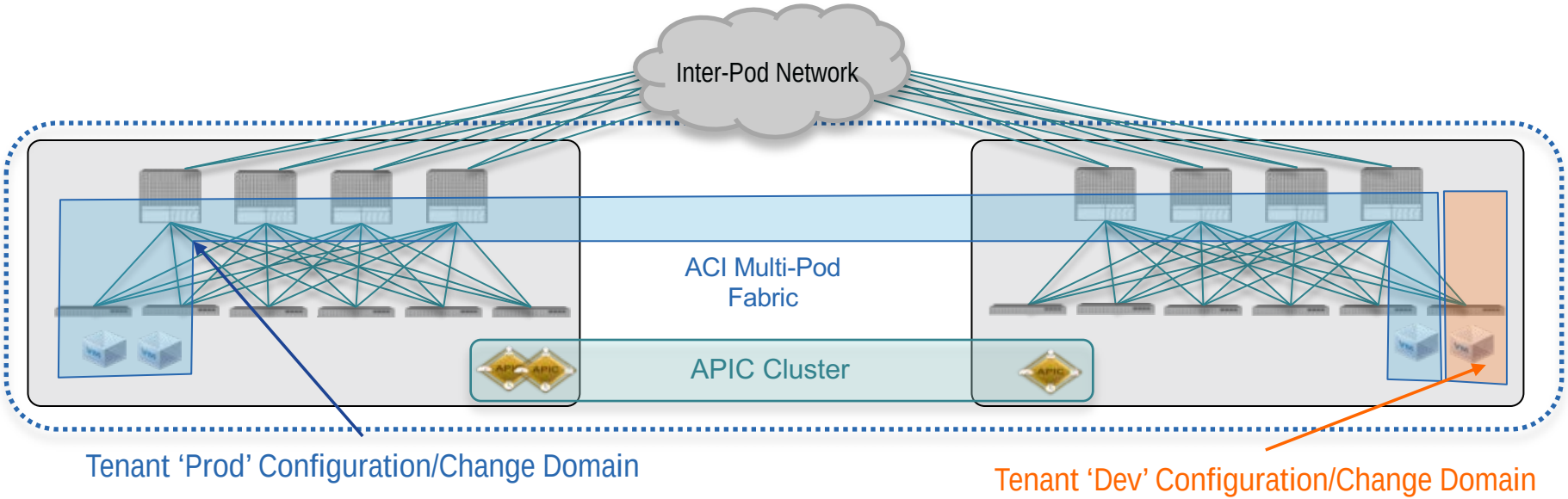
ACI Multi-Pod

ACI Multi-Pod
Overview



- 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

Single Availability Zone with Tenant Isolation
Isolation for 'Virtual Network Zone and Application' Changes

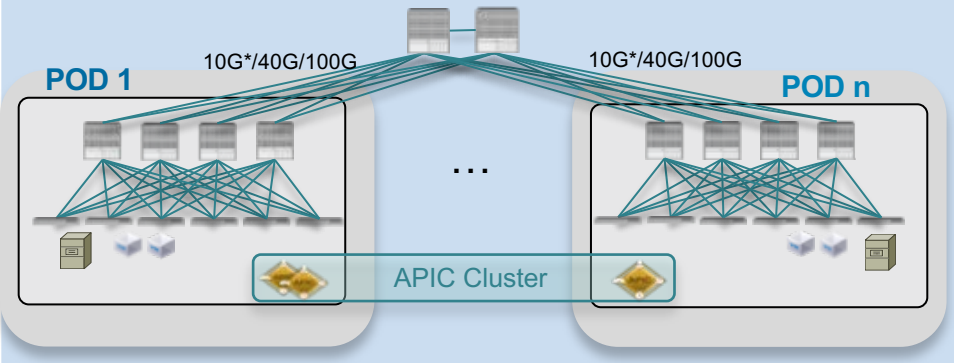


- The ACI 'Tenant' construct provide a **domain of application** and associated virtual network policy change
- Domain of operational change for an application (e.g. production vs. test)

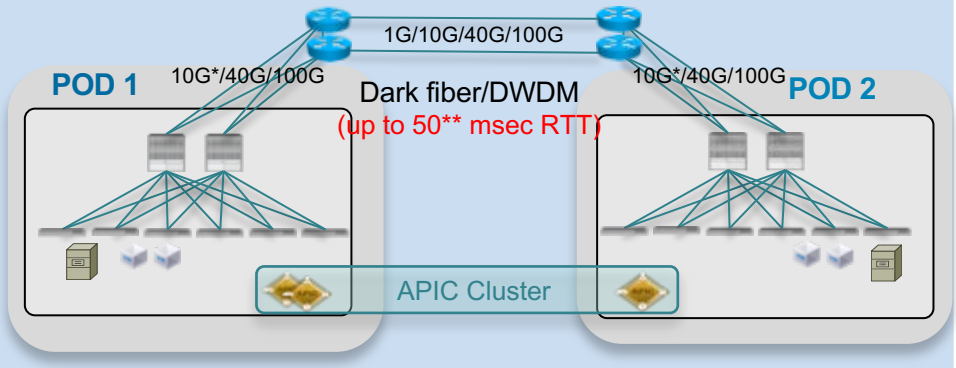
ACI Multi-Pod
Supported Topologies

Cisco Connect

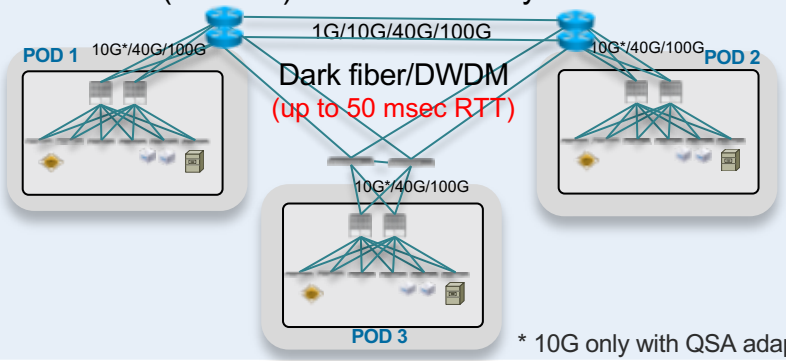
Intra-DC



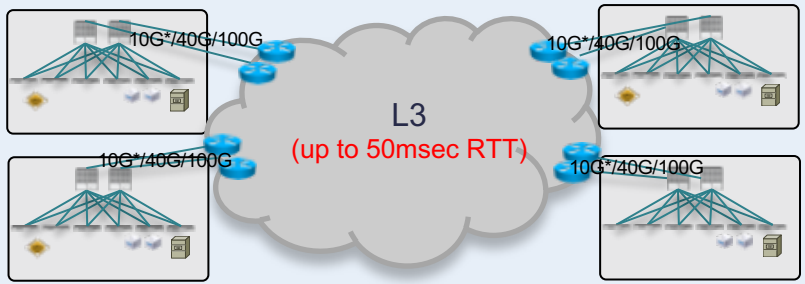
Two DC sites directly connected



3 (or more) DC Sites directly connected



Multiple sites interconnected by a generic L3 network



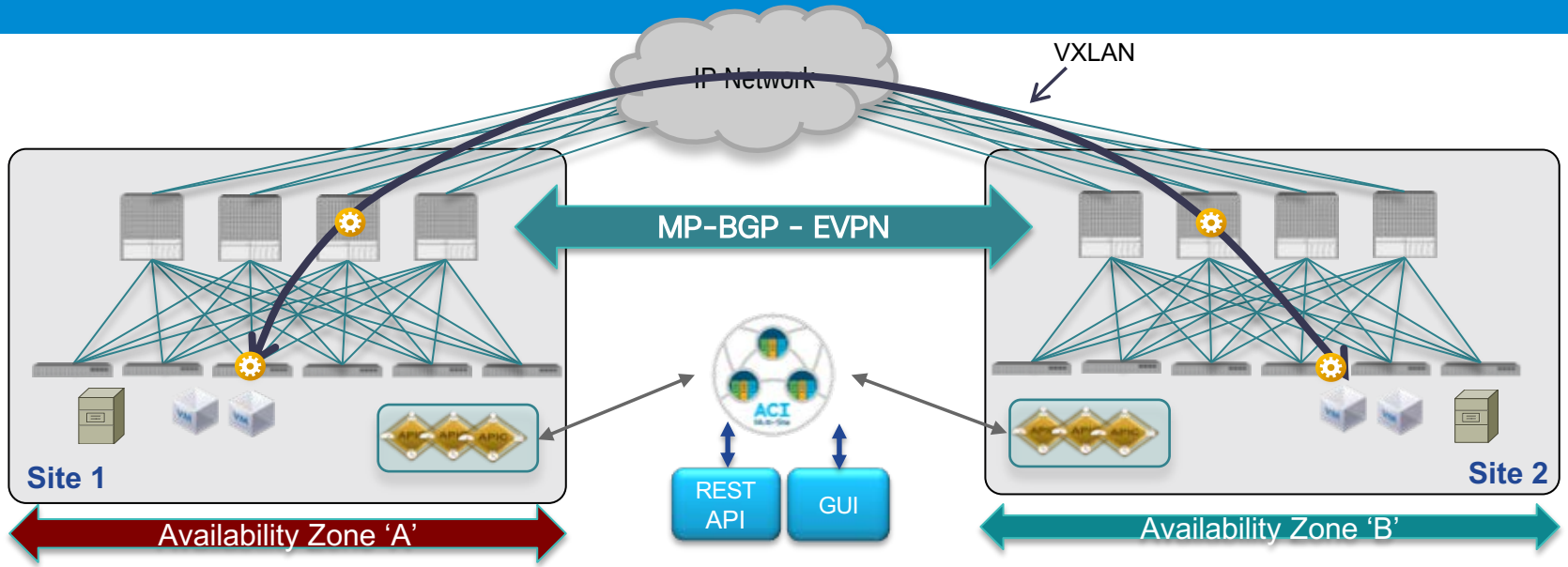
* 10G only with QSA adapters on EX/FX and 9364C spines

** 50 msec support added in SW release 2.3(1)

Cisco
Connect

ACI Multi-Site

ACI Multi-Site
Overview



- Separate ACI Fabrics with independent APIC clusters
- 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

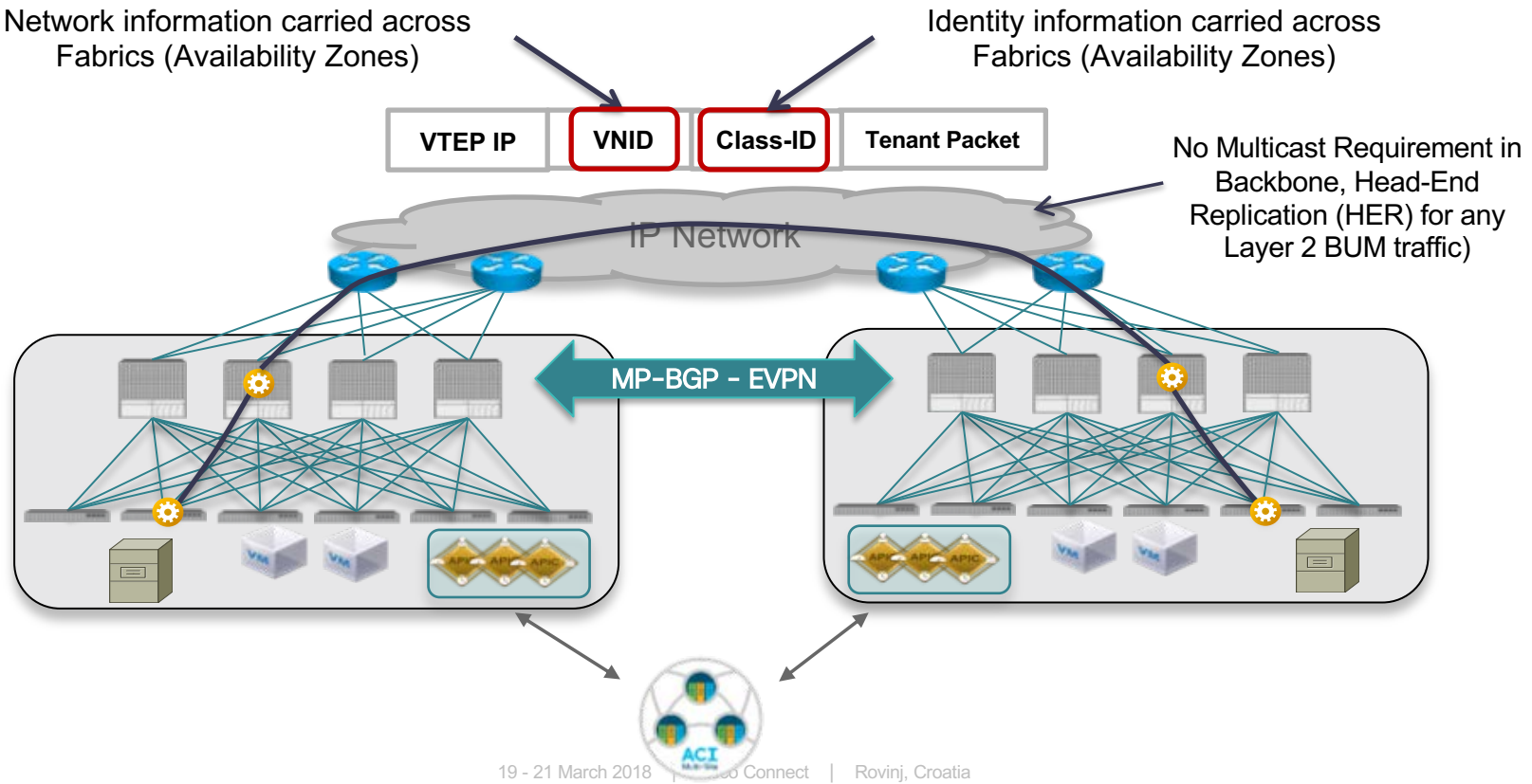
ACI Multi-Site
Main Use Cases

Cisco Connect

Scale-Up Model to Build a Large Intra-DC Network

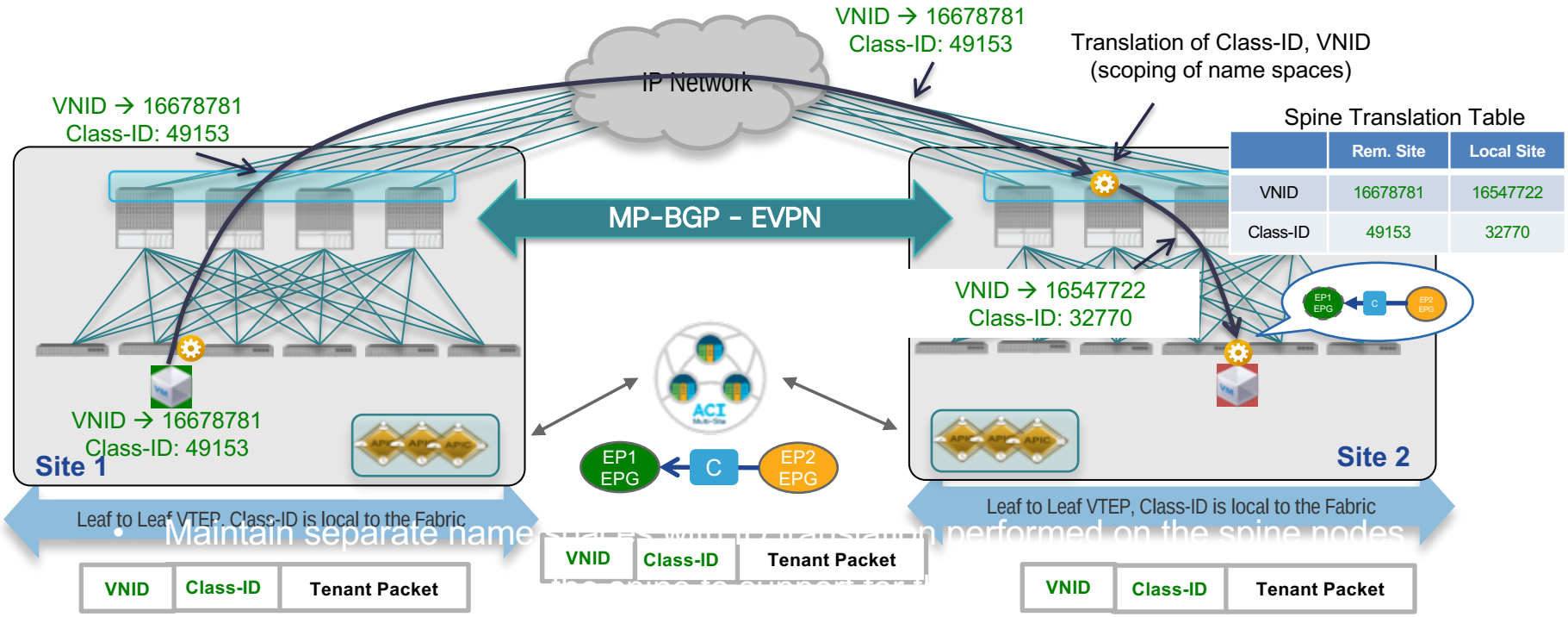
Data Center Interconnect (DCI)

ACI Multi-Site
Network and Identity Extended between Fabrics



ACI Multi-Site
Namespace Normalization

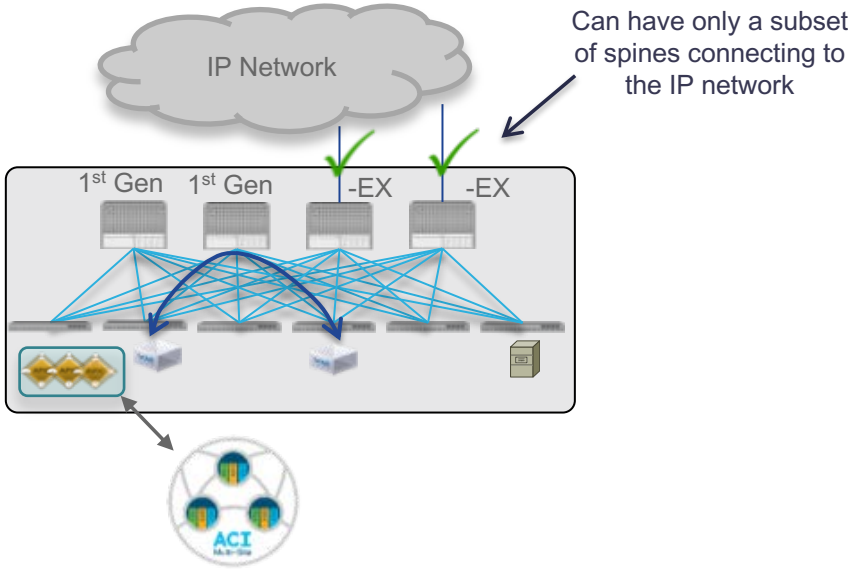
Cisco Connect



- Maintain separate name [illegible] performed on the spine nodes

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
- 1st generation spines (including 9336PQ) not supported
- Can still leverage those for intra-site leaf to leaf communication

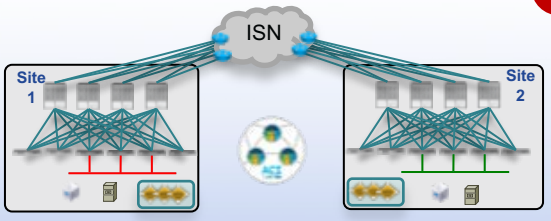


ACI Multi-Site Networking Options
Per Bridge Domain Behavior

Cisco Connect

Layer 3 only across sites

1

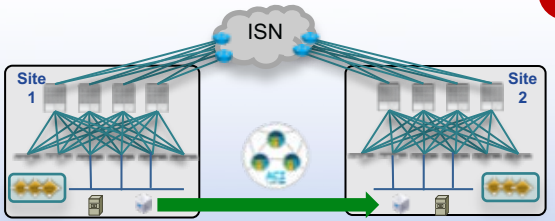


- Bridge Domains and subnets not extended across Sites
- Layer 3 Intra-VRF or Inter-VRF communication (shared services across VRFs/Tenants)

MSO GUI (BD)

IP Mobility without BUM flooding

2

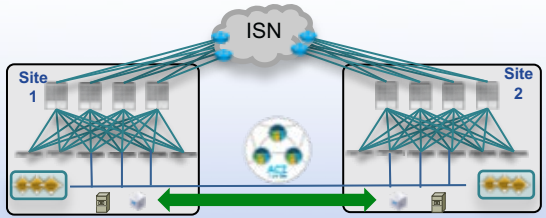


- Same IP subnet defined in separate Sites
- Support for IP Mobility ('cold' and 'live'* VM migration) and intra-subnet communication across sites
- **No Layer 2 BUM flooding across sites**

MSO GUI (BD)

Layer 2 adjacency across Sites

3

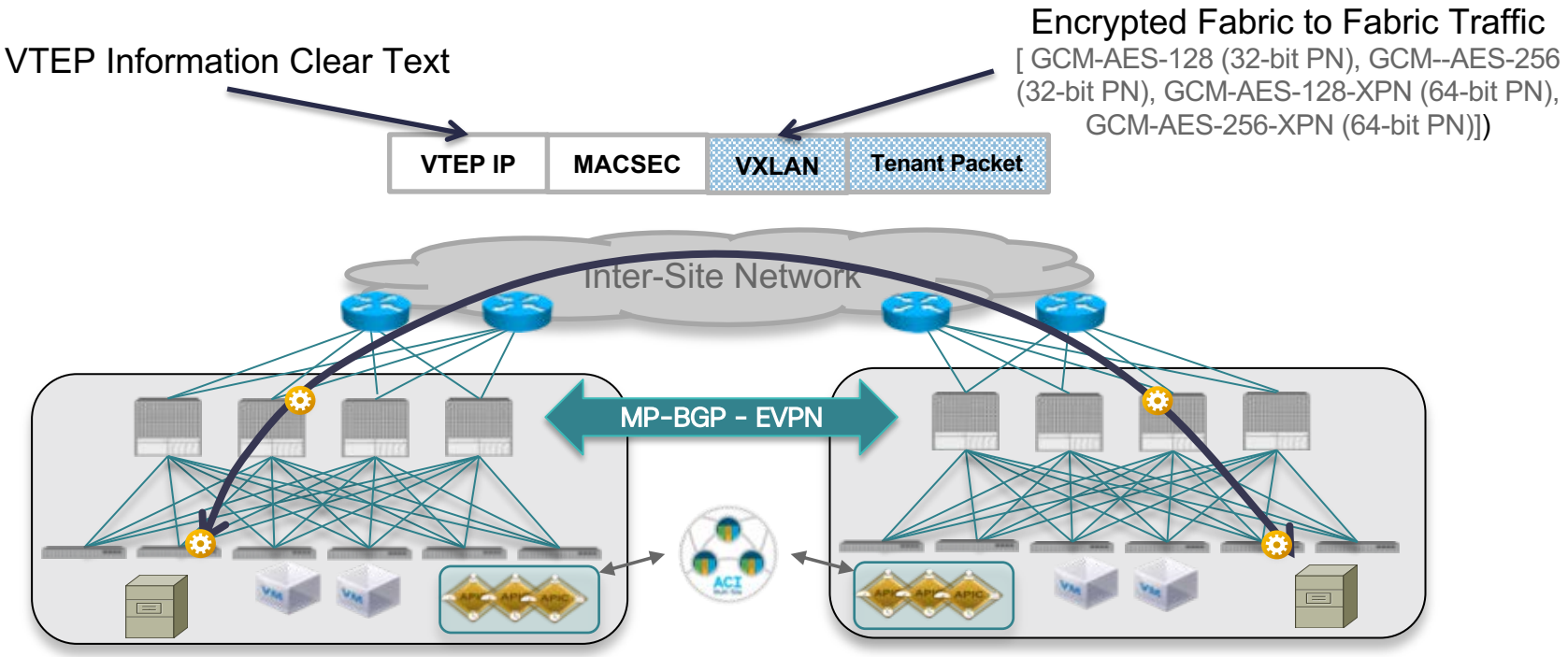


- Interconnecting separate sites for fault containment and scalability reasons
- Layer 2 domains stretched across Sites, support for 'live'* VM migration and application clustering
- **Layer 2 BUM flooding across sites**

MSO GUI (BD)

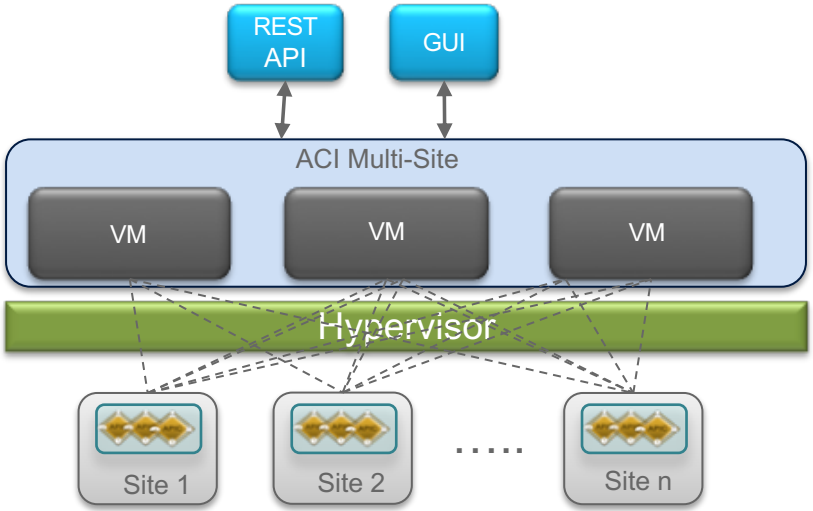
*'Live' migration officially supported from ACI release 3.2

ACI Multi-Site
CloudSec Encryption for VXLAN Traffic



Support planned for a future ACI release for FX line cards and 9364C platform

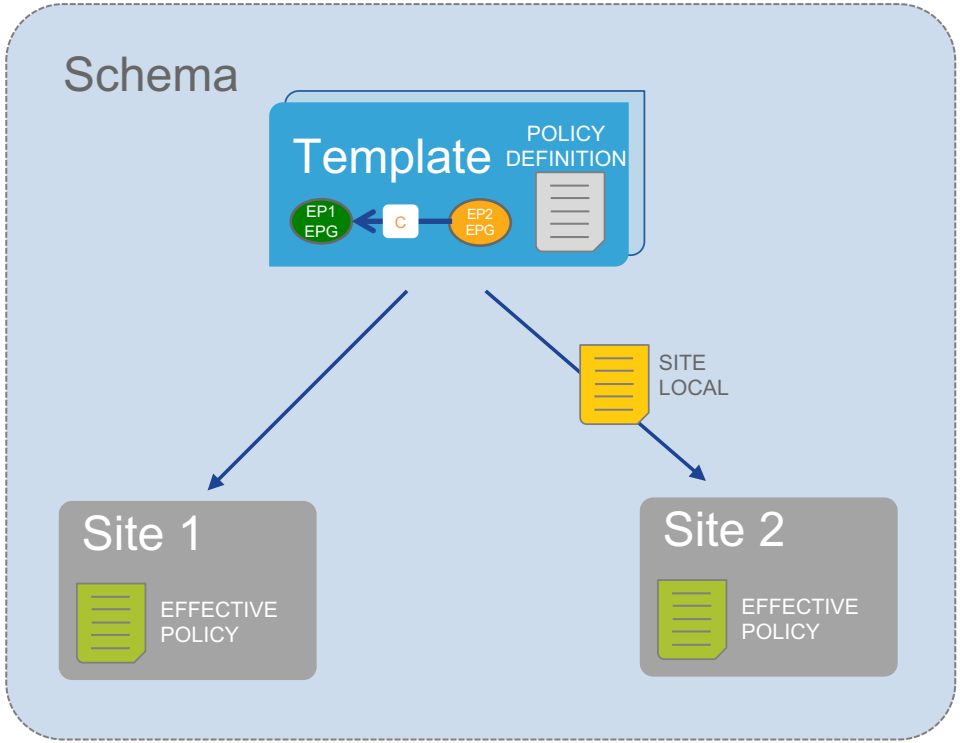
ACI Multi-Site
Multi-Site Orchestrator (MSO)



- Micro-services architecture
  - Multiple MSO nodes are created and run concurrently (active/active)
  - vSphere VM only form factor initially (physical appliance planned for a future ACI release)
- OOB Mgmt connectivity to the APIC clusters deployed in separate sites
  - Support for 500 msec to 1 sec RTT
- Main functions offered by MSO:
  - Monitoring the health-state of the different ACI Sites
  - Provisioning of day-0 configuration to establish inter-site EVPN control plane
  - Defining and provisioning policies across sites
  - Day-2 operation functionalities

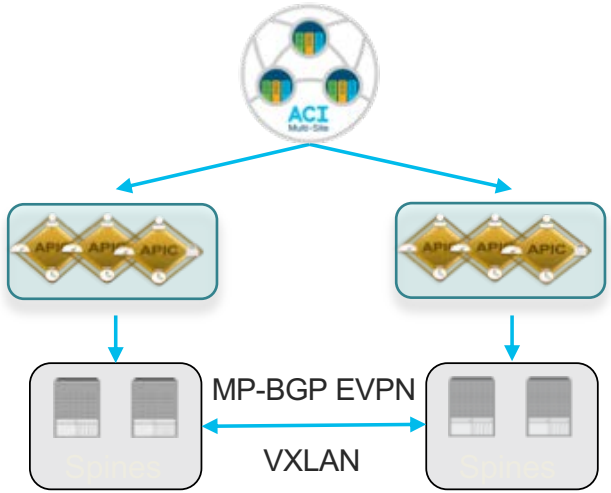
ACI Multi-Site
MSO Schema and Templates

- Template = APIC policy definition (App & Network)
- Template is the scope/granularity of what can be pushed to sites
- Template is associated to all managed sites or to a subset of sites
- Schema = container of Templates sharing a common use-case
- Scope of change: policies in different templates can be pushed to separate sites at different times



ACI Multi-Site
Day-2 Operations: Full-Stack Consistency Checker

ACI 3.2 Release



- Multi-Site Infra: Unicast, Multicast, BGP TEPs and Tunnel state
- Multi-Site Tenant and EPG granularity:
  - Inspect and validate full-stack programming: 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 will both be supported

Cisco
Connect

ACI Remote Physical Leaf

ACI Remote Physical Leaf
Business Value and Use Cases

Extending the ACI policy model outside the main datacenter to remote sites distributed over IP Backbone (Satellites DCs, CoLo locations, etc.)

Extending ACI fabric policy and L2/L3 connectivity to a small DR site without requiring the deployment of a full-blown ACI Fabric

Centralized Policy Management and Control Plane for remote locations

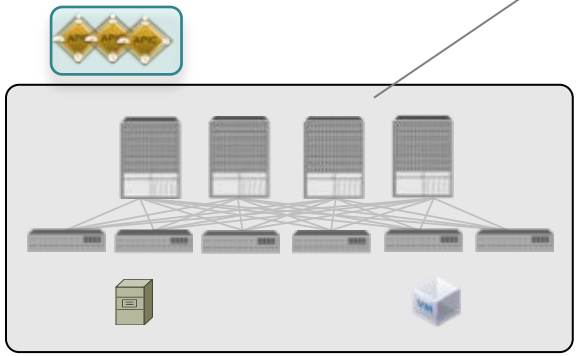
Small form factor solution at locations with space constraints

ACI Remote Physical Leaf
Conceptual Architecture

Cisco Connect

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

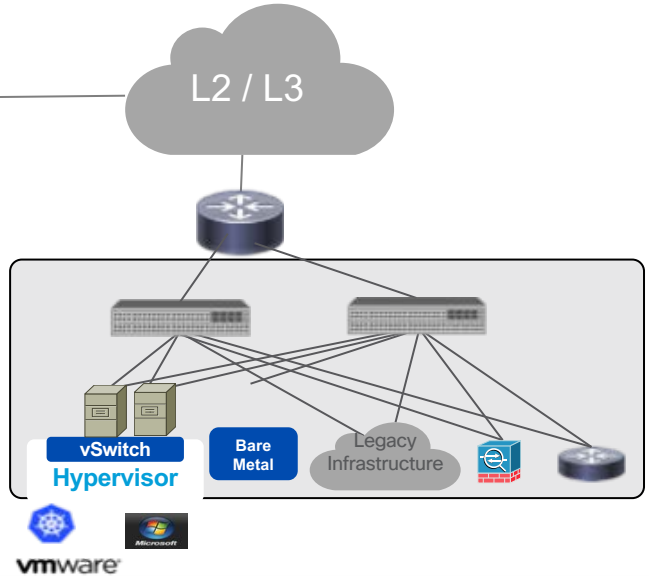
APIC and Spine Nodes (Proxy function) remain at primary Pod(s)



ACI Main DC

IPN Requirements (3.1)

- 150 msec maximum RTT
- 500 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

All hardware from -EX onwards is required for remote leaf nodes and the spines to which they get associated

ACI Main Pod

Supported Spines

Fixed
- N9364C

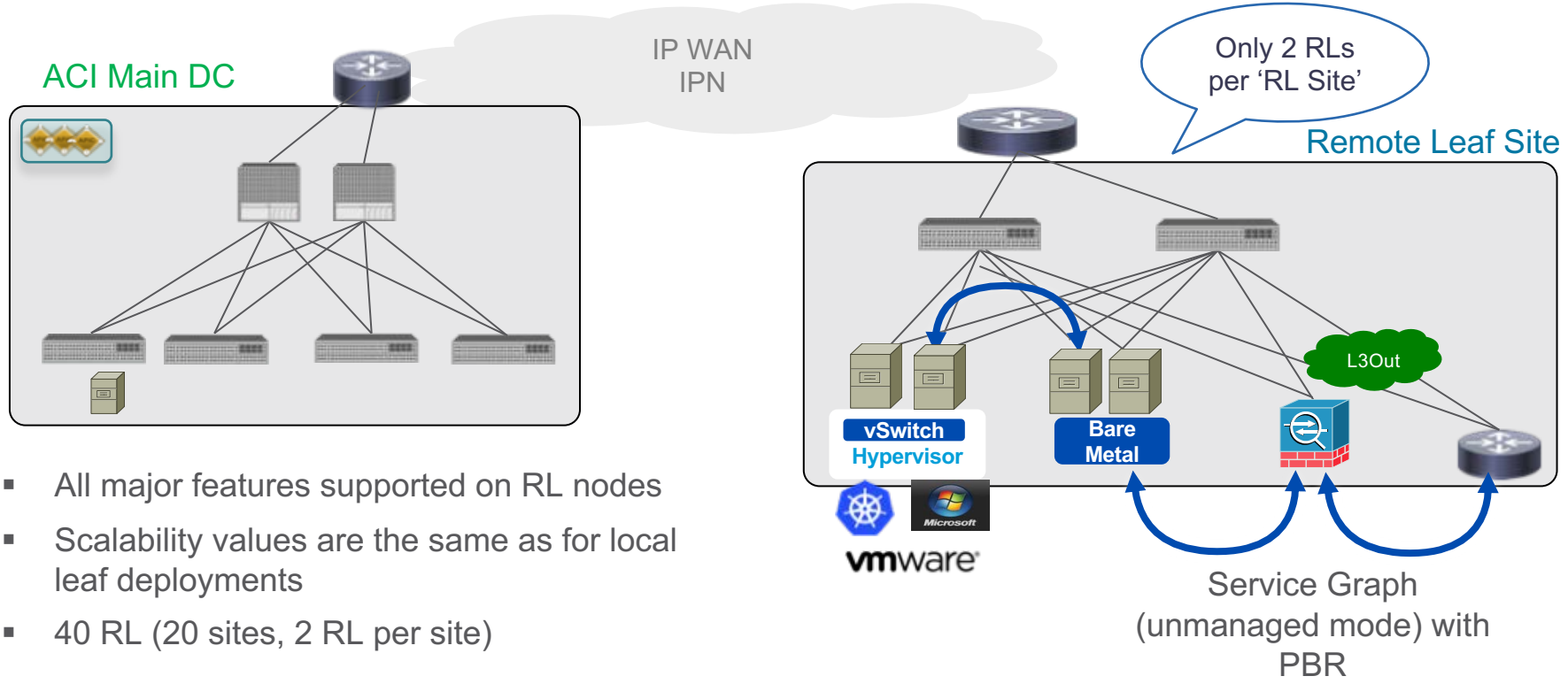
Modular
- N9732C-EX
- N9736C-FX

Remote Leaf Nodes

Supported Leaf
- N93180YC-EX
- N93108TC-EX
- N93180LC-EX
- N93180YC-FX
- N9K-C93108TC-FX
- N9K-C9348GC-FXP

ACI Remote Physical Leaf
Functionalities and Scale

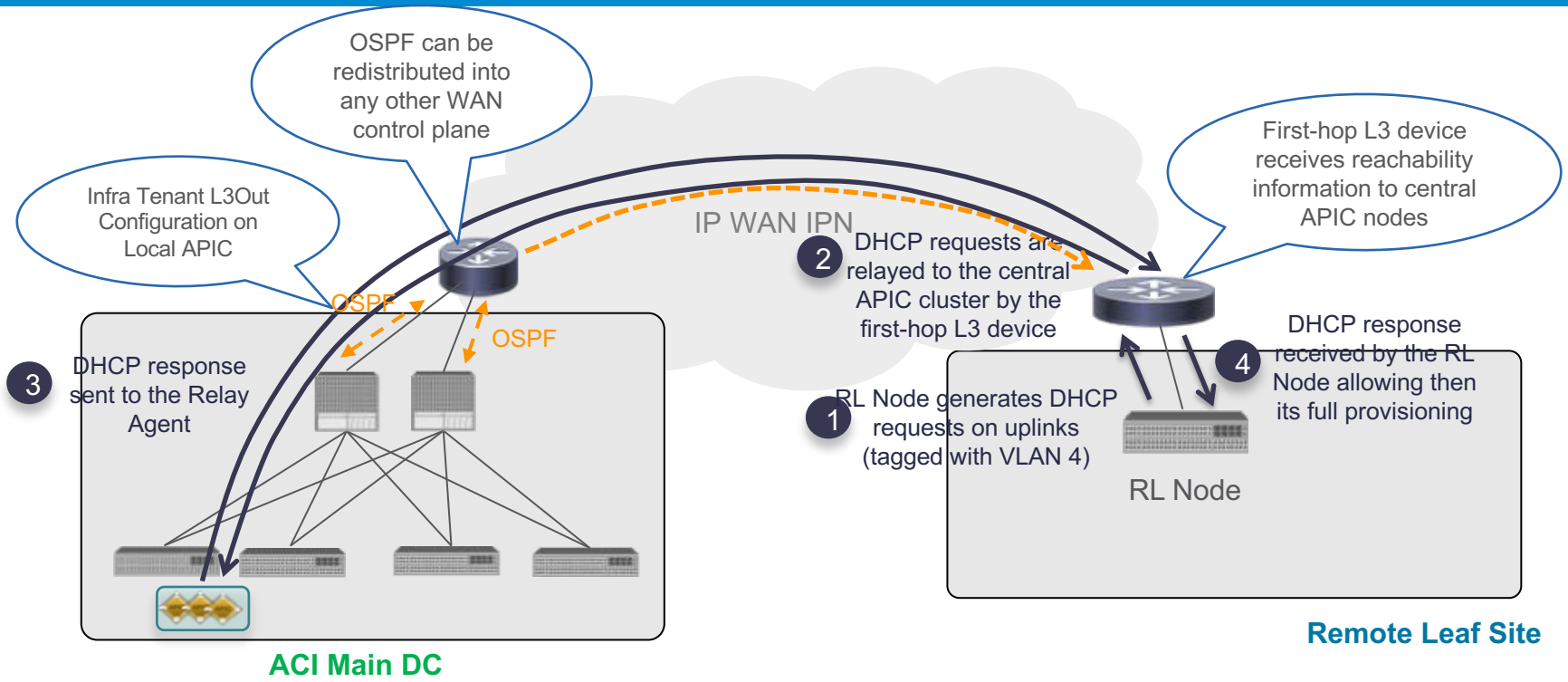
Cisco Connect



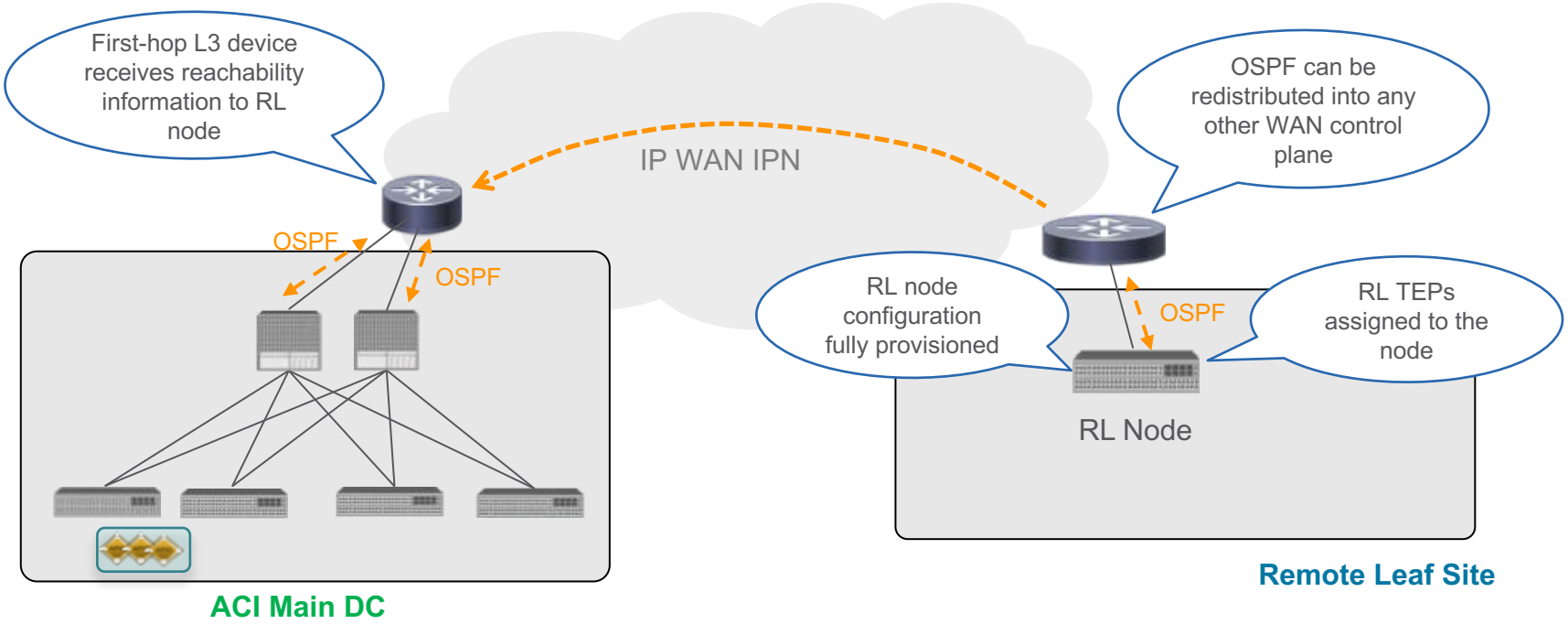
- All major features supported on RL nodes
- Scalability values are the same as for local leaf deployments
- 40 RL (20 sites, 2 RL per site)

ACI Remote Physical Leaf
Automatic RL Discovery

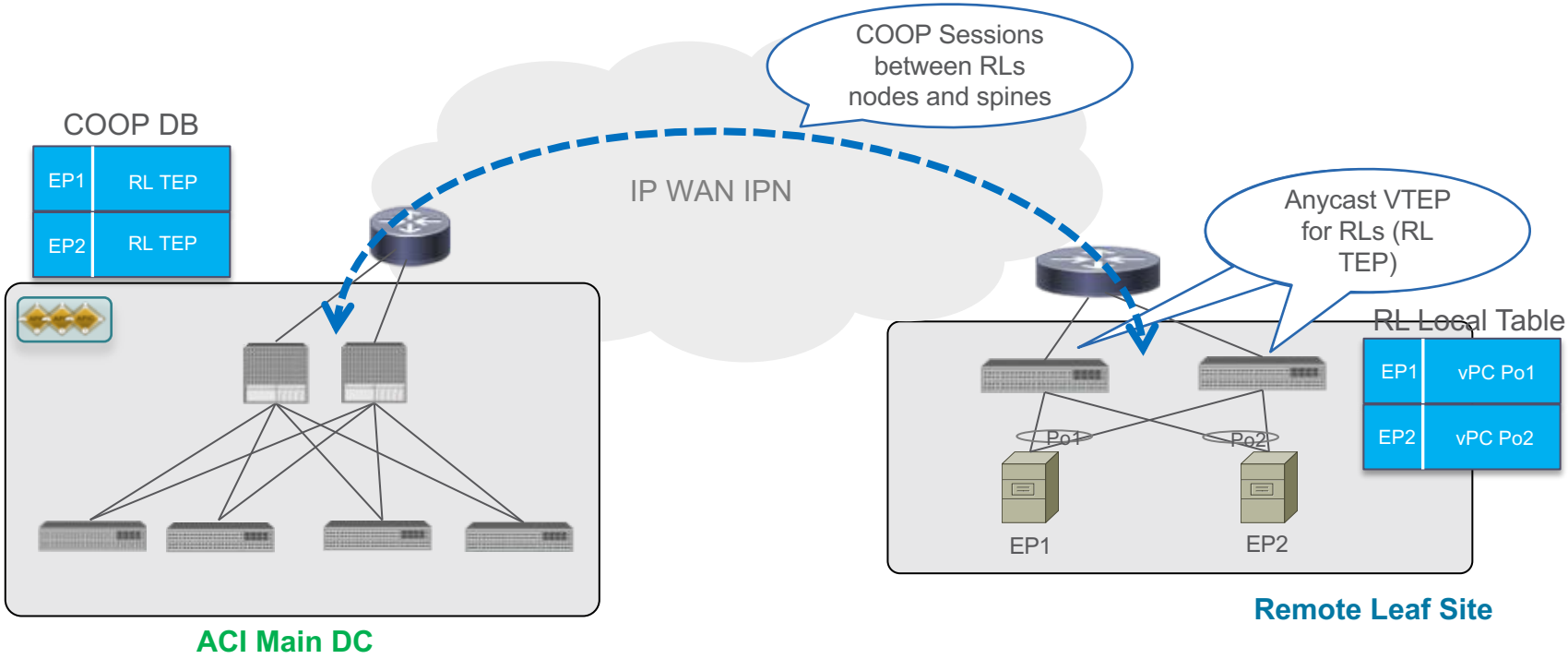
Cisco Connect



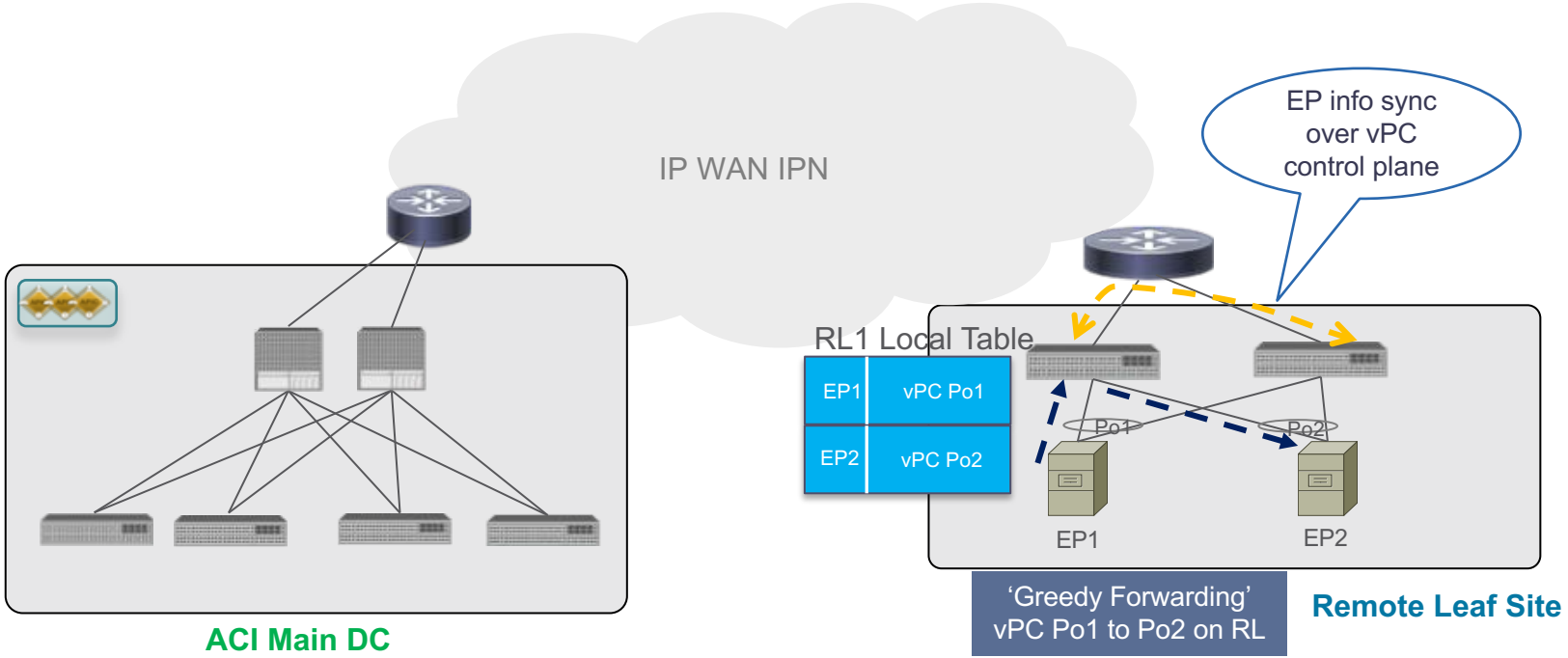
ACI Remote Physical Leaf
Establishing End-to-End IP Connectivity



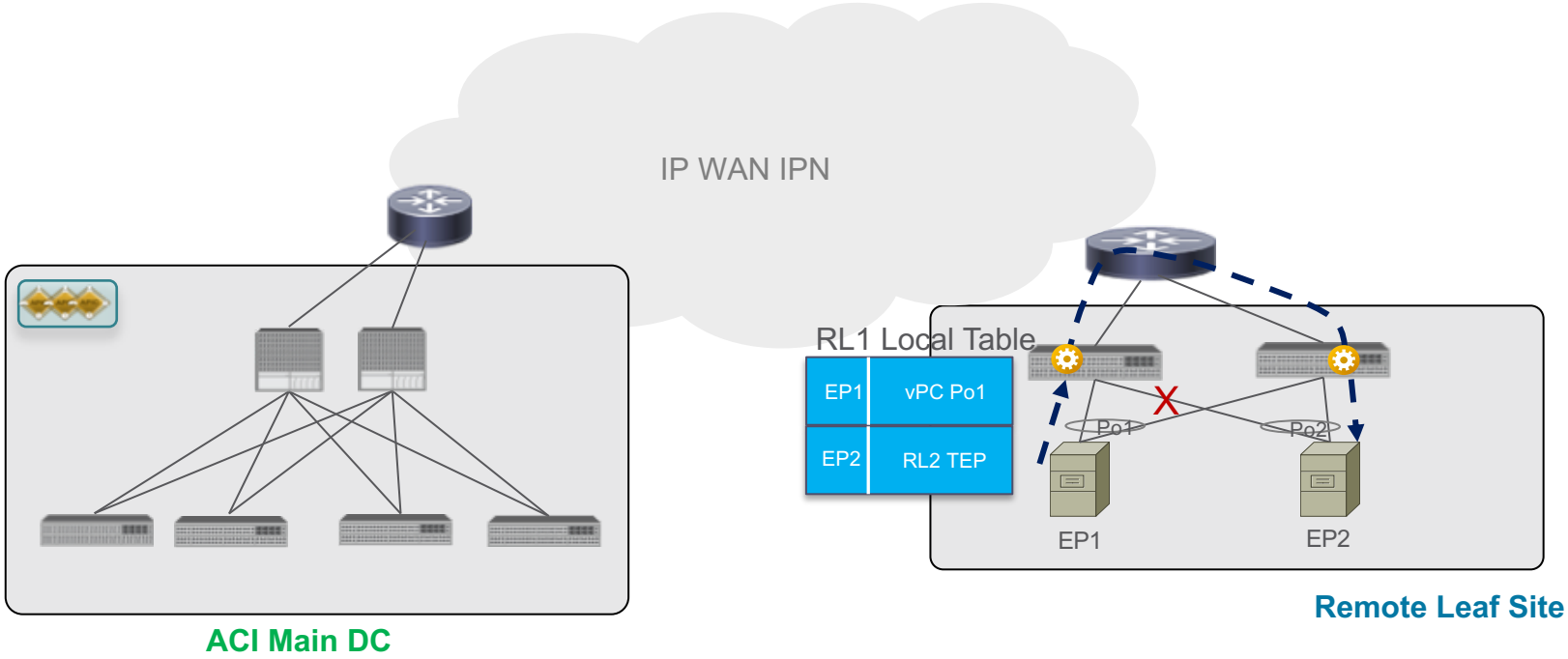
ACI Remote Physical Leaf
COOP for Announcing Remote Endpoint Information



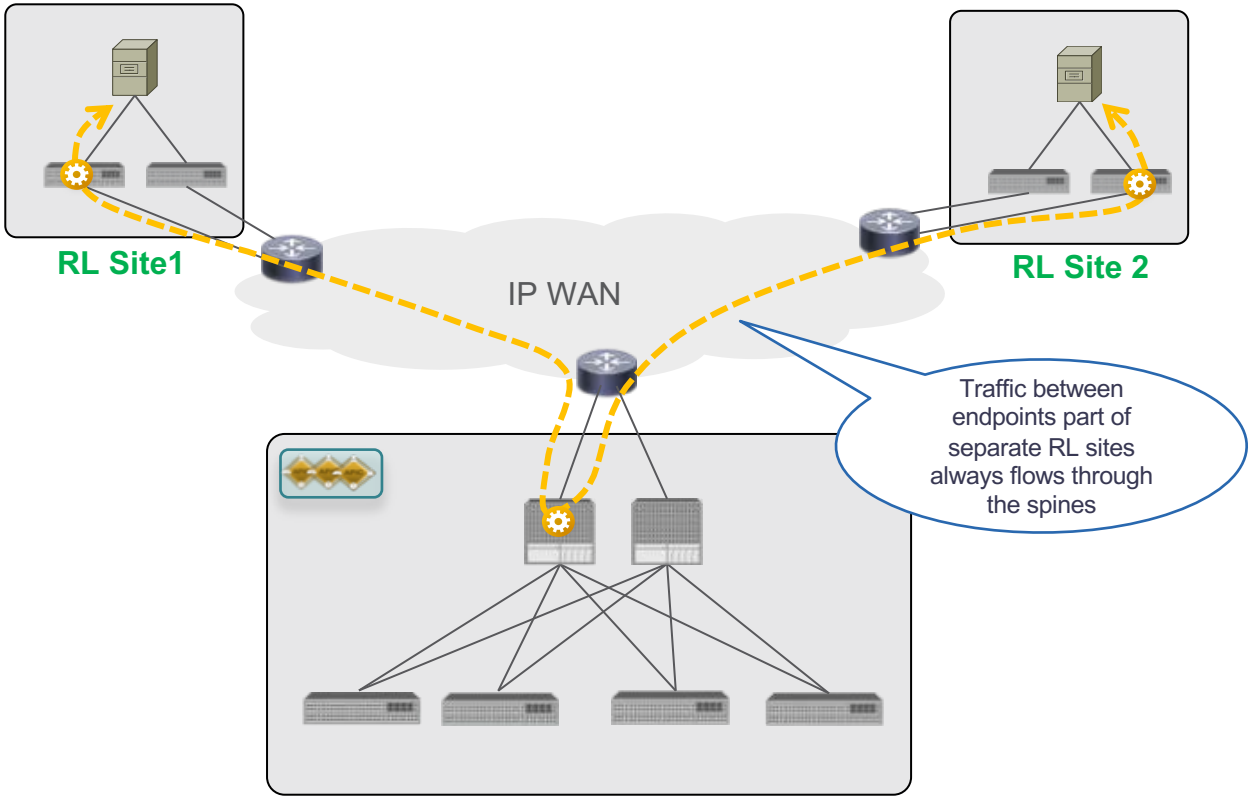
ACI Remote Physical Leaf
EP-to-EP Flow Local to RL Site (vPC)



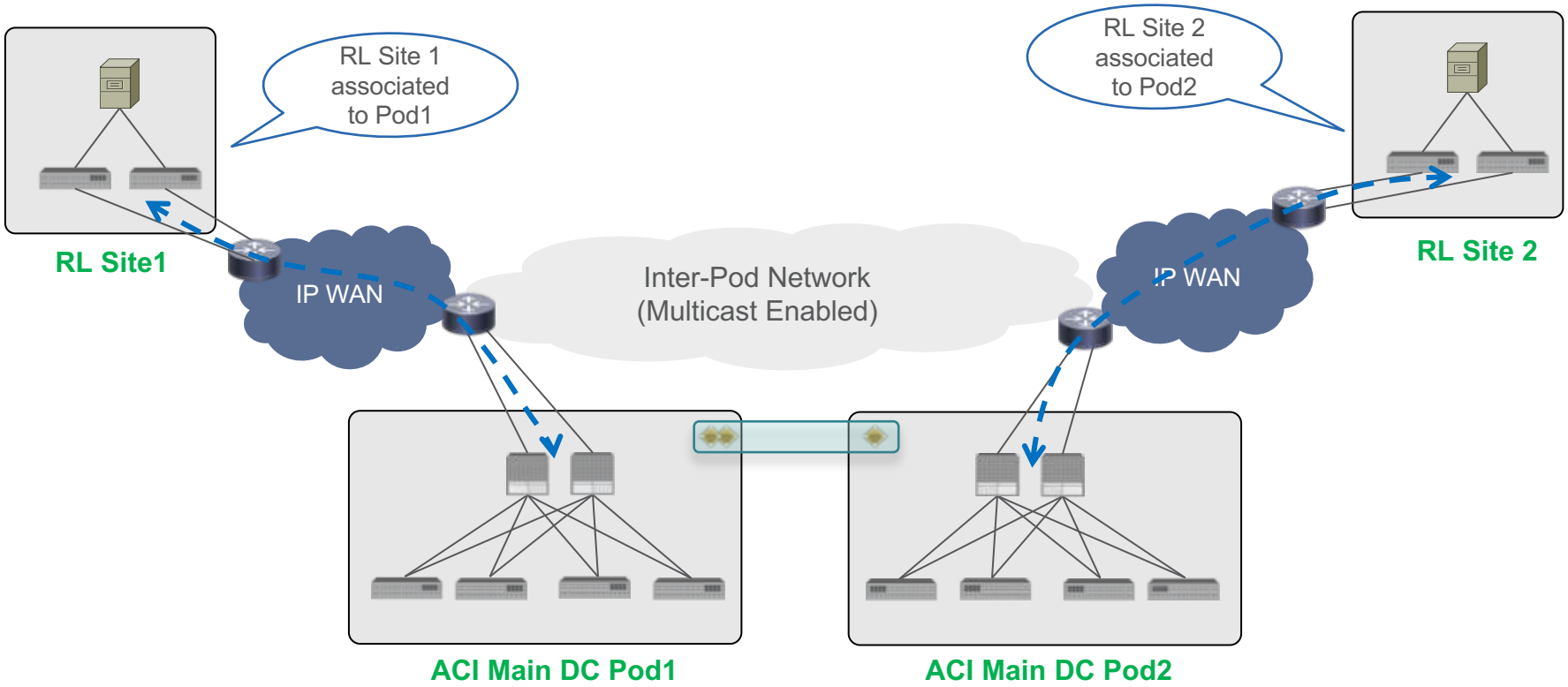
ACI Remote Physical Leaf
EP-to-EP Flow Local to RL Site (vPC Link Failure)



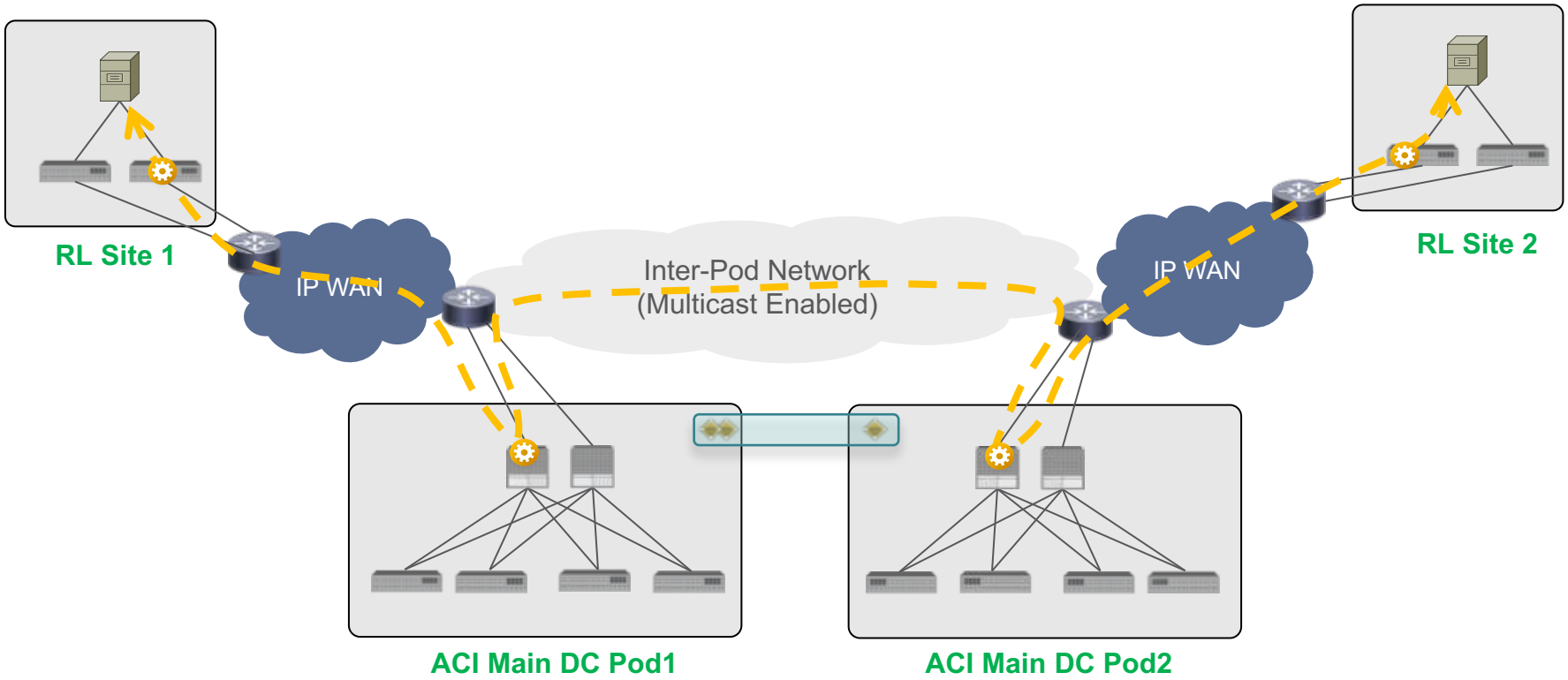
ACI Remote Physical Leaf
Communication between Endpoints in Separate RL Sites



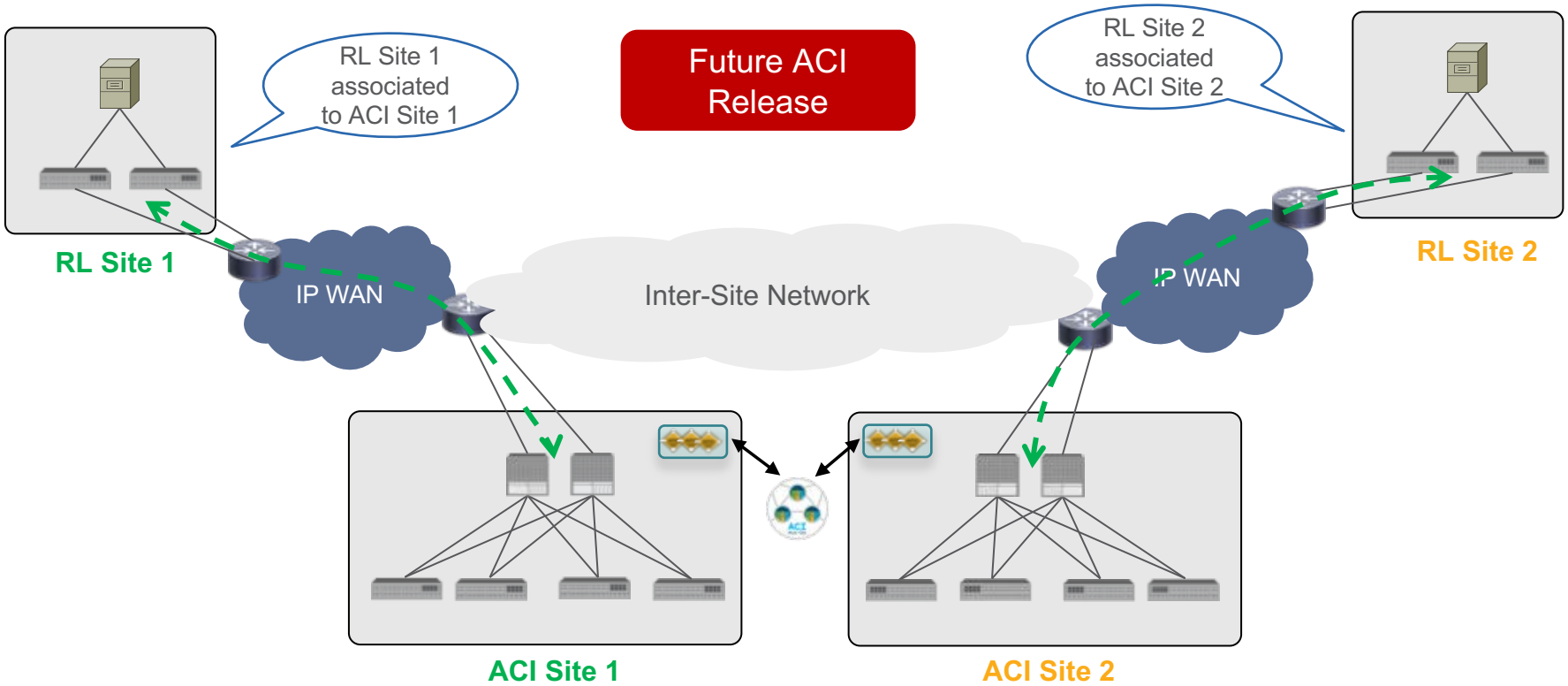
ACI Remote Physical Leaf and Multi-Pod
RL Sites Can Be Associated to Separate Pods



ACI Remote Physical Leaf and Multi-Pod
RL Sites Can Be Associated to Separate Pods (Data Plane)



ACI Remote Physical Leaf and Multi-Site
RL Sites Can Be Associated to Separate Pods

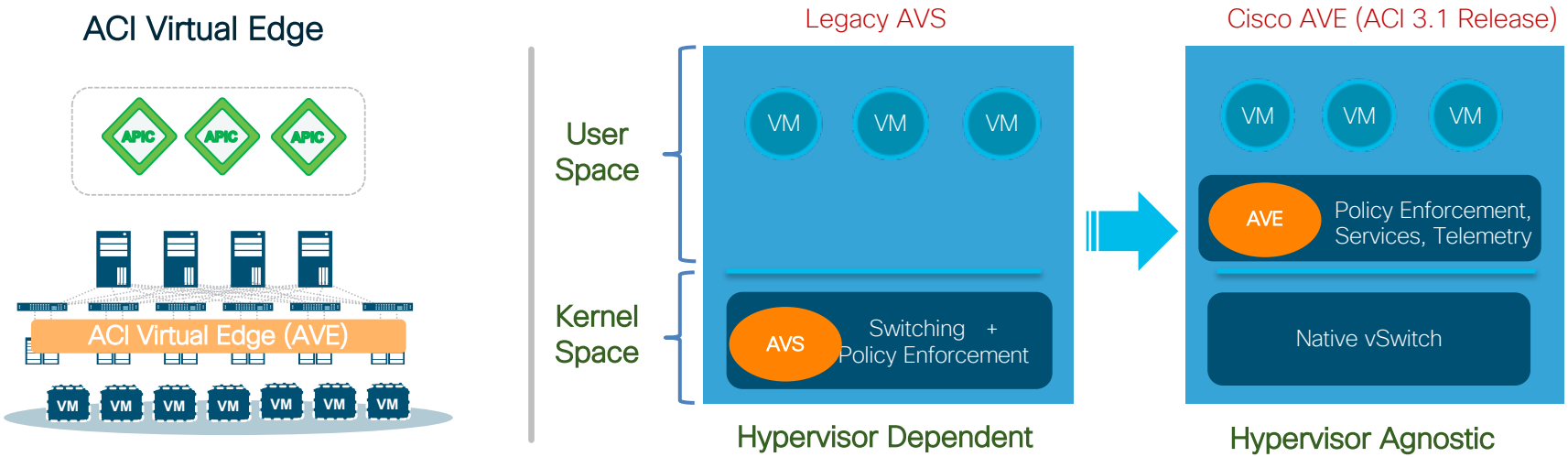


Cisco
Connect

ACI Remote Virtual Leaf (vPod)

Cisco ACI Virtual Edge
Decoupled From Hypervisor Kernel API Dependencies

Cisco Connect



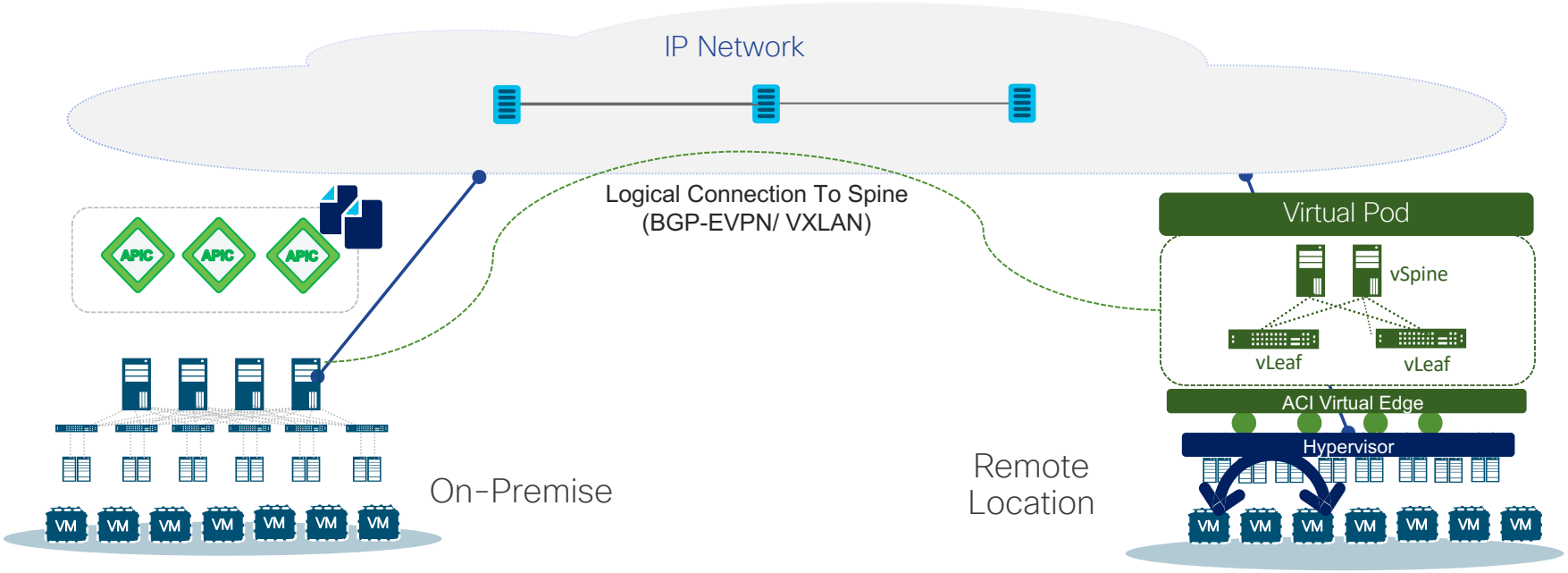
Maintain Existing Operational Models

Simple Transition/Migration AVS => AVE

Policy Consistency Across Multiple Hypervisors

AVS/AVE Feature Parity

Cisco ACI Virtual Pod (vPod)
Extend ACI To Bare-Metal Cloud



Bare Metal Clouds
(IBM BlueMix, AWS Elastic Metal etc.)

Remote Data Centers

Colo Facilities
(Equinix, CoreSite etc.)

BrownField Deployments

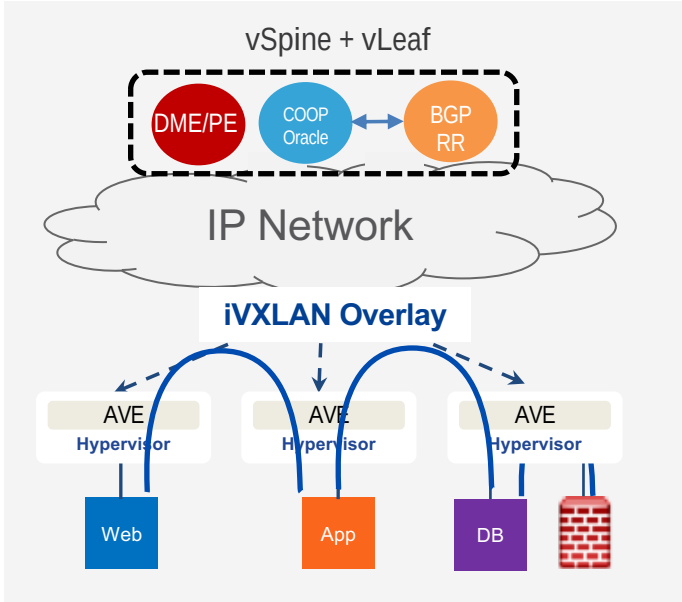
vPod Functional Components
vSpine, vLeaf, and AVE

vSpine + vLeaf

- Run as container services inside VMs at the vPod location (co-located for availability)
- **vLeaf:** Distribute APIC policies to AVE forwarders (DME/PE)
- **vSpine**: Centralized endpoint and LPM database (COOP and BGP)
- *Not* in forwarding data path

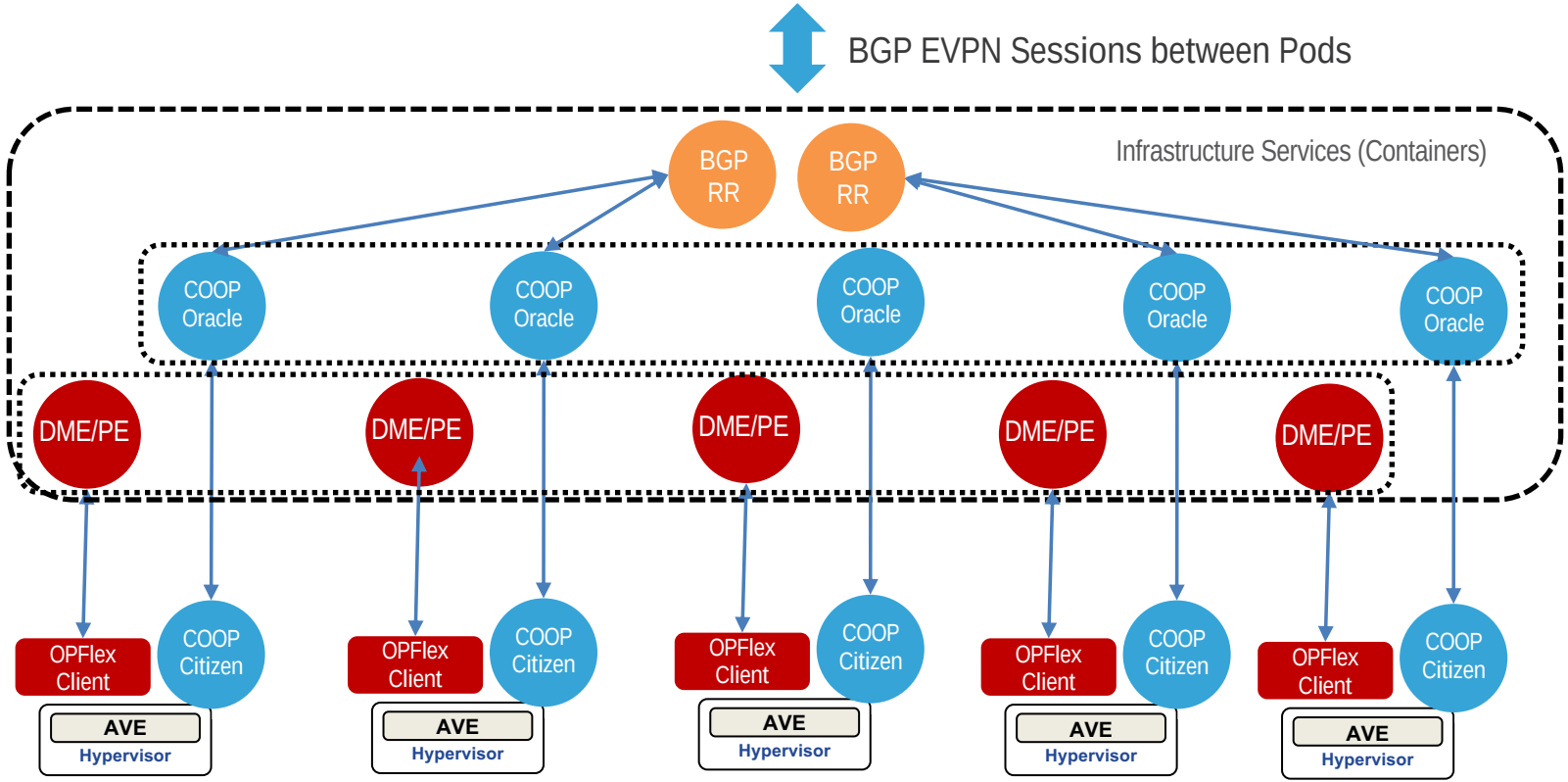
AVE

- Implements ACI data path functions
- Use iVXLAN for communication within Remote site as well as between the vPod and other Pods



vSpine, vLeaf, and AVE
Scale Out

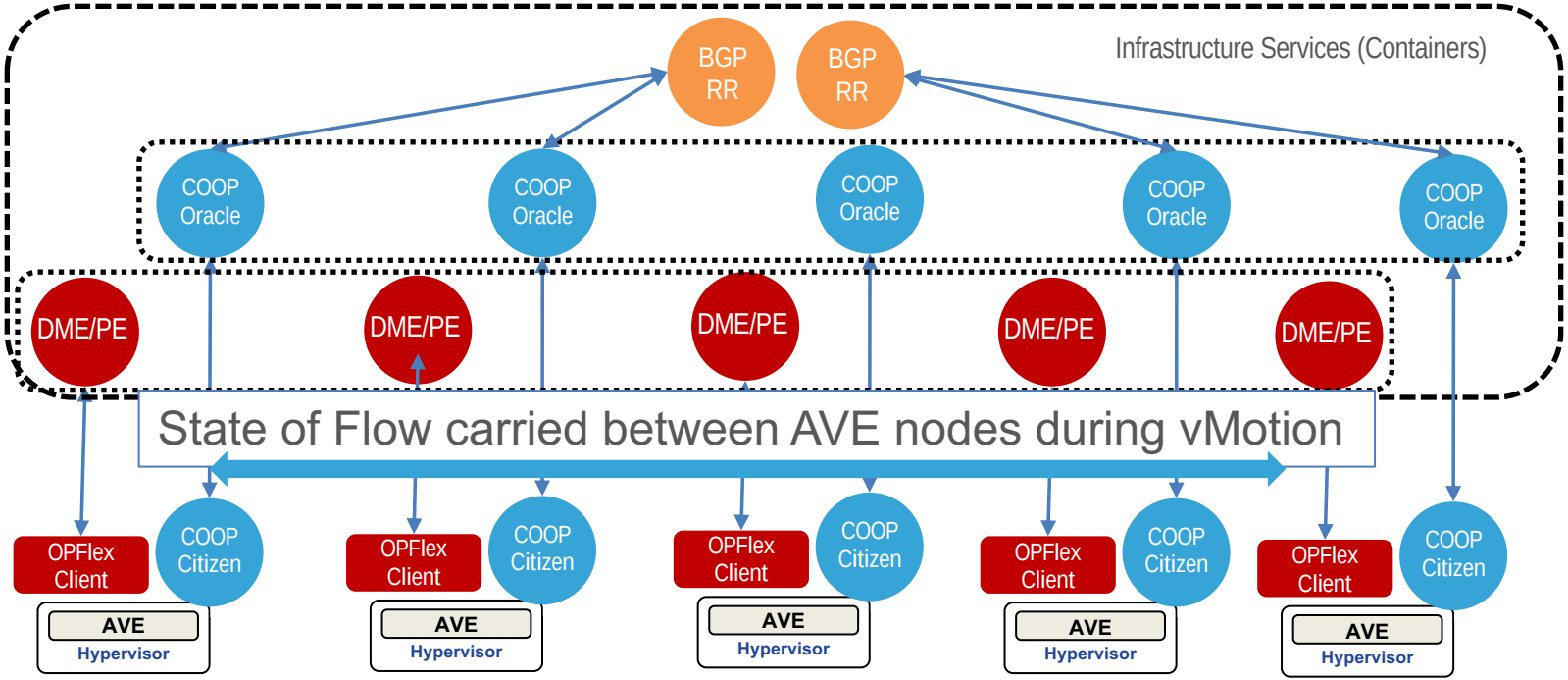
Cisco Connect



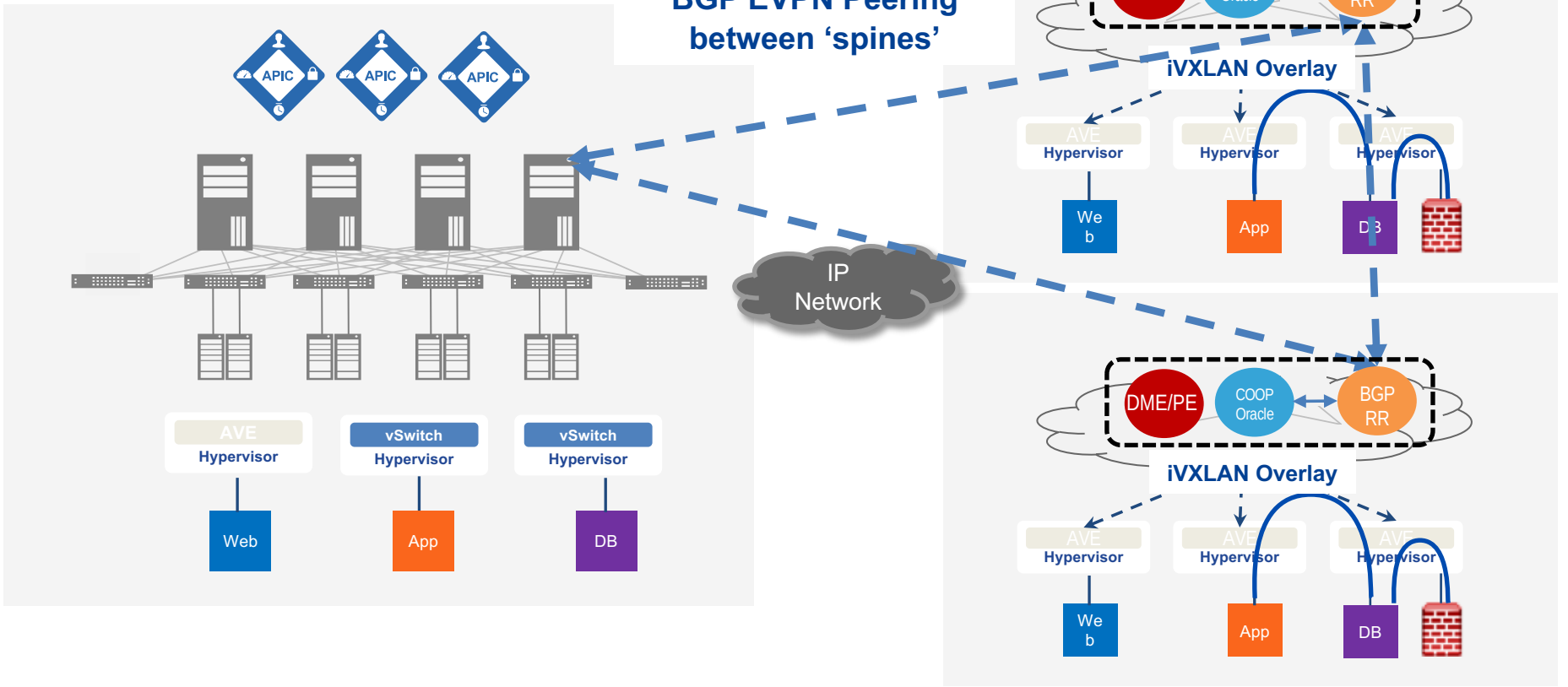
254 AVE per logical vPod

vSpine, vLeaf, and AVE
Flow State and vMotion

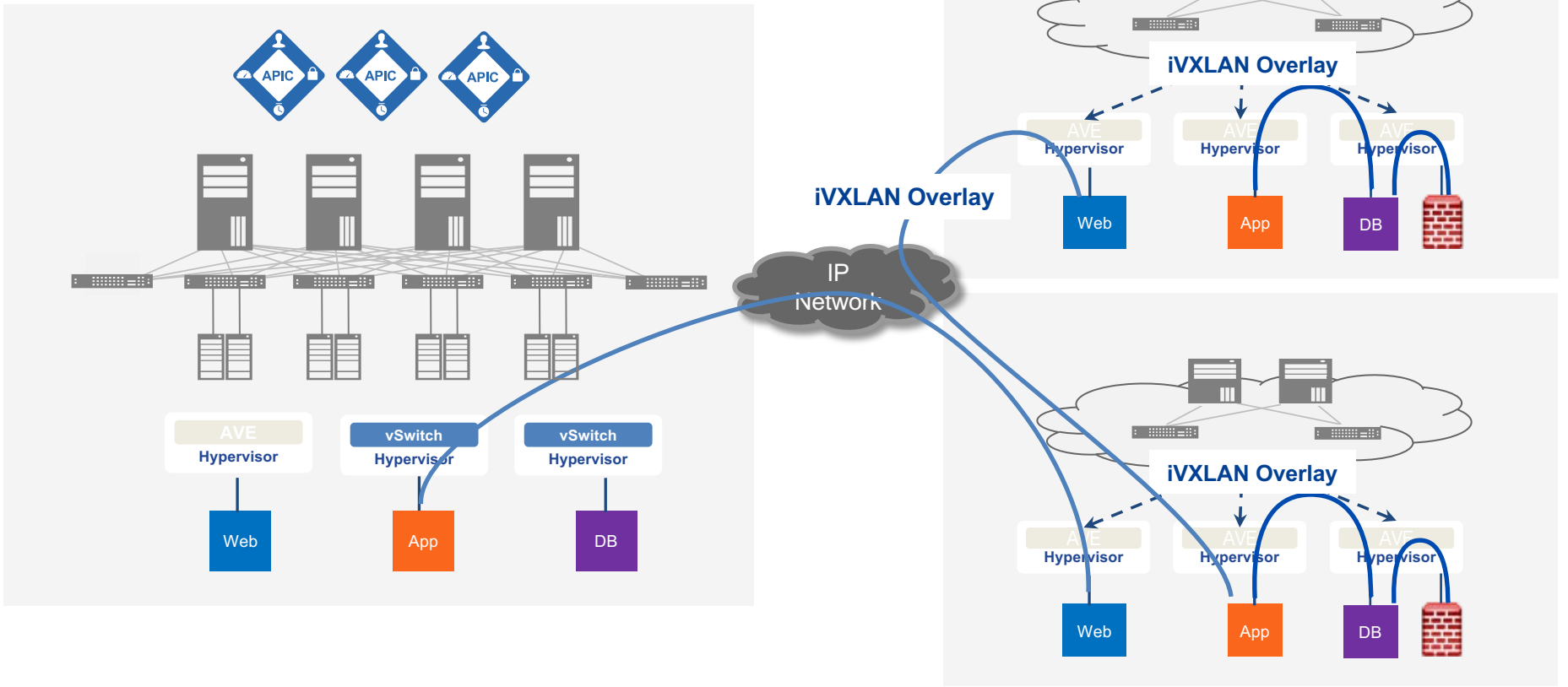
Cisco Connect



vPod
Inter-Pod Control Plane



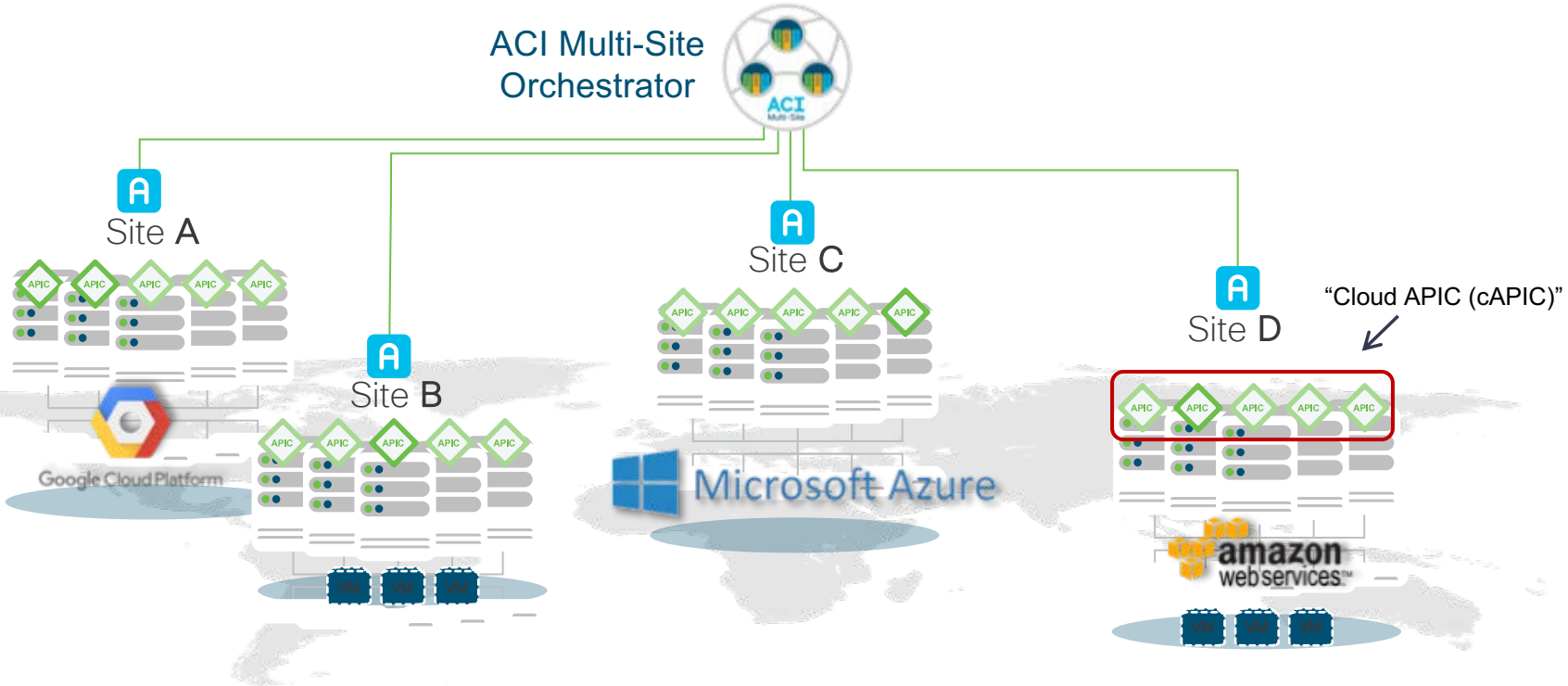
vPod
Inter-Pod Data Plane



Cisco
Connect

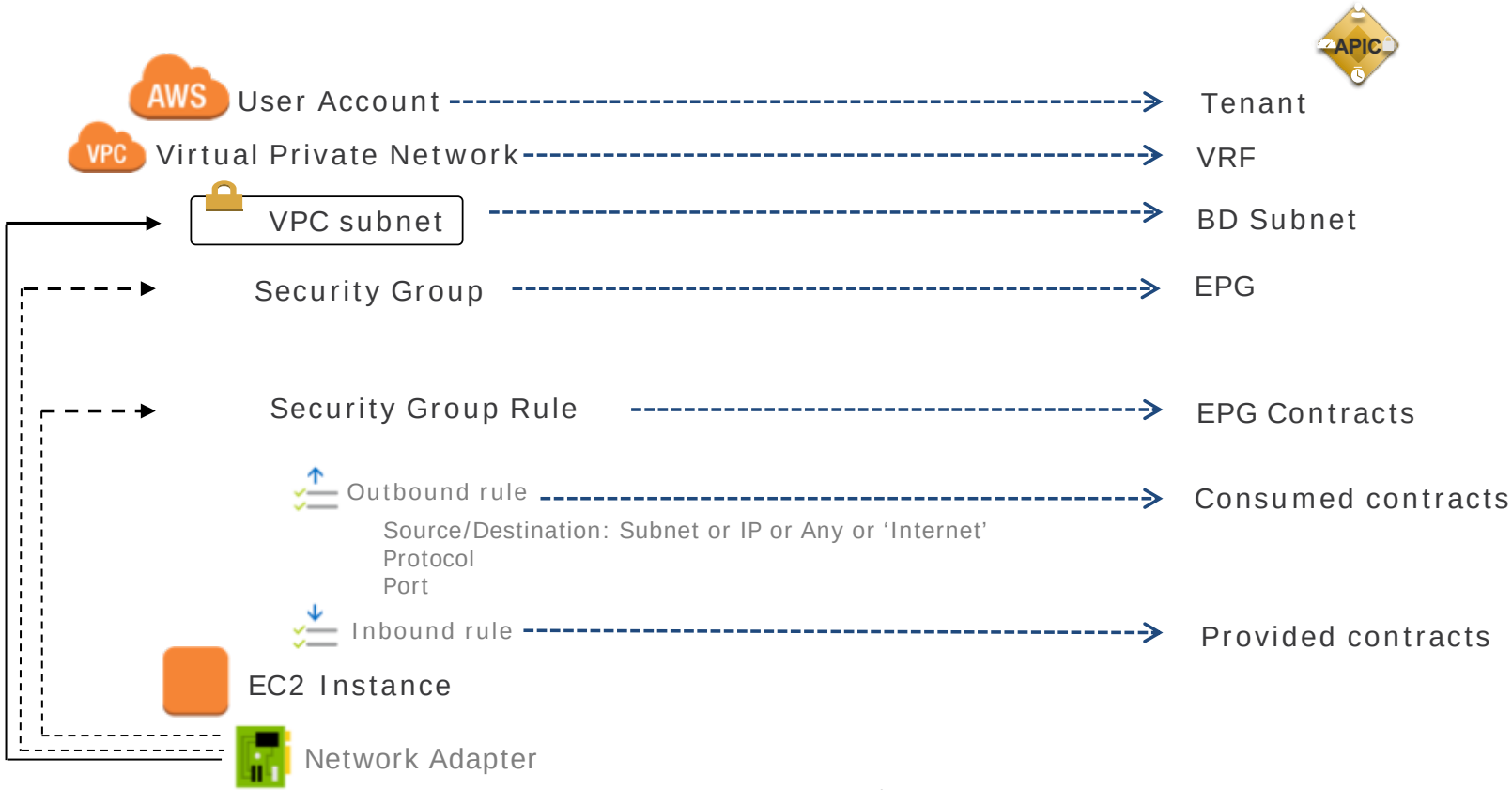
ACI Extensions to Multi-Cloud

ACI Extensions To Multi-Cloud

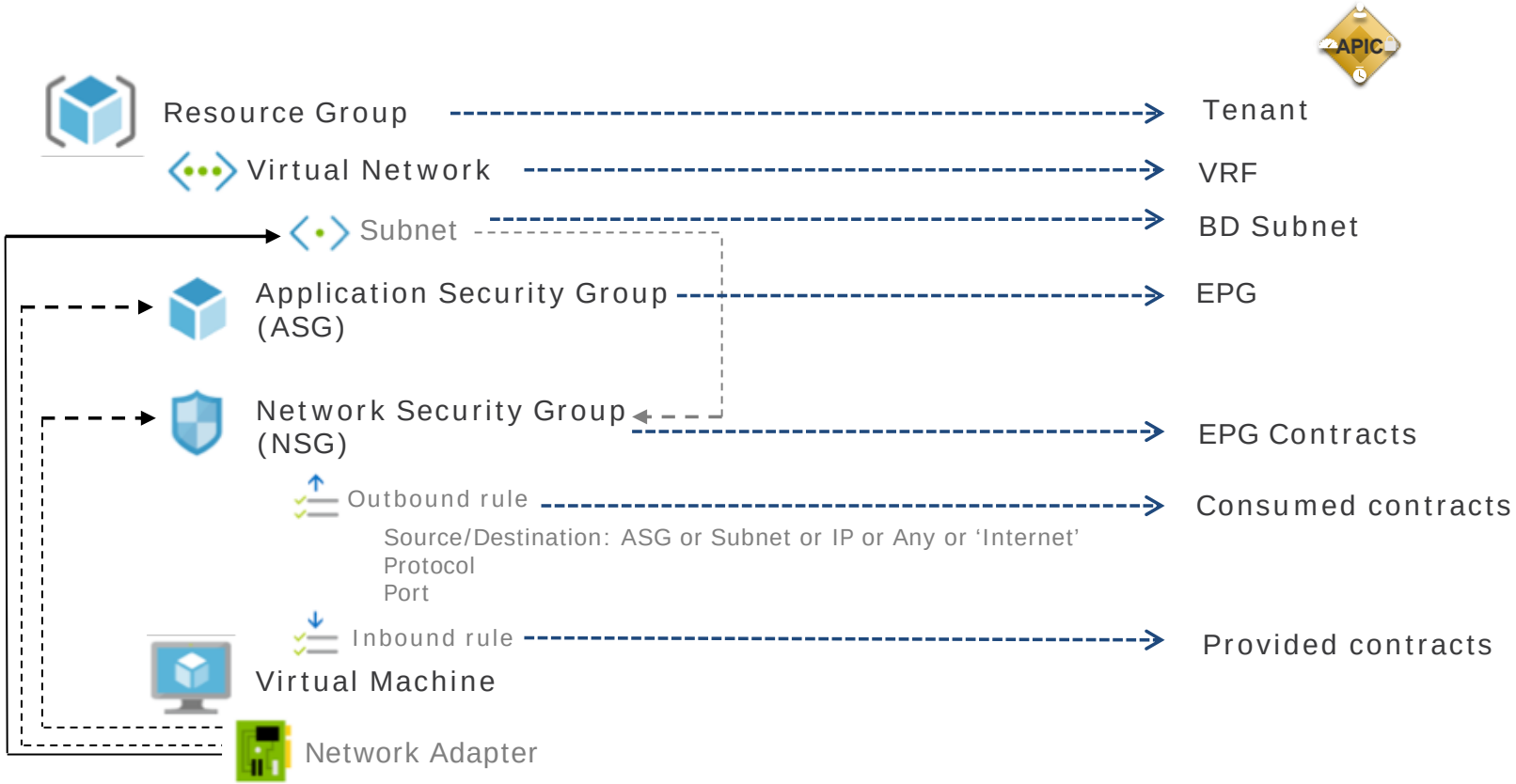


- Consistent Network and Policy across clouds
- Common Governance
- Single Point of Orchestration
- Secure Automated Connectivity

cAPIC and Policy Mapping
AWS Cloud Constructs



cAPIC and Policy Mapping
Azure Cloud Constructs



ACI Anywhere
Where to Go for More Information

Cisco Connect

- ✓ 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-Site White Paper

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

- 

✓ ACI Multi-Pod Cisco Live Barcelona 2018

  https://www.ciscolive.com/global/on-demand-library/?search=weston#/session/BRKACI-2003

- 

✓ ACI Multi-Site Cisco Live Barcelona 2018

  https://www.ciscolive.com/global/on-demand-library/?search=ardica#/session/BRKACI-2125

- ✓ ACI Physical Remote Leaf White Paper

  Coming soon!


CISCO