



Practice
Tests



Flash
Cards



Study
Planner



Video
Training

Official Cert Guide

Advance your IT career with hands-on learning

CCNP Data Center Application Centric Infrastructure

DCACI 300-620

Contents

	Introduction	xxv
Part I	Introduction to Deployment	
Chapter 1	The Big Picture: Why ACI?	2
	“Do I Know This Already?” Quiz	2
	Foundation Topics	4
	Understanding the Shortcomings of Traditional Networks	4
	Network Management	4
	Scalability and Growth	5
	Network Agility	8
	Security	8
	Network Visibility	9
	Recognizing the Benefits of Cisco ACI	9
	Network Management Touchpoints	9
	Traffic Flow Optimizations	10
	Scalability Optimizations	10
	Programmability	11
	Stateless Network	11
	Multitenancy	11
	Zero-Trust Security	14
	Cross-Platform Integrations	15
	New Architectural Possibilities	15
	Integrated Health Monitoring and Enhanced Visibility	16
	Policy Reuse	16
	Exam Preparation Tasks	16
	Review All Key Topics	16
	Complete Tables and Lists from Memory	17
	Define Key Terms	17
Chapter 2	Understanding ACI Hardware and Topologies	18
	“Do I Know This Already?” Quiz	18
	Foundation Topics	21
	ACI Topologies and Components	21
	Clos Topology	21
	Standard ACI Topology	22
	ACI Stretched Fabric Topology	24
	ACI Multi-Pod Topology	25
	ACI Multi-Site Topology	26

	ACI Multi-Tier Architecture	28
	Remote Leaf Topology	30
	APIC Clusters	32
	APIC Cluster Scalability and Sizing	33
	Spine Hardware	36
	First-Generation Spine Switches	37
	Second-Generation Spine Switches	37
	Leaf Hardware	38
	First-Generation Leaf Switches	38
	Second-Generation Leaf Switches	39
	Exam Preparation Tasks	41
	Review All Key Topics	41
	Complete Tables and Lists from Memory	41
	Define Key Terms	41
Chapter 3	Initializing an ACI Fabric	42
	“Do I Know This Already?” Quiz	42
	Foundation Topics	44
	Understanding ACI Fabric Initialization	44
	Planning Fabric Initialization	45
	Understanding Cabling Requirements	45
	Connecting APICs to the Fabric	46
	Initial Configuration of APICs	47
	APIC OOB Configuration Requirements	47
	Out-of-Band Versus In-Band Management	48
	Configuration Information for Fabric Initialization	48
	Switch Discovery Process	49
	Fabric Discovery Stages	51
	Switch Discovery States	51
	Initializing an ACI Fabric	52
	Changing the APIC BIOS Password	52
	Configuring the APIC Cisco IMC	52
	Initializing the First APIC	53
	Discovering and Activating Switches	55
	Understanding Graceful Insertion and Removal (GIR)	58
	Initializing Subsequent APICs	59
	Understanding Connectivity Following Switch Initialization	59
	Basic Post-Initialization Tasks	63
	Assigning Static Out-of-Band Addresses to Switches and APICs	63

	Applying a Default Contract to Out-of-Band Subnet	64
	Upgrading an ACI Fabric	66
	Understanding Schedulers	73
	Enabling Automatic Upgrades of New Switches	74
	Understanding Backups and Restores in ACI	75
	Making On-Demand Backups in ACI	76
	Making Scheduled Backups in ACI	79
	Taking Configuration Snapshots in ACI	80
	Importing Configuration Backups from Remote Servers	80
	Executing Configuration Rollbacks	82
	Pod Policy Basics	83
	Configuring Network Time Protocol (NTP) Synchronization	84
	Configuring DNS Servers for Lookups	90
	Verifying COOP Group Configurations	92
	Exam Preparation Tasks	93
	Review All Key Topics	93
	Complete Tables and Lists from Memory	94
	Define Key Terms	94
Chapter 4	Exploring ACI	96
	“Do I Know This Already?” Quiz	96
	Foundation Topics	98
	ACI Access Methods	98
	GUI	99
	CLI	100
	APIC CLI	100
	Switch CLI	102
	API	103
	Management Access Modifications	103
	Understanding the ACI Object Model	105
	Learning ACI Through the Graphical User Interface	107
	Exploring the Object Hierarchy by Using Visore	108
	Why Understand Object Hierarchy Basics for DCACI?	110
	Policy in Context	110
	Integrated Health Monitoring and Enhanced Visibility	110
	Understanding Faults	111
	The Life of a Fault	113
	Acknowledging Faults	115
	Faults in the Object Model	116

	Monitoring Policies in ACI	118
	Customizing Fault Management Policies	120
	Squelching Faults and Changing Fault Severity	121
	Understanding Health Scores	124
	Understanding Events	126
	Squelching Events	127
	Understanding Audit Logs	127
	Exam Preparation Tasks	128
	Review All Key Topics	128
	Complete Tables and Lists from Memory	129
	Define Key Terms	129
Part II	ACI Fundamentals	
Chapter 5	Tenant Building Blocks	130
	“Do I Know This Already?” Quiz	130
	Foundation Topics	132
	Understanding the Basic Objects in Tenants	132
	Tenants	133
	Predefined Tenants in ACI	134
	VRF Instances	135
	Bridge Domains (BDs)	137
	Endpoint Groups (EPGs)	137
	Application Profiles	138
	The Pain of Designing Around Subnet Boundaries	139
	BDs and EPGs in Practice	141
	Configuring Bridge Domains, Application Profiles, and EPGs	142
	Classifying Endpoints into EPGs	146
	APIC CLI Configuration of Tenant Objects	147
	Contract Security Enforcement Basics	148
	Contracts, Subjects, and Filters	148
	Contract Direction	149
	Contract Scope	150
	Zero-Trust Using EPGs and Contracts	151
	Objects Enabling Connectivity Outside the Fabric	151
	External EPGs	151
	Layer 3 Outside (L3Out)	153
	Tenant Hierarchy Review	153
	Exam Preparation Tasks	154
	Review All Key Topics	154

Complete Tables and Lists from Memory 154

Define Key Terms 154

Chapter 6 Access Policies 156

“Do I Know This Already?” Quiz 156

Foundation Topics 158

Pools, Domains, and AAEPs 158

VLAN Pools 159

Domains 160

Common Designs for VLAN Pools and Domains 161

Challenges with Overlap Between VLAN Pools 164

Attachable Access Entity Profiles (AAEPs) 165

Policies and Policy Groups 169

Interface Policies and Interface Policy Groups 169

Planning Deployment of Interface Policies 173

Switch Policies and Switch Policy Groups 174

Profiles and Selectors 176

Configuring Switch Profiles and Interface Profiles 179

Stateless Networking in ACI 182

Bringing It All Together 183

Access Policies Hierarchy in Review 183

Access Policies and Tenancy in Review 184

Exam Preparation Tasks 184

Review All Key Topics 184

Complete Tables and Lists from Memory 185

Define Key Terms 185

Chapter 7 Implementing Access Policies 186

“Do I Know This Already?” Quiz 186

Foundation Topics 188

Configuring ACI Switch Ports 188

Configuring Individual Ports 188

Configuring Port Channels 196

Configuring Virtual Port Channel (vPC) Domains 201

Configuring Virtual Port Channels 204

Configuring Ports Using AAEP EPGs 208

Implications of Initial Access Policy Design on Capabilities 210

Configuring Access Policies Using Quick Start Wizards 211

The Configure Interface, PC, and VPC Wizard 211

The Configure Interface Wizard 211

	Additional Access Policy Configurations	212
	Configuring Fabric Extenders	212
	Configuring Dynamic Breakout Ports	215
	Configuring Global QoS Class Settings	217
	Configuring DHCP Relay	219
	Configuring MCP	221
	Configuring Storm Control	223
	Configuring CoPP	225
	Modifying BPDU Guard and BPDU Filter Settings	230
	Modifying the Error Disabled Recovery Policy	231
	Configuring Leaf Interface Overrides	232
	Configuring Port Channel Member Overrides	232
	Exam Preparation Tasks	235
	Review All Key Topics	235
	Complete Tables and Lists from Memory	236
	Define Key Terms	236
Chapter 8	Implementing Tenant Policies	238
	“Do I Know This Already?” Quiz	238
	Foundation Topics	241
	ACI Endpoint Learning	241
	Lookup Tables in ACI	241
	Local Endpoints and Remote Endpoints	242
	Understanding Local Endpoint Learning	243
	Unicast Routing and Its Impact on Endpoint Learning	243
	Understanding Remote Endpoint Learning	244
	Understanding the Use of VLAN IDs and VNIDs in ACI	245
	Endpoint Movements Within an ACI Fabric	247
	Understanding Hardware Proxy and Spine Proxy	247
	Endpoint Learning Considerations for Silent Hosts	248
	Where Data Plane IP Learning Breaks Down	249
	Endpoint Learning on L3Outs	249
	Limiting IP Learning to a Subnet	249
	Understanding Enforce Subnet Check	250
	Disabling Data Plane Endpoint Learning on a Bridge Domain	250
	Disabling IP Data Plane Learning at the VRF Level	251
	Packet Forwarding in ACI	251
	Forwarding Scenario 1: Both Endpoints Attach to the Same Leaf	251
	Understanding Pervasive Gateways	252

Forwarding Scenario 2: Known Destination Behind Another Leaf	254
Verifying the Traffic Path Between Known Endpoints	254
Understanding Learning and Forwarding for vPCs	256
Forwarding Scenario 3: Spine Proxy to Unknown Destination	258
Forwarding Scenario 4: Flooding to Unknown Destination	261
Understanding ARP Flooding	262
Deploying a Multi-Tier Application	263
Configuring Application Profiles, BDs, and EPGs	264
Assigning Domains to EPGs	267
Policy Deployment Following BD and EPG Setup	267
<i>Mapping EPGs to Ports Using Static Bindings</i>	267
<i>Verifying EPG-to-Port Assignments</i>	269
<i>Policy Deployment Following EPG-to-Port Assignment</i>	270
Mapping an EPG to All Ports on a Leaf	270
Enabling DHCP Relay for a Bridge Domain	271
Whitelisting Intra-VRF Communications via Contracts	272
Planning Contract Enforcement	272
Configuring Filters for Bidirectional Application	273
<i>Configuring Subjects for Bidirectional Application of Filters</i>	275
<i>Understanding Apply Both Directions and Reverse Filter Ports</i>	277
<i>Verifying Subject Allocation to a Contract</i>	278
<i>Assigning Contracts to EPGs</i>	278
<i>Understanding the TCP Established Session Rule</i>	279
<i>Creating Filters for Unidirectional Application</i>	280
<i>Configuring Subjects for Unidirectional Application of Filters</i>	280
<i>Additional Whitelisting Examples</i>	282
<i>Verifying Contract Enforcement</i>	283
<i>Understanding the Stateful Checkbox in Filter Entries</i>	284
<i>Contract Scopes in Review</i>	284
Exam Preparation Tasks	285
Review All Key Topics	285
Complete Tables and Lists from Memory	287
Define Key Terms	287
Part III External Connectivity	
Chapter 9 L3Outs	288
“Do I Know This Already?” Quiz	288
Foundation Topics	291

L3Out Fundamentals	291
Stub Network and Transit Routing	291
Types of L3Outs	292
Key Functions of an L3Out	293
The Anatomy of an L3Out	293
Planning Deployment of L3Out Node and Interface Profiles	295
Understanding L3Out Interface Types	296
Understanding L3Out Bridge Domains	296
Understanding SVI Encap Scope	298
Understanding SVI Auto State	299
Understanding Prerequisites for Deployment of L3Outs	301
L3 Domain Implementation Examples	301
Understanding the Need for BGP Route Reflection	303
Implementing BGP Route Reflectors	304
Understanding Infra MP-BGP Route Distribution	305
Deploying L3Outs	307
Configuring an L3Out for EIGRP Peering	307
Deploying External EPGs	310
Verifying Forwarding Out an L3Out	312
Advertising Subnets Assigned to Bridge Domains via an L3Out	314
Enabling Communications over L3Outs Using Contracts	316
Deploying a Blacklist EPG with Logging	318
Advertising Host Routes Out an ACI Fabric	321
Implementing BFD on an EIGRP L3Out	321
Configuring Authentication for EIGRP	324
EIGRP Customizations Applied at the VRF Level	324
Configuring an L3Out for OSPF Peering	325
A Route Advertisement Problem for OSPF and EIGRP L3Outs	328
Implementing BFD on an OSPF L3Out	328
OSPF Customizations Applied at the VRF Level	329
Adding Static Routes on an L3Out	329
Implementing IP SLA Tracking for Static Routes	330
Configuring an L3Out for BGP Peering	334
Implementing BGP Customizations at the Node Level	337
Implementing Per-Neighbor BGP Customizations	339
Implementing BFD on a BGP L3Out	341
Implementing BGP Customizations at the VRF Level	342
Implementing OSPF for IP Reachability on a BGP L3Out	343

Implementing Hot Standby Router Protocol (HSRP)	344
IPv6 and OSPFv3 Support	344
Implementing Route Control	344
Route Profile Basics	344
Modifying Route Attributes to All Peers Behind an L3Out	346
Modifying Route Attributes to a Specific Peer Behind an L3Out	349
Assigning Different Policies to Routes at the L3Out Level	351
Configuring Inbound Route Filtering in ACI	352
Exam Preparation Tasks	353
Review All Key Topics	353
Complete Tables and Lists from Memory	356
Define Key Terms	356
Chapter 10 Extending Layer 2 Outside ACI	358
“Do I Know This Already?” Quiz	358
Foundation Topics	361
Understanding Network Migrations into ACI	361
Understanding Network-Centric Deployments	361
Understanding Full-Mesh Network-Centric Contracts	362
Understanding Any EPG	364
Understanding Preferred Group Members	365
Disabling Contract Enforcement at the VRF Instance Level	367
Flooding Requirements for L2 Extension to Outside Switches	368
Understanding GARP-Based Detection	370
Understanding Legacy Mode	371
Endpoint Learning Considerations for Layer 2 Extension	371
Preparing for Network-Centric Migrations	372
Implementing Layer 2 Connectivity to Non-ACI Switches	372
Understanding EPG Extensions	372
Understanding Bridge Domain Extensions	374
Comparing EPG Extensions and BD Extensions	374
Implementing EPG Extensions	375
Implementing L2Outs	380
Migrating Overlapping VLANs into ACI	385
Understanding ACI Interaction with Spanning Tree Protocol	386
Remediating Against Excessive Spanning Tree Protocol TCNs	386
Configuring MST Instance Mappings in ACI	387
Understanding Spanning Tree Protocol Link Types	388
Using MCP to Detect Layer 2 Loops	388

Exam Preparation Tasks	389
Review All Key Topics	389
Complete Tables and Lists from Memory	390
Define Key Terms	390

Part IV Integrations

Chapter 11 Integrating ACI into vSphere Using VDS 392

“Do I Know This Already?” Quiz	392
Foundation Topics	394
Understanding Networking in VMware vSphere	394
Understanding vSphere Standard Switches	395
Understanding vSphere Distributed Switches	397
Understanding vSphere System Traffic	397
Impact of vCenter Failure on Production Traffic	399
Understanding Port Bindings in vSphere	400
Understanding Teaming and Failover Policies	400
Understanding VMM Integration	403
Planning vCenter VMM Integrations	403
What Happens After VDS Deployment?	405
Understanding Immediacy Settings	405
Connecting ESXi Servers to the Fabric	407
Configuring Connectivity to ESXi in UCS Domains	407
Integrating ACI into vSphere Using VDS	407
Prerequisites for VMM Integration with vSphere VDS	408
Configuring a VMM Domain Profile	408
Adding ESXi Hosts to a VDS	411
Pushing EPGs to vCenter as Distributed Port Groups	415
Assigning VMs to Distributed Port Groups	417
Less Common VMM Domain Association Settings	418
Enhanced LACP Policy Support	419
Exam Preparation Tasks	422
Review All Key Topics	422
Complete Tables and Lists from Memory	423
Define Key Terms	423

Chapter 12 Implementing Service Graphs 424

“Do I Know This Already?” Quiz	424
Foundation Topics	426
Service Graph Fundamentals	426

Service Graphs as Concatenation of Functions	427
Service Graph Management Models	428
Understanding Network Policy Mode	428
Understanding Service Policy Mode	430
Understanding Service Manager Mode	432
When to Use Service Graphs	434
Choosing an L4–L7 Services Integration Method	435
Understanding Deployment Modes and the Number of BDs Required	435
Deploying Service Graphs for Devices in GoTo Mode	436
Deploying Service Graphs for Devices in GoThrough Mode	437
Deploying Service Graphs for One-Arm Load Balancers	437
Understanding Route Peering	438
Understanding Dynamic Endpoint Attach	439
Understanding Bridge Domain Settings for Service Graphs	439
Understanding Service Graph Rendering	440
Service Graph Implementation Workflow	441
Importing Device Packages	441
Identifying L4–L7 Devices to the Fabric	443
Creating Custom Function Profiles	444
Configuring a Service Graph Template	445
Configuring Device Selection Policies	446
Applying a Service Graph Template	446
Configuring Additional Service Graph Parameters	447
Monitoring Service Graphs and Devices	447
Service Graph Implementation Examples	447
Deploying an Unmanaged Firewall Pair in a Service Graph	447
Deploying Service Graphs for a Firewall in Managed Mode	453
Exam Preparation Tasks	460
Review All Key Topics	460
Complete Tables and Lists from Memory	461
Define Key Terms	461

Part V Management and Monitoring

Chapter 13 Implementing Management 462

“Do I Know This Already?” Quiz	462
Foundation Topics	464
Configuring Management in ACI	464
Understanding Out-of-Band Management Connectivity	464
Understanding In-Band Management Connectivity	465

	Deploying In-Band and OOB Management Side by Side	467
	Configuring In-Band Management	467
	Configuring Access Policies for APIC In-Band Interfaces	468
	Configuring the In-Band Management Bridge Domain	469
	Configuring In-Band Management IP Addressing	470
	Optionally Extending the In-Band Network Out of the Fabric	474
	Optionally Setting Up Additional Connectivity	476
	Whitelisting Desired Connectivity to and from an In-Band EPG	476
	Evaluating APIC Connectivity Preferences	478
	Out-of-Band Management Contracts in Review	479
	Exam Preparation Tasks	481
	Review All Key Topics	481
	Memory Tables	481
	Define Key Terms	481
Chapter 14	Monitoring ACI Using Syslog and SNMP	482
	“Do I Know This Already?” Quiz	482
	Foundation Topics	485
	Understanding System Messages	485
	Forwarding System Messages to Syslog Servers	487
	Apply Necessary Contracts to Allow Syslog Forwarding	487
	Configuring Syslog Monitoring Destination Groups	492
	Configuring Syslog Sources for Desired Monitoring Policies	494
	Verify Syslog Forwarding to Desired Syslog Servers	498
	Using SNMP in ACI	500
	ACI Support for SNMP	501
	ACI SNMP Configuration Caveats	502
	Configuring ACI for SNMP	502
	Apply Necessary Contracts for SNMP	503
	Associate an SNMP Policy with a Pod Policy	504
	Associate SNMP Contexts with Desired VRF Instances	506
	Configure SNMP Monitoring Destination Groups	507
	Configure SNMP Sources for All Desired Monitoring Policies	508
	Verify SNMP Forwarding to Desired SNMP Servers	509
	Exam Preparation Tasks	511
	Review All Key Topics	511
	Complete Tables and Lists from Memory	512
	Define Key Terms	512

Chapter 15 Implementing AAA and RBAC 514

- “Do I Know This Already?” Quiz 514
- Foundation Topics 516
- Implementing Role-Based Access Control (RBAC) 516
 - Understanding Security Domains 517
 - Understanding Privileges and Roles 519
 - Creating Local Users and Assigning Access 521
 - Tweaking Roles and User Access 525
 - Custom RBAC Rules 528
 - A Common RBAC Pitfall 531
- Integrating with External AAA Servers 532
 - Configuring ACI for TACACS+ 532
 - Configuring ISE to Authenticate and Authorize Users for ACI 536
 - Expected Cisco AV Pair Formatting for ACI 538
 - Configuring ACI for RADIUS 540
 - Configuring ACI for LDAP 541
 - AAA Authentication Policy Settings 547
 - Regaining Access to the Fabric via Fallback Domain 550
- Exam Preparation Tasks 550
- Review All Key Topics 550
- Complete Tables and Lists from Memory 551
- Define Key Terms 551

Part VI Operations

Chapter 16 ACI Anywhere 552

- “Do I Know This Already?” Quiz 552
- Foundation Topics 555
- ACI Multi-Site Fundamentals 555
 - Interconnecting ACI Fabrics with ACI Multi-Site 555
 - New ACI Multi-Site Constructs and Configuration Concepts 557
 - Locally Governed Versus MSO-Governed Configurations 557
 - Schemas and Templates in Practice 557
- Building Primary and Disaster Recovery
 - Data Centers with ACI 558
 - Centralized Orchestration and Management of Multiple Fabrics 559
 - Tweaking Broadcast and Stretch Settings on a Per-BD Basis 560
 - Cross-Data Center Ingress Routing Optimizations 561
 - Simultaneous or Independent Policy Deployment to Sites 561

	Building Active/Active Data Centers with ACI	562
	VMM Integrations Applicable to Multiple Data Centers	563
	Stateful-Services Integration in ACI Multi-Pod and Multi-Site	563
	Extending ACI to Remote Locations and Public Clouds	564
	Extending ACI into Public Clouds with ACI Multi-Site	564
	Extending ACI into Bare-Metal Clouds with vPod	564
	Integrating Remote Sites into ACI Using Remote Leaf Switches	564
	Exam Preparation Tasks	565
	Review All Key Topics	565
	Memory Tables	565
	Define Key Terms	565
Part VII	Final Preparation	
Chapter 17	Final Preparation	566
	Getting Ready	566
	Tools for Final Preparation	567
	Pearson Cert Practice Test Engine and Questions on the Website	567
	Accessing the Pearson Test Prep Software Online	567
	Accessing the Pearson Test Prep Software Offline	568
	Customizing Your Exams	568
	Updating Your Exams	569
	Premium Edition	569
	Suggested Plan for Final Review/Study	570
	Summary	570
Appendix A	Answers to the “Do I Know This Already?” Questions	572
Appendix B	CCNP Data Center Application Centric Infrastructure DCACI 300-620 Exam Updates	586
	Glossary	589
	Index	602
	Online Elements	
Appendix C	Memory Tables	
Appendix D	Memory Tables Answer Key	
Appendix E	Study Planner	
	Glossary	

CHAPTER 2

Understanding ACI Hardware and Topologies

This chapter covers the following topics:

ACI Topologies and Components: This section describes the key hardware components and acceptable topologies for ACI fabrics.

APIC Clusters: This section covers available APIC hardware models and provides an understanding of APIC cluster sizes and failover implications.

Spine Hardware: This section addresses available spine hardware options.

Leaf Hardware: This section outlines the leaf platforms available for deployment in ACI fabrics.

This chapter covers the following exam topics:

- 1.1 Describe ACI topology and hardware
- 6.1 Describe Multi-Pod
- 6.2 Describe Multi-Site

ACI is designed to allow small and large enterprises and service providers to build massively scalable data centers using a relatively small number of very flexible topologies.

This chapter details the topologies with which an ACI fabric can be built or extended. Understanding supported ACI topologies helps guide decisions on target-state network architecture and hardware selection.

Each hardware component in an ACI fabric performs a specific set of functions. For example, leaf switches enforce security rules, and spine switches track all endpoints within a fabric in a local database.

But not all ACI switches are created equally. Nor are APICs created equally. This chapter therefore aims to provide a high-level understanding of some of the things to consider when selecting hardware.

“Do I Know This Already?” Quiz

The “Do I Know This Already?” quiz allows you to assess whether you should read this entire chapter thoroughly or jump to the “Exam Preparation Tasks” section. If you are in doubt about your answers to these questions or your own assessment of your knowledge of the topics, read the entire chapter. Table 2-1 lists the major headings in this chapter and their corresponding “Do I Know This Already?” quiz questions. You can find the answers in Appendix A, “Answers to the ‘Do I Know This Already?’ Questions.”

Table 2-1 “Do I Know This Already?” Section-to-Question Mapping

Foundation Topics Section	Questions
ACI Topologies and Components	1–5
APIC Clusters	6
Spine Hardware	7, 8
Leaf Hardware	9, 10

CAUTION The goal of self-assessment is to gauge your mastery of the topics in this chapter. If you do not know the answer to a question or are only partially sure of the answer, you should mark that question as wrong for purposes of the self-assessment. Giving yourself credit for an answer you correctly guess skews your self-assessment results and might provide you with a false sense of security.

1. An ACI fabric is being extended to a secondary location to replace two top-of-rack switches and integrate a handful of servers into a corporate ACI environment. Which solution should ideally be deployed at the remote location if the deployment of new spines is considered cost-prohibitive and direct fiber links from the main data center cannot be dedicated to this function?
 - a. ACI Multi-Site
 - b. ACI Remote Leaf
 - c. ACI Multi-Tier
 - d. ACI Multi-Pod
2. Which of the following is a requirement for a Multi-Pod IPN that is not needed in an ACI Multi-Site ISN?
 - a. Increased MTU support
 - b. OSPF support on last-hop routers connecting to ACI spines
 - c. End-to-end IP connectivity
 - d. Multicast PIM-Bidir
3. Which of the following connections would ACI definitely block?
 - a. APIC-to-leaf cabling
 - b. Leaf-to-leaf cabling
 - c. Spine-to-leaf cabling
 - d. Spine-to-spine cabling
4. Which of the following are valid reasons for ACI Multi-Site requiring more specialized spine hardware? (Choose all that apply.)
 - a. Ingress replication of BUM traffic
 - b. IP fragmentation
 - c. Namespace normalization
 - d. Support for PIM-Bidir for multicast forwarding

5. Which of the following options best describes border leaf switches?
 - a. Border leaf switches provide Layer 2 and 3 connectivity to outside networks.
 - b. Border leaf switches connect to Layer 4–7 service appliances, such as firewalls and load balancers.
 - c. Border leaf switches are ACI leaf switches that connect to servers.
 - d. Border leaf switches serve as the border between server network traffic and FCoE storage traffic.
6. Which of the following statements is accurate?
 - a. A three-node M3 cluster of APICs can scale up to 200 leaf switches.
 - b. Sharding is a result of the evolution of what is called horizontal partitioning of databases.
 - c. The number of shards distributed among APICs for a given attribute is directly correlated to the number of APICs deployed.
 - d. A standby APIC actively synchronizes with active APICs and has a copy of all attributes within the APIC database at all times.
7. Out of the following switches, which are spine platforms that support ACI Multi-Site? (Choose all that apply.)
 - a. Nexus 93180YC-EX
 - b. Nexus 9364C
 - c. Nexus 9736C-FX line card
 - d. Nexus 9396PX
8. Which of the following is a valid reason for upgrading a pair of Nexus 9336PQ ACI switches to second-generation Nexus 9332C spine hardware? (Choose all that apply.)
 - a. Namespace normalization for ACI Multi-Site support
 - b. Support for 40 Gbps leaf-to-spine connectivity
 - c. Support for CloudSec
 - d. Support for ACI Multi-Pod
9. True or false: The Nexus 93180YC-FX leaf switch supports MACsec.
 - a. True
 - b. False
10. Which of the following platforms is a low-cost option for server CIMC and other low-bandwidth functions that rely on RJ-45 connectivity?
 - a. Nexus 9336C-FX2
 - b. Nexus 93180YC-FX
 - c. Nexus 9332C
 - d. Nexus 9348GC-FXP

Foundation Topics

ACI Topologies and Components

Like many other current data center fabrics, ACI fabrics conform to a Clos-based leaf-and-spine topology.

In ACI, leaf and spine switches are each responsible for different functions. Together, they create an architecture that is highly standardized across deployments. Cisco has introduced several new connectivity models and extensions for ACI fabrics over the years, but none of these changes break the core ACI topology that has been the standard from day one. Any topology modifications introduced in this section should therefore be seen as slight enhancements that help address specific use cases and not as deviations from the standard ACI topology.

Clos Topology

In his 1952 paper titled “A Study of Non-blocking Switching Networks,” Bell Laboratories researcher Charles Clos formalized how multistage telephone switching systems could be built to forward traffic, regardless of the number of calls served by the overall system.

The mathematical principles proposed by Clos also help address the challenge of needing to build highly scalable data centers using relatively low-cost switches.

Figure 2-1 illustrates a three-stage Clos fabric consisting of one layer for ingress traffic, one layer for egress traffic, and a central layer for forwarding traffic between the layers. Multi-stage designs such as this can result in networks that are not oversubscribed or that are very close to not being oversubscribed.

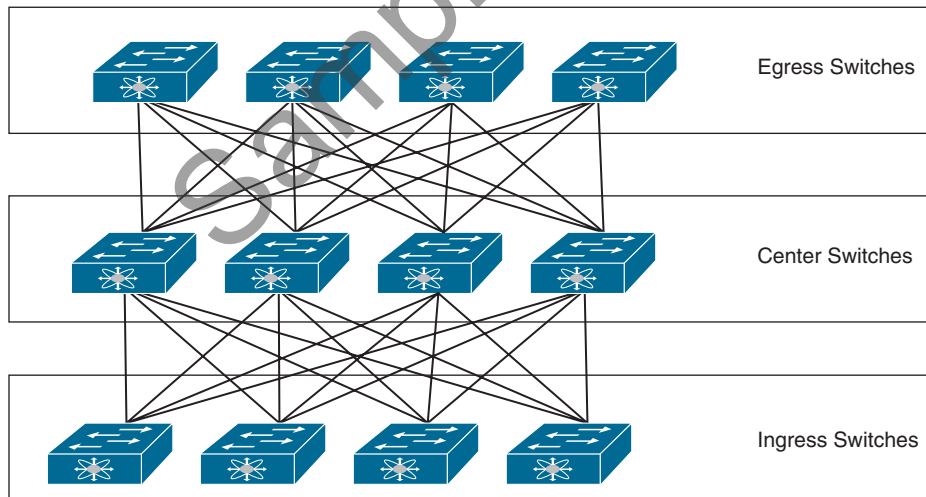


Figure 2-1 *Conceptual View of a Three-Stage Clos Topology*

Modern data center switches forward traffic at full duplex. Therefore, there is little reason to depict separate layers for ingress and egress traffic. It is possible to fold the top layer from the three-tier Clos topology in Figure 2-1 into the bottom layer to achieve what the industry refers to as a “folded” Clos topology, illustrated in Figure 2-2.

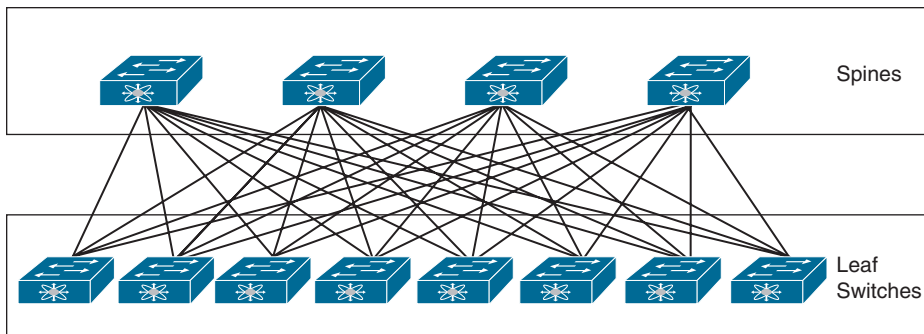


Figure 2-2 *Folded Clos Topology*

As indicated in Figure 2-2, a leaf switch is an ingress/egress switch. A spine switch is an intermediary switch whose most critical function is to perform rapid forwarding of traffic between leaf switches. Leaf switches connect to spine switches in a full-mesh topology.

NOTE At first glance, a three-tier Clos topology may appear to be similar to the traditional three-tier data center architecture. However, there are some subtle differences. First, there are no physical links between leaf switches in the Clos topology. Second, there are no physical links between spine switches. The elimination of cross-links within each layer simplifies network design and reduces control plane complexity.

Standard ACI Topology

An ACI fabric forms a Clos-based spine-and-leaf topology and is usually depicted using two rows of switches. Depending on the oversubscription and overall network throughput requirements, the number of spines and leaf switches will be different in each ACI fabric.

NOTE In the context of the Implementing Cisco Application Centric Infrastructure DCACI 300-620 exam, it does not matter whether you look at a given ACI fabric as a two-tiered Clos topology or as a three-tiered folded Clos topology. It is common for the standard ACI topology to be referred to as a two-tier spine-and-leaf topology.

Figure 2-3 shows the required components and cabling for an ACI fabric. Inheriting from its Clos roots, no cables should be connected between ACI leaf switches. Likewise, ACI spines being cross-cabled results in ACI disabling the cross-connected ports. While the topology shows a full mesh of cabling between the spine-and-leaf layers, a fabric can operate without a full mesh. However, a full mesh of cables between layers is still recommended.

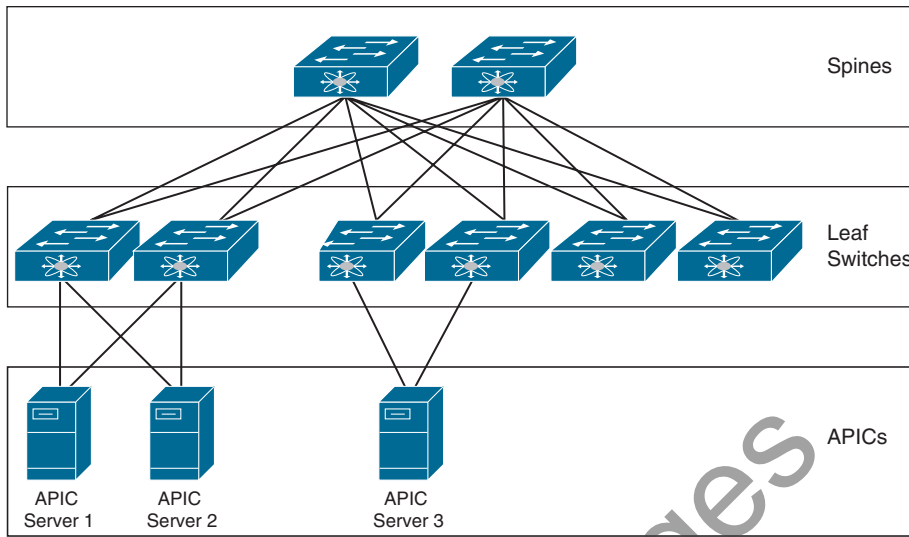


Figure 2-3 Standard ACI Fabric Topology

In addition to optics and cabling, the primary hardware components required to build an ACI fabric are as follows:

Key Topic

- Application Policy Infrastructure Controllers (APICs):** The APICs are the brains of an ACI fabric and serve as the single source of truth for configuration within the fabric. A clustered set of (typically three) controllers attaches directly to leaf switches and provides management, policy programming, application deployment, and health monitoring for an ACI fabric. Note in Figure 2-3 that APICs are not in the data path or the forwarding topology. Therefore, the failure of one or more APICs does not halt packet forwarding. An ACI fabric requires a minimum of one APIC, but an ACI fabric with one APIC should be used only for lab purposes.
- Spine switches:** ACI spine switches are Clos intermediary switches that have a number of key functions. They exchange routing updates with leaf switches via Intermediate System-to-Intermediate System (IS-IS) and perform rapid forwarding of packets between leaf switches. They provide endpoint lookup services to leaf switches through the Council of Oracle Protocol (COOP). They also handle route reflection to leaf switches using Multiprotocol BGP (MP-BGP), allowing external routes to be distributed across the fabric regardless of the number of tenants. (All three of these are control plane protocols and are covered in more detail in future chapters.) Spine switches also serve as roots for multicast trees within a fabric. By default, all spine switch interfaces besides the mgmt0 port are configured as fabric ports. *Fabric ports* are the interfaces that are used to interconnect spine and leaf switches within a fabric.
- Leaf switches:** Leaf switches are the ingress/egress points for traffic into and out of an ACI fabric. As such, they are the connectivity points for endpoints, including servers and appliances, into the fabric. Layer 2 and 3 connectivity from the outside world into an ACI fabric is also typically established via leaf switches. ACI security policy enforcement occurs on leaf switches. Each leaf switch has a number of high-bandwidth uplink ports preconfigured as fabric ports.

In addition to the components mentioned previously, optional hardware components that can be deployed alongside an ACI fabric include fabric extenders (FEX). Use of FEX solutions in ACI is not ideal because leaf hardware models currently on the market are generally low cost and feature heavy compared to FEX technology.

FEX attachment to ACI is still supported to allow for migration of brownfield gear into ACI fabrics. The DCACI 300-620 exam does not cover specific FEX model support, so neither does this book.

NOTE There are ways to extend an ACI fabric into a virtualized environment by using ACI Virtual Edge (AVE) and Application Virtual Switch (AVS). These are software rather than hardware components and are beyond the scope of the DCACI 300-620 exam.

Engineers may sometimes dedicate two or more leaf switches to a particular function. Engineers typically evaluate the following categories of leaf switches as potential options for dedicating hardware:

**Key
Topic**

- **Border Leaf:** *Border leaf* switches provide Layer 2 and 3 connectivity between an ACI fabric and the outside world. Border leaf switches are sometimes points of policy enforcement between internal and external endpoints.
- **Service Leaf:** *Service leaf* switches are leaf switches that connect to Layer 4–7 service appliances, such as firewalls and load balancers.
- **Compute Leaf:** *Compute leaf* switches are ACI leaf switches that connect to servers. Compute leaf switches are points of policy enforcement when traffic is being sent between local endpoints.
- **IP Storage Leaf:** *IP storage leaf* switches are ACI leaf switches that connect to IP storage systems. IP storage leaf switches can also be points of policy enforcement for traffic to and from local endpoints.

There are scalability benefits associated with dedicating leaf switches to particular functions, but if the size of the network does not justify dedicating leaf switches to a function, consider at least dedicating a pair of leaf switches as border leaf switches. Service leaf functionality can optionally be combined with border leaf functionality, resulting in the deployment of a pair (or more) of collapsed border/service leaf switches in smaller environments.

Cisco publishes a Verified Scalability Guide for each ACI code release. At the time of this writing, 500 is considered the maximum number of leaf switches that can be safely deployed in a single fabric that runs on the latest code.

ACI Stretched Fabric Topology

A *stretched ACI fabric* is a partially meshed design that connects ACI leaf and spine switches distributed in multiple locations. The stretched ACI fabric design helps lower deployment costs when full-mesh cable runs between all leaf and spine switches in a fabric tend to be cost-prohibitive.

Figure 2-4 shows a stretched ACI fabric across two sites.

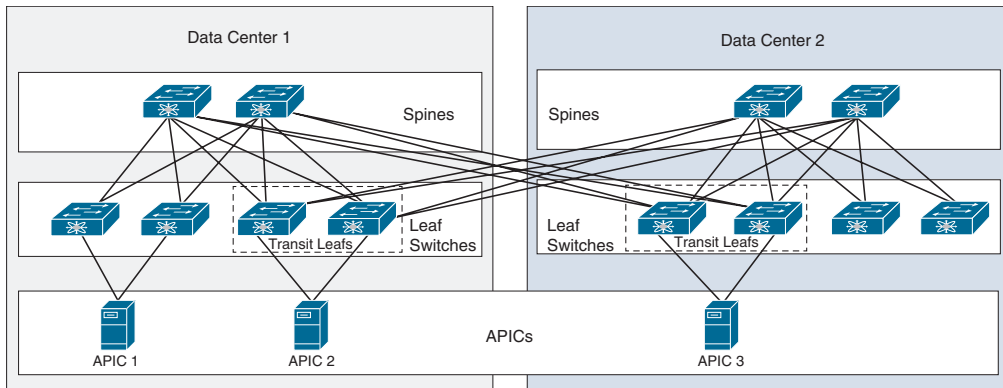


Figure 2-4 ACI Stretched Fabric Topology

A stretched fabric amounts to a single administrative domain and a single availability zone. Because APICs in a stretched fabric design tend to be spread across sites, cross-site latency is an important consideration. APIC clustering has been validated across distances of 800 kilometers between two sites.

A new term introduced in Figure 2-4 is *transit leaf*. A *transit leaf* is a leaf switch that provides connectivity between two sites in a stretched fabric design. Transit leaf switches connect to spine switches in both sites. No special configuration is required for transit leaf switches. At least one transit leaf switch must be provisioned in each site for redundancy reasons.

While stretched fabrics simplify extension of an ACI fabric, this design does not provide the benefits of newer topologies such as ACI Multi-Pod and ACI Multi-Site and stretched fabrics are therefore no longer commonly deployed or recommended.

ACI Multi-Pod Topology

Key Topic

The *ACI Multi-Pod* topology is a natural evolution of the ACI stretched fabric design in which spine and leaf switches are divided into pods, and different instances of IS-IS, COOP, and MP-BGP protocols run inside each pod to enable a level of control plane fault isolation.

Spine switches in each pod connect to an interpod network (IPN). Pods communicate with one another through the IPN. Figure 2-5 depicts an ACI Multi-Pod topology.

Key Topic

An ACI Multi-Pod IPN has certain requirements that include support for OSPF, end-to-end IP reachability, DHCP relay capabilities on the last-hop routers that connect to spines in each pod, and an increased maximum transmission unit (MTU). In addition, a Multi-Pod IPN needs to support forwarding of multicast traffic (PIM-Bidir) to allow the replication of broadcast, unknown unicast, and multicast (BUM) traffic across pods.

One of the most significant use cases for ACI Multi-Pod is active/active data center design. Although ACI Multi-Pod supports a maximum round-trip time latency of 50 milliseconds between pods, most Multi-Pod deployments are often built to achieve active/active functionality and therefore tend to have latencies of less than 5 milliseconds.

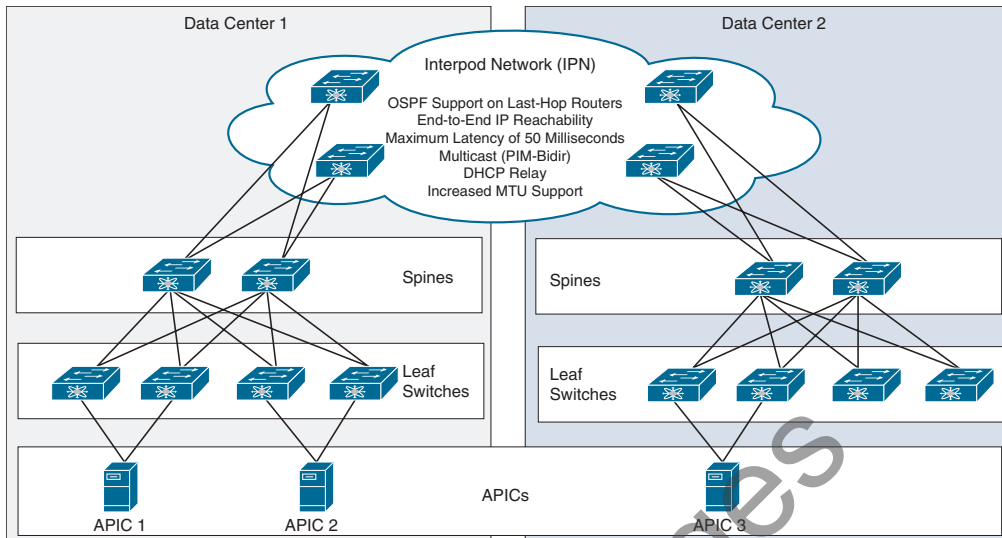


Figure 2-5 ACI Multi-Pod Topology

NOTE Another solution that falls under the umbrella of ACI Multi-Pod is Virtual Pod (vPod). ACI vPod is not a new topology per se. It is an extension of a Multi-Pod fabric in the form of a new pod at a remote location where at least two ESXi servers are available, and deployment of ACI hardware is not desirable. ACI vPod components needed at the remote site for this solution include virtual spine (vSpine) appliances, virtual leaf (vLeaf) appliances, and the Cisco ACI Virtual Edge. ACI vPod still requires a physical ACI footprint since vPod is managed by the overall Multi-Pod APIC cluster.

On the issue of scalability, it should be noted that as of the time of writing, 500 is the maximum number of leaf switches that can be safely deployed within a single ACI fabric. However, the Verified Scalability Guide for the latest code revisions specifies 400 as the absolute maximum number of leaf switches that can be safely deployed in each pod. Therefore, for a fabric to reach its maximum supported scale, leaf switches should be deployed across at least 2 pods within a Multi-Pod fabric. Each pod supports deployment of 6 spines, and each Multi-Pod fabric currently supports the deployment of up to 12 pods.

Chapter 16, “ACI Anywhere,” covers ACI Multi-Pod in more detail. For now, understand that Multi-Pod is functionally a single fabric and a single availability zone, even though it does not represent a single network failure domain.

ACI Multi-Site Topology



ACI Multi-Site is a solution that interconnects multiple ACI fabrics for the purpose of homogenous policy deployment across ACI fabrics, homogenous security policy deployment across on-premises ACI fabrics and public clouds, and cross-site stretched subnet capabilities, among others.

Key Topic

In an ACI Multi-Site design, each ACI fabric has its own dedicated APIC cluster. A clustered set of three nodes called Multi-Site Orchestrator (MSO) establishes API calls to each fabric independently and can configure tenants within each fabric with desired policies.

NOTE Nodes forming an MSO cluster have traditionally been deployed as VMware ESXi virtual machines (VMs). Cisco has recently introduced the ability to deploy an MSO cluster as a distributed application (.aci format) on Cisco Application Services Engine (ASE). Cisco ASE is a container-based solution that provides a common platform for deploying and managing Cisco data center applications. ASE can be deployed in three form factors: a physical form factor consisting of bare-metal servers, a virtual machine form factor for on-premises deployments via ESXi or Linux KVM hypervisors, and a virtual machine form factor deployable within a specific Amazon Web Services (AWS) region.

Figure 2-6 shows an ACI Multi-Site topology that leverages a traditional VM-based MSO cluster.

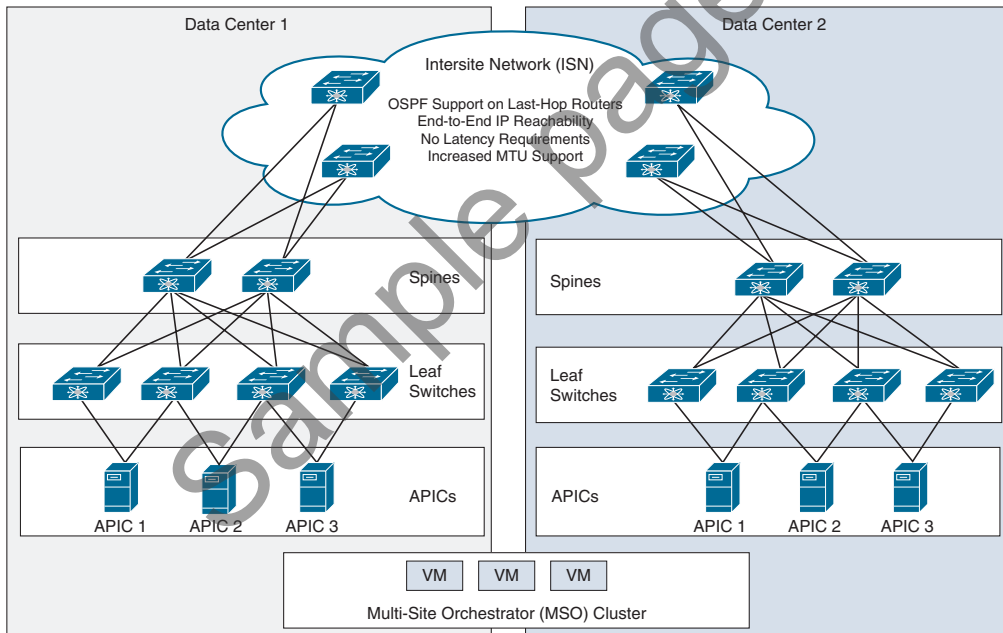


Figure 2-6 ACI Multi-Site Topology

Key Topic

As indicated in Figure 2-6, end-to-end communication between sites in an ACI Multi-Site design requires the use of an intersite network (ISN). An ACI Multi-Site ISN faces less stringent requirements compared to ACI Multi-Pod IPNs. In an ISN, end-to-end IP connectivity between spines across sites, OSPF on the last-hop routers connecting to the spines, and increased MTU support allowing VXLAN-in-IP encapsulation are all still required. However, ACI Multi-Site does not dictate any cross-site latency requirements, nor does it require support for multicast or DHCP relay within the ISN.

ACI Multi-Site does not impose multicast requirements on the ISN because ACI Multi-Site has been designed to accommodate larger-scale ACI deployments that may span the globe. It is not always feasible or expected for a company that has a global data center footprint to also have a multicast backbone spanning the globe and between all data centers.

**Key
Topic**

Due to the introduction of new functionalities that were not required in earlier ACI fabrics, Cisco introduced a second generation of spine hardware. Each ACI fabric within an ACI Multi-Site design requires at least one second-generation or newer piece of spine hardware for the following reasons:

- **Ingress replication of BUM traffic:** To accommodate BUM traffic forwarding between ACI fabrics without the need to support multicast in the ISN, Multi-Site-enabled spines perform ingress replication of BUM traffic. This function is supported only on second-generation spine hardware.
- **Cross-fabric namespace normalization:** Each ACI fabric has an independent APIC cluster and therefore an independent brain. When policies and parameters are communicated between fabrics in VXLAN header information, spines receiving cross-site traffic need to have a way to swap remotely significant parameters, such as VXLAN network identifiers (VNIDs), with equivalent values for the local site. This function, which is handled in hardware and is called *namespace normalization*, requires second-generation or newer spines.

Note that in contrast to ACI Multi-Site, ACI Multi-Pod *can* be deployed using first-generation spine switches.

For ACI Multi-Site deployments, current verified scalability limits published by Cisco suggest that fabrics with stretched policy requirements that have up to 200 leaf switches can be safely incorporated into ACI Multi-Site. A single ACI Multi-Site deployment can incorporate up to 12 fabrics as long as the total number of leaf switches in the deployment does not surpass 1600.

Each fabric in an ACI Multi-Site design forms a separate network failure domain and a separate availability zone.

ACI Multi-Tier Architecture

Introduced in Release 4.1, ACI Multi-Tier provides the capability for vertical expansion of an ACI fabric by adding an extra layer or tier of leaf switches below the standard ACI leaf layer.

With the Multi-Tier enhancement, the standard ACI leaf layer can also be termed the Tier 1 leaf layer. The new layer of leaf switches that are added to vertically expand the fabric is called the Tier 2 leaf layer. Figure 2-7 shows these tiers. APICs, as indicated, can attach to either Tier 1 or Tier 2 leaf switches.

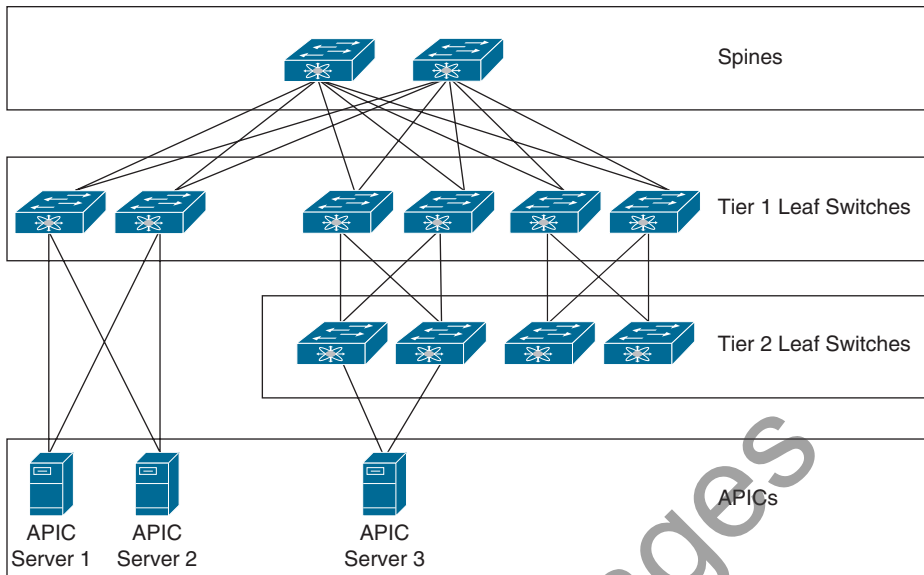


Figure 2-7 ACI Multi-Tier Topology

NOTE The topology shown in Figure 2-7 goes against the requirement outlined earlier in this chapter, in the section “Standard ACI Topology,” *not* to cross-connect leaf switches. The ACI Multi-Tier architecture is an exception to this rule. Leaf switches within each tier, however, still should never be cross-connected.

An example of a use case for ACI Multi-Tier is the extension of an ACI fabric across data center halls or across buildings that are in relatively close proximity while minimizing long-distance cabling and optics requirements. Examine the diagram in Figure 2-8. Suppose that an enterprise data center has workloads in an alternate building. In this case, the company can deploy a pair of Tier 1 leaf switches in the new building and expand the ACI fabric to the extent needed within the building by using a Tier 2 leaf layer. Assuming that 6 leaf switches would have been required to accommodate the port requirements in the building, as Figure 2-8 suggests, directly cabling these 6 leaf switches to the spines as Tier 1 leaf switches would have necessitated 12 cross-building cables. However, the use of an ACI Multi-Tier design enables the deployment of the same number of switches using 4 long-distance cable runs.

ACI Multi-Tier can also be an effective solution for use within data centers in which the cable management strategy is to minimize inter-row cabling and relatively low-bandwidth requirements exist for top-of-rack switches. In such a scenario, Tier 1 leaf switches can be deployed end-of-row, and Tier 2 leaf switches can be deployed top-of-rack.

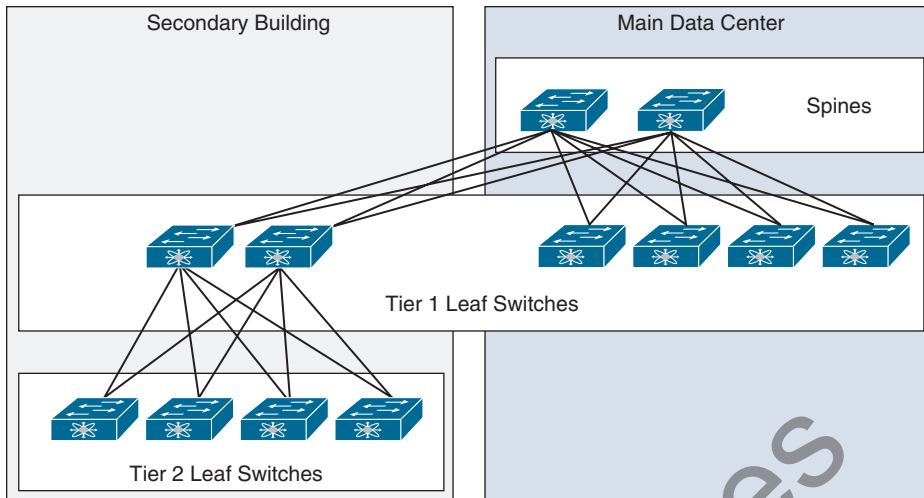


Figure 2-8 Extending an ACI Fabric by Using ACI Multi-Tier in an Alternative Location

NOTE ACI Multi-Tier *might not* be a suitable solution if the amount of bandwidth flowing upstream from Tier 2 leaf switches justifies the use of dedicated uplinks to spines.

Not all ACI switch platforms support Multi-Tier functionality.

Remote Leaf Topology

Key
Topic

For remote sites in which data center endpoints may be deployed but their number and significance do not justify the deployment of an entirely new fabric or pod, the ACI *Remote Leaf* solution can be used to extend connectivity and ensure consistent policies between the main data center and the remote site. With such a solution, leaf switches housed at the remote site communicate with spines and APICs at the main data center over a generic IPN. Each Remote Leaf switch can be bound to a single pod.

There are three main use cases for Remote Leaf deployments:

- **Satellite/small colo data centers:** If a company has a small data center consisting of several top-of-rack switches and the data center may already have dependencies on a main data center, this satellite data center can be integrated into the main data center by using the Remote Leaf solution.
- **Data center extension and migrations:** Cross-data center migrations that have traditionally been done through Layer 2 extension can instead be performed by deploying a pair of Remote Leafs in the legacy data center. This approach often has cost benefits compared to alternative Layer 2 extension solutions if there is already an ACI fabric in the target state data center.
- **Telco 5G distributed data centers:** Telecom operators that are transitioning to more distributed mini data centers to bring services closer to customers but still desire centralized management and consistent policy deployment across sites can leverage Remote Leaf for these mini data centers.

In addition to these three main use cases, disaster recovery (DR) is sometimes considered a use case for Remote Leaf deployments, even though DR is a use case more closely aligned with ACI Multi-Site designs.

In a Remote Leaf solution, the APICs at the main data center deploy policy to the Remote Leaf switches as if they were locally connected.

Figure 2-9 illustrates a Remote Leaf solution.

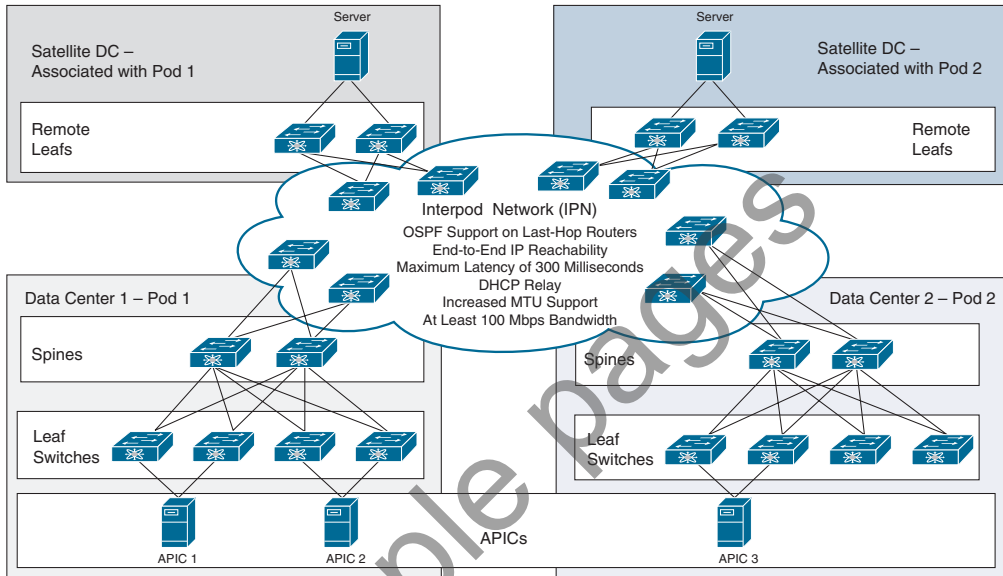


Figure 2-9 Remote Leaf Topology and IPN Requirements

IPN requirements for a Remote Leaf solution are as follows:

- **MTU:** The solution must support an end-to-end MTU that is at least 100 bytes higher than that of the endpoint source traffic. Assuming that 1500 bytes has been configured for data plane MTU, Remote Leaf can be deployed using a minimum MTU of 1600 bytes. An IPN MTU this low, however, necessitates that ACI administrators lower the ACI fabricwide control plane MTU, which is 9000 bytes by default.
- **Latency:** Up to 300 milliseconds latency between the main data center and remote location is acceptable.
- **Bandwidth:** Remote Leaf is supported with a minimum IPN bandwidth of 100 Mbps.
- **VTEP reachability:** A Remote Leaf switch logically associates with a single pod if integrated into a Multi-Pod solution. To make this association possible, the Remote Leaf should be able to route traffic over the IPN to the VTEP pool of the associated pod. Use of a dedicated VRF for IPN traffic is recommended where feasible.

- **APIC infra IP reachability:** A Remote Leaf switch needs IP connectivity with all APICs in a Multi-Pod cluster at the main data center. If an APIC has assigned itself IP addresses from a VTEP range different than the pod VTEP pool, the additional VTEP addresses need to also be advertised over the IPN.
- **OSPF support on upstream routers:** Routers northbound of both the Remote Leaf switches and the spine switches need to support OSPF and must be able to encapsulate traffic destined to directly attached ACI switches using VLAN 4. This requirement exists only for directly connected devices and does not extend end-to-end in the IPN.
- **DHCP relay:** The upstream router directly connected to Remote Leaf switches needs to enable DHCP relay to relay DHCP packets to the APIC IP addresses in the infra tenant. The DHCP relay configuration needs to be applied on the VLAN 4 subinterface or SVI.

Note that unlike a Multi-Pod IPN, a Remote Leaf IPN does not require Multicast PIM-Bidir support. This is because the Remote Leaf solution uses headend replication (HER) tunnels to forward BUM traffic between sites.

In a Remote Leaf design, traffic between known local endpoints at the remote site is switched directly, whether physically or virtually. Any traffic whose destination is in ACI but is unknown or not local to the remote site is forwarded to the main data center spines.

NOTE Chapter 16 details MTU requirements for IPN and ISN environments for ACI Multi-Pod and ACI Multi-Site. It also covers how to lower control plane and data plane MTU values within ACI if the IPN or ISN does not support high MTU values. Although it does not cover Remote Leaf, the same general IPN MTU concepts apply.

Not all ACI switches support Remote Leaf functionality. The current maximum verified scalability number for Remote Leaf switches is 100 per fabric.

APIC Clusters

The ultimate size of an APIC cluster should be directly proportionate to the size of the Cisco ACI deployment. From a management perspective, any active APIC controller in a cluster can service any user for any operation. Controllers can be transparently added to or removed from a cluster.

Key Topic

APICs can be purchased either as physical or virtual appliances. Physical APICs are 1 rack unit (RU) Cisco C-Series servers with ACI code installed and come in two different sizes: M for medium and L for large. In the context of APICs, “size” refers to the scale of the fabric and the number of endpoints. Virtual APICs are used in ACI mini deployments, which consist of fabrics with up to two spine switches and four leaf switches.

Key Topic

As hardware improves, Cisco releases new generations of APICs with updated specifications. At the time of this writing, Cisco has released three generations of APICs. The first generation of APICs (M1/L1) shipped as Cisco UCS C220 M3 servers. Second-generation APICs (M2/L2) were Cisco UCS C220 M4 servers. Third-generation APICs (M3/L3) are shipping as UCS C220 M5 servers.