

CISCO *Live!*



#CiscoLive



The bridge to possible

ACI Multi-Site Architecture and Deployment

Part 1

Max Ardica, Distinguished Engineer

@maxardica

BRKDCN-2480a



CISCO *Live!*

#CiscoLive

Cisco Webex App

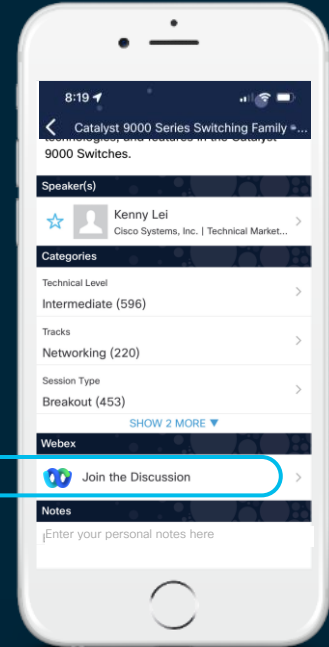
Questions?

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

How

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

Webex spaces will be moderated by the speaker until June 17, 2022.



<https://ciscolive.ciscoevents.com/ciscolivebot/#BRKDCN-2480a>

Session Objectives



- At the end of the session, the participants should be able to:
 - ✓ Articulate the different deployment options to interconnect Cisco ACI networks (Multi-Pod and Multi-Site) and when to choose one vs. the other
 - ✓ Understand the functionalities and specific design considerations associated to the ACI Multi-Site architecture
- Initial assumption:
 - ✓ The audience already has a good knowledge of ACI main concepts (Tenant, BD, EPG, L2Out, L3Out, etc.)



Agenda

- Introduction
- Nexus Dashboard Orchestrator (NDO) Architecture
- Provisioning Policies on NDO
- Inter-Site Connectivity Deployment Considerations
- ACI Multi-Site Control and Data Plane
- Connecting to the External L3 Domain
- Network Services Integration

BRKDCN-2480b

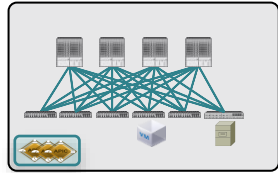
Introduction



ACI Anywhere

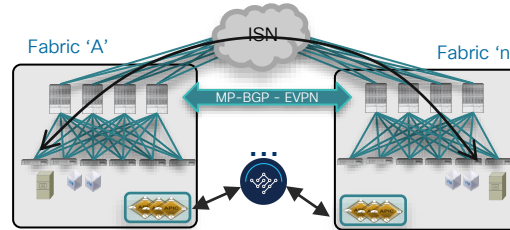
Fabric and Policy Domain Evolution

ACI Single Pod Fabric



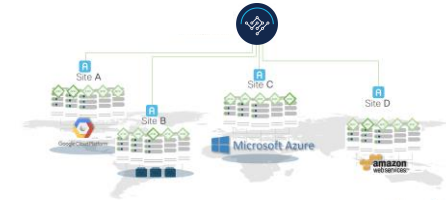
ACI 2.0 - Multiple Networks (Pods or Availability Zones) in a single Fabric (Region)

ACI Multi-Site



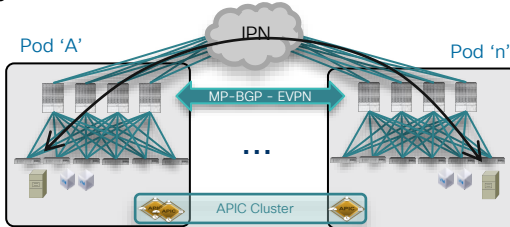
ACI 3.1/4.0 - Remote Leaf and vPod extends a Fabric to remote locations

Cloud ACI



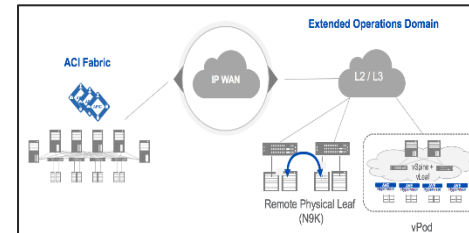
ACI 1.0 - Leaf/Spine Single Pod Fabric

ACI Multi-Pod Fabric



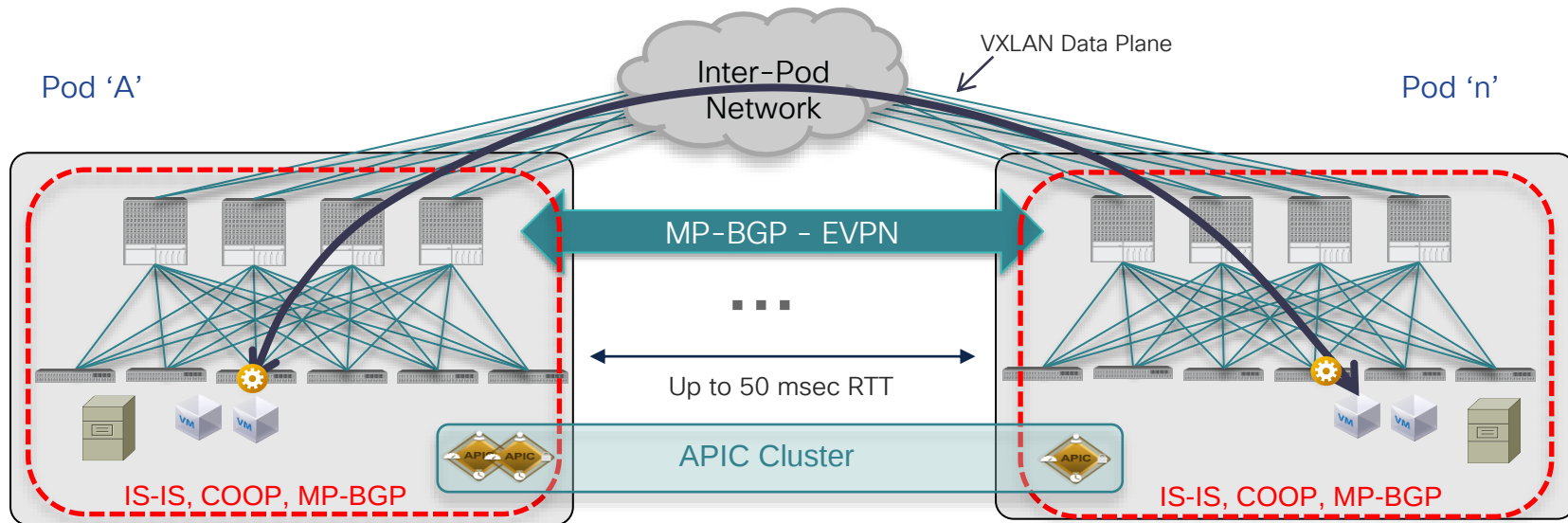
ACI 3.0 - Multiple Fabrics (Regions) interconnected in the same Multi-Site Orchestrator domain

ACI Remote Leaf



ACI 4.1 & 4.2 - ACI Extensions to Public Cloud

ACI Multi-Pod Overview

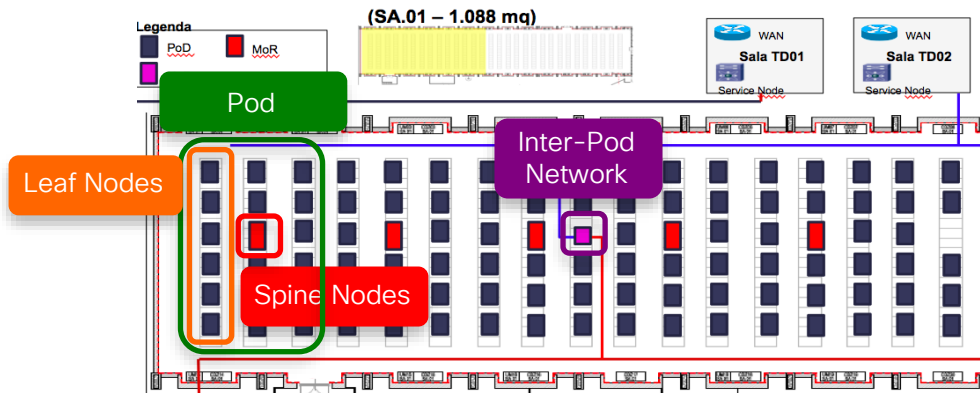


- Multiple ACI Pods connected by an IP Inter-Pod L3 network, each Pod consists of leaf and spine nodes
- Up to 50 msec RTT supported between Pods
- Managed by a single APIC Cluster
- Single Management and Policy Domain
- Forwarding control plane (IS-IS, COOP) fault isolation
- Data Plane VXLAN encapsulation between Pods
- End-to-end policy enforcement

ACI Multi-Pod

Most Common Use Cases

- Need to scale up a single ACI fabric above the number of leaf nodes supported in a single Pod
- Desire to divide an ACI fabric into smaller network fault domains (AZs)
- Handling 3-tiers physical cabling layout (for example traditional N7K/N5K/N2K deployments)

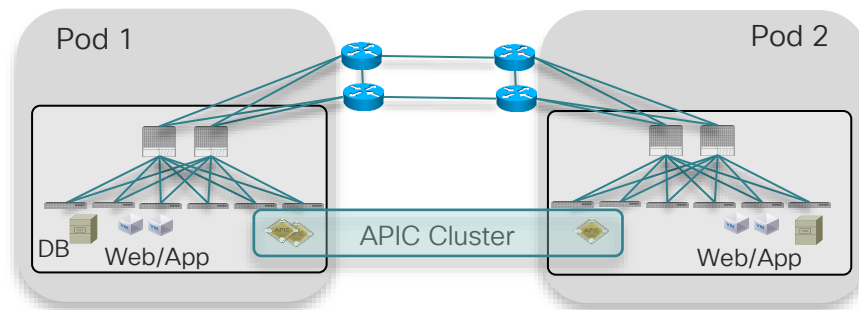


True Active/Active DC deployments

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

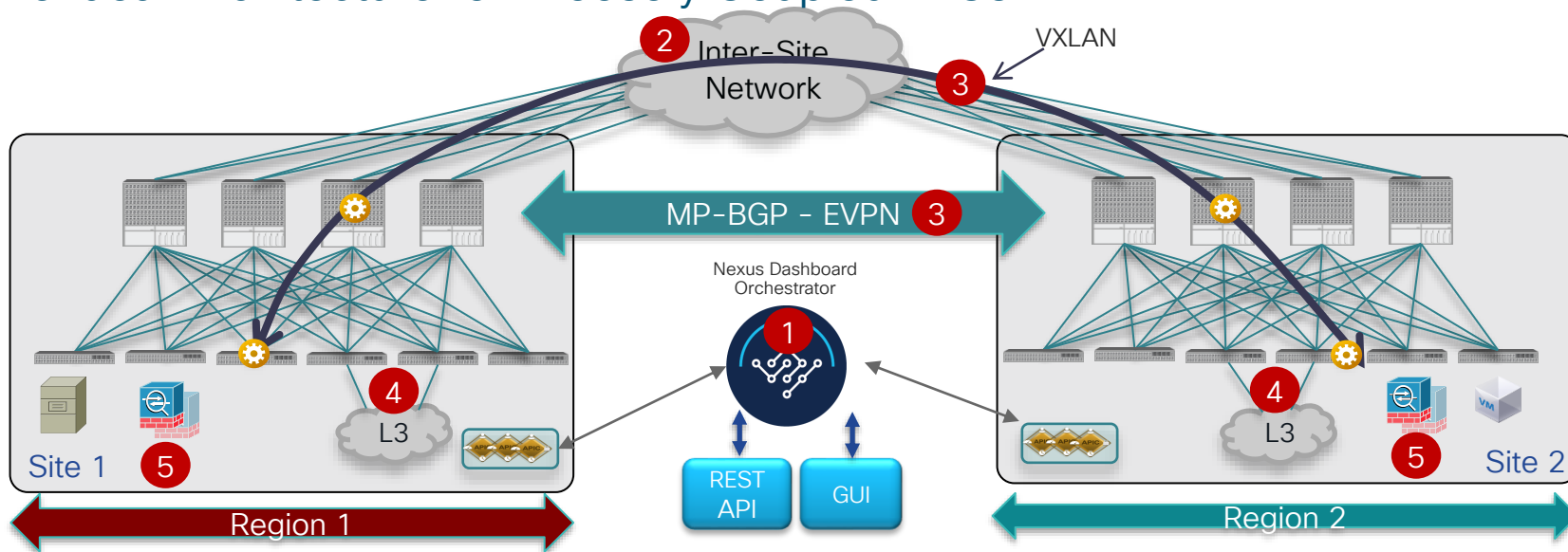
Deployment of Active/Standby or Active/Active clustered network services (FWs, SLBs) across DCs

Application clustering (lots of L2 BUM extension required across Pods)



ACI Multi-Site

The Ideal Architecture for “Loosely Coupled” DCs



- 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

Multi-Pod or Multi-Site?

That is the question...



And the answer is...

BOTH!

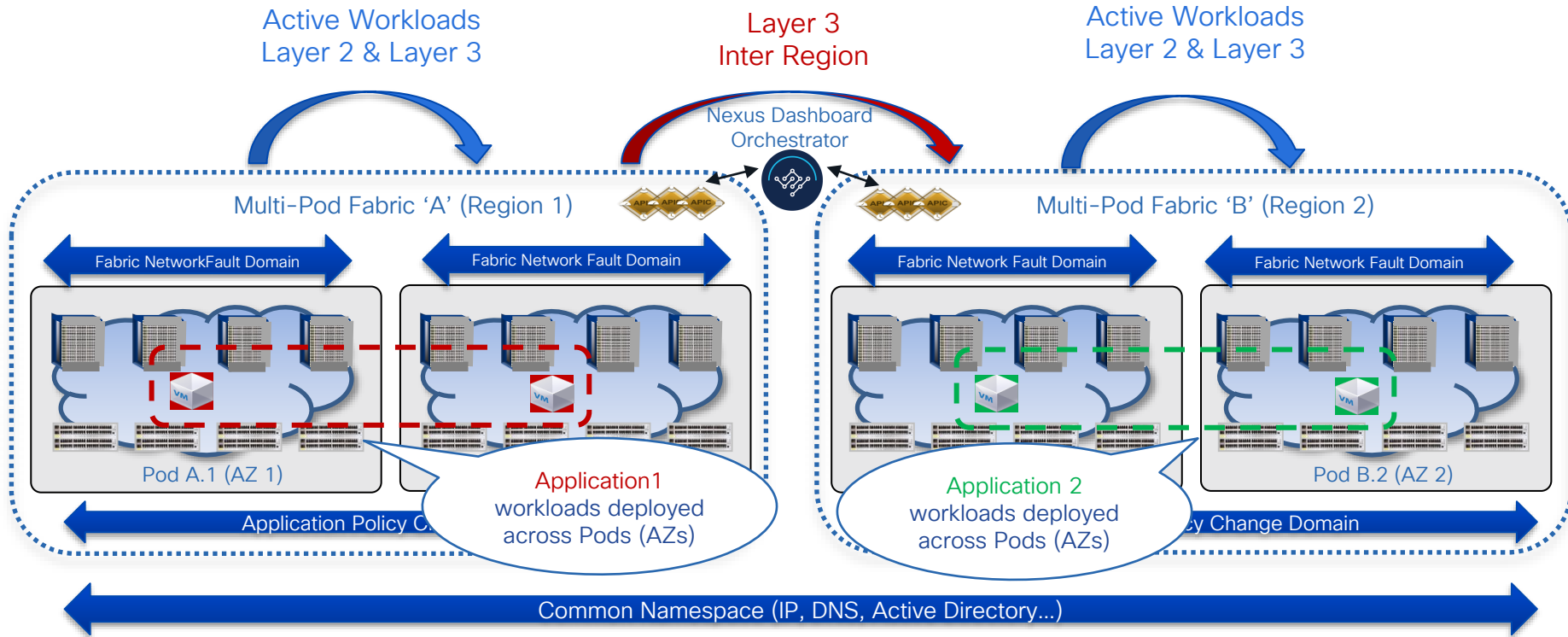


Terminology

- Pod – A Leaf/Spine network sharing a common control plane (ISIS, BGP, COOP, ...)
 - Pod == Availability Zone
- Fabric – Scope of an APIC Cluster, it can be one or more PODs
 - Fabric == Region
- Multi-Pod – Single APIC Cluster with multiple leaf spine networks
 - Multi-Pod == Multiple Availability Zones within a Single Region (Fabric)
- Multi-Fabric – Multiple APIC Clusters + associated Pods (you can have Multi-Pod with Multi-Fabric)*
 - Multi-Fabric == Multi-Site == a DC infrastructure with multiple regions

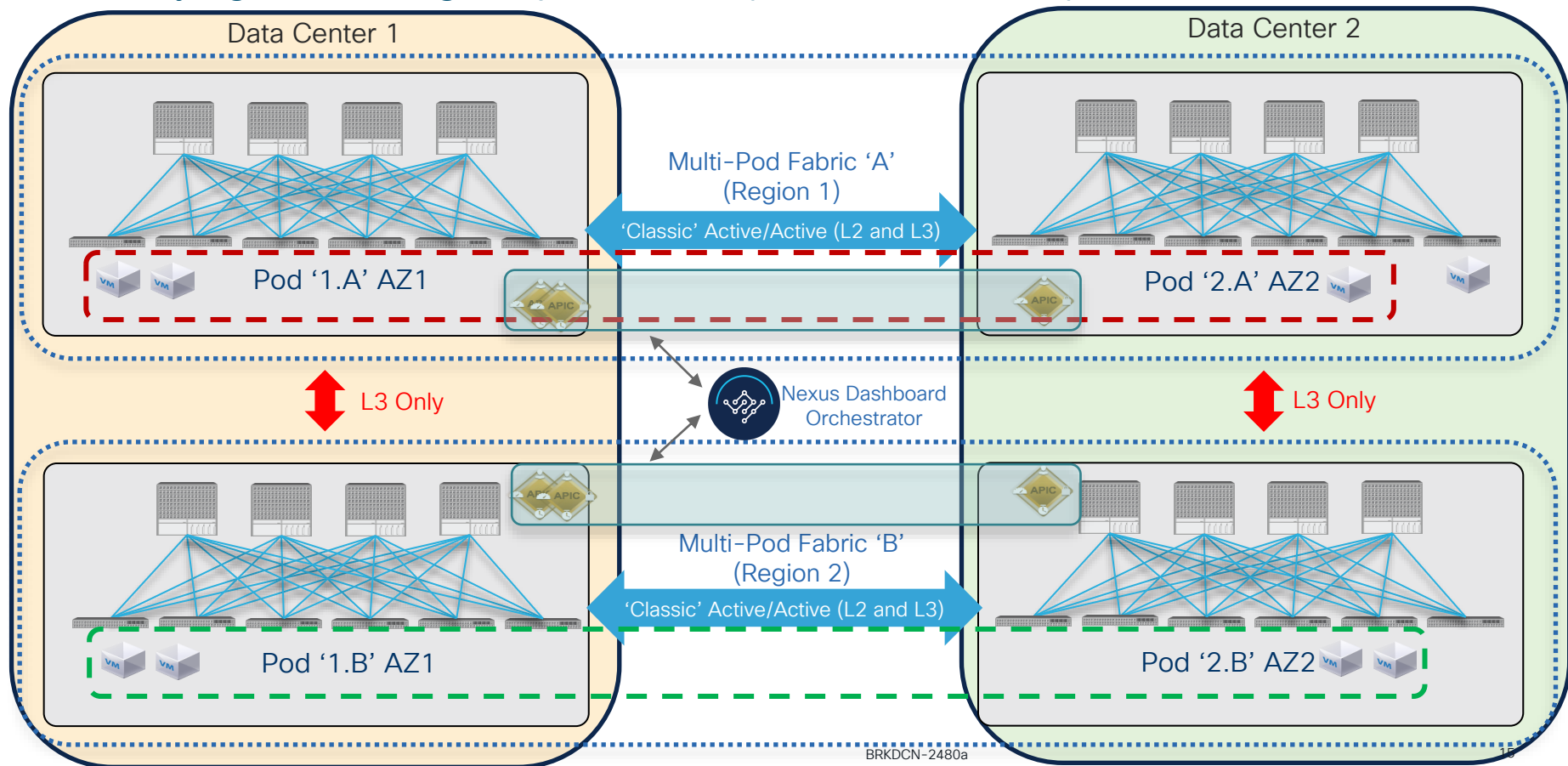
Systems View (How do these things relate)

Change and Network Fault Domain Isolation



Multi-Pod + Multi-Site

Satisfying Conflicting Requirements (A/A DCs and DR)

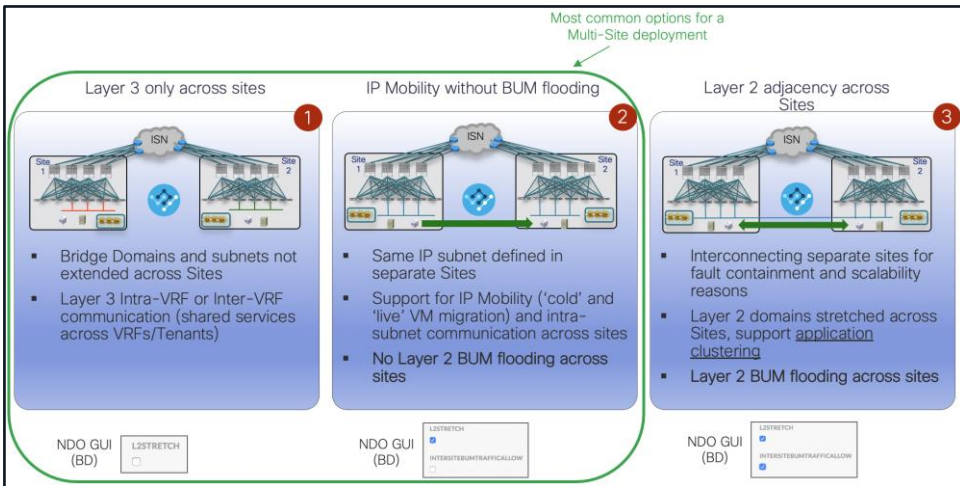


But wait! Couldn't I
deploy Multi-Site also to
handle more typical
Multi-Pod use cases?



Multi-Site for Active/Active Application Deployments?

Multi-Site for Active/Active Application Deployments



- ACI Multi-Site allows to extend connectivity and policies between separate APIC domains
 - Layer 3 only across sites
 - Layer 2 with and without BUM flooding
- Keep in mind some specific considerations before deploying Multi-Site for “classic” Active/Active application deployments (i.e. same application components deployed across sites)
 - Loss of change and network fault domain isolation across separate ACI domains
 - Creation of separate VMM domains by design (loss of intra-cluster functionalities like DRS, vSphere FT/HA, ...)
 - Specific service node insertion deployment considerations (use of separate service nodes per fabric, limited support for service nodes clustering across sites, no support for vAny + PBR, ...)

ACI Multi-Site

Most Common Use Cases

- Scale-up DC network model

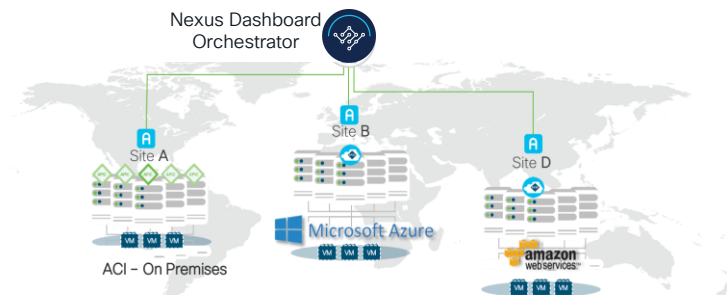
Building a very large intra-DC network (above 500* leaf nodes)



*500 leaf nodes supported in a Multi-Pod fabric starting with ACI 4.2(4) release

- Cloud ACI

Integration between on-prem and public clouds (AWS, Azure, GCP)



- Data Center Interconnect (DCI)

Extend connectivity/policy between 'loosely coupled' DC sites

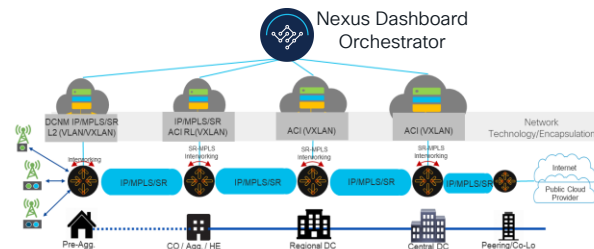
Disaster Recovery and IP mobility use cases



- SP 5G Telco DC/Cloud

Centralized DC Orchestration for "Isolated Fabrics"

SR-MPLS/MPLS Handoff on Border Leaf nodes



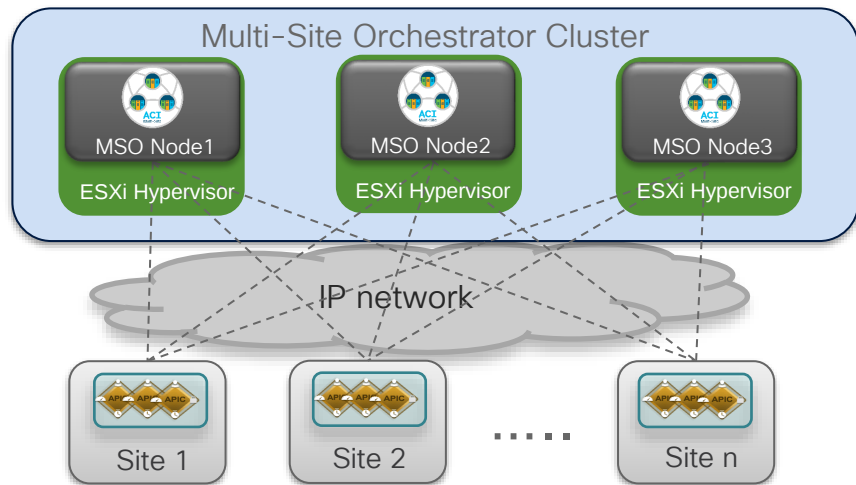
Nexus Dashboard Orchestrator (NDO) Architecture





Original Multi-Site Orchestrator Option

VM Based MSO Cluster (OVA)



- Supported from the beginning (MSO release 1.0(1))
- Each Cisco Multi-Site Orchestrator node is packaged in a VMware vSphere virtual appliance (OVA)
- For high availability, you should deploy each Cisco 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), 48 Gbps of memory, and 100 GB of disk space
- MSO 3.1(1) is the last supported release with this form factor

Cisco Multi-Site Orchestrator has become Cisco Nexus Dashboard Orchestrator



Up to release 3.1(1)



From release 3.3(1)

Cisco Nexus Dashboard

Simple to Automate, Simple to Consume

Powering **automation**
Unified **agile** platform



Cisco Nexus
Dashboard

Insights

Fabric Discovery

Fabric Controller

Orchestrator

Data Broker

SAN Controller

Consume all services in one place



Private cloud

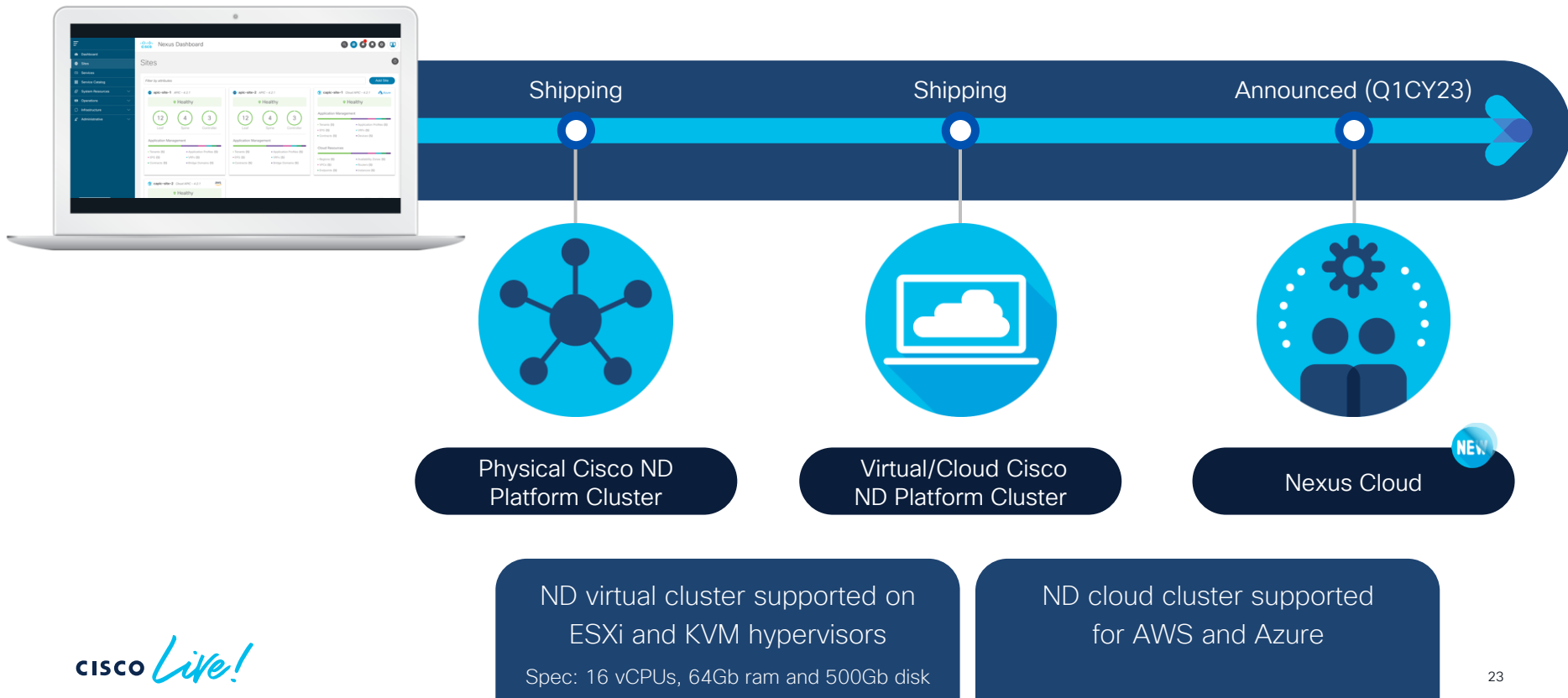
Public cloud



Custom/third-party

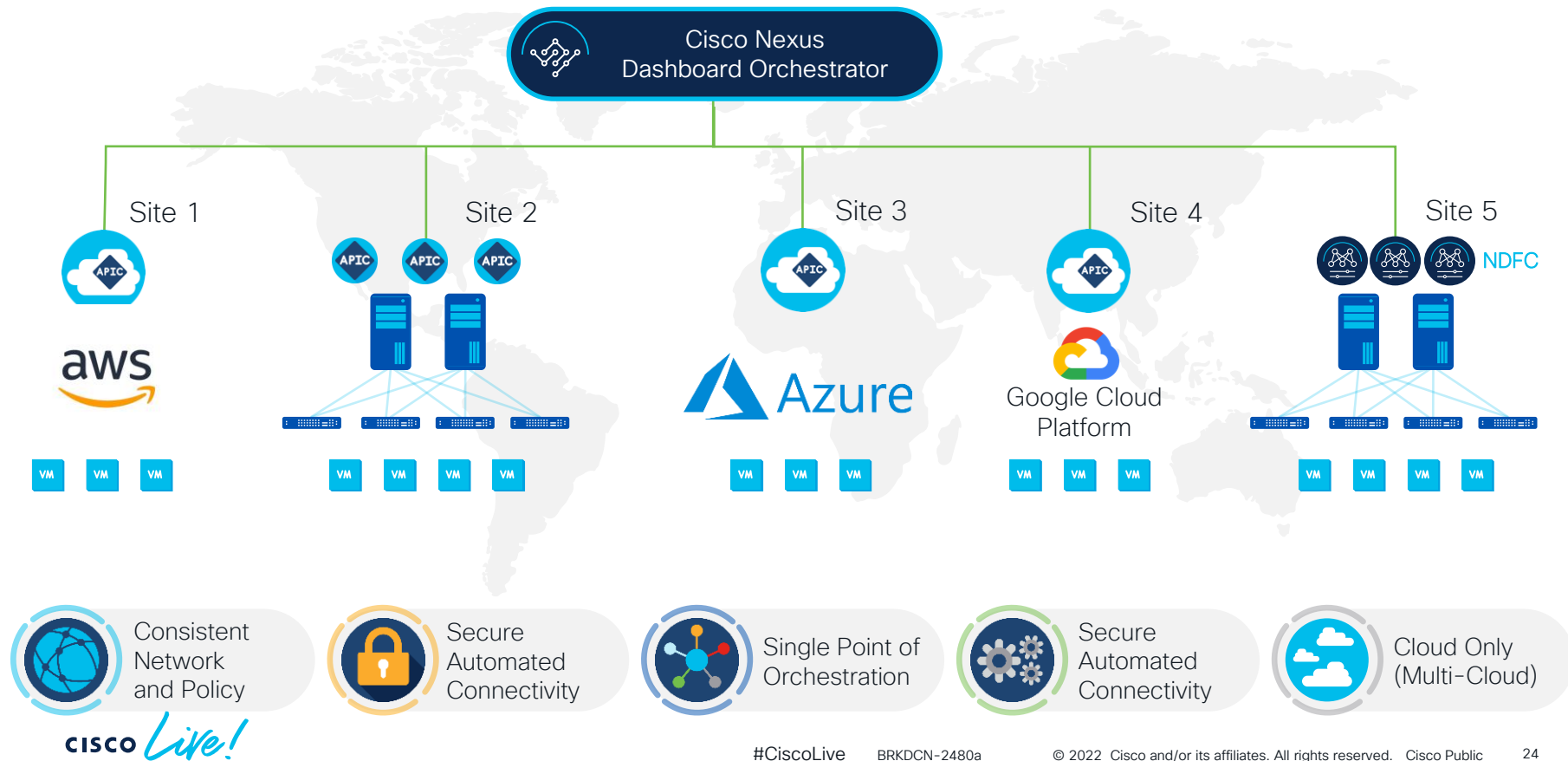


Cisco Nexus Dashboard Deployment Evolution

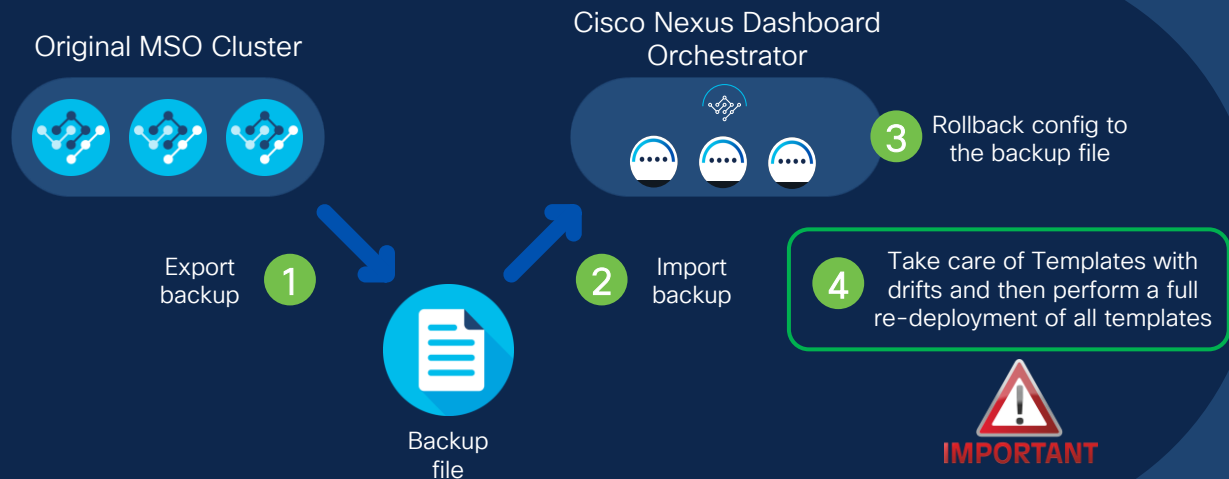


Cisco Nexus Dashboard Orchestrator

Evolution of Cisco Hybrid Cloud and Multi-Cloud Architectures



Migrating the MSO Cluster to Cisco NDO



Note

This procedure is only supported for SW upgrades but not for SW downgrades

<https://www.cisco.com/c/en/us/td/docs/dcn/ndo/3x/deployment/cisco-nexus-dashboard-orchestrator-deployment-guide-371/ndo-deploy-migrate-37x.html>

Provisioning Policies on NDO

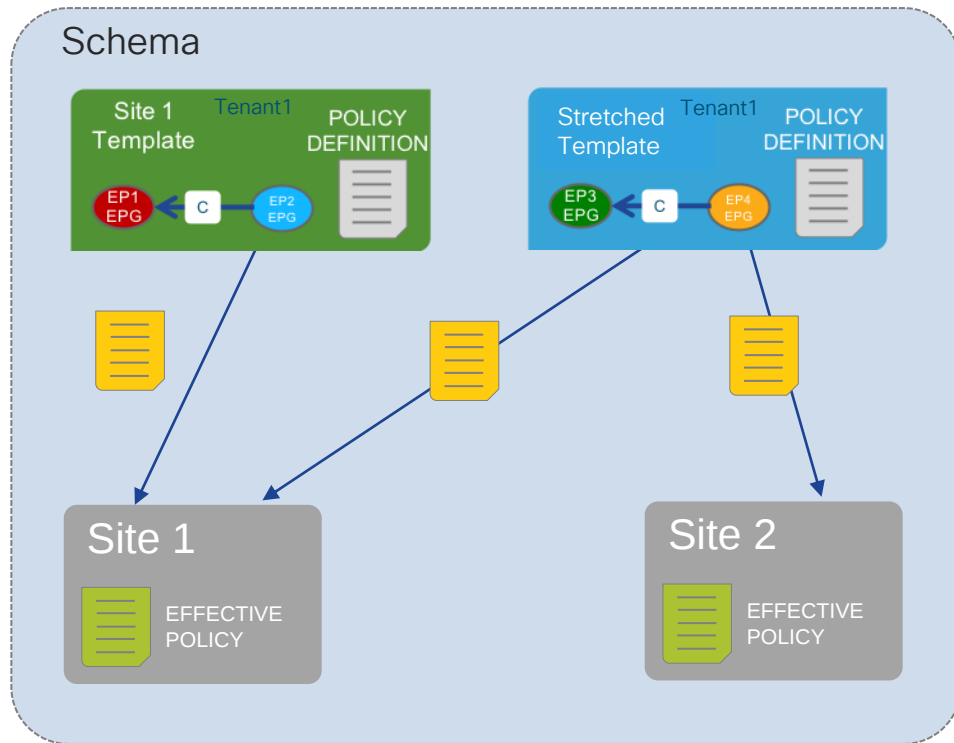


Provisioning Application Policies on NDO

ACI Multi-Site

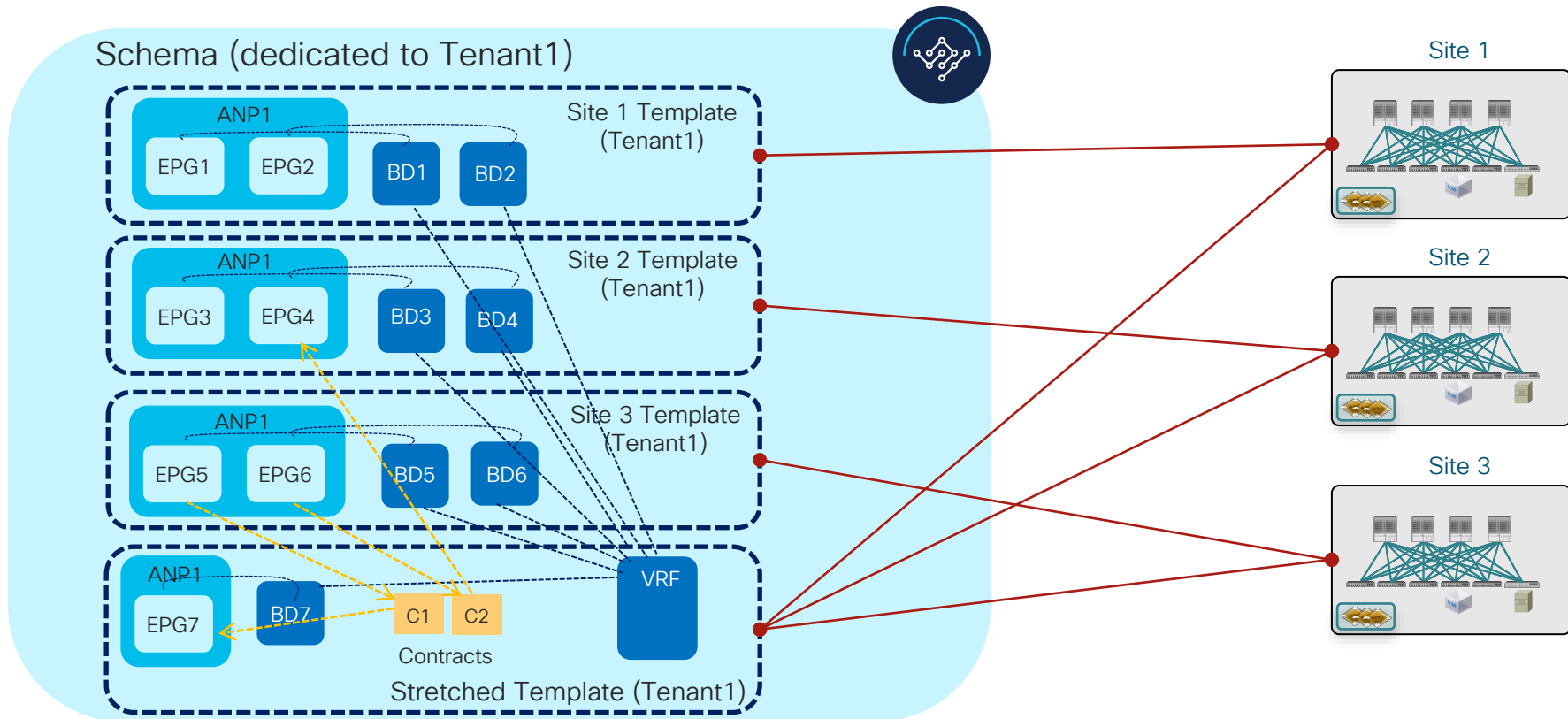
NDO Schema and Templates

- Template = ACI policy definition (ANP, EPGs, BDs, VRFs, etc.)
- Schema = container of Templates sharing a common use-case
 - As an example, a schema can and should be dedicated to a Tenant
- The template is 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



Best Practices for Designing a Schema

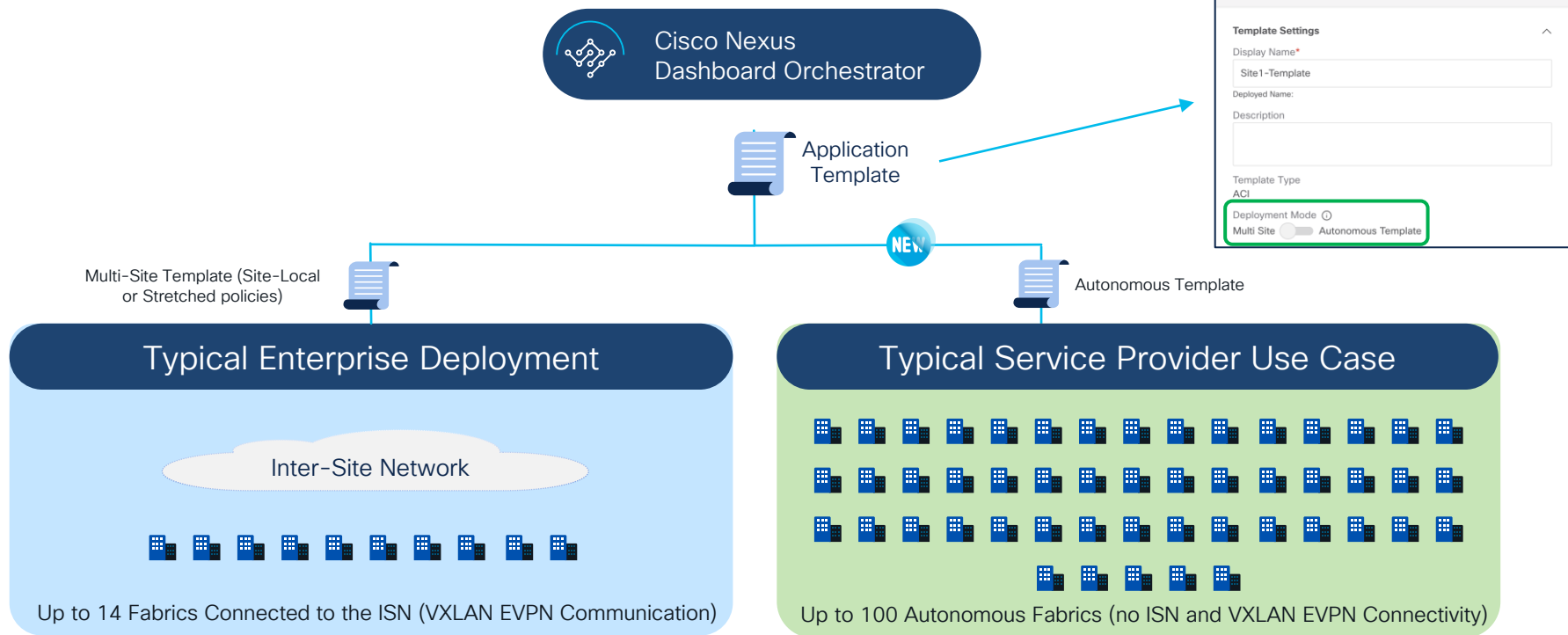
One Template per Site, plus a 'Stretched' Template



Cisco Nexus Dashboard Orchestrator Application Templates



Cisco Nexus Dashboard
Orchestrator 4.0(1)

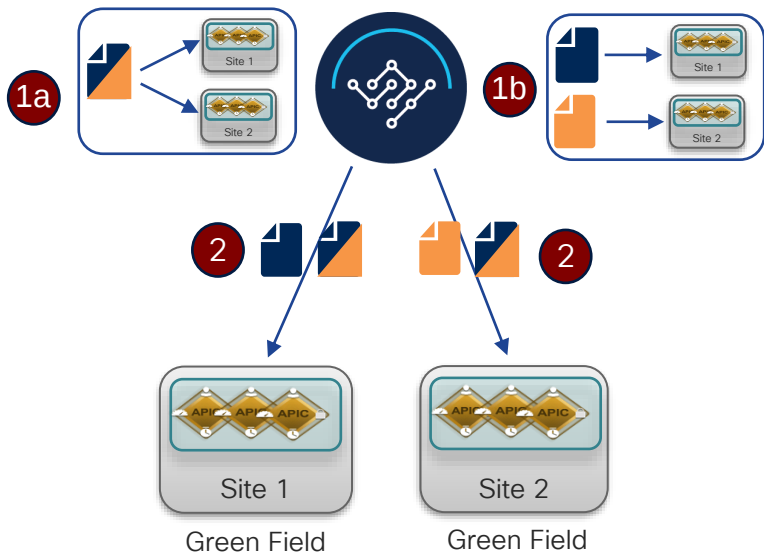


Build feature template once and deploy to many fabrics

ACI Multi-Site Orchestrator

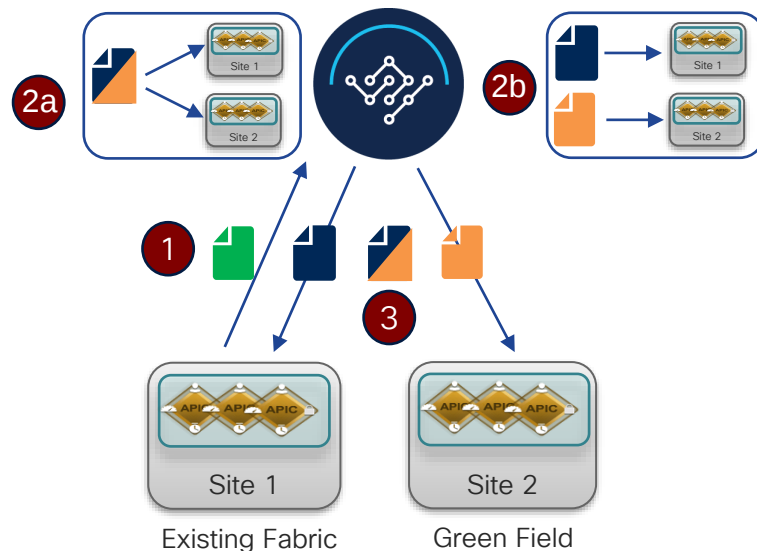
Defining Policies in a Template

Green Field Deployment



- 1a. Model new tenant and policies to a common template on MSO and associate the template to both sites (for stretched objects)
- 1b. Model new tenant and policies to site-specific templates and associate them to each site
2. Push policies to the ACI sites

Import Policies from an Existing Fabric

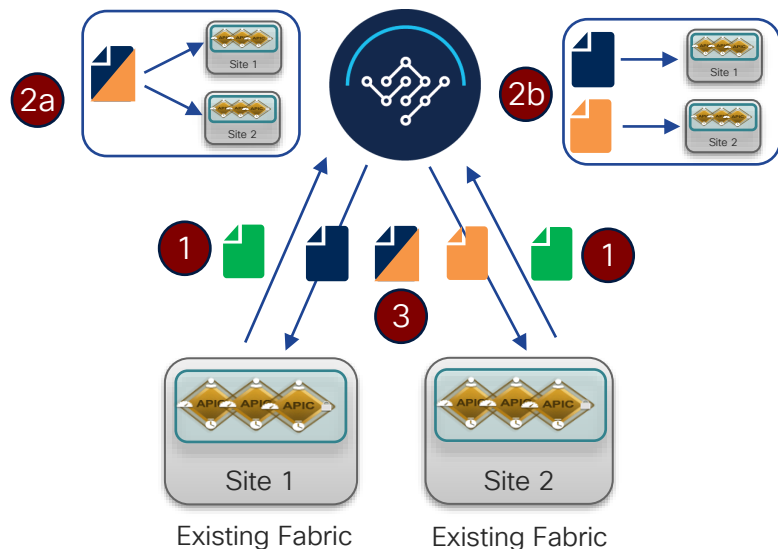


1. Import existing tenant policies from site 1 to new common and site-specific templates on MSO
- 2a. Associate the common template to both sites (for stretched objects)
- 2b. Associate site-specific templates to each site
3. Push the policies back to the ACI sites

ACI Multi-Site Orchestrator

Defining Policies in a Template (2)

Import Policies from Multiple Existing Fabrics



1. Import existing tenant policies from site 1 and site 2 to new common and site-specific templates on ACI MSO
- 2a. Associate the common template to both sites (for stretched objects)
- 2b. Associate site-specific templates to each site
3. Push the policies back to the ACI sites

- In the current implementation, MSO does not allow diff/merge operations on policies from different APIC domains
- It is still possible to import policies for the same tenant from different APIC domains, under the assumption those are no conflicting
 - Tenant defined with the same Name
 - Name and policies for existing stretched objects are also common

Operational Enhancements on NDO

NDO Operational Enhancements

Features

NDO 3.4(1)
Template versioning
and rollback

NDO 3.4(1)
Template deployment
plan visibility

NDO 3.4(1)
Change
control workflow

NDO 3.4(1)
Detach templates
from Sites

NDO 3.6(1)
Configuration drift
reconciliation workflow

Support rollback of template
from newer to older version-id
Label a template as Golden

Shows preview of what NDO
is going to provisioning
to each site

New personas for
management and
provisioning of configuration

Configuration is not
removed from the
APIC/NDFC domains

NDO workflow that synchronizes
and merges any config changes
made in APIC or NDFC domains



Granular roll back of
templates specific
configuration

Better visibility to reduce
errors and seize the
impact of a
template's deployment

More structured
deployments
which enables
increased flexibility

Ease of use
for migration

Simplify the understanding and
reconciliation of config drifts between
NDO and APIC/NDFC

Benefits

NDO Operational Enhancements

For more information and demonstrations of all those NDO operational enhancements:

- Template Versioning
<https://video.cisco.com/video/6277140235001>
- Template Deployment Plan Visibility
<https://video.cisco.com/video/6277137504001>
- Change Control Workflow
<https://video.cisco.com/video/6277140011001>



Provisioning Fabric and Tenant Policies on NDO

NDO 4.0(1) Release

Supporting Different Types of Policies



Application Management Policies

Used to define tenant policies
(Application Network Profiles,
EPGs, BDs, VRFs, etc.)

Fabric Management Policies

Used to define fabric access
policies, interface and monitoring
policies

Nexus Dashboard

Orchestrator

Dashboard

Application Management

- Tenants
- Schemas
- Tenant Policies
- Cloud Policies

Fabric Management

- Fabric Policies
- Fabric Resources Policies
- Monitoring Policy

Operations

Infrastructure

Integration

Sites

Filter by attributes

Health	Name
Warning	Site1 (ACI) 5.2(4d)
Minor	Site2 (ACI) 5.2(4d)

Cisco Nexus Dashboard Orchestrator

Fabric Features Template



Cisco Nexus Dashboard
Orchestrator 4.0(1)



Cisco Nexus
Dashboard Orchestrator

NEW

Feature
Template

NDO Fabric Feature Capabilities

- SPAN (Access, ERSPAN, v4/v6 destination)
- PTP profiles, SyncE, NTP
- DHCP
- QoS (Marking, policing, queuing)
- Multicast v4/v6
- Physical Domains/VLAN Pools

Typical Enterprise Deployment

Inter-Site Network



Up to 124 Fabrics Connected to the ISN (VXLAN EVPN Communication)

Typical Service Provider Use Case



Up to 100 Autonomous Fabrics (no ISN and VXLAN EVPN Connectivity)

Build feature template once and deploy to many fabrics

Supporting Different Types of Templates



Application Template

- Used for provisioning of tenant 'overlay' configuration
- Contains ANPs, EPGs, BDs, VRFs, contracts, etc.

Tenant Policy Template

- Used for provisioning of global tenant policies
- Contains route-maps and multicast route-maps, IGMP/DHCP-Relay/QoS policies, etc.

Fabric Policies Template

- Used for provisioning of specific fabric policies (VLAN pools, physical domains, NTP, PTP, etc.) and interface level policies (LLDP, CDP, LACP, etc.)
- These policies are referenced from fabric resources policies templates

Fabric Resources Policies Template

- Used for the configuration of interfaces, physical or logical (port-channels, vPCs)
- Allows provisioning of node and Pod specific policies (SyncE, PTP, MACsec, etc.)

Monitoring Policies Template

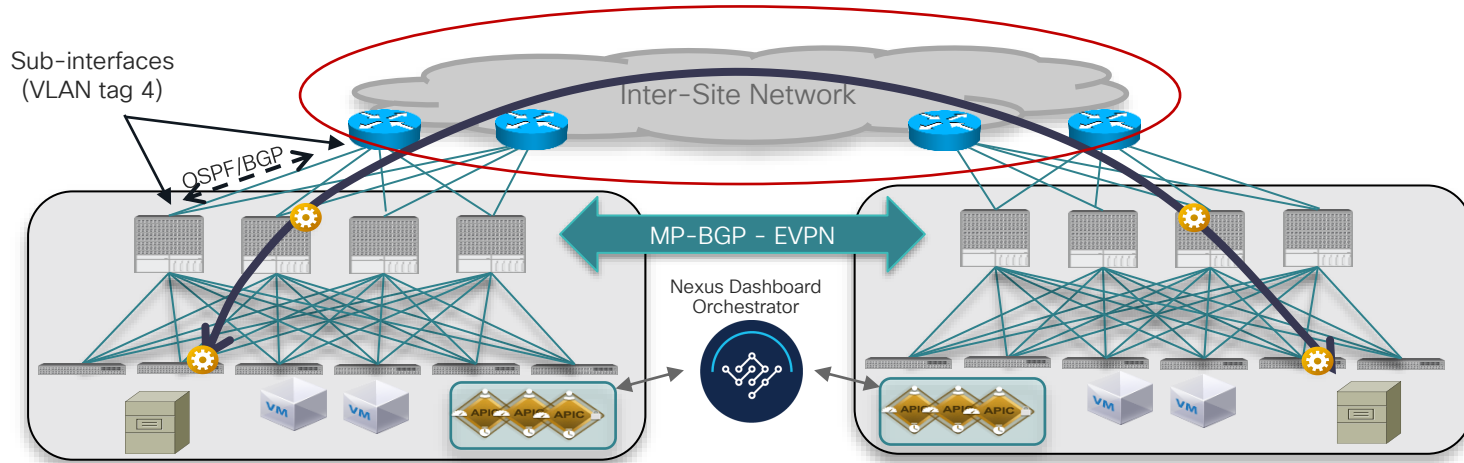
- Used for centralized provisioning of SPAN sessions
- Support for SPAN Tenant and SPAN Access session types

Leverage Operational Enhancements for All Types of Templates

Inter-Site Connectivity Deployment Considerations



Inter-Site Network (ISN) Functional Requirements



- Not managed by APIC or NDO, must be independently configured (day-0 configuration)
- IP topology can be arbitrary, not mandatory to connect all the spine nodes to the ISN
- ISN main functional requirements:
 - ✓ OSPF/BGP* to peer with the spine nodes and exchange TEP address reachability
Must use sub-interfaces (with VLAN tag 4) toward the spines
 - ✓ No multicast requirement for BUM traffic forwarding across sites
 - ✓ Increased end-to-end MTU support (at least 50/54 extra Bytes)

Inter-Site Connectivity

Frequently Asked Questions



1 What platforms can or should I deploy in the ISN?



- Any network device capable of routing traffic and supporting packets with increased MTU size can be deployed in the ISN
- Need sub-interfaces support for the ISN devices directly connected to the spines

2 Do I need to run L3 multicast inside the ISN?



- No, ingress replications is performed by the ACI spine nodes to forward BUM traffic across sites
- This function is only required for the BDs that are stretched across sites with BUM flooding enabled

3 Can I use a Layer 2 only infrastructure as ISN?



- No, the only officially supported configuration consists in deploying the ISN nodes as L3 network devices (particularly the ISN devices connected to the spines)

Inter-Site Connectivity

Frequently Asked Questions (2)



4 Do I need to deploy a dedicated infrastructure as ISN?



- No, the network providing ISN services for Multi-Site could also be used for other functions
- It is recommended (but not mandatory) to use a dedicated VRF for providing ISN connectivity

5 Is there a minimum bandwidth I should deploy between sites?



- No, the bandwidth required between sites mostly depends on the amount of east-west connectivity expected between sites

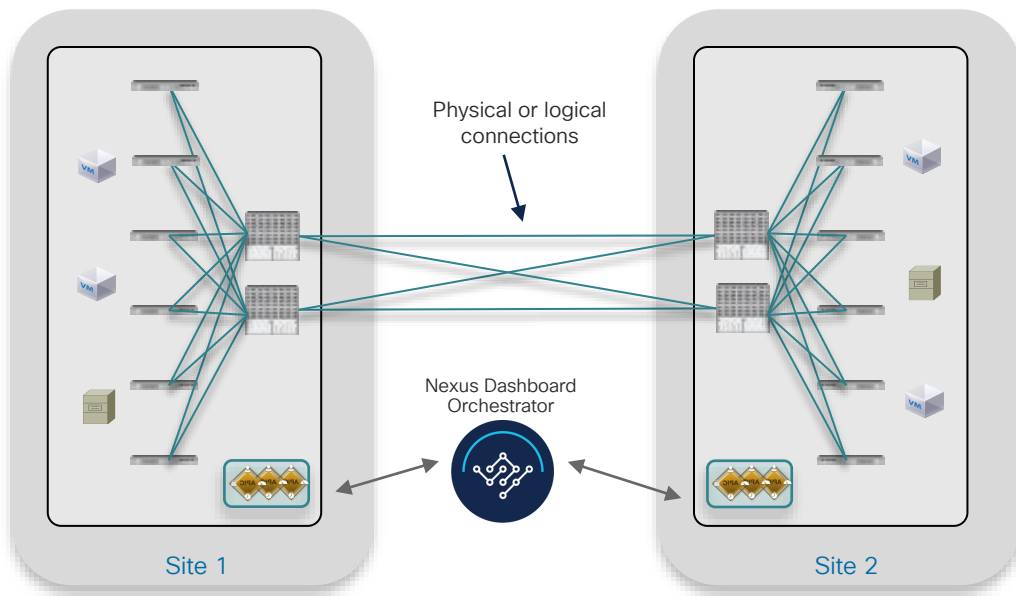
6 Is OSPF the only protocol supported to peer with the ISN network?



- No, from ACI release 5.2(1) and NDO release 3.5(1) we introduced support also for BGP peering between the spines and the first L3 hop ISN devices

ACI Multi-Site

Spines in Separate Sites Connected Back-to-Back



- Back-to-back connections only supported between 2 sites from ACI 3.2(1) release
 - A site cannot be 'transit' for communication between other sites
 - Requires physical or logical (i.e. use of PWs) point-to-point connections (no support for an intermediate Layer 2 infrastructure)
 - Best practice recommendation is to create a full mesh topology
 - 'Square' topology is also supported, yet less optimized (less bandwidth, higher convergence, etc.)

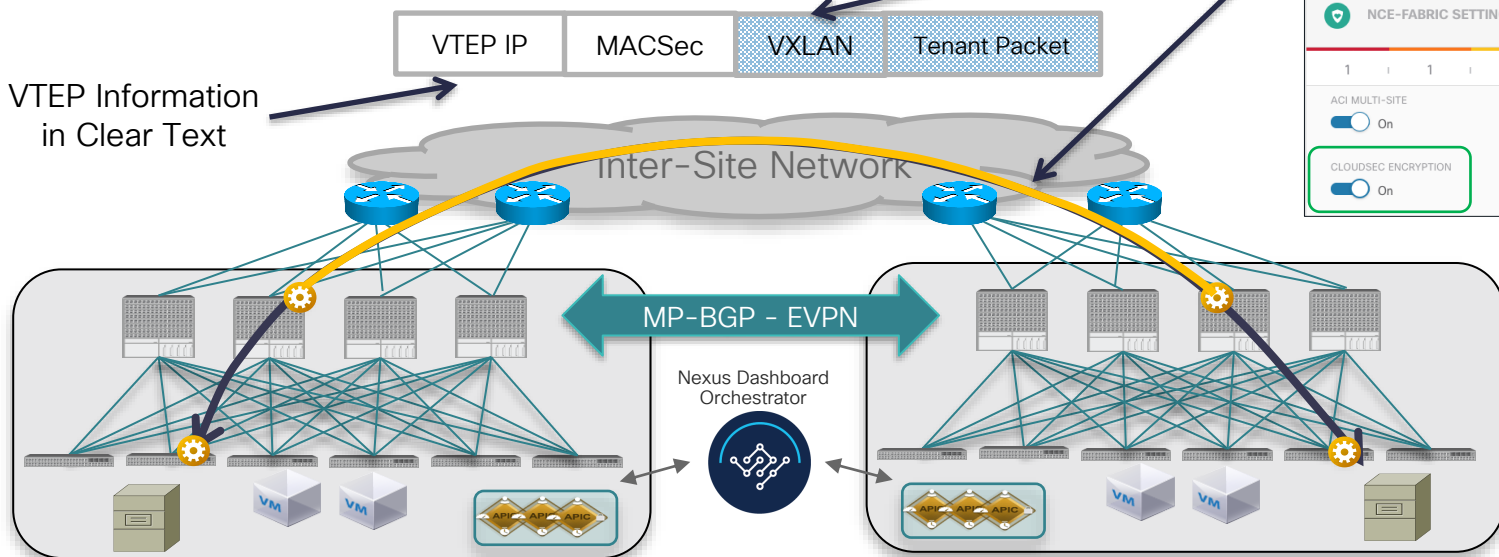
ACI Multi-Site

CloudSec Encryption for VXLAN Traffic

CloudSec = “TEP-to-TEP MACSec”

Encrypted Fabric to Fabric Traffic
[GCM-AES-256-VPN (64-bit PN)]

VTEP Information
in Clear Text



<https://www.cisco.com/c/en/us/td/docs/dcn/ndo/3x/configuration/cisco-nexus-dashboard-orchestrator-configuration-guide-aci-371/ndo-configuration-aci-infra-cloudsec-37x.html>

See You for BRKDCN-2480b
(Part 2)

Today @ 2.30 pm
Location Mandalay Bay E

ACI Multi-Site

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>
- ✓ 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>
- ✓ Cisco Multi-Site Deployment Guide for ACI Fabrics
<https://www.cisco.com/c/en/us/td/docs/dcn/whitepapers/cisco-multi-site-deployment-guide-for-aci-fabrics.html>
- ✓ ACI Multi-Site and Service Node Integration White Paper
<https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-743107.html>
- ✓ ACI Multi-Site Training Sessions
<https://www.cisco.com/c/en/us/solutions/data-center/learning.html#~nexus-dashboard>

Technical Session Surveys

- Attendees who fill out a minimum of four session surveys and the overall event survey will get Cisco Live branded socks!
- Attendees will also earn 100 points in the Cisco Live Game for every survey completed.
- These points help you get on the leaderboard and increase your chances of winning daily and grand prizes.



Cisco Learning and Certifications

From technology training and team development to Cisco certifications and learning plans, let us help you empower your business and career. www.cisco.com/go/certs

Pay for Learning with Cisco Learning Credits

(CLCs) are prepaid training vouchers redeemed directly with Cisco.



Learn

Cisco U.

IT learning hub that guides teams and learners toward their goals

Cisco Digital Learning

Subscription-based product, technology, and certification training

Cisco Modeling Labs

Network simulation platform for design, testing, and troubleshooting

Cisco Learning Network

Resource community portal for certifications and learning



Train

Cisco Training Bootcamps

Intensive team & individual automation and technology training programs

Cisco Learning Partner Program

Authorized training partners supporting Cisco technology and career certifications

Cisco Instructor-led and Virtual Instructor-led training

Accelerated curriculum of product, technology, and certification courses



Certify

Cisco Certifications and Specialist Certifications

Award-winning certification program empowers students and IT Professionals to advance their technical careers

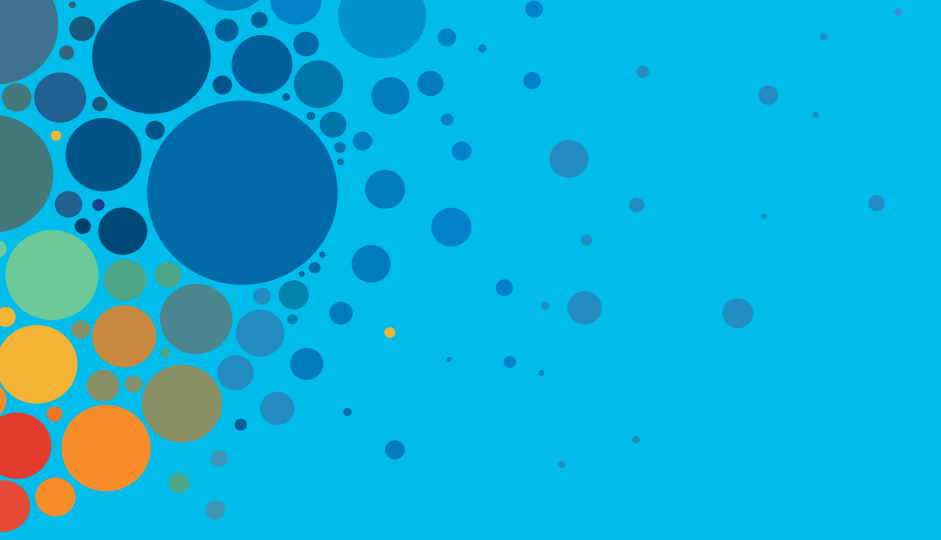
Cisco Guided Study Groups

180-day certification prep program with learning and support

Cisco Continuing Education Program

Recertification training options for Cisco certified individuals

Here at the event? Visit us at **The Learning and Certifications lounge at the World of Solutions**



Continue your education

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



The bridge to possible

Thank you

CISCO *Live!*



#CiscoLive