

Cisco IT ACI Design



This white paper is the first in a series of case studies that explains how Cisco IT deployed ACI to deliver improved business performance. These in-depth case studies cover the Cisco IT ACI data center design, migration to ACI, network security, the ACI NetApp storage area network deployment, and virtualization with AVS, UCS, KVM, and VMware. These white papers will enable field engineers and customer IT architects to assess the product, plan deployments, and exploit its application centric properties to flexibly deploy and manage robust highly scalable integrated data center and network resources.

Table of Contents

Cisco IT ACI Fabric Design Goals.....	3
Uniform ACI Fabric Infrastructure Topologies.....	5
ACI Fabric Logical Constructs	15
ACI VLAN Automation Contributes to Near-Zero Downtime and Lower Operating Costs	18
Enhanced Security.....	18
Virtual Compute Integration	23
Reporting and Alerting.....	25
Automation	26
Conclusion.....	27

Cisco IT ACI Fabric Design Goals

The Cisco® IT deployment of Application Centric Infrastructure (ACI) enables its global data center network to deliver the enhanced business value they must have – compelling total cost of ownership, near 100% availability, and agility that includes letting business applications developers directly provision the infrastructure resources they need in a self-service fashion.

Worldwide Data Centers



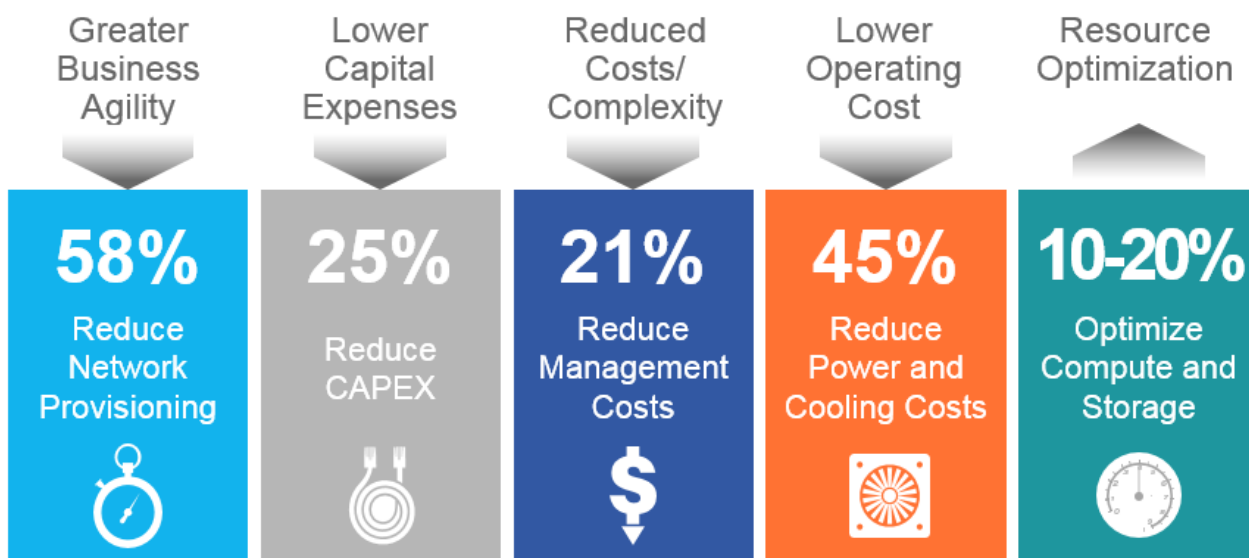
The Cisco IT organization operates multiple business application and engineering development data centers distributed throughout the world. The infrastructure for each data center (DC) is big. For example, the Allen, Texas DC is just one of 30 worldwide. The 856 network devices in the Allen DC support 2300 traditional and private-cloud applications, run 8000 virtual machines, include 1700 Cisco Unified Computing System™ (Cisco UCS®) blades and 710 bare metal servers, with 14.5PB of NAS storage and 12PB of SAN storage. As Cisco's data centers grow, quick and agile application deployment becomes increasingly challenging.

Cisco ACI enables Cisco IT to use a common application-aware policy-based operating model across their physical and virtual environments. The ACI deployment high level design objectives include the following:

- Provision anything anywhere within a data center
- Manage compute, storage, and network resource pools within virtual boundaries
- Cost effectively deliver near-zero application down time
- Take advantage of the ACI policy-driven model to more easily design for reuse and automation
- Enhance network access security and domain based role based user access control

Realizing these objectives enables Cisco IT to deliver the enhanced business value to the enterprise summarized in the illustration below (refer to this [IDC business value brief](#)).

Cisco IT Projected ACI Benefits



Benny Van De Voorde, Cisco IT Architect explains, "One of the unique design opportunities in ACI is for us to specify core infrastructure services once for the entire fabric then let applications developers directly consume them according to their application requirements." This white paper details how Cisco IT designed its ACI deployment to do just that.

Uniform ACI Fabric Infrastructure Topologies

While standardization and reuse as a data center design strategy is not new, provisioning data center infrastructure according to software defined standardized constructs is transformative. The combination of standardized data center ACI fabric topologies and software defined standardized constructs enables seamless dynamic provisioning of any data center workload anywhere.

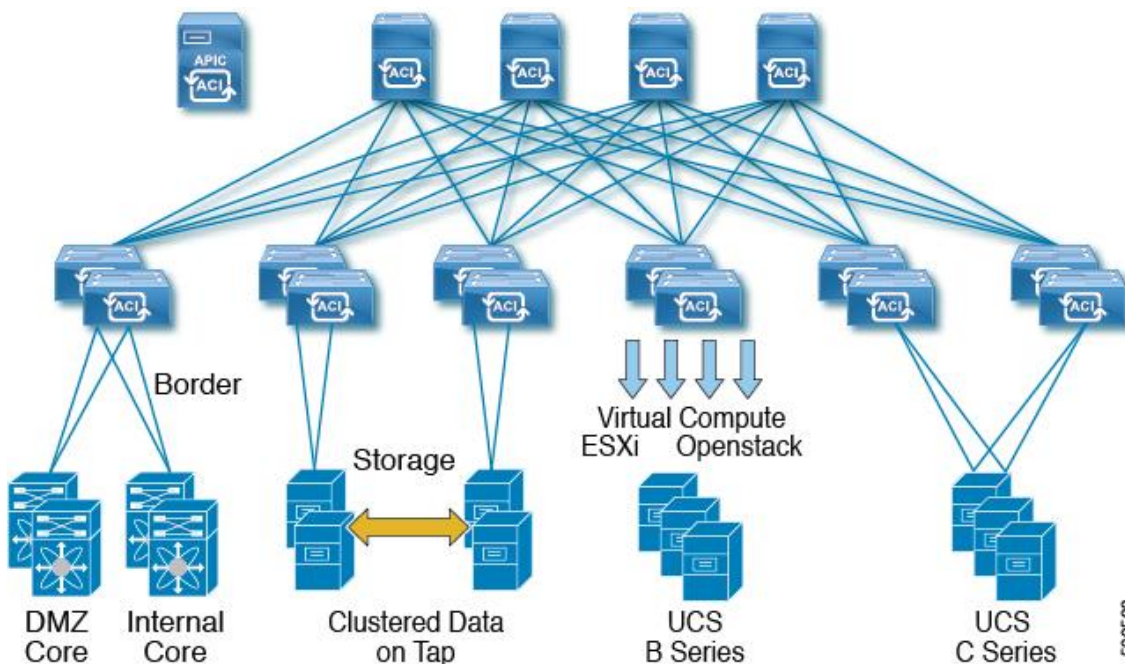
Template Driven Data Center Standard Topologies

Depending on the size of the workload requirement, Cisco IT deploys uniform ACI data center topologies.

Standard Cisco IT ACI Data Center Fabrics

The standard data center (DC) has four spine switches, one pair of border leaf switches for external connectivity, two or more pairs of leaf switches for end point connectivity, and the minimum supported number of three APIC controllers.

Standard Data Center



The scale out capacity is 288 leaf switches with up to 12 40GB links between each spine and leaf switch. Cisco IT uses the standard DC ACI fabric topology in production data

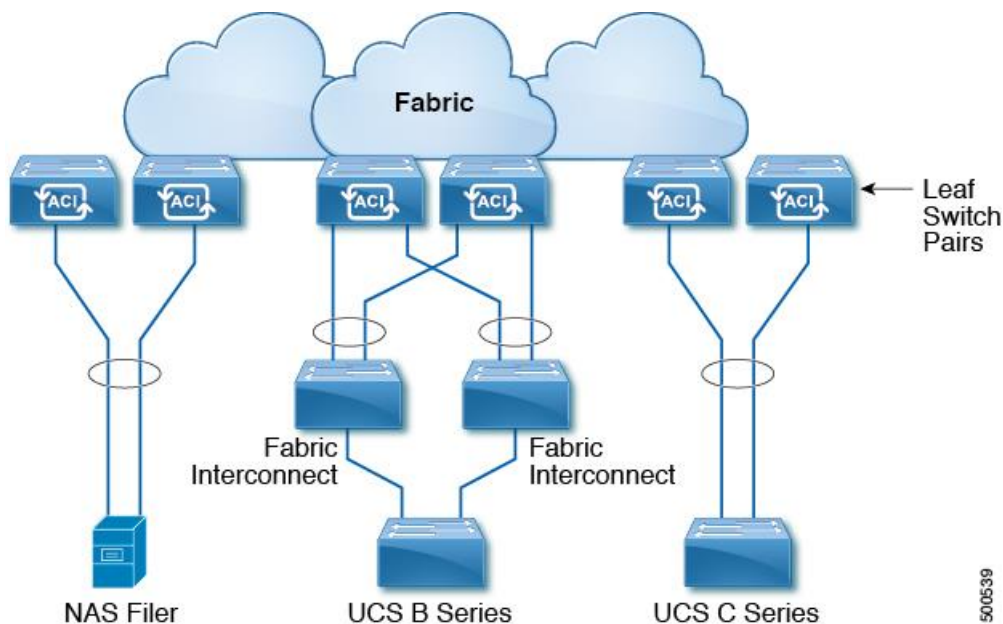
centers such as those in Research Triangle Park, North Carolina, Richardson, Texas, and Allen, Texas.

The primary difference between the standard and small DC is the model of the spine switch. The small DC ACI fabric is suitable for a small-to-medium sized DC such as Amsterdam in the Netherlands. The small DC has four spine switches, one pair of border leaf switches, one pair of leaf switches for end point connectivity, and three APIC controllers. The scale out capacity goes to 36 leaf switches with up to 12 40GB links between each spine and leaf switch.

Virtual Port Channel Templates

In Cisco IT ACI deployments, a pod is a pair of leaf switches that provides virtual port channel (vPC) connectivity to end points, although it is not mandatory for an end port to be connected via vPC.

vPC Connectivity



Connecting devices such as a UCS Fabric Interconnect (FI), NAS filer, or Fabric Extender (FEX) to a leaf switch pair using a vPC provides increased resiliency and redundancy. Unlike a vPC on the Nexus 5/6/7K platforms, an ACI vPC leaf switch pair does not need direct physical connectivity peer links to each other.

Compute and IP Storage Templates

The Cisco IT standardized compute and storage pod templates enable applications to

flexibly tap into any available compute or storage resources.

UCS B Series Compute Template

Cisco UCS B series clusters provide the majority of compute resources in a DC. A UCS B series compute pod has up to 3 UCS domains (clusters). A typical domain has 5 chassis, with up to 8 blades per chassis (120 physical servers per pod). Each fabric interconnect has four uplinks, two to each leaf switch. When both intra and inter-rack very low latency high bandwidth is required, ACI leaf switches are placed directly in the server cabinet and the servers connect directly to them via 10 gigabit Ethernet.

Each UCS B series domain has dual fabric interconnects (A and B side), with each FI having four 10GE uplinks, spread equally between the two leaf switches in the pod pair. The links are setup in vPC mode and both FIs are active. This arrangement provides a total of 80Gbps for every UCS cluster.

Using four 10GE uplinks from each UCS B series domain to each leaf switch is a total of 4x10GE interfaces required on the leaf switches. The leaf switches can support two more UCS domains, but the remaining 10GE interfaces on the leaf switches are left available for monitoring systems, etc.

UCS C Series High Density Compute Template

New applications and solutions that follow a horizontal scale out philosophy such as Hadoop and Ceph storage are driving a new type of pod where the goal is to have as many UCS C series servers as possible within a rack. In this topology, the C series servers connect directly to the ACI leaf switches.

Legacy Compute Template

Although UCS B series servers are the current standard and most prevalent compute platform, there are still many legacy servers supported in the ACI fabric. The connectivity required for these servers ranges from 10Gbps down to 100Mbps, some copper, some fiber. The leaf switches support as low as 1/10Gbps classical Ethernet. To support the older required Ethernet connections, fabric extenders (FEX) are used. For consistency, all legacy servers connect to the fabric via a FEX. That is, no legacy server connects directly to a leaf switch, even if it has 1Gbps capability.

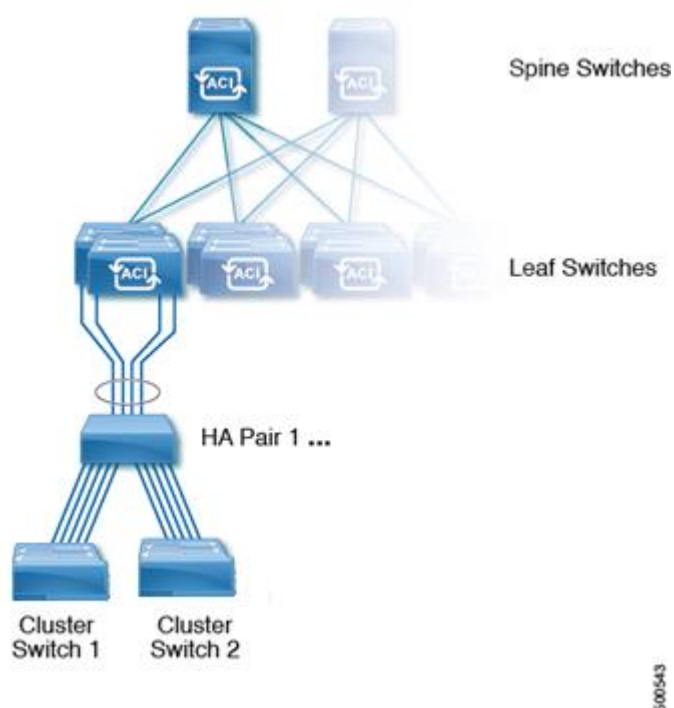
Each FEX uplink connection to a single leaf switch is via four 10GE uplinks arranged in a port channel. If downstream legacy switches require a vPC, this is configured. However,

server connectivity is more often set up with primary and standby interfaces spread between FEXs on different leaf switches.

IP Storage Template

Cisco IT has moved from a filer per pod NAS model to a consolidated/centralized NAS model. NAS filers are run on dedicated leaf switch pairs.

NetApp cDOT Storage Cluster Template



Each filer head has a primary link made up of four 10GE interfaces in a virtual port channel (two 10GE to each leaf switch in the pair).

Cisco's existing NetApp NAS implementation uses the FAS80xx all flash platforms Clustered Data ONTAP (cDOT) based virtual arrays presented to Cisco Enterprise Linux (CEL) hosts. NetApp storage efficiency features such as de-duplication are widely used at Cisco. Unlike most de-dup technology, NetApp single instance store (SIS) can be used with the primary data storage and structured data formats such as Oracle databases. Cisco has multiple copies of several moderate to large Oracle databases aligned into development tracks. These instances today occupy multiple Petabytes (PB) of storage, consuming a large amount of the data center resources in Research Triangle Park, NC (RTP), Richardson, TX (RCDN), and Allen, TX (ALLN).

The change from 7-Mode NAS to cDOT allows a filer IP to failover between two physical NAS filer heads. The cDOT NAS cluster shares the load between the two physical NAS filers by making one half of the IP addresses active on one leaf switch pair and the other half on a second leaf pair. Should one filer fail, then the IP addresses that were active on that filer come up automatically on the other filer in the pair.

Border Leaf Template

The Cisco IT ACI border switch topology is a pair of leaf switches configured for connecting to networks outside the ACI fabric. Ethernet ports on an ACI leaf switch connect to upstream data center core switches. Any ACI leaf switch in the fabric can be a border leaf. Cisco IT dedicates a pair of leaf switches to this function because they are located physically closer to the upstream network than the rest of the leaf switches in the fabric. While it is not a requirement that border leaf switches be dedicated to external network connectivity, a large data center that supports high volume traffic between the ACI fabric and the core network might choose to dedicate leaf switches to providing these services.

The Cisco IT border leaf switches run EIGRP to the upstream network switches/routers. The data center core advertises the routes learned from the border leaf switches to the rest of the Cisco internal network.

L4-L7 Services

L4-L7 services can be integrated in ACI Fabric in two ways:

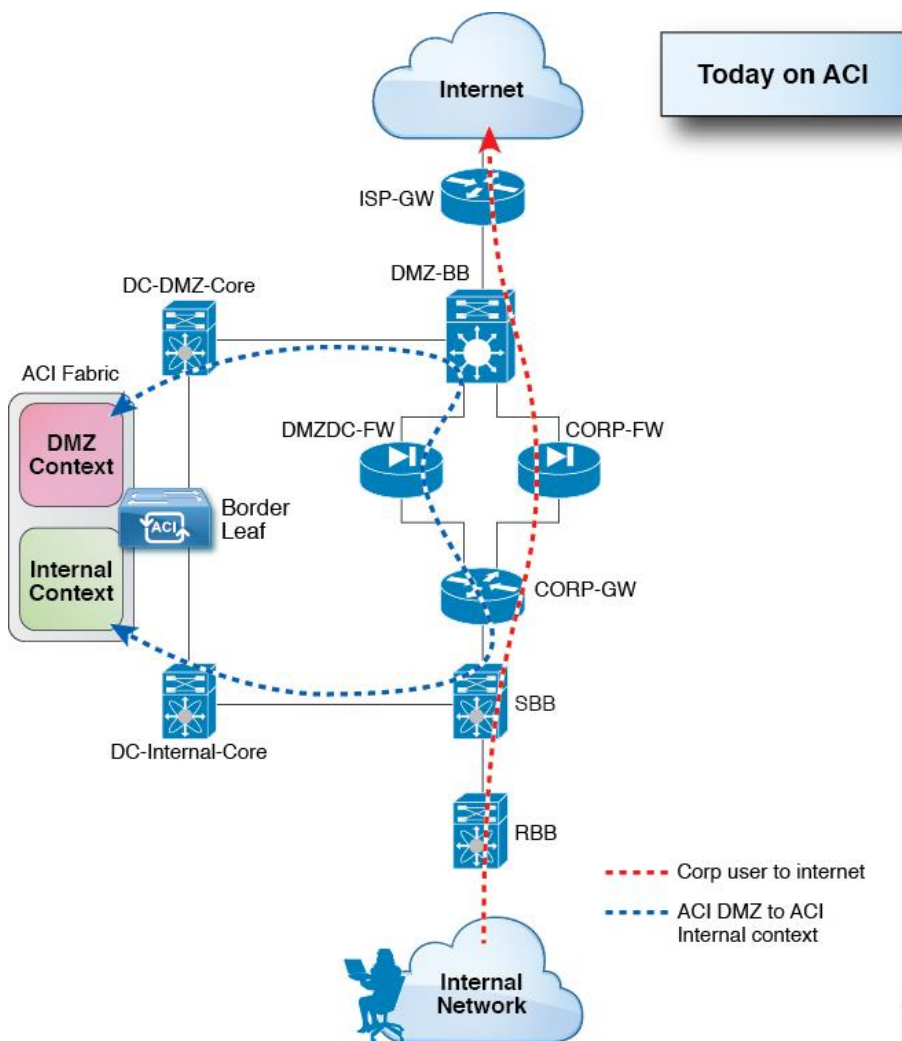
- Service Graphs
- Directly on the L4-L7 device

Today, Cisco IT runs enhanced network service appliances – firewalls, load balancers, and so forth – on physical appliances but is migrating to virtual appliance firewalls that run on top of a hypervisor.

Current DMZ Template

The ACI fabric does not provide firewall services such as stateful session inspection or unified threat management deep packet inspection. This level of security is satisfied with an external firewall. Today, Cisco IT uses the firewall services solution illustrated in the following figure.

Current Firewall solution



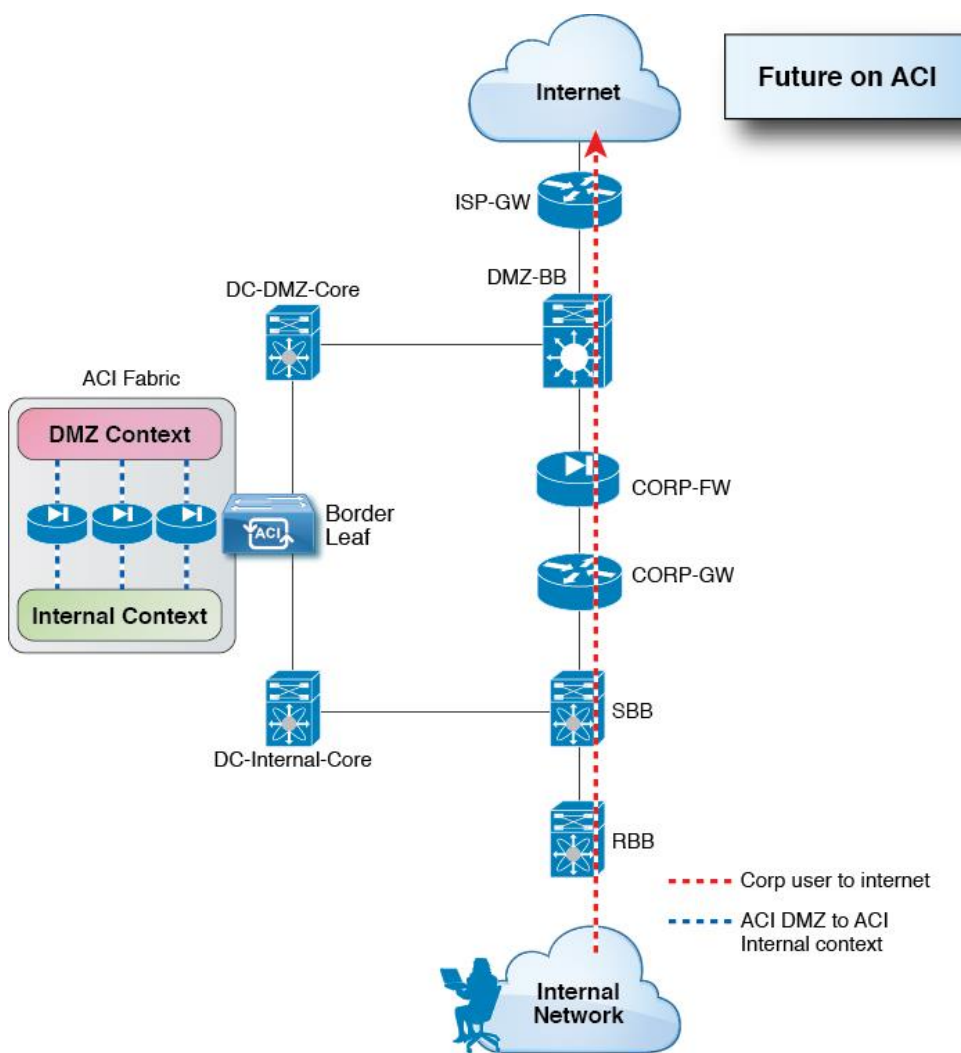
This solution locates physical Cisco ASA 5500 Series Adaptive Security Appliances (ASA) outside the ACI fabric. The dedicated pair of border leaf switches are configured with uplink connections to both the DMZ and internal networks. Both fabric connections are required to uplink to the data center network core. This makes the fabric look like another DC pod from a layer 3 routing perspective. In the case where the ACI fabric is the only DC network in the facility, the fabric can uplink directly to the network core for that site. Fabric to DMZ routing is done in the same way as any other DC pod. The subnets in the DMZ fabric context (VRF) are advertised to the DMZ.

This solution will be replaced shortly with the more flexible solution discussed below.

Target DMZ Template

Cisco IT prefers firewall services to be delivered using virtualized appliances. In cases where a single instance of a network service device needs to have high levels of performance, then a physical appliance is still used. Virtualized appliances scale out easily, adding capacity quickly only when actually needed. Another advantage of many smaller network service devices over fewer bigger ones is that the impact of a fault on a network service appliance is smaller.

Target ACI Firewall solution

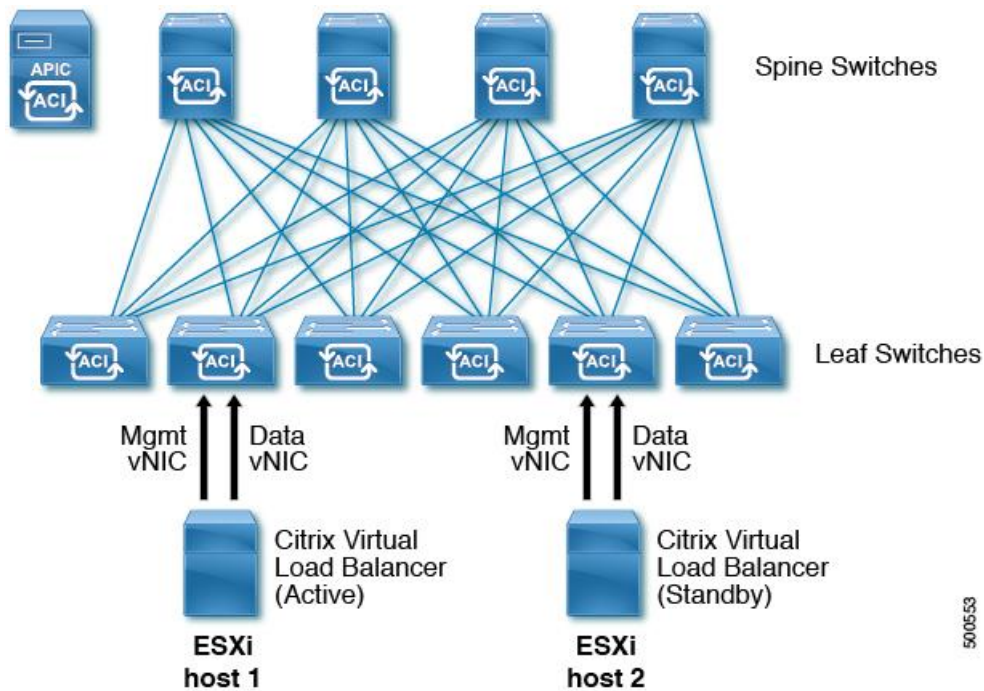


The Cisco IT target firewall solution uses ACI L4-L7 service graphs to place multiple virtual ASA appliances inside the ACI fabric. This solution provides simple automation that enables smaller firewalls that can be deployed per application.

Server Load Balancer Template

Cisco IT uses CITRIX virtual server load balancers across its ACI data center deployments.

Citrix Virtual Load Balancer

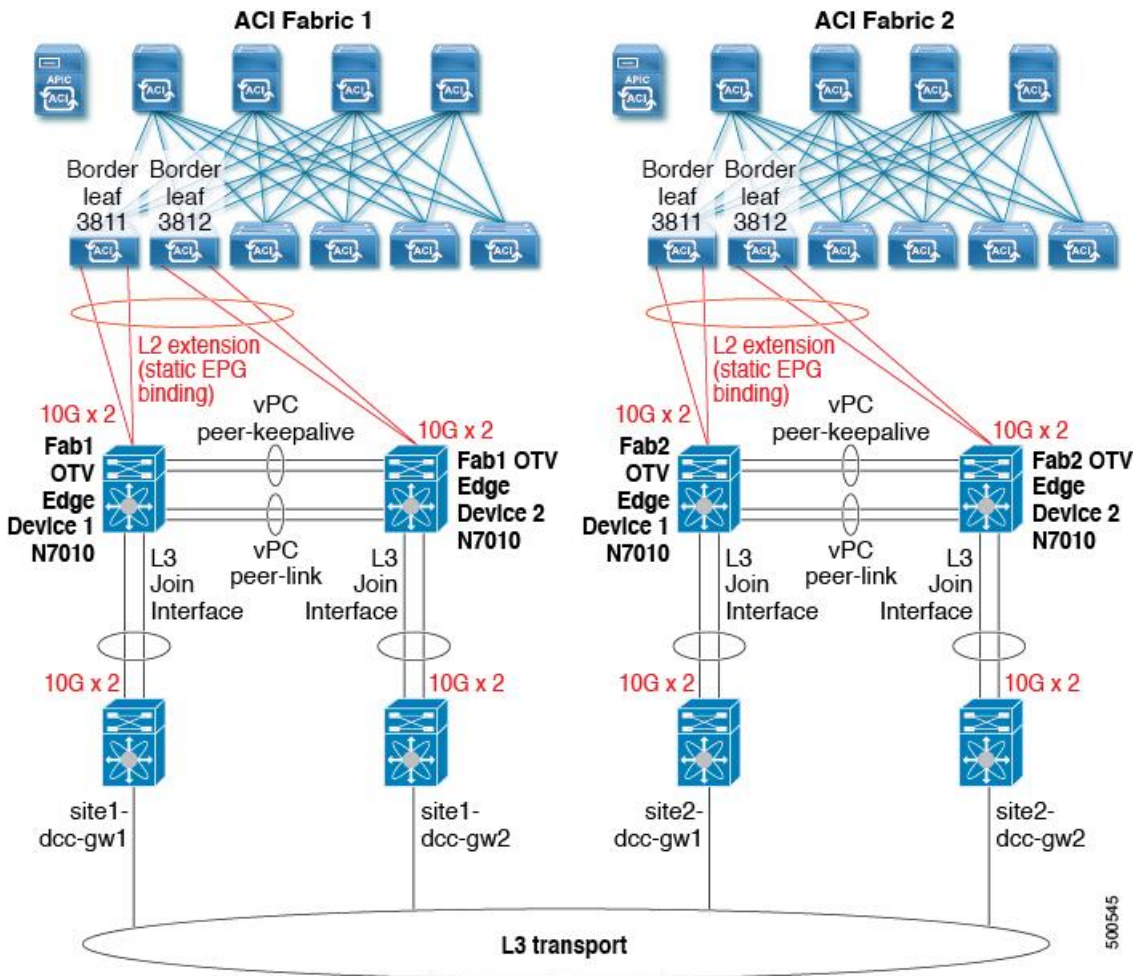


OTV Layer 2 Extensions

Layer 2 extensions enable multiple layer 2 bridge domains to be joined over a layer 3 transport network. Cisco IT uses Overlay Transport Virtualization (OTV) in the traditional networks to provide Layer 2 extensions between data centers. OTV is a protocol designed specifically for Data Center Interconnection (DCI). It offers many built-in functions that require no configuration, such as fault isolation and loop prevention. Built-in features include the elimination of L2 unknown unicast flooding and controlled ARP flooding over the overlay, as well as providing a boundary and site isolation of the local STP domain.

The primary ACI OTV use case is the storage team's implementation of NetApp MetroCluster for high availability within and between data centers.

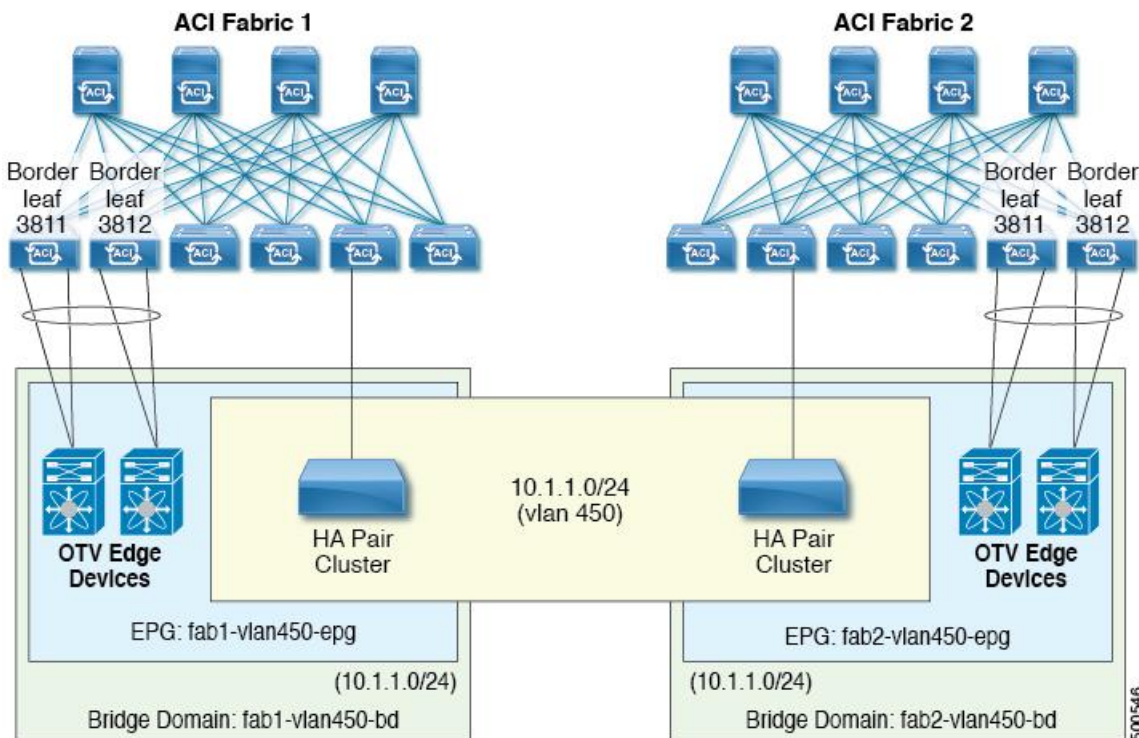
OTV Topology



OTV is deployed on 2 Nexus 7010s dedicated per fabric. Each N7010 is equipped with dual supervisors (N7K-SUP2E) and dual line cards (N7K-M132XP-12L).

L2 connectivity between the ACI fabric and the OTV edge devices is via a double-sided vPC. L3 connectivity between OTV edge devices and upstream data center network core (DCC) gateways is via a traditional 2 member port-channel. The OTV edge device pair for each fabric has two separate port-channels directly between them for vPC peer-keepalive and vPC peer-link configurations.

OTV ACI Logical Constructs



On the ACI fabric side, the OTV L2 connections are to border leaf switch pairs. In ACI, the endpoint group (EPG) is a collection of endpoints (physical or virtual) that are connected directly or indirectly to the network. The Cisco IT OTV border leaf interfaces are mapped to an EPG via static VLAN to EPG bindings.

In ACI, the bridge domain (BD) defines a unique Layer 2 MAC address space and a Layer 2 flood domain if such flooding is enabled. When interconnecting two ACI fabrics, the associated BD MAC addresses must be unique per fabric so that ARP broadcasts work properly. The ACI default BD MAC address is used for the BD in one of the fabrics; the BD MAC address in the other fabric is configured to be different. The ACI fabric default is for BD ARP flooding to be disabled, but the Cisco IT ACI/OTV configuration requires it to be enabled while keeping the ACI default of L2 unknown unicast flooding being disabled.

An external BD must be associated with an EPG that is used with OTV. The OTV gateway vPCs must have BPDU Filter enabled so as to provide high availability during failover scenarios and avoid lengthy periods of traffic loss during these periods.

The Nexus 7010 OTV edge devices use the intermediate system to intermediate system protocol (IS-IS) hello interval on the OTV join interface set to a tested value that enables fast re-convergence during failover. The site-VLAN is added to the allowed VLAN list on

the ACI facing port-channels, along with the extended VLANs, to enable the OTV edge device to become the active forwarder (AED) in the event the other OTV edge device in a site fails. Spanning-tree is enabled on the N7Ks, however BPDUs are filtered at the ACI fabric leaf switch ports.

Extended BD VLANs in ACI are set to `public` so that their subnets are advertised from ACI to the DCC gateways. Routes for the extended VLAN subnets must be filtered at the appropriate DCC gateways in order to preference ingress traffic coming into the DC towards the *home* of the extended VLAN subnet. This configuration is used today for OTV in the traditional network. An EIGRP distribute-list is configured on the DCC interfaces towards the SBB gateways, filtering the extended VLAN subnets only. The `DENY_OTV` prefix-list is updated accordingly on the DCC gateways.

ACI Fabric Logical Constructs

The ACI policy model is the basis for managing the entire fabric, including the infrastructure, authentication, security, services, applications, and diagnostics. Logical constructs in the policy model define how the fabric meets the needs of any of the functions of the fabric. From the point of view of data center design, the following three broad portions of the policy model are most relevant:

- Infrastructure policies that govern the operation of the equipment.
- Tenant policies that enable an administrator to exercise domain-based access control over the traffic within the fabric and between the fabric and external devices and networks.
- Virtual Machine Manager (VMM) domain policies that group VM controllers with similar networking policy requirements.

Tenant policies are the core ACI construct that enable business application deployment agility. Tenants can map to logical segmentation/isolation constructs of public cloud service providers. Tenants can be isolated from one another or can share resources.

Within a tenant, bridge domains define a unique Layer 2 MAC address space and a Layer 2 flood domain if such flooding is enabled. A bridge domain must be linked to a context (VRF) and have at least one subnet that is associated with it. While a context (VRF) defines a unique IP address space, that address space can consist of multiple subnets. Those subnets are defined in one or more bridge domains that reference the

corresponding context (VRF). Subnets in bridge domains can `public` (exported to routed connections), `private` (used only within the tenant) or `shared` across contexts (VRFs) and across tenants.

The endpoint group (EPG) is the most important object in the policy model. Endpoints are devices that are connected to the network directly or indirectly. EPGs are fully decoupled from the physical and logical topology. Endpoint examples include servers, virtual machines, network-attached storage, external Layer 2 or Layer 3 networks, or clients on the Internet. Policies apply to EPGs, never to individual endpoints. An EPG can be statically configured by an administrator, or dynamically configured by an automated system such as vCenter or OpenStack.

EPGs and bridge domains are associated with networking domains. An ACI fabric administrator creates networking domain policies that specify ports, protocols, VLAN pools, and encapsulation. These policies can be used exclusively by a single tenant, or shared. The following networking domain profiles can be configured:

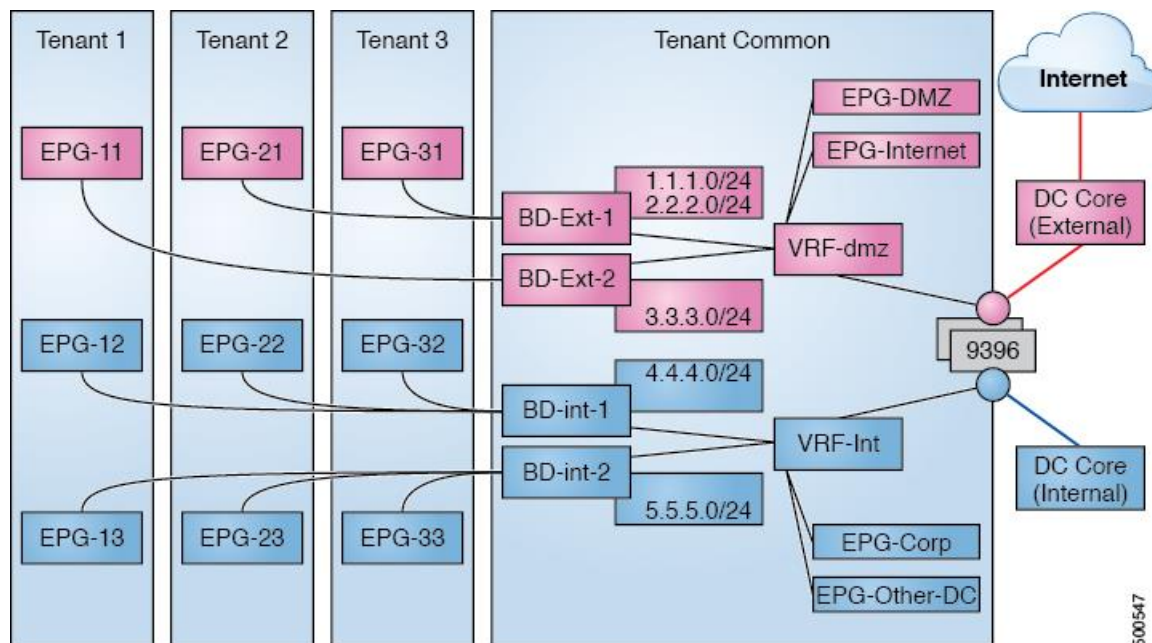
- VMM domain profiles are required for virtual machine hypervisor integration.
- Physical domain profiles are typically used for bare metal server attachment and management access.
- Bridged outside network domain profiles are typically used to connect a bridged external network trunk switch to a leaf switch in the ACI fabric.
- Routed outside network domain profiles are used to connect a router to a leaf switch in the ACI fabric.

A domain is configured to be associated with a VLAN pool. EPGs are then configured to use the VLANs associated with a domain.

Virtual machine management connectivity to a hypervisor is an example of a configuration that uses a dynamic EPG. Once the virtual machine management domain is configured in the fabric, the hypervisor triggers the dynamic configuration of EPGs that enable virtual machine endpoints to start up, move, and shut down as needed.

The following figure provides an overview of the Cisco IT implementation of ACI tenant constructs.

Networking Design Logical View: EPG to BD Subnets to VRFs to External (L3Out)



In the ACI fabric, a context is a VRF. Cisco IT uses two routing contexts (VRFs) within the fabric, one for DMZ/external and one for internal. This assures that there is complete isolation between the DMZ and internal security zones. Cisco IT minimizes the number of ACI contexts (VRFs) they deploy for the following reasons:

- Simplicity – lots of cross talk among the thousands of production applications.
- Avoid IP overlap.
- Avoid route leaking.

There are important differences between VLANs and BDs.

- BDs, by default, do not flood broadcast, multicast, or unknown unicast packets.
- The policy model does not rely on VLANs to segment and control traffic between hosts.
- Hosts in different subnets can be in the same BD.

IP subnets are configured in the network by adding them to BDs. Many IP subnets can be configured per BD.

The ACI fabric can support a single BD per fabric with all subnets configured onto that single BD. Alternatively, the ACI fabric can be configured with a 1:1 mapping from BD to subnet. Depending on the size of the subnet, Cisco IT configures one to five subnets per BD.

It is important to note that from a forwarding perspective, the fabric is completely self-managing. That is, the ACI fabric does not need any specific configuration for L2/3 forwarding within the fabric.

ACI VLAN Automation Contributes to Near-Zero Downtime and

Lower Operating Costs

Cisco, in partnership with other leading vendors, proposed the Virtual Extensible LAN (VXLAN) standard to the IETF as a solution to the data center network challenges posed by traditional VLAN technology. The VXLAN standard provides for elastic workload placement and higher scalability of Layer2 segmentation.

The ACI fabric VXLAN technology enables highly automated deployment of VLANs that are decoupled from the underlying physical infrastructure. The ACI fabric automatically provisions static or dynamic VLAN allocations from specified VLAN pools within the scope of a specified networking domain. This not only frees Cisco IT from the chore of managing the details of VLAN configurations, it also enables Cisco IT to evacuate a compute or IP storage system for maintenance purposes. This enables completing network, storage, compute upgrades (software or hardware), or infrastructure upgrades in data centers without application downtime.

Enhanced Security

By default, endpoints can communicate freely within a single EPG but are not permitted to talk to any device in any other EPG. If necessary, ACI microsegmentation and intra-EPG deny policies that restrict endpoint communications within an EPG can provide granular endpoint security enforcement to any virtual or physical endpoint(s) in a tenant. Traffic between EPGs must be explicitly permitted (ie: a whitelist security model) via the use of contracts. The contract is able to match application traffic through layer 3-4 matching and permit or drop appropriately.

A Cisco ACI fabric is inherently secure because it uses a zero-trust model and relies on

many layers of security.

All user system access or API calls require AAA and role-based access control that restricts the read/write access of tenant sub-tree read or write. Northbound interfaces utilize certificates and encryption. Rogue or counterfeit devices cannot access fabric resource because the ACI fabric uses a hardware key store and requires certificate based authentication. Within the fabric, the infrastructure VLAN (used for APIC to switch communication) is an isolated space and all messages are encrypted. All software images and binaries are signed and verified before they can boot up a device within the ACI fabric.

All management interfaces (representational state transfer [REST], command-line interface [CLI] and GUI) are authenticated in Cisco ACI using authentication, authorization, and accounting (AAA) services (LDAP and Microsoft Active Directory, RADIUS, and TACACS+) and role based access control (RBAC) policies, which map users to roles and domains.

Cisco IT configures ACI RBAC using TACACS+ user authentication that assigns each user to their corresponding domain and role within that domain.

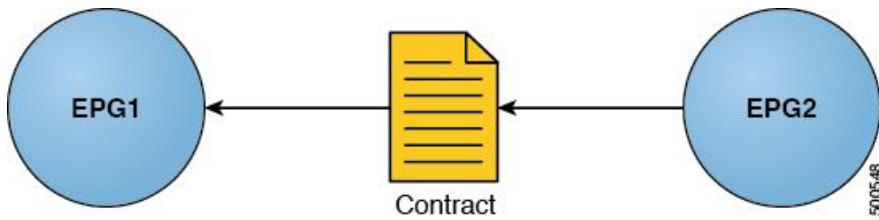
Contracts Govern Communications: Without a Contract, Data Does Not Flow

In the ACI security model, contracts contain the policies that govern the communication between EPGs. The contract specifies what can be communicated and the EPGs specify the source and destination of the communications.

Endpoints in EPG 1 can communicate with endpoints in EPG 2 and vice versa if the contract allows it. This policy construct is very flexible. There can be many contracts between EPG 1 and EPG 2, there can be more than two EPGs that use a contract, and contracts can be reused across multiple sets of EPGs.

This providing/consuming relationship is typically shown graphically with arrows between the EPGs and the contract. There is directionality in the relationship between EPGs and contracts. EPGs can either provide or consume a contract. Note the direction of the arrows in the following illustration.

Application Contracts



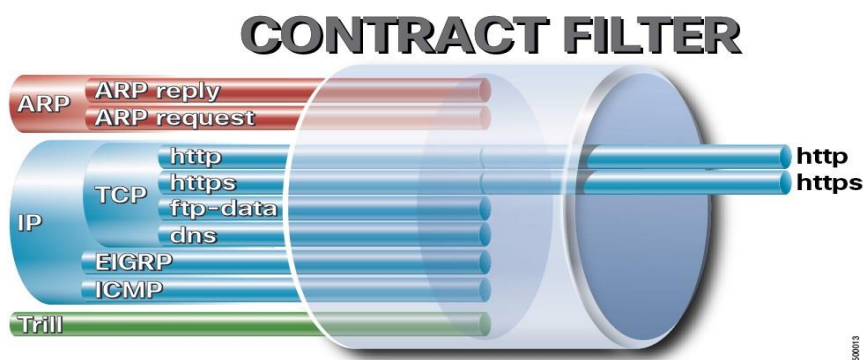
In this example, EPG1 is the *consumer* and EPG2 is the *provider*.

An EPG that provides a contract is typically a set of endpoints that provide a service to a set of client devices, such as web servers or DNS servers. The protocols used by that service are defined in the contract. An EPG that consumes a contract is typically a set of endpoints that are clients of that service. When the client endpoint (consumer) tries to connect to a server endpoint (provider), the contract specifies if that connection is allowed. Unless otherwise specified, that contract would not allow a server to initiate a connection to a client. However, another contract between the EPGs could easily allow a connection in that direction.

A contract consists of filters and actions. A contract is constructed in a hierarchical manner. It consists of one or more subjects, each subject contains one or more filters, and each filter can define one or more protocols. The contract action Cisco IT uses most is *allow* (forward).

The filter is used to match traffic based on layer 3 and layer 4 information. For example, a web server might provide the contract filter illustrated below that specifies http and https traffic.

Contract Filters

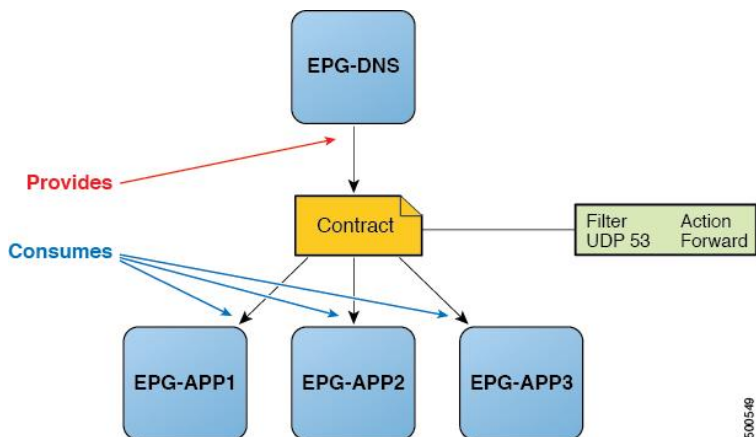


In this example, the contract would *allow* http and https traffic. So, only http and http

traffic would be allowed between EPGs that use this contract.

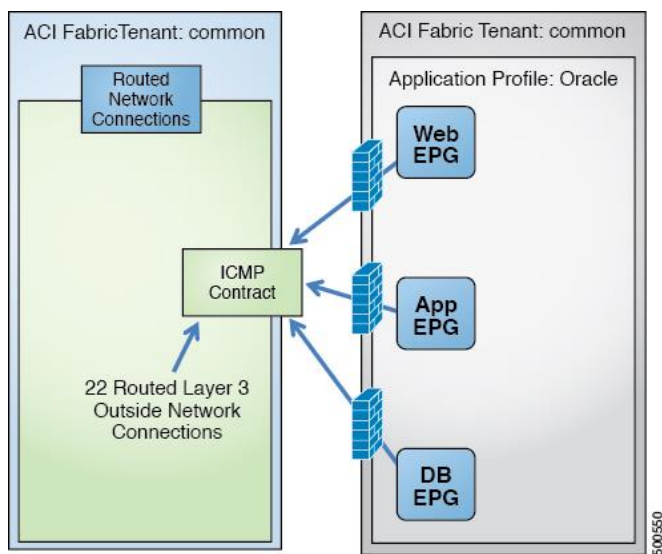
Another example would be an EPG that has DNS servers in it provides a contract allowing UDP port 53. As shown in the following illustration, this contract enables EPGs to consume the DNS service from this EPG.

Provided and Consumed Contracts



Standard network management protocols such as SSH and SNMP that most servers would expose are set up once in individual contracts that are reused across the fabric. For example, a contract specifies that the Internet Control Message Protocol (ICMP) is allowed.

Contract Example: One ICMP Contract for Many L3 Out Connections



From then on, a single contract that uses the vzAny wildcard feature of contracts can be reused for all routed connections automatically by virtue of the host being a member of a

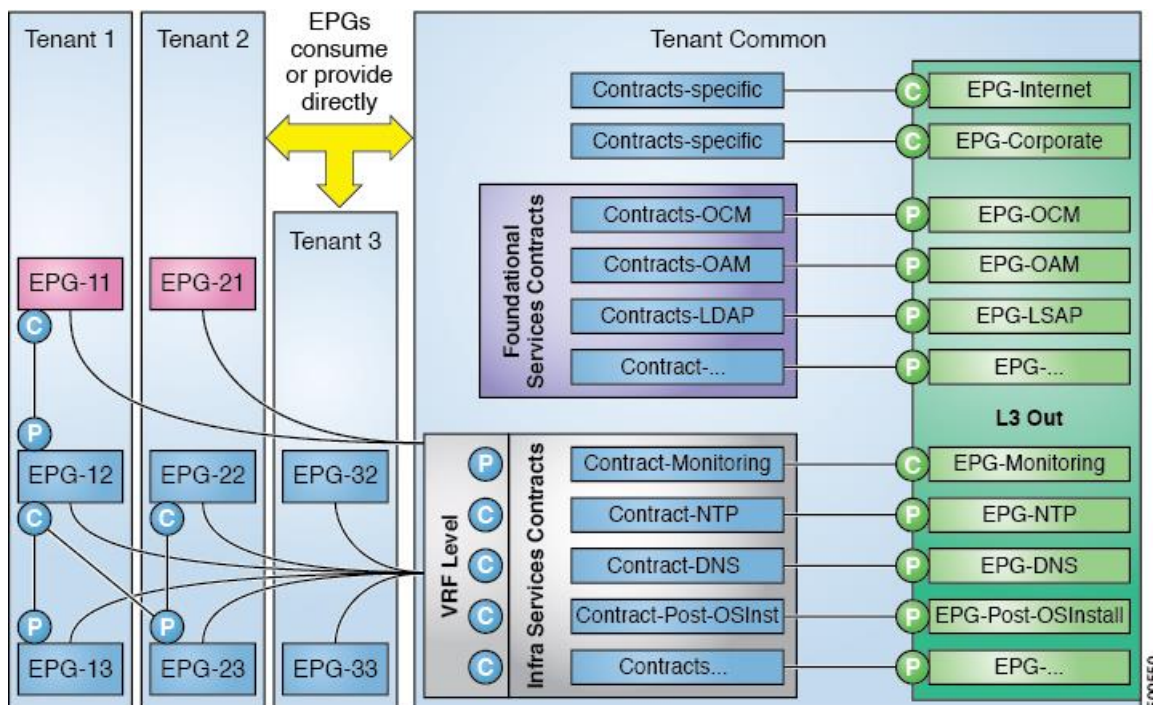
logical group that must comply with the rules of the contract. By dynamically applying contract rules to all EPGs in a context (VRF), vzAny automates the process of configuring EPG contract relationships. Whenever a new EPG is added to a context (VRF), vzAny contract rules automatically apply. The zAny one-to-all EPG relationship is the most efficient way of enabling the ACI leaf switches to apply contract rules to all EPGs in a context (VRF).

Cisco IT ACI Shared Contract Architecture

Contracts specify what is allowed in the ACI fabric function as a kind of template that can be reused whenever a new service or EPG needs access to standard data center functions. Compared with ACLs, contracts are much simpler to manage, and remove the need for most ACLs.

Prior to ACI, 80% of Cisco IT's ACL entries were set up enabling communication to shared infrastructure services and shared application middleware services. Cisco IT has opted to present these as contracts in Tenant Common within ACI, as such they will be easily consumable by any EPG within ACI.

Cisco IT Shared Services Contracts Architecture



Shared infrastructure services contracts include the following:

- DNS

-
- NTP
 - Monitoring Systems
 - Security Monitoring Systems
 - Syslog
 - Puppet

Shared middleware (foundational) services contracts include the following:

- Active Directory
- LDAP
- Authentication Systems
- Oracle Connection Manager
- Messaging middleware
- Web Security Gateway

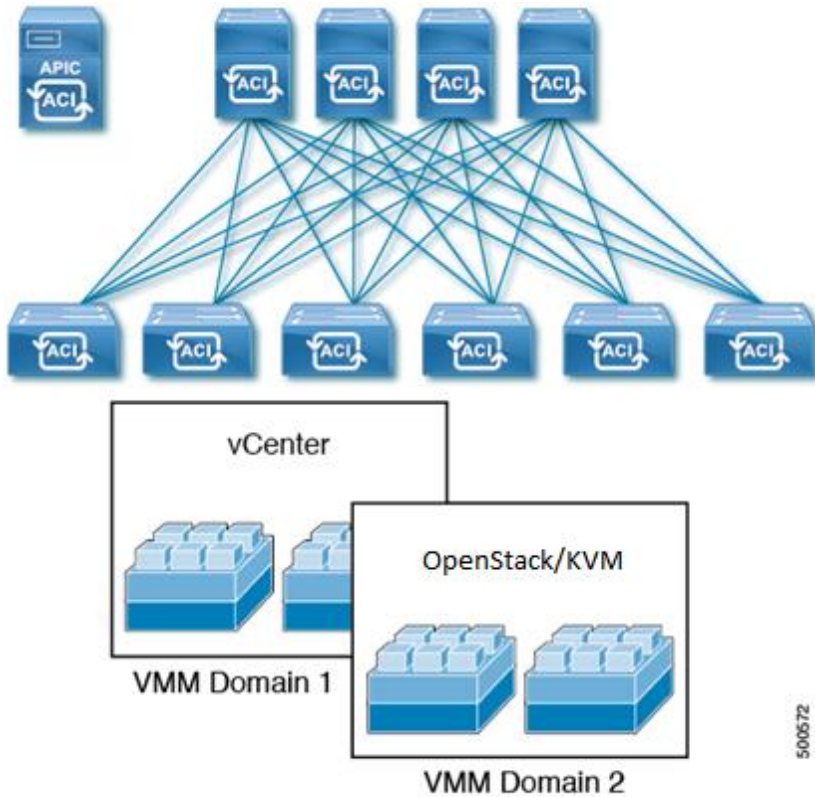
Virtual Compute Integration

Cisco ACI virtual machine networking provides hypervisors from multiple vendors programmable and automated access to high-performance scalable virtualized data center infrastructure. ACI VM networking enables consistent enforcement of policies across both virtual and physical workloads managed by hypervisors from multiple vendors. The ACI APIC controller provides centralized troubleshooting, application health score, and virtualization monitoring.

ACI fabric virtual machine manager (VMM) domains enable an administrator to configure connectivity policies for virtual machine controllers.

The VMM domain policy is created in the APIC and pushed into the leaf switches.

ACI VM Controller Integration (rough update; final image coming)



VMM domains provide the following:

- A common layer in the ACI fabric that enables scalable fault-tolerant support for multiple VM controller platforms.
- VMM support for multiple tenants within the ACI fabric.
- Automated static or dynamic VLAN allocations from specified VLAN pools.

Non-exclusive Use of Compute Pods by Multiple VM Controllers

The Cisco IT ACI solution integrates Cisco IT’s virtual compute controllers. Initially, most virtualized compute infrastructure is on VMWare. However, OpenStack/KVM is being aggressively pursued, which ACI can also integrate. Multiple VM hypervisors from different vendors can be run concurrently on the ACI fabric, regardless of which switches are associated with the ACI VMM domains, and where the compute pods are connected to the ACI fabric. A single ACI leaf can be connected to both VMware VMs, and OpenStack/KVM VMs that are all running on a UCS B compute pod.

Cisco IT leverages the next generation of Cisco’s Nexus 1000v (N1Kv) distributed virtual

switch (DVS), currently called Cisco Advanced Virtual Switch (AVS), which has been modified to integrate into the ACI model. The most significant difference between the existing N1Kv and AVS, is that the virtual supervisor function has moved into the APIC controller. This provides a single point of management, configuration and control for all VMs.

Automated Provisioning of Static or Dynamic VLANs from Specified VLAN Pools

Another significant difference between the N1Kv and AVS is that the APIC uses the new OpFlex protocol to control the VEM, both control and data channels. This is carried over the "infrastructure" VLAN – 4093. As such, only a single VLAN needs to be trunked down from the ACI leaf switch to the ESXi/KVM server.

VXLAN encapsulation is used between the ESXi/KVM server and the leaf switch to identify the different EPGs/port groups that the VMs reside on. Each EPG will have a unique VXLAN ID on the leaf to physical server link. This is locally significant only. That is the VXLAN ID for that particular EPG within the fabric will be different to that used on the downlink to the server. The leaf switch will automatically handle the mapping of one VXLAN ID into the other.

Reporting and Alerting

The ACI application policy infrastructure controller (APIC) maintains a comprehensive, current run-time representation of the administrative and operational state of the entire ACI Fabric system in the form of a collection of managed objects. The system generates faults, errors, events, and audit log data according to the run-time state of the system and the policies that the system and user create to manage these processes.

- Faults
- Events
- Errors
- Audit Logs

ACI statistics enable trend analysis and troubleshooting. Statistics gathering can be configured for ongoing or on-demand collection. Statistics provide real-time measures of

observed objects. Statistics can be collected in cumulative counters and gauges. Policies define what statistics are gathered, at what intervals, and what actions to take. For example, a policy could raise a fault on an EPG if a threshold of dropped packets on an ingress VLAN is greater than 1000 per second.

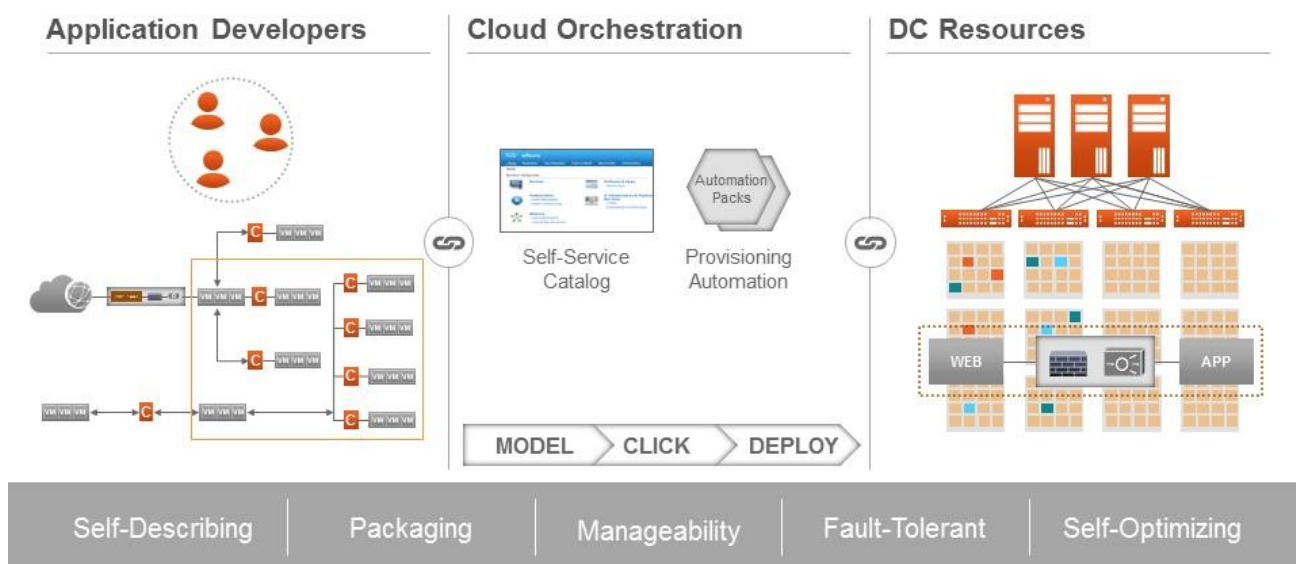
Cisco IT uses the ACI API to integrated the monitoring, alerting, and statistics capabilities of the ACI fabric with the Cisco in-house data center monitoring system.

Automation

The Cisco ACI architecture allows application requirements to define the network. This architecture simplifies, optimizes, and accelerates the entire application deployment life cycle. The separation of logical data center constructs from the physical equipment enables dynamic automated placement of any workload anywhere in the fabric.

Cisco IT Application Centric Cloud

Application-Centric Cloud: Target

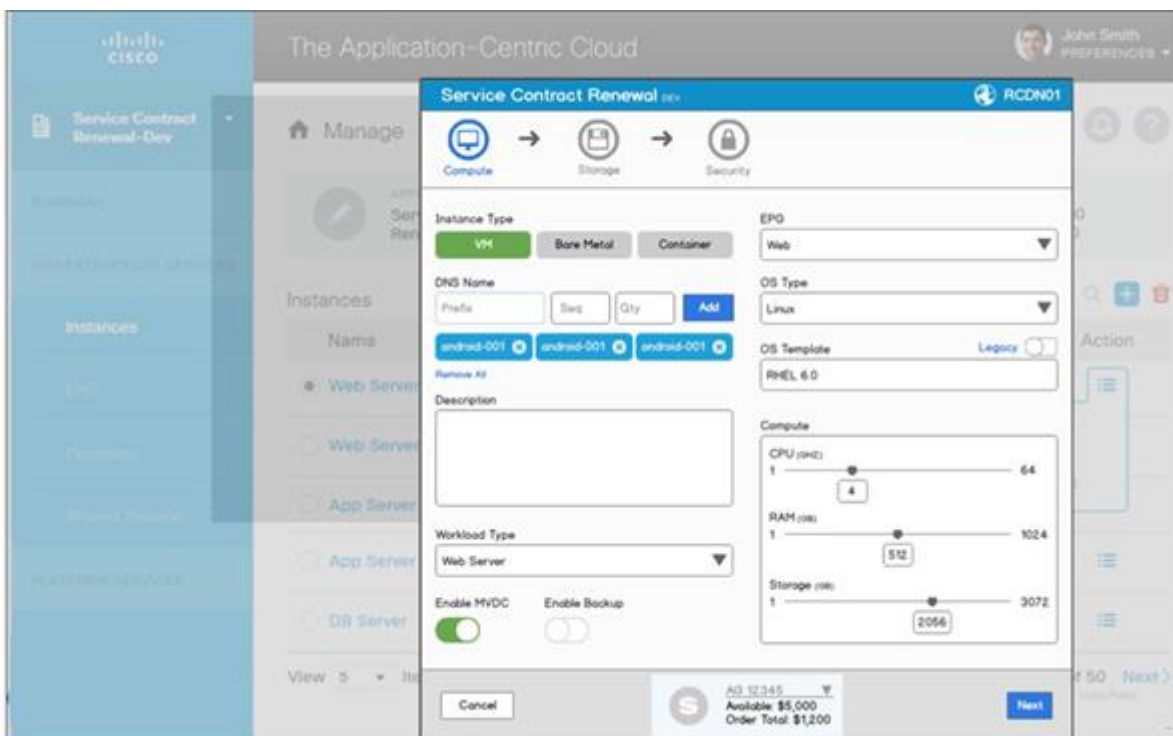


ACI enables Cisco IT operations to provide application developers direct access to standardized fabric infrastructure in a highly automated fashion through an open API while enforcing the security and governance requirements of the organization. The entire

ACI fabric is accessible through an open REST API that enables application developers to directly access the ACI fabric. The API accepts and returns HTTP or HTTPS messages that contain JavaScript Object Notation (JSON) or Extensible Markup Language (XML) documents. Any programming language can be used to generate the messages and the JSON or XML documents that contain the API methods or managed object (MO) descriptions. Various tools provided with the ACI fabric and in the ACI Toolkit enable developers to quickly develop applications that directly interact with the ACI fabric.

Cisco IT has enhanced its private cloud model by integrating ACI into its existing automation framework.

Cisco IT Cloud Automation Framework



Cisco IT's Application Centric Cloud enables infrastructure consumers to provision EPGs, endpoints (virtual machines), storage and connectivity between EPGs, and to infrastructure and middleware applications in a self-service manner.

Conclusion

The Cisco® IT deployment of ACI enables its global data center network to deliver the

enhanced business value they must have – compelling total cost of ownership, near 100% availability, enhanced network access security, and agility that includes letting business applications developers directly provision the infrastructure resources they need in a self-service fashion.