Software-Defined Access

Solution Design Guide

June 2020

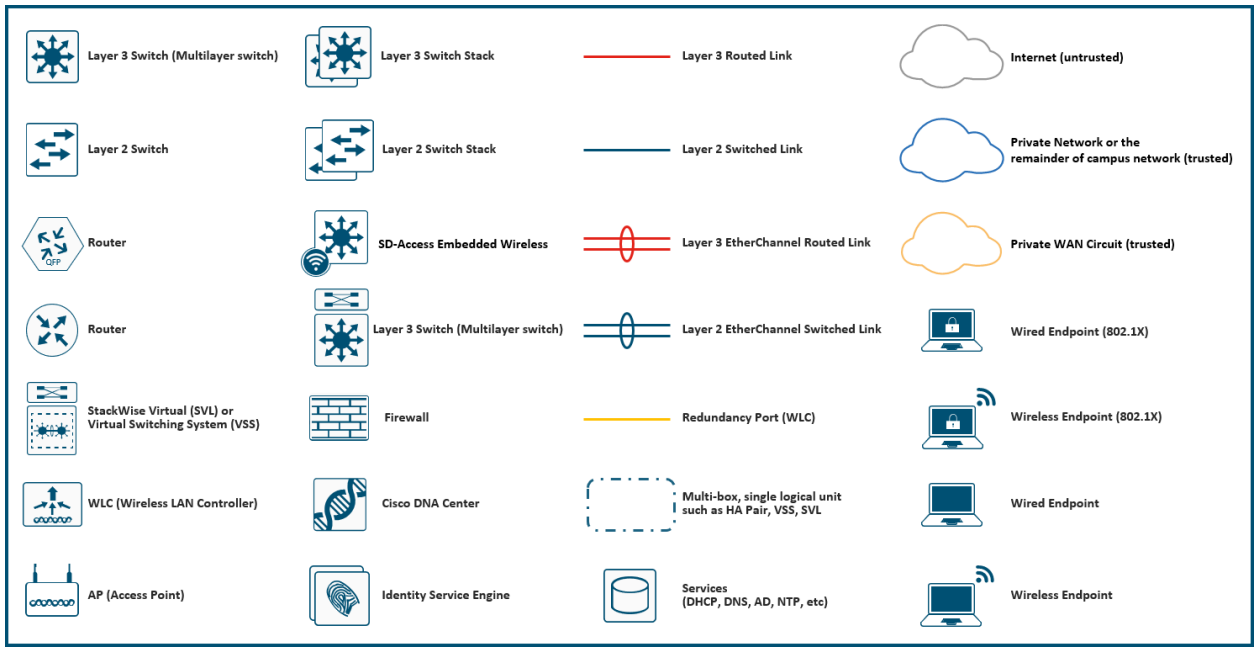
Contents

Document Organization

This document is organized into the following chapters:

Chapter	Description
Cisco Digital Network Architecture	Introduction and Campus Network Evolution
SD-Access Solution Components	Key Components of the SD-Access Solution
SD-Access Operational Planes	Control Plane, Data Plane, Policy Plane, and Management Plane Technologies
SD-Access Architecture Network Components	Fabrics, Underlay Networks, Overlay Networks, and Shared Services
SD-Access Fabric Roles and Terminology	Control Plane Node, Border Node, Edge Node, and other Fabric elements
SD-Access Design Considerations	LAN Design Principles, Layer 3 Routed Access, Role Considerations, and Feature Considerations
SD-Access Site Reference Models	Site Size Reference Models and Topologies
SD-Access Migration	Migration Support and Strategies
Appendices	Additional References and Resources

Icons Used in this Document



Cisco Digital Network Architecture and Software-Defined Access

Cisco® Software-Defined Access (SD-Access) is the evolution from traditional campus designs to networks that directly implement the intent of an organization. SD-Access is software application running on Cisco DNA Center hardware that is used to automate wired and wireless campus networks.

Fabric technology, an integral part of SD-Access, provides wired and wireless campus networks with programmable overlays and easy-to-deploy network virtualization, permitting a physical network to host one or more logical networks to meet the design intent. In addition to network virtualization, fabric technology in the campus network enhances control of communications, providing software-defined segmentation and policy enforcement based on user identity and group membership. Software-defined segmentation is seamlessly integrated using Cisco TrustSec® technology, providing micro-segmentation for groups within a virtual network using scalable group tags (SGTs). Using Cisco DNA Center to automate the creation of virtual networks with integrated security and segmentation reduces operational expenses and reduces risk. Network performance, network insights, and telemetry are provided through the Assurance and Analytics capabilities.

This design guide provides an overview of the requirements driving the evolution of campus network designs, followed by a discussion about the latest technologies and designs that are available for building a SD-Access network to address those requirements. It is a companion to the associated deployment guides for SD-Access, which provide configurations explaining how to deploy the most common implementations of the designs described in this guide. The intended audience is a technical decision maker who wants to understand Cisco's campus offerings, learn about the available technology options, and use leading practices for designing the best network for the needs of an organization.

Companion Resources

Find the companion guides Cisco DNA Center & ISE Management Infrastructure Deployment Guide, SD-Access Fabric Provisioning Prescriptive Deployment Guide, SD-Access for Distributed Campus Prescriptive Deployment Guide, related deployment guides, design guides, and white papers, at the following pages:

- https://www.cisco.com/go/designzone
- https://cs.co/en-cvds

If you didn't download this guide from Cisco Community or Design Zone, you can check for the latest version of this guide.

Scale Metrics and Latency Information

For current scale metrics and latency information, please see the SD-Access Resources and Latency Design Guidance on Cisco.com Technology & Support Community.

Evolution of Campus Network Designs for Digital-Ready Organizations

With digitization, software applications are evolving from simply supporting business processes to becoming, in some cases, the primary source of business revenue and competitive differentiation. Organizations are now constantly challenged by the need to scale their network capacity to react quickly to application demands and growth. Because the campus network is used by people with different levels of access and their BYOD devices to access these applications, the wired and wireless LAN capabilities should be enhanced to support those changing needs.

Network Requirements for the Digital Organization

The following are the key requirements driving the evolution of existing campus networks.

Flexible Ethernet Foundation for Growth and Scale

- **Simplified deployment and automation**—Network device configuration and management through a centralized controller using open APIs allows for very fast, lower-risk deployment of network devices and services.

- **Increased bandwidth needs**—Bandwidth needs are doubling potentially multiple times over the lifetime of a network, resulting in the need for new networks to aggregate using 10 Gbps Ethernet to 40 Gbps to 100 Gbps capacities over time.
- **Increased capacity of wireless access points**—The bandwidth demands on wireless access points (APs) with the latest 802.11ac Wave 2 and 802.11ax (Wi-Fi 6) technology now exceed 1 Gbps, and the IEEE has now ratified the 802.3bz standard that defines 2.5 Gbps and 5 Gbps Ethernet.
- **Additional power requirements from Ethernet devices**—New devices, such as lighting, surveillance cameras, virtual desktop terminals, remote access switches, and APs, may require higher power to operate. The access layer design should have the ability to support Power over Ethernet (PoE) with 60W per port, offered with Cisco Universal Power Over Ethernet (UPOE), and the access layer should also provide PoE perpetual power during switch upgrade and reboot events. As power demands continue to increase with new endpoints, IEEE 802.3bt and Cisco UPOE-Plus (UPOE+) can provide power up to 90W per port.

Integrated Services and Security

- **Consistent wired and wireless security capabilities**—Security capabilities, described below, should be consistent whether a user is connecting to a wired Ethernet port or connecting over the wireless LAN.
- **Network assurance and analytics**—The deployment should proactively predict network-related and security-related risks by using telemetry to improve the performance of the network, devices, and applications, even with encrypted traffic.
- **Identity services**—Identifying users and devices connecting to the network provides the contextual information required to implement security policies for access control, network segmentation by using scalable group membership, and mapping of devices into virtual networks.
- **Network virtualization**—The capability to share a common infrastructure while supporting multiple VNs with isolated data and control planes enables different sets of users and applications to be isolated securely.
- **Group-based policies**—Creating access and application policies based on user group information provides a much easier and scalable way to deploy and manage security policies. Traditional access control lists (ACLs) can be difficult to implement, manage, and scale because they rely on network constructs such as IP addresses and subnets rather than group membership. Group membership is an IP-agnostic approach to policy creation which provides ease of operation for the network operator and a more scalable approach to ACLs.
- **Software-defined segmentation**—Scalable group tags assigned from group-based policies can be used to segment a network to achieve data plane isolation within physical and virtual networks.

SD-Access Use Case for Healthcare Networks: Macro-Segmentation

Our healthcare records are just as valuable to attackers as our credit card numbers and online passwords. Hospitals are required to have HIPAA-compliant wired and wireless networks that can provide complete and constant visibility into their network traffic to protect sensitive medical devices (such as servers for electronic medical records, vital signs monitors, or nurse workstations) so that a malicious device cannot compromise the networks.

A patient's mobile device, when compromised by malware, can change network communication behavior to propagate and infect other endpoints. It is considered abnormal behavior when a patient's mobile device communicates with any medical device. SD-Access can address the need for complete isolation between

patient devices and medical facility devices by using macro-segmentation and putting devices into different overlay networks, enabling the isolation.

SD-Access Use Case for University Networks: Micro-Segmentation

In a University example, students and faculty machines may both be permitted to access printing resources, but student machines should not communicate directly with faculty machines, and printing devices should not communicate with other printing devices.

SD-Access can address the need for isolation of devices in the same virtual network through micro-segmentation. By using Scalable Group Tags (SGTs), users can be permitted access to printing resources, though the printing resources cannot directly communicate with each other.

SD-Access Use Case for Enterprise Networks: Macro- and Micro-Segmentation

In the Enterprise, users, devices, and applications all utilize the network to access resources. Building control systems such as badge readers and physical security systems such as video surveillance devices need access to the network in order to operate, though these devices are segmented into different overlay networks than where the users resides. Guest network access is common for visitors to the enterprise and for employee BYOD use. However, the Guest network can remain completely isolated from the remainder of the corporate network and the building management network using different overlay networks.

Users and devices on the corporate overlay network have different access needs. These users and devices may need access to printing and internal web servers such as corporate directory. However, not all will need access to development servers, employee and payroll data from human resources, and other department-specific resources. Using SGTs, users and device within the overlay network can be permitted access to specific resources and denied access to others based on their group membership.

Deploying these intended outcomes for the needs of the organization is simplified by using the automation capabilities built into Cisco DNA Center, and those simplifications span both the wired and wireless domains.

Other organizations may have business requirements where secure segmentation and profiling are needed:

- **Education**—College campus divided into administrative and student residence networks.
- **Retail**—Isolation for point-of-sale machines supporting payment card industry compliance (PCI DSS).
- **Manufacturing**—Isolation for machine-to-machine traffic in manufacturing floors.

SD-Access Solution Components

This chapter is organized into the following sections:

Chapter	Section
SD-Access Solution Components	Cisco DNA Center Hardware Appliance Cisco DNA Center Software Identity Services Engine

The SD-Access solution is provided through a combination of Cisco DNA Center, the Identity Services Engine (ISE), and wired and wireless device platforms which have fabric functionality. As described later in the Fabric Roles section, the wired and wireless device platforms are utilized to create the elements of a fabric site. This section describes the functionality of the remaining two components for SD-Access: Cisco DNA Center and the Identity Services Engine.

Cisco DNA Center Hardware Appliance

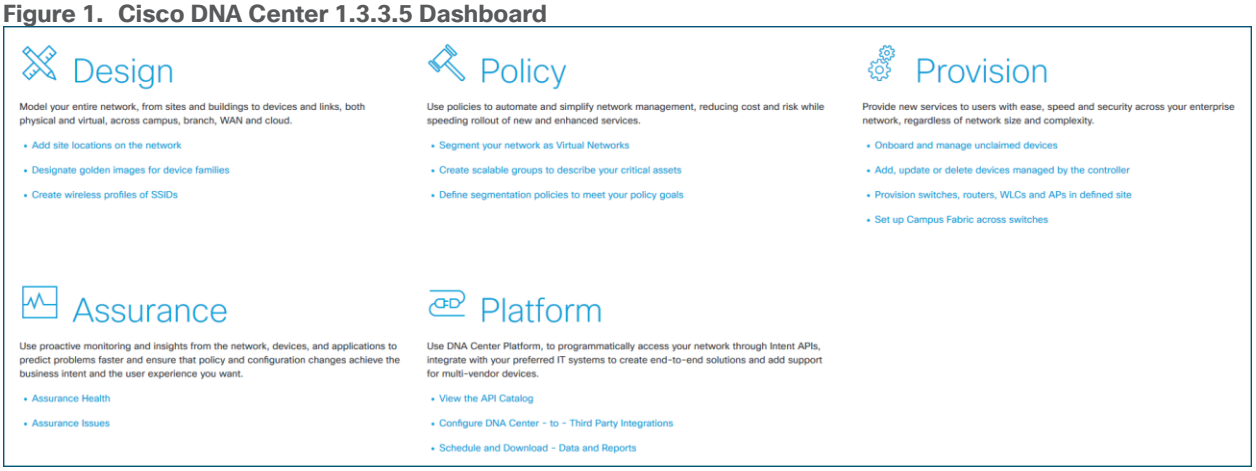
Cisco DNA Center software, including the SD-Access application package, run on Cisco DNA Center hardware appliance. The appliance is available in form factors sized to support not only the SD-Access application but also network Assurance and Analytics, Software image management (SWIM), Wide-Area Bonjour, and new capabilities as they are available.

Tech tip
For additional information about the Cisco DNA Center Appliance capabilities, see the data sheet on Cisco.com.

Cisco DNA Center Software

Cisco DNA Center is the centralized manager running a collection of application and services powering the Cisco Digital Network Architecture (Cisco DNA). Cisco DNA begins with the foundation of a digital-ready infrastructure that includes routers, switches, access-points, and Wireless LAN controllers. Automation, Analytics, Visibility, and management of the Cisco DNA network is enabled through Cisco DNA Center Software. SD-Access is part of this software and is used to design, provision, apply policy, and facilitate the creation of an intelligent wired and wireless campus network with assurance. In addition to automation for SD-Access, Cisco DNA Center provides applications to improve an organization's efficiency such as network device health dashboards.

Figure 1. Cisco DNA Center 1.3.3.5 Dashboard



Cisco DNA Center centrally manages major configuration and operations workflow areas.

- **Design**—Configures device global settings, network site profiles for physical device inventory, DNS, DHCP, IP addressing, SWIM repository, device templates, and telemetry configurations such as Syslog, SNMP, and NetFlow.
- **Policy**—Defines business intent including creation of virtual networks, assignment of endpoints to virtual networks, policy contract definitions for groups, and configures application policies (QoS).
- **Provision**—Provisions devices and adds them to inventory for management, supports Cisco Plug and Play, creates fabric sites along with other SD-Access components, and provides service catalogs such as Stealthwatch Security Analytics and Application Hosting on the Cisco Catalyst 9000 Series Switches.
- **Assurance**—Enables proactive monitoring and insights to confirm user experience meets configured intent, using network, client, and application health dashboards, issue management, sensor-driven testing, and Cisco AI Network Analytics.

- **Platform**—Allows programmatic access to the network and system integration with third-party systems via APIs by using feature set bundles, configurations, a runtime dashboard, and a developer toolkit.

Identity Services Engine

Cisco Identity Services Engine (ISE) is a secure network access platform enabling increased management awareness, control, and consistency for users and devices accessing an organization's network. ISE is an integral and mandatory component of SD-Access for implementing network access control policy. ISE performs policy implementation, enabling dynamic mapping of users and devices to scalable groups, and simplifying end-to-end security policy enforcement. Within ISE, users and devices are shown in a simple and flexible interface. ISE integrates with Cisco DNA Center by using Cisco Platform Exchange Grid (pxGrid) and REST APIs (Representational State Transfer Application Programming Interfaces) for endpoint event notifications and automation of policy configurations on ISE.

The SD-Access solution integrates Cisco TrustSec by supporting end-to-end group-based policy with Scalable Group Tags (SGTs). Scalable Group Tags are a metadata value that is transmitted in the header of fabric-encapsulated packets. While SGTs are administered by Cisco ISE through the tightly integrated REST APIs, Cisco DNA Center is used as the pane of glass to manage and create SGTs and define their policies. Group and policy services are driven by ISE and orchestrated by Cisco DNA Center's policy authoring workflows. Policy management with identity services is enabled in an SD-Access network using ISE integrated with Cisco DNA Center for dynamic mapping of users and devices to scalable groups. This simplifies end-to-end security policy management and enforcement at a greater scale than traditional network policy implementations relying on IP access-lists.

ISE Personas

A Cisco ISE node can provide various services based on the *persona* that it assumes. Personas are simply the services and specific feature set provided by a given ISE node. The four primary personas are PAN, MnT, PSN, and pxGrid.

- **Policy Administration Node (PAN)**— A Cisco ISE node with the Administration persona allows performs all administrative operations on Cisco ISE. It handles all system-related configurations that are related to functionality such as authentication, authorization, and auditing.
- **Monitor and Troubleshooting Node (MnT)**— A Cisco ISE node with the Monitoring persona functions as the log collector and stores log messages from all the administration and Policy Service nodes in the network. This persona provides advanced monitoring and troubleshooting tools that used to effectively manage the network and resources. A node with this persona aggregates and correlates the data that it collects to provide meaningful information in the form of reports.
- **Policy Service Node (PSN)**— A Cisco ISE node with the Policy Service persona provides network access, posture, guest access, client provisioning, and profiling services. This persona evaluates the policies and makes all the decisions. Typically, there would be more than one PSN in a distributed deployment. All Policy Service nodes that reside in the same high-speed Local Area Network (LAN) or behind a load balancer can be grouped together to form a node group.
- **Platform Exchange Grid (pxGrid)**—A Cisco ISE node with pxGrid persona shares the context-sensitive information from Cisco ISE session directory with other network systems such as ISE ecosystem partner systems and Cisco platforms. The pxGrid framework can also be used to exchange policy and configuration data between nodes like sharing tags and policy objects. TrustSec information like tag definition, value, and description can be passed from Cisco ISE to other Cisco management platforms such as Cisco DNA Center and Cisco Stealthwatch.

ISE supports standalone and distributed deployment models. Multiple, distributed nodes can be deployed together to provide failover resiliency and scale. The range of deployment options allows support for hundreds of thousands of endpoint devices. Minimally, a basic two-node ISE deployment is recommended for SD-Access single site deployments with each ISE node running all services (personas) for redundancy.

SD-Access fabric nodes send authentication requests to the Policy Services Node (PSN) service persona running in ISE. In the case of a standalone deployment, the PSN persona is referenced by a single IP address. An ISE distributed model uses multiple, active PSN personas, each with a unique address. All PSN addresses are learned by Cisco DNA Center, and the Cisco DNA Center user associates the fabric sites to the applicable PSN.

Tech tip
For additional details on ISE personas and services, please see Cisco Identity Services Engine Administrator Guide, Chapter: Set Up Cisco ISE in a Distributed Environment. For additional ISE deployment and scale details, please see ISE Performance & Scale on Cisco.com Security Community.

SD-Access Operational Planes

This chapter is organized into the following sections:

Chapter	Section
SD-Access Operational Planes	Control Plane – LISP Data Plane – VXLAN Policy Plane – Cisco TrustSec Management Plane – Cisco DNA Center

There are four key technologies, that make up the SD-Access solution, each performing distinct activities in different network planes of operation: control plane, data plane, policy plane, and management plane.

- **Control Plane**—Messaging and communication protocol between infrastructure devices in the fabric.
- **Data Plane**—Encapsulation method used for the data packets.
- **Policy Plane**—Used for security and segmentation.
- **Management Plane**—Orchestration, assurance, visibility, and management.

In SD-Access the control plane is based on LISP (Locator/ID Separation Protocol), the data plane is based on VXLAN (Virtual Extensible LAN), the policy plane is based on Cisco TrustSec, and the management plane is enabled and powered by Cisco DNA Center.

Control Plane – LISP

In many networks, the IP address associated with an endpoint defines both its identity and its location in the network. In these networks, the IP address is used for both network layer identification (who the device is on the network) and as a network layer locator (where the device is at in the network or to which device it is connected). This is commonly referred to as *addressing following topology*. While an endpoint's location in the network will change, who this device is and what it can access should not have to change. The Locator/ID Separation Protocol (LISP) allows the separation of identity and location though a mapping relationship of these two *namespaces*: an endpoint's identity (EID) in relationship to its routing locator (RLOC).

The LISP control plane messaging protocol is an architecture to communicate and exchange the relationship between these two *namespaces*. This relationship is called an *EID-to-RLOC mapping*. This EID and RLOC combination provide all the necessary information for traffic forwarding, even if an endpoint uses an unchanged IP address when appearing in a different network location (associated or mapped behind different RLOCs).

Simultaneously, the decoupling of the endpoint identity from its location allows addresses in the same IP subnetwork to be available behind multiple Layer 3 gateways in disparate network locations (such as multiple wiring closets), versus the one-to-one coupling of IP subnetwork with network gateway in traditional networks. This provides the benefits of a Layer 3 Routed Access network, described in a later section, without the requirement of a subnetwork to only exist in a single wiring closet.

Instead of a typical traditional routing-based decision, the fabric devices query the control plane node to determine the routing locator associated with the destination address (EID-to-RLOC mapping) and use that RLOC information as the traffic destination. In case of a failure to resolve the destination routing locator, the traffic is sent to the default fabric border node. The response received from the control plane node is stored in the LISP map-cache, which is merged to the Cisco Express Forwarding (CEF) table and installed in hardware.

Data Plane – VXLAN

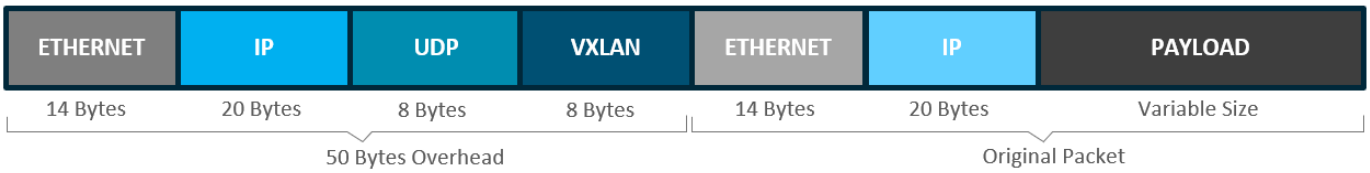
VXLAN is an encapsulation technique for data packets. When encapsulation is added to these data packets, a *tunnel* network is created. Tunneling encapsulates data packets from one protocol inside a different protocol and transports the original data packets, unchanged, across the network. A lower-layer or same-layer protocol (from the OSI model) can be carried through this tunnel creating an *overlay*. In SD-Access, this overlay network is referred to as the *fabric*.

VXLAN is a MAC-in-IP encapsulation method. It provides a way to carry lower-layer data across the higher Layer 3 infrastructure. Unlike routing protocol tunneling methods, VXLAN preserves the original Ethernet header from the original frame sent from the endpoint. This allows for the creation of an overlay at Layer 2 and at Layer 3 depending on the needs of the original communication. For example, Wireless LAN communication (IEEE 802.11) uses Layer 2 datagram information (MAC Addresses) to make bridging decisions without a direct need for Layer 3 forwarding logic.

SD-Access also places additional information in the fabric VXLAN header including alternative forwarding attributes that can be used to make policy decisions by identifying each overlay network using a VXLAN network identifier (VNI). Layer 2 overlays are identified with a VLAN to VNI correlation (L2 VNI), and Layer 3 overlays are identified with a VRF to VNI correlation (L3 VNI).

Any encapsulation method is going to create additional MTU (maximum transmission unit) overhead on the original packet. As show in Figure 2, VXLAN encapsulation uses a UDP transport. Along with the VXLAN and UDP headers used to encapsulate the original packet, an outer IP and Ethernet header are necessary to forward the packet across the wire. At minimum, these extra headers add 50 bytes of overhead to the original packet.

Figure 2. Fabric VXLAN (VNI) Encapsulation Overhead



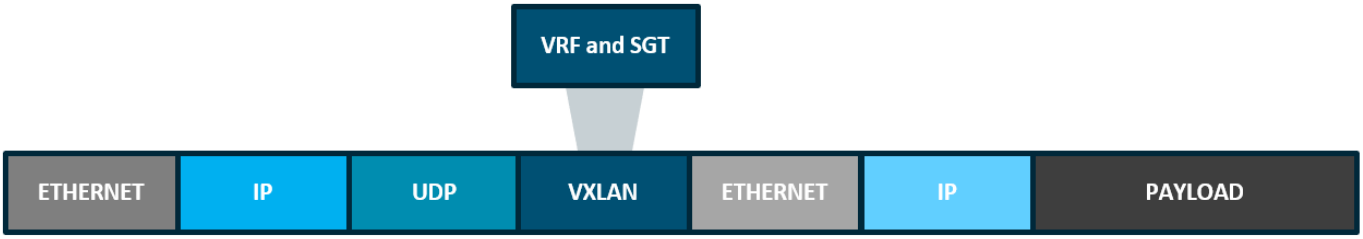
Policy Plane – Cisco TrustSec

Cisco TrustSec decouples access that is based strictly on IP addresses and VLANs by using logical groupings in a method known as Group-Based Access Control (GBAC). The goal of Cisco TrustSec technology is to assign an SGT value to the packet at its ingress point into the network. An access policy elsewhere in the network is then enforced based on this tag information.

An SGT is a form of metadata and is a 16-bit value assigned by ISE in an authorization policy when user, device, or application connects to the network.

The fabric VXLAN encapsulation method is actually used by both the data plane and policy plane. In the policy plane, the alternative forwarding attributes (the SGT value and VRF values) are encoded into the header, and carried across the overlay.

Figure 3. Fabric VXLAN Alternative Forwarding Attributes



Tech tip
A bit-level diagram of the VXLAN encapsulation method used in SD-Access fabric along with low-level details on policy constructs insertion into the header can be found in Appendix A.

Management Plane – Cisco DNA Center

Cisco DNA Center is a foundational component of SD-Access, enabling automation of device deployments and configurations into the network to provide the speed and consistency required for operational efficiency. Through its automation capabilities, the control plane, data plane, and policy plane for the fabric devices is easily, seamlessly, and consistently deployed. Through Assurance, visibility and context are achieved for both the infrastructure devices and endpoints.

A full understanding of LISP and VXLAN is not required to deploy the fabric in SD-Access, nor is there a requirement to know the details of how to configure each individual network component and feature to create the consistent end-to-end behavior offered by SD-Access. Cisco DNA Center is an intuitive, centralized management system used to design, provision, and apply policy across the wired and wireless SD-Access network. It takes the user's intent and programmatically applies it to network devices.

SD-Access Architecture Network Components

This chapter is organized into the following sections:

Chapter	Section
SD-Access Architecture Network Components	What is a Fabric? Underlay Network Overlay Network

Chapter	Section
	Shared Services

The SD-Access architecture is supported by fabric technology implemented for the campus, enabling the use of virtual networks (*overlay networks*) running on a physical network (*underlay network*) creating alternative topologies to connect devices. This section describes and defines the word *fabric*, discusses the SD-Access fabric underlay and overlay network, and introduces shared services which are a shared set of resources accessed by devices in the overlay. This section provides an introduction for these *fabric*-based network terminologies used throughout the rest of the guide. Design consideration for these are covered in a later section.

What is a Fabric?

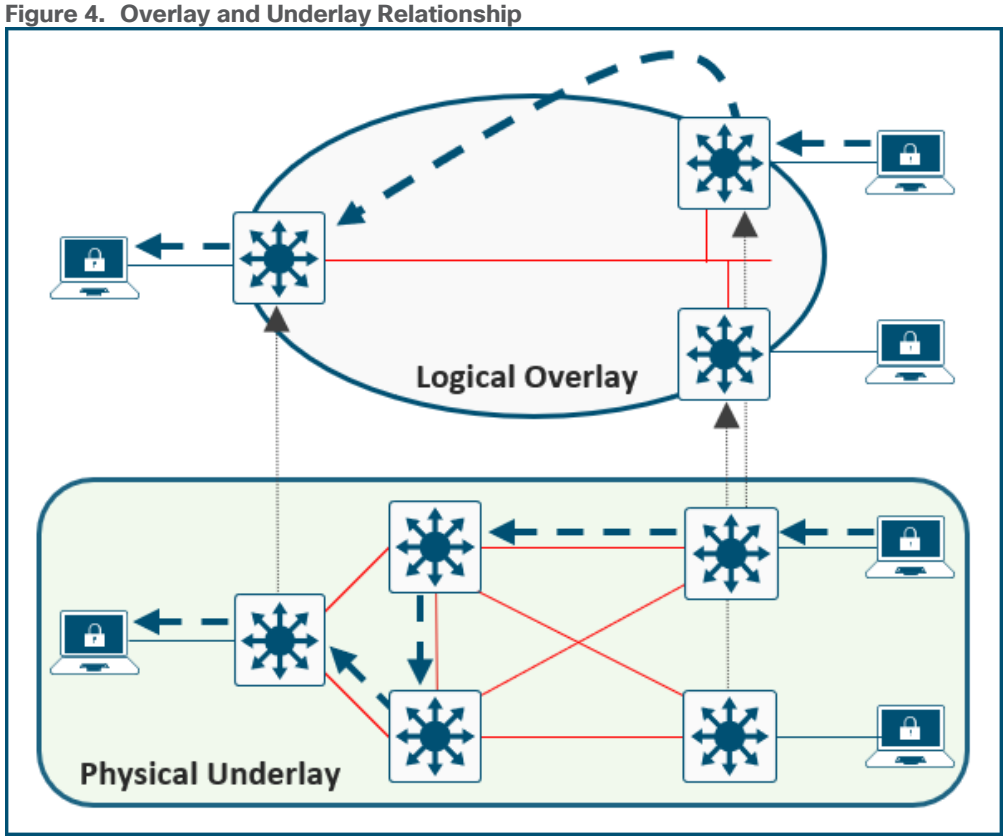
A fabric is simply an overlay network. Overlays are created through encapsulation, a process which adds additional header(s) to the original packet or frame. An overlay network creates a logical topology used to virtually connect devices that are built over an arbitrary physical *underlay* topology. In an idealized, theoretical network, every device would be connected to every other device. In this way, any connectivity or topology imagined could be created. While this theoretical network does not exist, there is still a technical desire to have all these devices connected to each other in a full mesh. This is where the term *fabric* comes from: it is a cloth where everything is connected together. In networking, an overlay (or tunnel) provides this logical full-mesh connection.

Underlay Network

The *underlay* network is defined by the physical switches and routers that are used to deploy the SD-Access network. All network elements of the underlay must establish IP connectivity via the use of a routing protocol. Instead of using arbitrary network topologies and protocols, the underlay implementation for SD-Access uses a well-designed Layer 3 foundation inclusive of the campus edge switches which is known as a Layer 3 Routed Access design. This ensures performance, scalability, and resiliency, and deterministic convergence of the network.

In SD-Access, the underlay switches (edge nodes) support the physical connectivity for users and endpoints. However, end-user subnets and endpoints are not part of the underlay network—they are part of the automated overlay network.

Figure 4. Overlay and Underlay Relationship

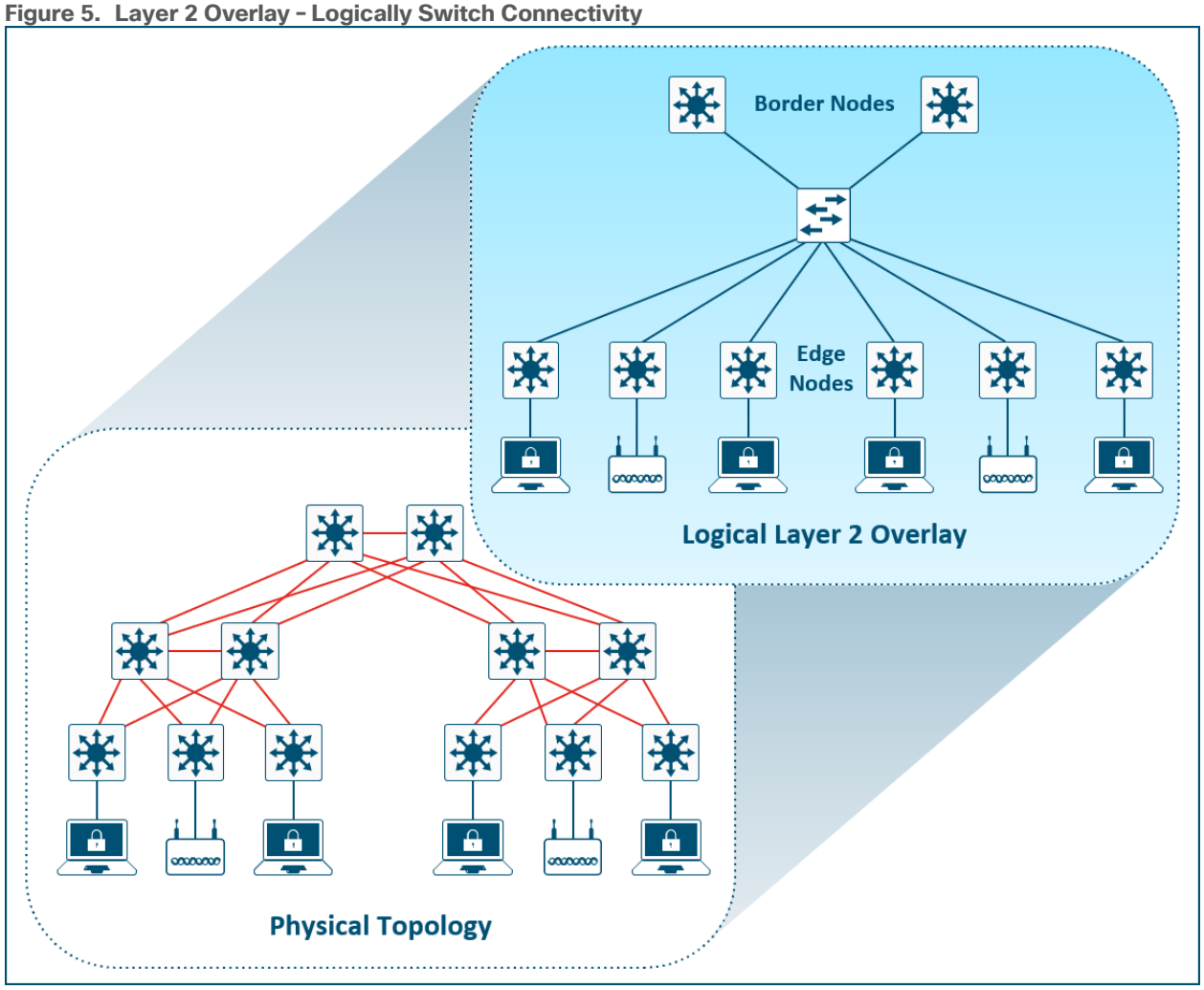


Overlay Network

An *overlay* network is created on top of the underlay network through virtualization (virtual networks). The data plane traffic and control plane signaling are contained within each virtualized network, maintaining isolation among the networks and an independence from the underlay network. Multiple overlay networks can run across the same underlay network through virtualization. In SD-Access, the user-defined overlay networks are provisioned as a virtual routing and forwarding (VRF) instances that provide separation of routing tables.

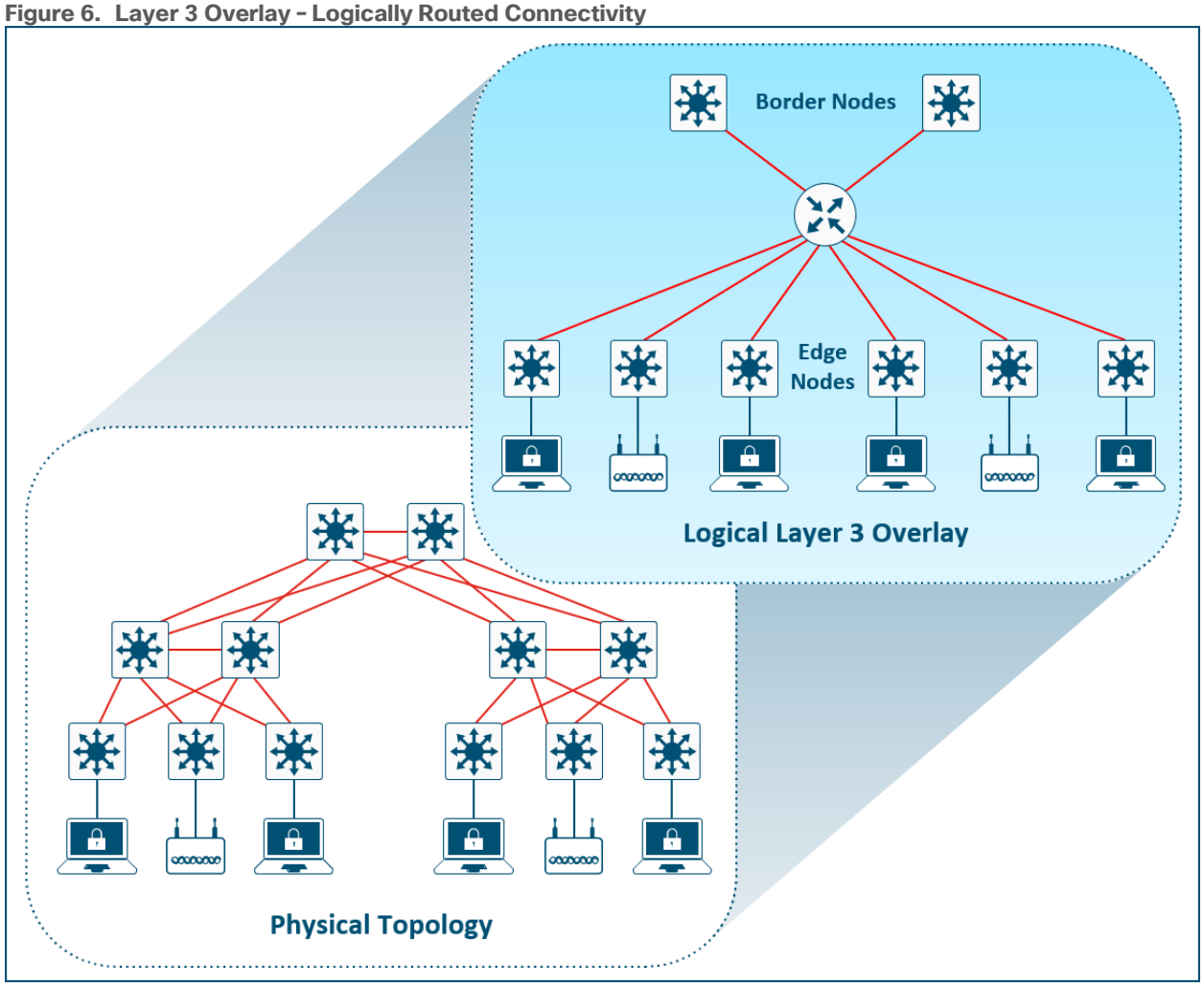
SD-Access allows for the extension of Layer 2 and Layer 3 connectivity across the overlay through the services provided by through LISP. Layer 2 overlay services emulate a LAN segment to transport Layer 2 frames by carrying a subnet over the Layer 3 underlay as shown in Figure 5.

Figure 5. Layer 2 Overlay – Logically Switch Connectivity



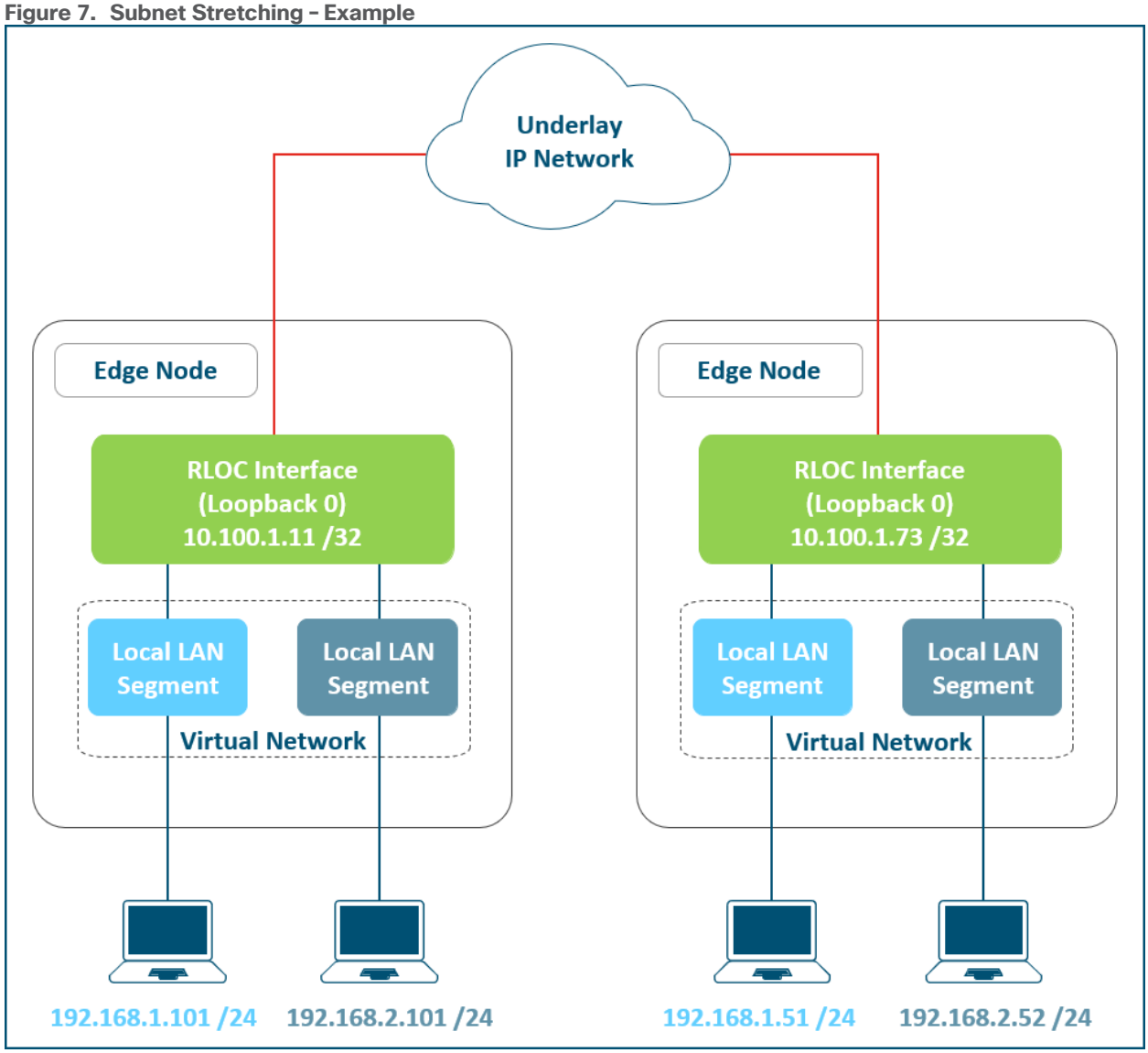
Layer 3 overlays abstract the IP-based connectivity from the physical connectivity as shown in Figure 6. This can allow multiple IP networks to be part of each virtual network. Each Layer 3 overlay, its routing tables, and its associated control planes are completely isolated from each other.

Figure 6. Layer 3 Overlay – Logically Routed Connectivity



The following diagram shows an example of two subnets that are part of the overlay network. The subnets stretch across physically separated Layer 3 devices–two edge nodes. The RLOC interfaces, or Loopback 0 interfaces in SD-Access, are the only underlay routable address that are required to establish connectivity between endpoints of the same or different subnet within the same VN.

Figure 7. Subnet Stretching – Example



Shared Services

Networks need some form of shared services that can be reused across multiple virtual networks. It is important that those shared services are deployed correctly to preserve the isolation between different virtual networks accessing those services. The use of a VRF-Aware Peer directly attached outside of the fabric provides a mechanism for route leaking of shared services prefixes across multiple networks, and the use of firewalls provides an additional layer of security and monitoring of traffic between virtual networks. Examples of shared services include:

- **Wireless infrastructure**—Radio frequency performance and cost efficiency is improved using common wireless LANs (single SSID) versus previous inefficient strategies of using multiple SSIDs to separate endpoint communication. Traffic isolation is achieved by assigning dedicated VLANs and using dynamic VLAN assignment using 802.1X authentication to map wireless endpoints into their corresponding VNs.
- **DHCP, DNS, IP address management (IPAM), and Active Directory (AD)**—The same set of infrastructure services can be reused if they have support for virtualized networks. Special capabilities such as advanced DHCP scope selection criteria, multiple domains, and support for overlapping address space are some of the capabilities required to extend the services beyond a single network.

- **Internet access**—The same set of Internet firewalls can be used for multiple virtual networks. If firewall policies need to be unique for each virtual network, the use of a multi-context firewall is recommended.
- **IP voice/video collaboration services**—When IP phones and other unified communications devices are connected in multiple virtual networks, the call control signaling to the communications manager and the IP traffic between those devices needs to be able to traverse multiple VNs in the infrastructure.
- **Servers and Critical Systems**—NTP servers, Building Management Systems (BMS), network orchestrators, management appliances, support systems, administrative applications, databases, payroll systems, and other critical applications may be required for access by one or many virtual networks.

SD-Access Fabric Roles and Terminology

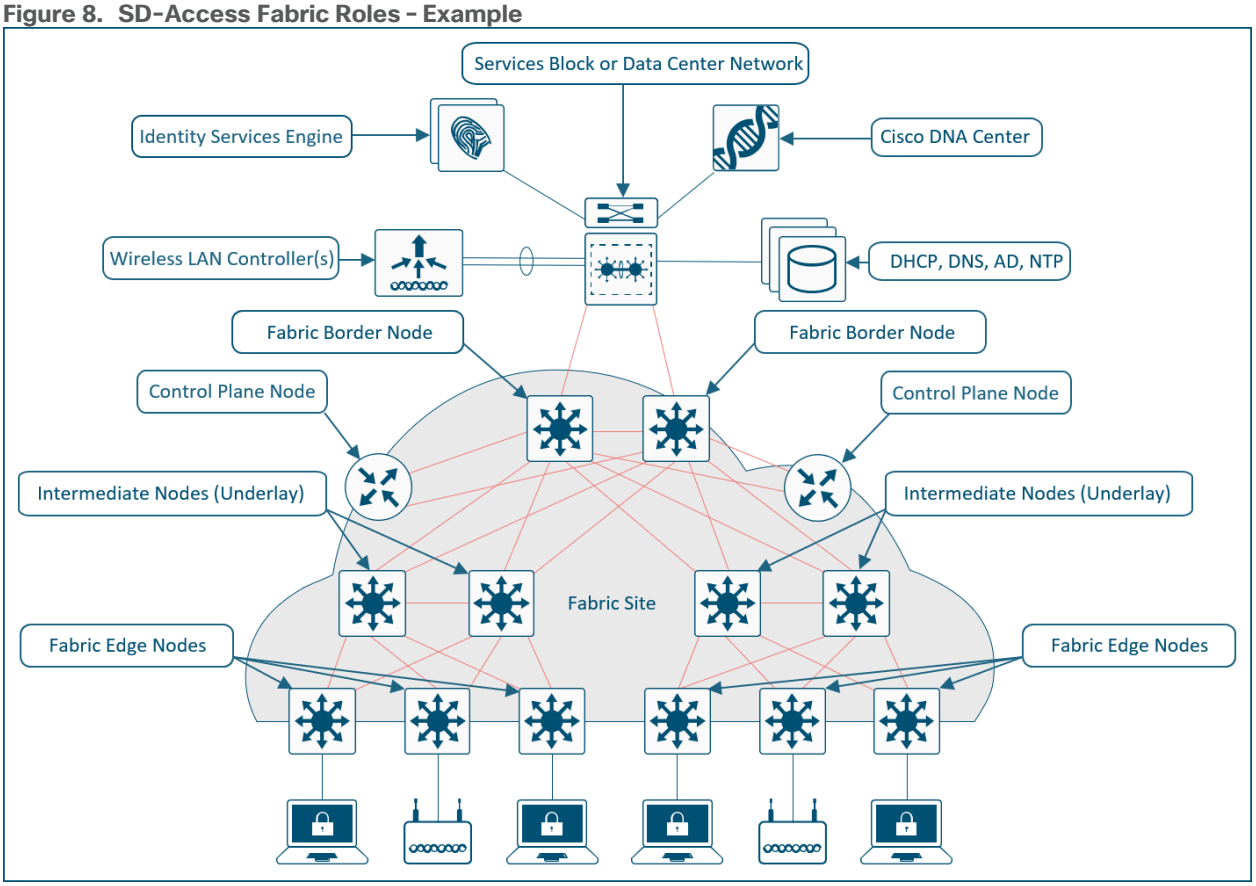
This chapter is organized into the following sections:

Chapter	Section
SD-Access Fabric Roles and Terminology	Control Plane Node Edge Node Intermediate Node Border Node Fabric in a Box Extended Node Fabric WLC Fabric-Mode Access Point SD-Access Embedded Wireless Transit and Peer Networks Transit Control Plane Node Fabric Domain Fabric Site

The SD-Access solution is provided through a combination of Cisco DNA Center, the Identity Services Engine (ISE), and wired and wireless device platforms which have fabric functionality. The wired and wireless device platforms are utilized to create the elements of a fabric site. A fabric site is defined as location that has its own control plane node and an edge node. For wireless, a fabric-mode WLC is dedicated to the site, and for policy, an ISE Policy Service Node (PSN) is used. A fabric border node is required to allow traffic to egress and ingress the fabric site.

A fabric role is an SD-Access software construct running on physical hardware. These software constructs were designed with modularity and flexibility in mind. For example, a device can run a single role, or a device can also run multiple roles. Care should be taken to provision the SD-Access fabric roles in the same way the underlying network architecture is built: *distribution of function*. Separating roles onto different devices provides the highest degree of availability, resilience, deterministic convergence, and scale.

Figure 8. SD-Access Fabric Roles – Example



Control Plane Node

The SD-Access fabric control plane node is based on the LISP Map-Server and Map-Resolver functionality combined on the same node. The control plane node's database tracks all endpoints in the fabric site and associates the endpoints to fabric nodes, decoupling the endpoint IP address or MAC address from the location (closest router) in the network.

The control plane node enables the following functions:

- **Host tracking database**—The host tracking database (HTDB) is a central repository of Endpoint ID to Routing Locator (EID-to-RLOC) bindings where the RLOC is simply the IP address of the Loopback 0 interface on a fabric node. The HTDB is equivalent to a *LISP site,* in traditional LISP, which includes what endpoint ID can be and have been registered.
- **Endpoint identifiers (EID)**—The endpoint identifier is an address used for numbering or identifying an endpoint device in the network. The SD-Access solution supports MAC Address, IPv4 Address, and IPv6 addresses as EIDs.
- **Map-Server**—The LISP Map-Server (MS) receives endpoint registrations indicating the associated RLOC and uses this to populate the HTDB.
- **Map-resolver**—The LISP Map-Resolver (MR) responds to queries from fabric devices requesting RLOC mapping information from the HTDB in the form of an EID-to-RLOC binding. This tells the requesting device to which fabric node an endpoint is connected and thus where to direct traffic.

Edge Node

The SD-Access fabric edge nodes are the equivalent of an access layer switch in a traditional campus LAN design. The edge node functionality is based on the Ingress and Egress Tunnel Routers (xTR) in LISP. The edge nodes must be implemented using a Layer 3 routed access design. The provide the following fabric functions:

- **Endpoint registration**—Each edge node has a LISP control-plane session to all control plane nodes. After an endpoint is detected by the edge node, it is added to a local database called the *EID-table*. Once the host is added to this local database, the edge node also issues a LISP map-register message to inform the control plane node of the endpoint so the central HTDB is updated.
- **Anycast Layer 3 gateway**—A common gateway (IP and MAC addresses) is used at every edge node that shares a common EID subnet providing optimal forwarding and mobility across different RLOCs. On edge nodes, the Anycast Layer 3 gateway is instantiated as a Switched Virtual Interface (SVI) with a hard-coded MAC address that is uniform across all edge nodes within a fabric site.
- **Mapping of user to virtual network**—Endpoints are placed into virtual networks by assigning the endpoint to a VLAN associated to an SVI that is forwarding for a VRF. Together, these make up the Layer 2 and Layer 3 LISP VNIs, respectively, which maintain fabric segmentation even at the control plane communication level.
- **AAA Authenticator**—The mapping of endpoints into VLANs can be done statically or dynamically using an Authentication Server. Operating as a Network Access Device (NAD), the edge node is an integral part of the IEEE 802.1X port-based authentication process by collecting authentication credentials from connected devices, relaying the to the Authentication Server, and enforcing the authorization result.
- **VXLAN encapsulation/de-encapsulation**—Packets and frames received from endpoint, either directly connected to an edge node or through it by way of an extended node or access point, are encapsulated in fabric VXLAN and forwarded across the overlay. Traffic is either sent to another edge node or to the border node, depending on the destination.

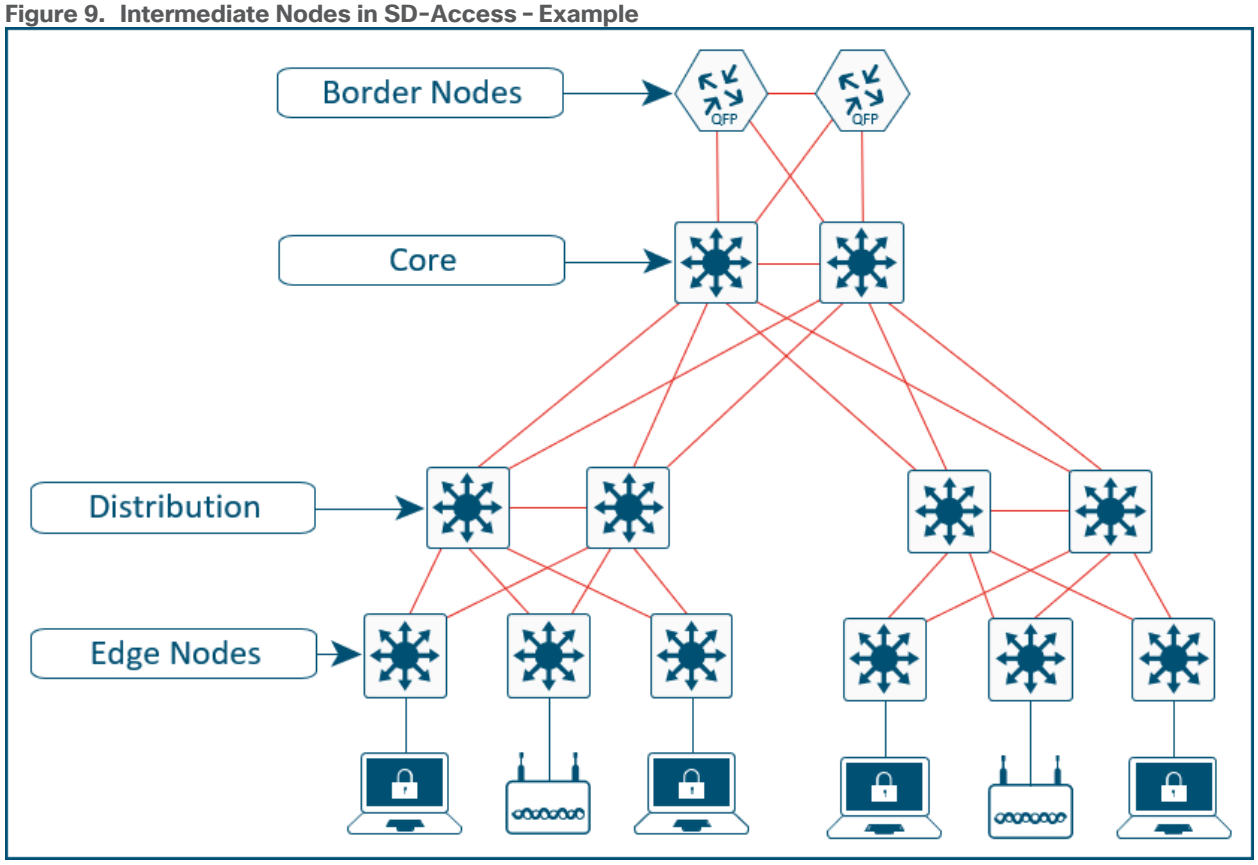
  When fabric encapsulated traffic is received for the endpoint, such as from a border node or from another edge node, it is de-encapsulated and sent to that endpoint. This encapsulation and de-encapsulation of traffic enables the location of an endpoint to change, as the traffic can be encapsulated towards different edge nodes in the network, without the endpoint having to change its address.

Intermediate Node

Intermediate nodes are part of the Layer 3 network used for interconnections among the devices operating in a fabric role such as the interconnections between border nodes and edge nodes. These interconnections are created in the Global Routing Table on the devices and is also known as the *underlay network*. For example, if a three-tier campus deployment provisions the core switches as the border nodes and the access switches as the edge nodes, the distribution switches are the intermediate nodes.

The number of intermediate nodes is not limited to a single layer of devices. For example, borders nodes may be provisioned on an enterprise edge routers resulting in the intermediate nodes being the core and distribution layers as shown in Figure 9.

Figure 9. Intermediate Nodes in SD-Access – Example



Intermediate nodes do not have a requirement for VXLAN encapsulation/de-encapsulation, LISP control plane messaging support, or SGT awareness. Their requirement is to provide IP reachability, physical connectivity, and to support the additional MTU requirement to accommodate the larger-sized IP packets encapsulated with fabric VXLAN information. Intermediate nodes simply route and transport IP traffic between the devices operating in fabric roles.

Tech tip
VXLAN adds 50 bytes to the original packet. The common denominator and recommended MTU value available on devices operating in a fabric role is 9100. Network should have a minimum starting MTU of at least 1550 bytes to support the fabric overlay. MTU values between 1550 and 9100 are supported along with MTU values larger than 9100 though there may be additional configuration and limitations based on the original packet size. MTU 9100 is provisioned as part of LAN Automation. Devices in the same routing domain and Layer 2 domain should be configured with a consistent MTU size to support routing protocol adjacencies and packet forwarding without fragmentation.

Border Node

The fabric border nodes serve as the gateway between the SD-Access fabric site and the networks external to the fabric. The border node is responsible for network virtualization interworking and SGT propagation from the fabric to the rest of the network.

Border nodes implement the following functions:

- **Advertisement of EID subnets**—BGP (Border Gateway Protocol) is the routing protocol provisioned to advertise the coarse-aggregate endpoint prefix space outside the fabric. This is also necessary so that traffic from outside of the fabric destined for endpoints in the fabric is attracted back to the border nodes.

- **Fabric site exit point**—The external border node is the gateway of last resort for the fabric edge nodes. This is implemented using LISP Proxy Tunnel Router (PxTR) functionality. Also possible is the internal border node which registers known networks (IP subnets) with the fabric control plane node.
- **Network virtualization extension to the external world**—The border node can extend network virtualization from inside the fabric to outside the fabric by using VRF-lite and VRF-aware routing protocols to preserve the segmentation.
- **Policy mapping**—The border node maps SGT information from within the fabric to be appropriately maintained when exiting that fabric. Discussed further in the Micro-segmentation section, when the fabric packet is de-encapsulated at border, SGT information can be propagated using SGT Exchange Protocol (SXP) or by directly mapping SGTs into the Cisco metadata field in a packet using inline tagging.
- **VXLAN encapsulation/de-encapsulation**—Packets and frames received from outside the fabric and destined for an endpoint inside of the fabric are encapsulated in fabric VXLAN by the border node. Packets and frames sourced from inside the fabric and destined outside of the fabric are de-encapsulated by the border node. This is similar to the behavior used by an edge node except, rather than being connected to endpoints, the border node connects a fabric site to a non-fabric network.

Fabric in a Box

Fabric in a Box is an SD-Access construct where the border node, control plane node, and edge node are running on the same fabric node. This may be a single switch, a switch with hardware stacking, or a StackWise Virtual deployment. Fabric in a Box is discussed further in Fabric in a Box Site Reference Model section.

Extended Node

SD-Access Extended Nodes provide the ability to extend the enterprise network by providing connectivity to non-carpeted spaces of an enterprise – commonly called the *Extended Enterprise*. This allows network connectivity and management of IoT devices and the deployment of traditional enterprise end devices in outdoor and non-carpeted environments such as distribution centers, warehouses, or Campus parking lots.

This feature extends consistent, policy-based automation to Cisco Industrial Ethernet, Catalyst 3560-CX Compact, and Digital Building Series switches and enables segmentation for user endpoints and IoT devices connected to these nodes. Using Cisco DNA Center automation, switches in the extended node role are onboarded to their connected edge node using an 802.1Q trunk over an EtherChannel with one or multiple physical link members. Extended nodes are discovered using zero-touch Plug-and-Play.

Extended nodes offer a Layer 2 port extension to a fabric edge node while providing segmentation and group-based polices to the endpoints connected to these switches. Endpoints, including fabric-mode APs, can connect directly to the extended node. VLANs and SGTs are assigned using host onboarding as part of fabric provisioning.

The benefits of extending fabric capabilities using extended nodes are operational simplicity for IoT using Cisco DNA Center-based automation, consistent policy across IT and OT (Operational Technology) systems, and greater network visibility of IoT (Internet of Things) devices.

Tech tip
Additional design details and supported platforms are discussed in Extended Node Design section below. For further details on Cisco IoT solutions and the associated Cisco Validated Designs, please see the Cisco Extended Enterprise Non-Fabric and SD-Access Fabric Design Guide, Connected Communities Infrastructure Solution Design Guide, and visit https://www.cisco.com/go/iot.

Fabric WLC

Both fabric WLCs and non-fabric WLCs provide AP image and configuration management, client session management, and mobility services. Fabric WLCs provide additional services for fabric integration such as registering MAC addresses of wireless clients into the host tracking database of the fabric control plane nodes during wireless client join events and supplying fabric edge node RLOC-association updates to the HTDB during client roam events.

In a traditional Cisco Unified Wireless network, or *non-fabric* deployment, both control traffic and data traffic are tunneled back to the WLC using CAPWAP (Control and Provisioning of Wireless Access Points). From a CAPWAP control plane perspective, AP management traffic is generally lightweight, and it is the client data traffic that is generally the larger bandwidth consumer. Wireless standards have allowed larger and larger data rates for wireless clients, resulting in more and more client data that is tunneled back to the WLC. The requires a larger WLC with multiple high-bandwidth interfaces to support the increase in client traffic.

In *non-fabric* wireless deployments, wired and wireless traffic have different enforcement points in the network. Quality of service and security are addressed by the WLC when it bridges the wireless traffic onto the wired network. For wired traffic, enforcement is addressed by the first-hop access layer switch. This paradigm shifts entirely with SD-Access Wireless. In SD-Access Wireless, the CAPWAP tunnels between the WLCs and APs are used for control traffic only. Data traffic from the wireless endpoints is tunneled to the first-hop fabric edge node where security and policy can be applied at the same point as with wired traffic.

Typically, fabric WLCs connect to a shared services network though a distribution block or data center network that is connected outside the fabric and fabric border, and the WLC management IP address exists in the global routing table. For wireless APs to establish a CAPWAP tunnel for WLC management, the APs must be in a VN that has access to this external device. This means that the APs are deployed in the global routing table and that the WLC's address must be present in the GRT within the fabric site.

In the SD-Access solution, Cisco DNA Center configures wireless APs to reside within an overlay VN named *INFRA_VN* which maps to the global routing table. This avoids the need for route leaking or fusion routing (a multi-VRF device selectively sharing routing information) to establish connectivity between the WLCs and the APs. Each fabric site must have a WLC unique to that site. Most deployments place the WLC in the local fabric site itself, not across a WAN, because of latency requirements for local mode APs. Further latency details are covered in the section below.

Tech tip
Strategies on connecting the fabric to shared services and details on route leaking and fusion routing are discussed in the External Connectivity and VRF-Aware Peer sections below.

Fabric access points operate in local mode. This requires an RTT (round-trip time) of 20ms or less between the AP and the WLC. This generally means that the WLC is deployed in the same physical site as the access points. If this latency requirement is meant through dedicated dark fiber or other very low latency circuits between the physical sites and the WLCs deployed physically elsewhere such as in a centralized data center, WLCs and APs may be in different physical locations as shown later in Figure 42.

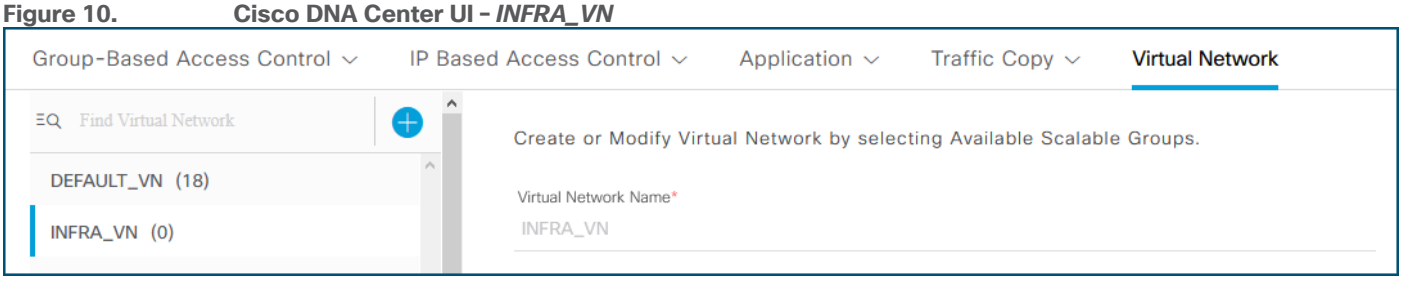
This deployment type, with fabric APs in a separate physical location than their fabric WLCs, is commonly deployed in metro area networks and in SD-Access for Distributed Campus. APs should not be deployed across the WAN or other high latency circuits from their WLCs in an SD-Access network. A maximum RTT of 20ms between these devices is crucial.

Fabric-Mode Access Points

The fabric-mode APs are Cisco Wi-Fi 6 (802.11ax) and 802.11ac Wave 2 APs associated with the fabric WLC that have been configured with one or more fabric-enabled SSIDs. Fabric-mode APs continue to support the same wireless media services that traditional APs support such as applying AVC, quality of service (QoS), and other wireless policies. Fabric APs establish a CAPWAP control plane tunnel to the fabric WLC and join as local-mode APs. They must be directly connected to the fabric edge node or extended node switch in the fabric site. For their data plane, Fabric APs establish a VXLAN tunnel to their first-hop fabric edge switch where wireless client traffic is terminated and placed on the wired network.

Fabric APs are considered a special case *wired host*. Edge nodes use Cisco Discovery Protocol (CDP) to recognize APs as these *wired hosts*, apply specific port configurations, and assign the APs to a unique overlay network called *INFRA_VN*. As a *wired host*, access points have a dedicated EID-space and are registered with the control plane node. This EID-space is associated with a predefined overlay network called *INFRA_VN* in the Cisco DNA Center UI as shown in Figure 10. It is a common EID-space (prefix space) and common virtual network for all fabric APs within a fabric site. The assignment to this overlay virtual network allows management simplification by using a single subnet to cover the AP infrastructure at a fabric site.

Figure 10. Cisco DNA Center UI – *INFRA_VN*



Tech tip
For additional information and details on wireless operations and communications with SD-Access Wireless, Fabric WLCs, and Fabric APs, please see the SD-Access Wireless Design and Deployment Guide.

SD-Access Embedded Wireless

To enable wireless controller functionality without a hardware WLC in distributed branches and small campuses, the Cisco Catalyst 9800 Embedded Wireless Controller is available for Catalyst 9000 Series switches as a software package on switches running in Install mode. The wireless control plane of the embedded controller operates like a hardware WLC. CAPWAP tunnels are initiated on the APs and terminate on the Cisco Catalyst 9800 Embedded Wireless Controller. The data plane uses VXLAN encapsulation for the overlay traffic between the APs and the fabric edge node.

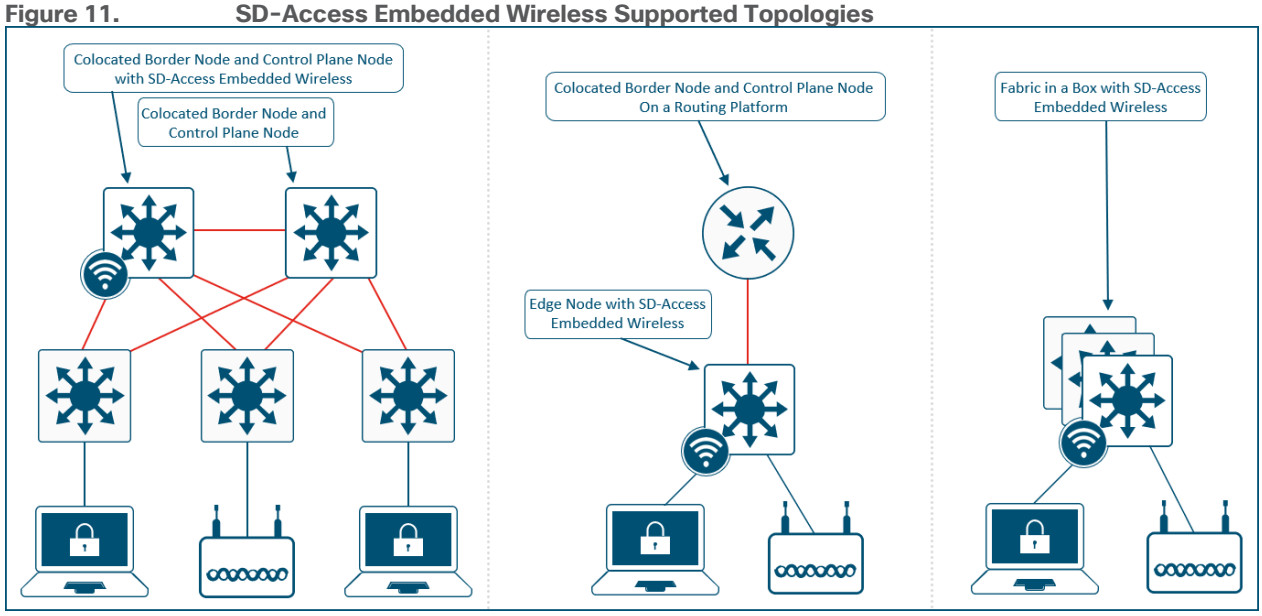
The Catalyst 9800 Embedded Wireless Controller for Catalyst 9000 Series switches is supported for SD-Access deployments with three topologies:

- Cisco Catalyst 9000 Series switches functioning as colocated border and control plane.
- Cisco Catalyst 9000 Series switches functioning as an edge node when the border and control plane node are on a routing platform.
- Cisco Catalyst 9000 Series switches functioning as a Fabric in a Box.

Tech tip
All Catalyst 9000 Series switches support the SD-Access Embedded Wireless functionality except for the Catalyst 9200,

9200L, and 9600 Series Switches. Fabric in a Box deployments operating in StackWise Virtual do not support the embedded wireless controller functionality and should use a hardware-based or virtual WLC (Catalyst 9800-CL).

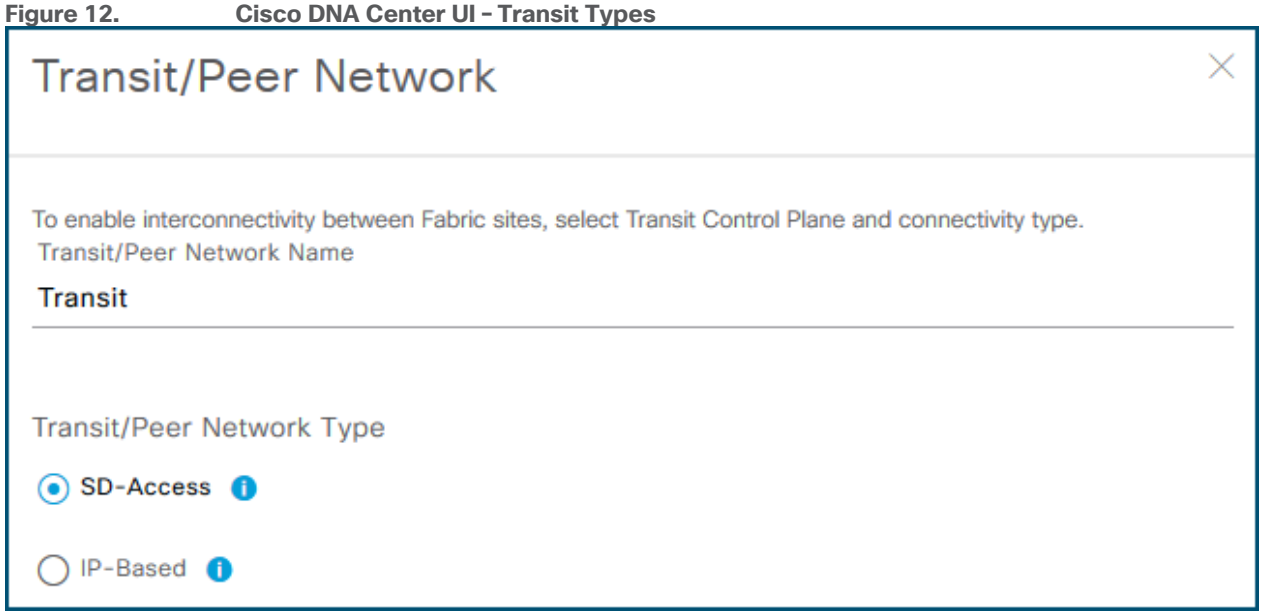
Figure 11. SD-Access Embedded Wireless Supported Topologies



Transit and Peer Network

Transits, referred to as *Transit/Peer Networks* in Cisco DNA Center, connect multiple fabric site together. They are an SD-Access construct that defines how Cisco DNA Center will automate the border node configuration for the connections between fabric sites or between a fabric site and the external world. This connectivity may be MAN, WAN, or Internet. The WAN could be MPLS, SD-WAN, IWAN, or other WAN variations. As shown in Figure 12, the Cisco DNA Center user interface refers to the transits as *IP-Based* or *SD-Access* transit/peer network types.

Figure 12. Cisco DNA Center UI - Transit Types



- **IP-Based Transits**—Packets are de-encapsulated from the fabric VXLAN into native IP. Once in native IP, they are forwarded using traditional routing and switching modalities. IP-based transits are provisioned with VRF-lite to connect to the upstream device. IP-Based transits are commonly used to connect to shared services using a VRF-Aware Peer and connecting to upstream routing infrastructure or firewall for connectivity to WAN and Internet.
- **SD-Access Transits**—SD-Access transits are exclusive used in SD-Access for Distributed Campus. Using the SD-Access transit, packets are encapsulated between sites using the fabric VXLAN encapsulation. This natively carries the macro (VRF) and micro (SGT) policy constructs between fabric sites.

Transit Control Plane Nodes

Transit control planes nodes are a fabric role construct supported in SD-Access for Distributed Campus. It operates in the same manner as a site-local control plane node except it services the entire fabric. Transit control plane nodes are only required when using SD-Access transits.

Each fabric site will have their own site-local control plane nodes for intra-site communication, and the entire domain will use the transit control plane nodes for inter-site communication. Transit control plane nodes provide the following functions:

- **Site aggregate prefix registration**—Border nodes connected to the SD-Access Transit use LISP map-register message to inform the transit control plane nodes of the aggregate prefixes associated with the fabric site. This creates an aggregate HTDB for all fabric sites connected to the transit. Rather than a host route being associated with a routing locator (EID-to-RLOC binding) which is what occurs in a site-local control plane node, the transit control plane node associated the aggregate prefix with a border node's RLOC.
- **Control Plane signaling**—Once aggregate prefixes are registered for each fabric site, control-plane signaling is used to direct traffic between the sites. When traffic from an endpoint in one fabric site needs to send traffic to an endpoint in another site, the transit control plane node is queried to determine to which site's border node this traffic should be sent.

Fabric Domain

A fabric domain is a Cisco DNA Center UI construct. It is an organization scope that consists of multiple fabric sites and their associated transits. The concept behind a fabric domain is to show certain geographic portions of the network together on the screen. For example, an administrator managing a fabric site in San Jose, California, USA and another fabric site in Research Triangle Park, North Carolina, USA, which are approximately 3,000 miles (4,800 kilometers) apart, would likely place these fabric sites in different fabric domains unless they were connected to each other with the same transit.

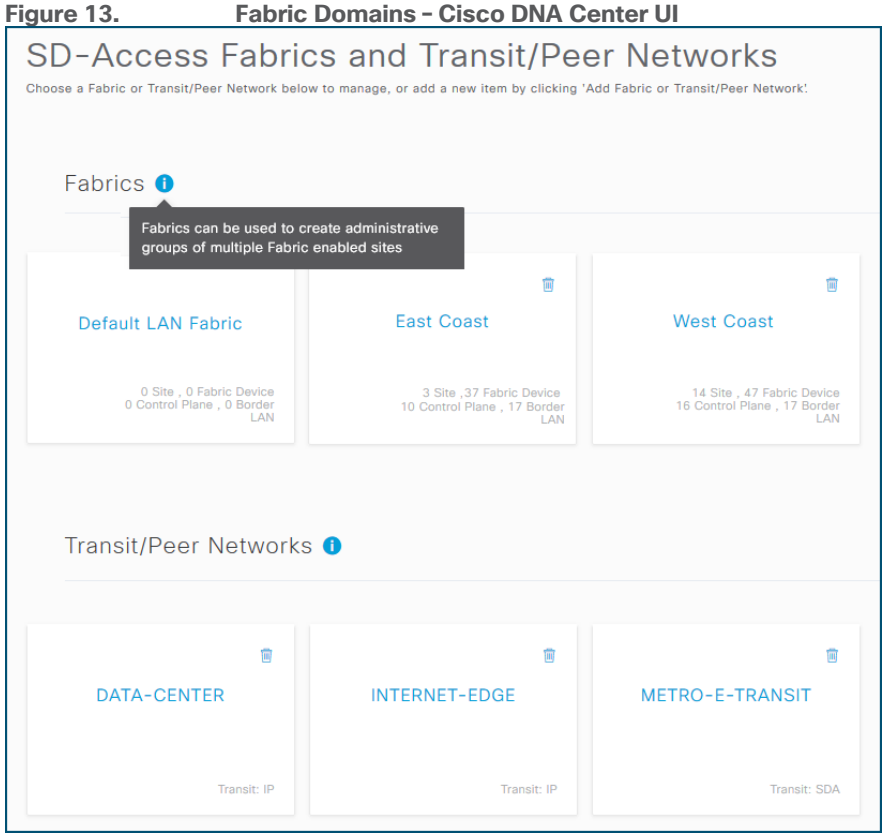
Figure 13 shows three fabric domains. *Default LAN Fabric* is created by default, though is not required to be used, and *East Coast* and *West Coast* are user-defined. The large text ***Fabrics*** represents fabric domains and not fabric sites which are shown Figure 14. Also shown are three different *Transit/Peer Networks*. *DATA-CENTER* and *INTERNET-EDGE* are both IP-based transits, and *METRO-E-TRANSIT* is an SD-Access transit used for Distributed Campus.

Both *East Coast* and *West* Coast have a number of fabric sites, three (3) and fourteen (14) respectively, in their domain along with a number of control plane nodes and borders nodes. It is not uncommon to have hundreds of sites under a single fabric domain.

Tech tip

For the number of supported fabric domains based on appliance size, please reference the Cisco DNA Center Data Sheet Appliance Scale and Hardware Specifications and Cisco DNA Center and SD-Access 1.3 Scale Metrics on Cisco Communities.

Figure 13. Fabric Domains – Cisco DNA Center UI



Tech tip

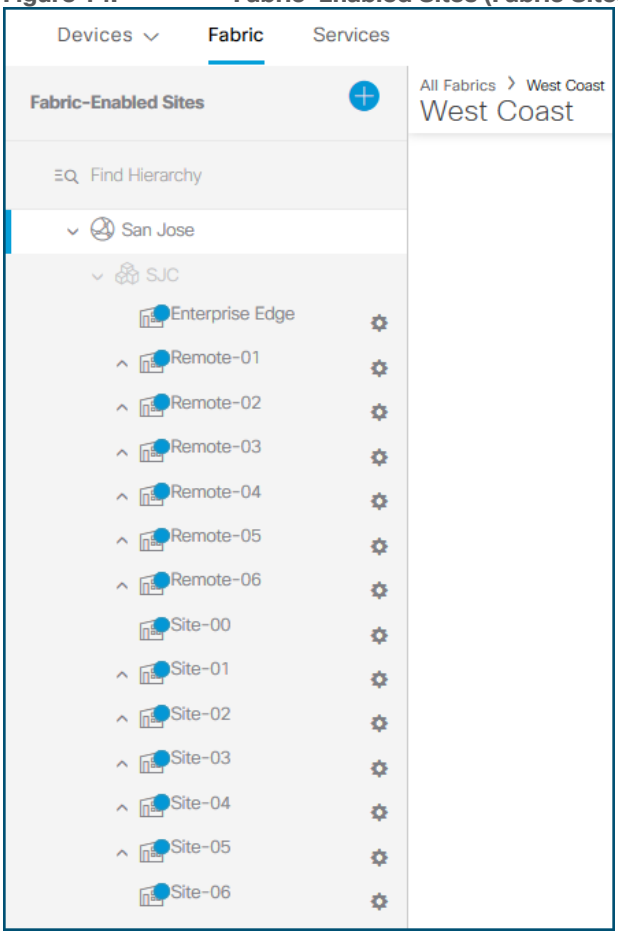
For additional details on fabric domains, please see BRKCRS-2810-Cisco SD-Access - Under the Hood (2019, Cancun) and SD-Access for Distributed Campus Deployment Guide.

Fabric Site

A fabric site is composed of a unique set of devices operating in a fabric role along with the intermediate nodes used to connect those devices. At minimum, a fabric site must have a control plane node and an edge node, and to allow communication to other destinations outside of the fabric site, a border node. A fabric site generally has an associated WLC and potentially an ISE Policy Service Node (PSN).

Figure 14 shows the ***Fabric-Enabled Sites***, or fabric sites, under the *West Coast* domain from Figure 13. Fourteen (14) fabric sites have been created. Each site has its own independent set of control plane nodes, border nodes, and edge nodes along with a WLC.

Figure 14. Fabric-Enabled Sites (Fabric Sites) – Cisco DNA Center UI



SD-Access Design Considerations

This chapter is organized into the following sections:

Chapter	Section
SD-Access Design Considerations	LAN Design Principles Device Role Design Principles Feature-Specific Design Requirements Wireless Design External Connectivity Security Policy Considerations Multidimensional Considerations

Any successful design or system is based on a foundation of solid design theory and principles. Designing an SD-Access network or fabric site as a component of the overall enterprise LAN design model is no different than designing any large networking system. The use of a guiding set of fundamental engineering principles ensures that the design provides a balance of availability, security, flexibility, and manageability required to meet current and future technology needs.

This section provides design guidelines that are built upon these balanced principles to allow an SD-Access network architect to build the fabric using next-generation products and technologies. These principles allow

for simplified application integration and the network solutions to be seamlessly built on a modular, extensible, and highly-available foundation design that can provide continuous, secure, and deterministic network operations.

This section will begin by discussing LAN design principles, discusses design principles covering specific device roles, feature-specific design considerations, wireless design, external connectivity, security policy design, and multidimensional considerations.

LAN Design Principles

This major section is organized into the following subsections:

Section	Subsection
LAN Design Principles	Underlay Network Design Overlay Network Design Shared Services Design Fabric DHCP Overview and Design Latency

The following LAN design principles apply to networks of any size and scale. This section looks at underlay network, overlay network, shared services and services blocks, DHCP in the Fabric along with latency requirements for the network.

Underlay Network Design

This section is organized into the following subsections:

Section	Subsection
Underlay Network Design	Layer 3 Routed Access Introduction Enterprise Campus Architecture Layer 2 (Switched) Access – Traditional Campus Design About Layer 3 Routed Access Layer 3 Routed Access and SD-Access Network Design StackWise Virtual in SD-Access and Layer 3 Routed Access Networks

Having a well-designed underlay network ensures the stability, performance, and efficient utilization of the SD-Access network. Automation for deploying the underlay is available using Cisco DNA Center using the LAN Automation capability which is discussed in a later section.

Whether using LAN Automation or deploying the network manually, the underlay networks for the fabric have the following general design requirements:

- **Layer 3 Routed Access**—The use of a Layer 3 routed access network for the fabric provides the highest level of availability without the need to use loop avoidance protocols such as Spanning-Tree (STP), interface bundling techniques using link aggregation technologies such as EtherChannel, and Layer 2 redundancy technologies like StackWise Virtual (SVL), Virtual Switching System (VSS), or Nexus Virtual Port-Channels (vPCs).

- **Increase default MTU**—The VXLAN header adds 50 bytes of encapsulation overhead. Enabling a campus and branch wide MTU of 9100 ensures that Ethernet jumbo frames can be transported without fragmentation inside the fabric.
- **Point-to-point links**—Point-to-point links provide the quickest convergence times because they eliminate the need to wait for the upper layer protocol timeouts typical of more complex topologies. Combining point-to-point links with the recommended physical topology design provides fast convergence in the event of a link failure.

  The fast convergence is a benefit of quick link failure detection triggering immediate use of alternate topology entries preexisting in the routing and forwarding table. Implement the point-to-point links using optical technology as optical (fiber) interfaces are not subject to the same electromagnetic interference (EMI) as copper links. Copper interfaces can be used, though optical ones are preferred.
- **ECMP**—Equal-cost multi-path routing is a routing strategy where next-hop packet forwarding to a single destination can occur over multiple best paths. Load balancing between these ECMP paths is performed automatically using Cisco Express Forwarding (CEF). ECMP-aware routing protocols should be used to take advantage of the parallel-cost links and to provide redundant forwarding paths for resiliency.
- **BFD**—Bidirectional Forwarding Detection enhances fault detection and convergence characteristics of routing protocols. Routing protocols use the absence of *Hello* packets to determine if an adjacent neighbor is down (commonly called *Hold Timer* or *Dead Timer*). Thus, the ability to detect liveliness in a neighbor is based on the frequency of *Hello* packets.

  Each *Hello* packet is processed by the routing protocol adding to the overhead and rapid *Hello* messages creates an inefficient balance between liveliness and churn. BFD provides low-overhead, sub-second detection of failures in the forwarding path between devices and can be set a uniform rate across a network using different routing protocols that may have variable *Hello* timers.
- **NSF**—Non-stop forwarding, or *graceful restart*, works with SSO (stateful switchover) to provide continued forwarding of packets in the event of a route processor (RP) switchover. NSF-aware IGP routing protocols should be used to minimize the amount of time that a network is unavailable following a switchover.
- **SSO**—Stateful Switchover maintains stateful feature information, such as user session, by synchronizing state information between a primary and backup route processor such as an RPs in routing platforms or supervisor engines in switching platforms. SSO should be enabled in concert with NSF on supported devices.
- **IGP process for the fabric**—While IS-IS is recommended and required for LAN Automation, as described below, other classless routing protocols such as OSPF and EIGRP are supported and are both ECMP and NSF-aware.
- **Loopback propagation**—The loopback addresses assigned to the underlay devices need to propagate outside of the fabric to establish connectivity to infrastructure services such as fabric control plane nodes, DNS, DHCP, and AAA.

  Loopback 0 interfaces (RLOC) require a /32 subnet mask. These addresses also be propagated throughout the fabric site.

  Reachability between loopback address (RLOCs) cannot use the default route. They must use a /32 route.

- **WLC reachability**—Connectivity to the WLC should be treated like reachability to the loopback addresses. A default route in the underlay cannot be used by the APs to reach the WLCs. A specific route (non-default route) to the WLC IP address must exist in the Global Routing Table at each switch where the APs are physically connected. This can be a host route (/32) or summarized route.
- **LAN Automation for deployment**—The configuration of the underlay can be orchestrated by using LAN Automation services in Cisco DNA Center. The LAN Automation feature is an alternative to manual underlay deployments for new networks and uses an IS-IS routed access design. Although there are many alternative routing protocols, the IS-IS routing protocol offers operational advantages such as neighbor establishment without IP protocol dependencies, peering capability using loopback addresses, and agnostic treatment of IPv4, IPv6, and non-IP traffic. Manual underlays are also supported and allow variations from the automated underlay deployment (for example, a different IGP could be chosen), though the underlay design principles still apply.

Layer 3 Routed Access Introduction

For campus designs requiring simplified configuration, common end-to-end troubleshooting tools, and the fastest convergence, a design using Layer 3 switches in the access layer (routed access) in combination with Layer 3 switching at the distribution layer and core layers provides the most rapid convergence of data and control plane traffic flows.

This section describes the Enterprise Campus hierarchical network structure followed by traditional campus designs that use the distribution layer as the Layer 2/Layer 3 boundary (switched access). This traditional design is then contrasted against moving the Layer 2/Layer 3 boundary to the access layer (routed access), a requirement for SD-Access, and finally discusses design considerations for Layer 3 routed access.

Enterprise Campus Architecture Introduction

Hierarchical network models are the foundation for modern network architectures. This allows network systems, both large and small, simple and complex, to be designed and built using modularized components. These components are then assembled in a structured and hierarchical manner while allowing each piece (component, module, and hierarchical point) in the network to be designed with some independence from overall design. Modules (or blocks) can operate semi-independently of other elements, which in turn provides higher availability to the entire system. By dividing the Campus system into subsystems and assembling them into a clear order, a higher degree of stability, flexibility, and manageability is achieved for the individual pieces of the network and the campus deployment as a whole.

These hierarchical and modular networks models are referred to as the *Cisco Enterprise Architecture Model* and have been the foundation for building highly available, scalable, and deterministic networks for nearly two decades. The Enterprise Architecture Model separates the network into different functional areas called modules or blocks designed with hierarchical structures. The Enterprise Campus is traditionally defined with a three-tier hierarchy composed of the Core, Distribution, and Access Layers. In smaller networks, two-tiers are common with core and distribution collapsed into a single layer (collapsed core). The key idea is that each element in the hierarchy has a specific set of functions and services that it offers. The same key idea is referenced later in the fabric control plane node and border node design section.

The access layer represents the network edge where traffic enters or exits the campus network towards users, devices, and endpoints. The primary function of an access layer switch is to provide network access to the users and endpoint devices such as PCs, printers, access points, telepresence units, and IP phones.

The distribution layer is the interface between the access and the core providing multiple, equal cost paths to the core, intelligent switching and routing, and aggregation of Layer 2 and Layer 3 boundaries.

The Core layer is the backbone interconnecting all the layers and ultimately providing access to the compute and data storage services located in the data center and access to other services and modules throughout the network. It ties the Campus together with high bandwidth, low latency, and fast convergence.

Tech tip
For additional details on the Enterprise Campus Architecture Model, please see: • Hierarchical Network Design Overview • Cisco Enterprise Architecture Model • Enterprise Campus 3.0 Architecture Overview • High Availability Design Guide • Medium Enterprise Design Profile Reference Guide • Small Enterprise Design Profile Reference Guide

Layer 2 (Switched) Access – Traditional Campus Design

In typical hierarchical design, the access layer switch is configured as a Layer 2 switch that forwards traffic on high speed trunk ports to the distribution switches. The distribution switches are configured to support both Layer 2 switching on their downstream trunks and Layer 3 switching on their upstream ports towards the core of the network. The function of the distribution switch in this design is to provide boundary functions between the bridged Layer 2 portion of the campus and the routed Layer 3 portion, including support for the default gateway, Layer 3 policy control, and all required multicast services.

The distribution block would typically span VLANs across the layer with the default gateway provided through SVI (Switched Virtual Interfaces) and distribution peer switches running first-hop redundancy protocols (FHRP) such as HSRP (Hot Standby Router Protocol). Alternatively, distribution switch peers may run Virtual Switching System (VSS) or Stackwise Virtual (SVL) to act as a single, logical entity and provide Multichassis EtherChannel (MEC) to access layer switches.

Layer 2 access networks provide the flexibility to allow applications that require Layer 2 connectivity to extend across multiple wiring closets. This design does come with the overhead of Spanning-Tree Protocol (STP) to ensure loops are not created when there are redundant Layer 2 paths in the network.

The stability of and availability for the access switches is layered on multiple protocol interactions in a Layer 2 switched access deployment. For example, in a common Layer 2 access network, the HSRP gateway for a VLAN should be the STP root bridge. Link Aggregation (LAG) is provided via LACP (Link Aggregation Control Protocol) or PAgP (Port Aggregation Protocol) to connect to upstream switches using MEC. These upstream switches are often configured with VSS / SVL, separate protocols themselves from LAG, to provide a logical entity across two physical devices. Trunking protocols ensure VLANs are spanned and forwarded to the proper switches throughout the system. While all of this can come together in an organized, deterministic, and accurate way, there is much overhead involved both in protocols and administration, and ultimately, spanning-tree is the protocol pulling all the desperate pieces together. All the other protocols and their interactions rely on STP to provide a loop-free path within the redundant Layer 2 links. If a convergence problem occurs in STP, all the other technologies listed above can be impacted.

About Layer 3 Routed Access

The hierarchical Campus, whether Layer 2 switched or Layer 3 routed access, calls for a full mesh equal-cost routing paths leveraging Layer 3 forwarding in the core and distribution layers of the network to provide the most reliable and fastest converging design for those layers. An alternative to Layer 2 access model described above is to move the Layer 3 demarcation boundary to the access layer. Layer 2 uplink trunks on the Access

switches are replaced with Layer 3 point-to-point routed links. This brings the advantages of equal cost path routing to the Access layer.

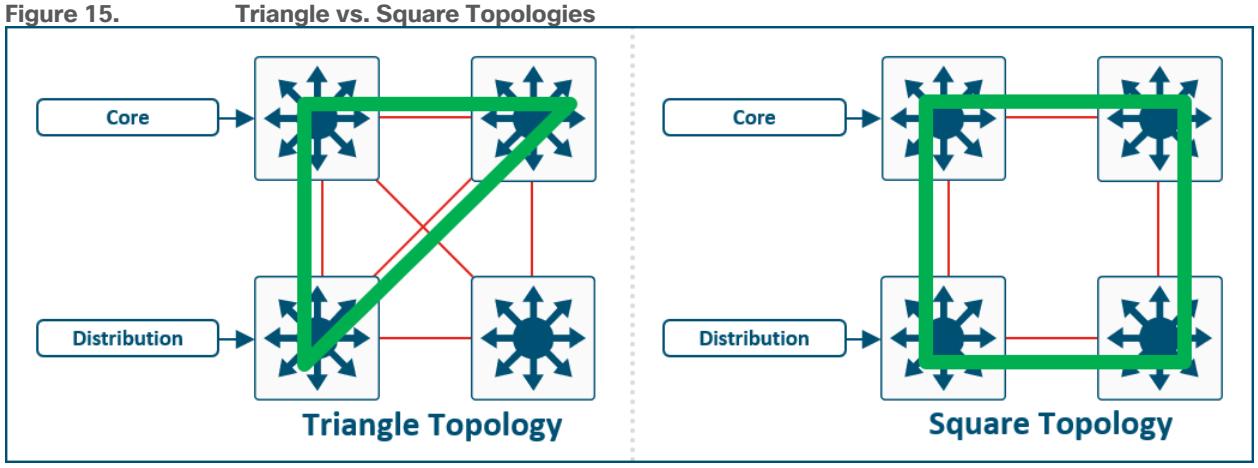
Using routing protocols for redundancy and failover provides significant convergence improvement over spanning-tree protocol used in Layer 2 designs. Each switch has two routes and two associated hardware Cisco Express Forwarding (CEF) forwarding adjacency entries. Traffic is forwarded with both entries using equal-cost multi-path (ECMP) routing. In the event of a failure of an adjacent link or neighbor, the switch hardware and software immediately remove the forwarding entry associated with the lost neighbor. However, the switch still has a remaining valid route and associated CEF forwarding entry. With an active and valid route, traffic is still forwarded. The result is a simpler overall network configuration and operation, dynamic load balancing, faster convergence, and a single set of troubleshooting tools such as *ping* and *traceroute*.

Tech tip
Layer 3 routed access is defined by Layer 3 point-to-point routed links between devices in the Campus hierarchy. In an SD-Access network, Access and distribution switches should not peer with their upstream neighbors using SVIs and trunk ports. SVIs and trunk ports between the layers still have an underlying reliance on Layer 2 protocol interactions.

Layer 3 Routed Access and SD-Access Network Design

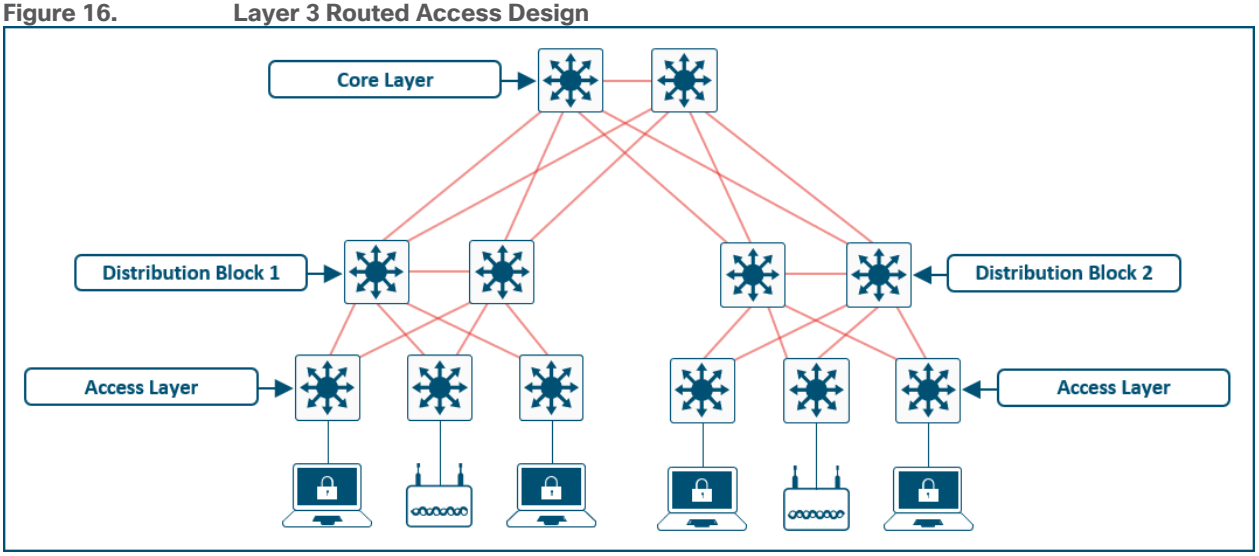
SD-Access networks start with the foundation of a well-design, highly available Layer 3 routed access foundation. For optimum convergence at the core and distribution layer, build triangles, not squares, to take advantage of equal-cost redundant paths for the best deterministic convergence. In Figure 15, the graphic on the left shows triangle topologies which are created by devices crosslinking with each other and with their upstream/downstream peers. The graphic on the right shows square topologies that are created when devices are not connected to both upstream/downstream peers. Square topologies should be avoided.

Figure 15. Triangle vs. Square Topologies



As illustrated in Figure 16, Core switch peer devices should be cross linked to each other. Distribution switches within the same distribution block should be crosslinked to each other and connected to each core switch. Access switches should be connected to each distribution switch within a distribution block, though they do not need to be cross-linked to each other.

Figure 16. Layer 3 Routed Access Design



The interior gateway routing (IGP) routing protocol should be fully featured and support Non-Stop Forwarding, Bidirectional Forwarding Detection, and equal cost multi-path. IS-IS, EIGRP, and OSPF each support these features and can be used as an IGP to build a Layer 3 routed access network. Point-to-point links should be optimized with BFD, a hard-coded carrier-delay and load-interval, enabled for multicast forwarding, and CEF should be optimized to avoid polarization and under-utilized redundant paths.

Tech tip
For more information on Layer 3 routed access design methodology and high availability tuning, please see: Routed Access Layer Design Guide, Tuning for Optimized Convergence Guide , and Routed Access Layer Assurance Guide.

StackWise Virtual in SD-Access and Layer 3 Routed Access Networks

StackWise Virtual (SVL), like its predecessor Virtual Switching System (VSS), is designed to address and simplify Layer 2 operations. It is the virtualization of two physical switches into a single logical switch from a control and management plane perspective. It provides the potential to eliminate spanning tree, first hop redundancy protocol needs, along with multiple touch points to configure those technologies. Using Multichassis EtherChannel (MEC), bandwidth can be effectively doubled with minimized convergence timers using stateful and graceful recovery. In traditional networks, StackWise virtual is positioned in the distribution layer and in collapsed core environments to help VLANs span multiple access layer switches, to provide flexibility for applications and services requiring Layer 2 adjacency, and to provide Layer 2 redundancy.

Layer 3 routed access moves the Layer 2/Layer 3 boundary from the distribution layer to the access layer. The distribution and collapsed core layers are no longer required to service the Layer 2 adjacency and Layer 2 redundancy needs with the boundary shifted. While StackWise Virtual can provide an operational simplicity for control plane protocols and physical adjacencies, it is at the expense of additional protocols designed to solve Layer 2 challenges, and, when leveraged in a Layer 3 routed network, can result in the loss of a redundant IGP/EGP control plane instance. In a Layer 3 routed access environment, two separate, physical switches are best used in all situations except those that may require Layer 2 redundancy.

For example, at the access layer, if physical hardware stacking is not available in the deployed platform, StackWise Virtual can be used to provide Layer 2 redundancy to the downstream endpoints. In SD-Access, StackWise Virtual is best positioned in three places:

- **Edge Node**—Extended nodes or downstream servers hosting virtual endpoints often require Layer 2 high availability. StackWise Virtual can provide multiple, redundant 1- and 10-Gigabit Ethernet connections common on downstream devices.
- **Fabric in a Box**—When deploying a Fabric in a Box, if the given platform does not support hardware stacking, StackWise Virtual can provide redundancy and high availability.
- **Layer 2 Border Handoff**—To support the appropriate scale and physical connectivity when using the Layer 2 handoff feature, StackWise virtual can provide multiple multichassis 10-, 25-, 40-, and even 100-Gigabit Ethernet connections as a handoff connection to an external entity.

Overlay Network Design

In the SD-Access fabric, the overlay networks are used for transporting user traffic across the fabric. The fabric encapsulation also carries scalable group information used for traffic segmentation inside the overlay VNs. Consider the following in the design when deploying virtual networks:

- **Virtual Networks (Macro-segmentation)**—Use virtual networks when requirements dictate isolation at both the data plane and control plane. In general, if devices need to communicate with each other, they should be placed in the same virtual network. If communication is required between different virtual networks, use an external firewall or other device to enable inter-VN communication. Virtual Network provides the same behavior and isolation as VRFs.
- **SGTs (Micro-segmentation)**—Segmentation using SGTs allows for simple-to-manage group-based policies and enables granular data plane isolation between groups of endpoints within a virtualized network. Using SGTs also enables scalable deployment of policy without having to do cumbersome updates for these policies based on IP addresses.
- **Reduce subnets and simplify DHCP management**—In the overlay, IP subnets can be stretched across the fabric without flooding issues that can happen on large Layer 2 networks. Use fewer subnets and DHCP scopes for simpler IP addressing and DHCP scope management. Subnets are sized according to the services that they support, versus being constrained by the location of a gateway. Enabling the optional broadcast flooding (Layer 2 flooding) feature can limit the subnet size based on the additional bandwidth and endpoint processing requirements for the traffic mix within a specific deployment.
- **Avoid overlapping IP subnets**—Different overlay networks can support overlapping address space, but be aware that most deployments require shared services across all VNs and some may use inter-VN communication. Avoid overlapping address space so that the additional operational complexity of adding a network address translation (NAT) device is not required for shared services communication.

Tech tip
The underlay network uses IPv4 address for the Loopback 0 (RLOC) interfaces on the devices operating in a Fabric Role. Connectivity in the underlay should use IPv4 routing to propagate the /32 RLOC routes as discussed in the Underlay Network design section. Endpoints in the overlay space can use IPv4 addresses or dual-stack IPv4/IPv6 addresses.

Shared Services Design

This section is organized into the following subsections:

Section	Subsection
Shared Services Design	Services Block Design

Section	Subsection
	Shared Services Routing Table

Services Block Design

As campus network designs utilize more application-based services, migrate to controller-based WLAN environments, and continue to integrate more sophisticated Unified Communications, it is essential to integrate these services into the campus smoothly while providing for the appropriate degree of operational change management and fault isolation. And this must be done while continuing to maintain a flexible and scalable design.

A services block provides for this through the centralization of servers and services for the Enterprise Campus. The services block serves a central purpose in the campus design: it isolates or separates specific functions into dedicated services switches allowing for cleaner operational processes and configuration management. It also provides a centralized location for applying network security services and policies such as NAC, IPS, or firewall.

The services block is not necessarily a single entity. There might be multiple services blocks depending on the scale of the network, the level of geographic redundancy required, and other operational and physical factors. One services block may service an entire deployment, or each area, building, or site may have its own block.

The services block does not just mean putting more boxes in the network. It is the purpose-built linkage between the campus network and the end user services such as DHCP, DNS, Active Directory (AD), servers, and critical systems and the endpoint services such as the WLC and Unified Communication Systems.

Services blocks are delineated by the services block switch. The services block switch can be a single switch, multiple switches using physical hardware stacking, or be a multi-box, single logical entity such as StackWise Virtual (SVL), Virtual Switching System (VSS), or Nexus Virtual Port-Channels (vPCs). The goal of the services block switch is to provide Layer 3 access to the remainder of the enterprise network and Layer 2 redundancy for the servers, controllers, and applications in the services block. This allows the services block to keep its VLANs distinct from the remainder of the network stack such as the access layer switches which will have different VLANs.

WLCs, Unified Communication Services, and other compute resources should be interconnected with the service block switch using link aggregation (LAG). These Ethernet connections should be distributed among different modular line cards or switch stack members as much as possible to ensure that the failure of a single line card or switch does not result in total failure of the services to remainder of the network. Terminating on different modules within a single Catalyst and Nexus modular switch or different switch stack members provides redundancy and ensures that connectivity between the services block switch and the service block resources are maintained in the rare event of a failure.

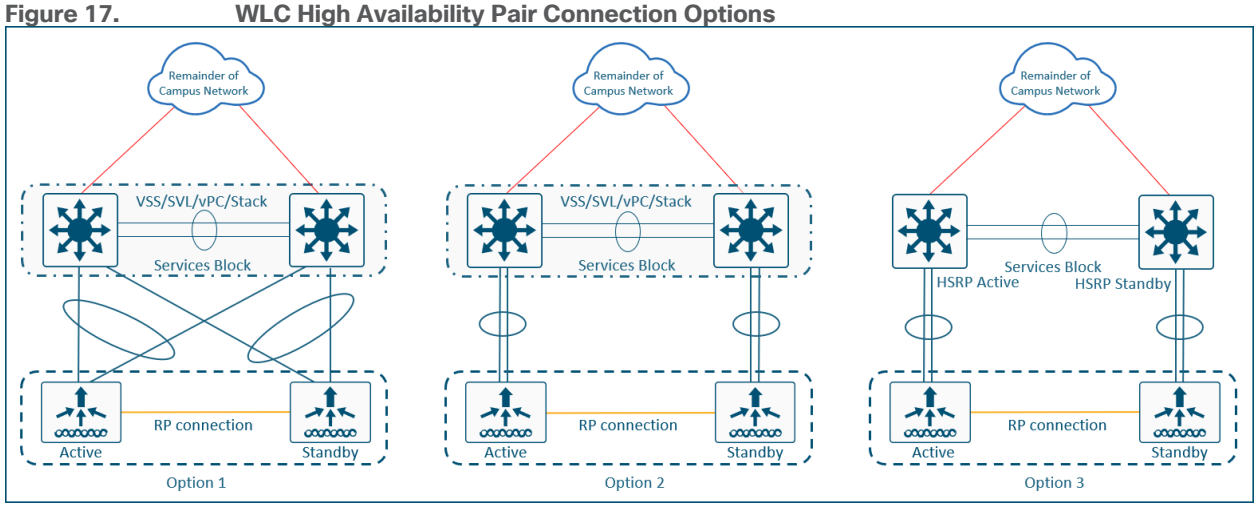
The key advantage of using link aggregation is design performance, reliability, and simplicity. With the Ethernet bundle comprising up to eight links, link aggregation provides very high traffic bandwidth between the controller, servers, applications, and the remainder of the network. If any of the individual ports fail, traffic is automatically migrated to one of the other ports. If at least one port is functioning, the system continues to operate, remain connected to the network, and is able to continue to send and receive data.

When connecting wireless controllers to the services block using link aggregation, one of three approaches can be used:

- **Option 1**—The WLCs are connected to the services block with a Layer 2 port-channel on each WLC connecting to each upstream switch. The links are spread across the physical switches. This is the recommended option.

- **Option 2**—The WLCs are connected to the services block with a Layer 2 port-channel on each WLC without spreading the links across the physical switches. This is a variation of first option and is recommended only if the existing physical wiring will not allow for Option 1.
- **Option 3**—If the services block is not operating in a logical configuration such as VSS, SVL, vPC, or a switch stack, then the first hop redundancy protocol (FHRP) HSRP should be used between the two devices in the services block. One WLC is connected via a port-channel trunk to the HSRP Active switch, and the other WLC is connected via a port-channel trunk to the HSRP Standby switch.

Figure 17. WLC High Availability Pair Connection Options



Tech tip

AireOS WLCs should connect the Redundancy Ports (RPs) back to back on all releases supported in SD-Access. Catalyst 9800 WLCs operating on code before Cisco IOS XE 17.1 (Amsterdam) should connect their RPs through the upstream switch and not back to back.

For additional information regarding RP design and RP connectivity on code after Cisco IOS XE 17.1 on the Catalyst 9800s WLC, please see: High Availability SSO Deployment Guide for Cisco Catalyst 9800 Series Wireless Controllers, Cisco IOS XE Amsterdam 17.1.

Enterprise Campus deployments may span a large geographic area and be separated by MAN, WAN, or even public Internet circuits. If the survivability requirements for these locations necessitate network access, connectivity, and services in the event of egress circuit failure or unavailability, then a services block should be deployed at each physical location with these requirements. Commonly, medium to large deployments will utilize their own services block for survivability, and smaller locations will use centralized, rather than local services.

In very small sites, small branches, and remote sites, services are commonly deployed and subsequently accessed from a central location, generally a headquarters (HQ). However, due to the latency requirements for Fabric APs which operate in local mode, WLCs generally need to be deployed at each location. A maximum round trip time (RTT) of 20ms is required between a local mode access point and the WLC. For these very small or branch locations, a services block may not be needed if the **only** local service is the wireless LAN controller. Some deployments may be able to take advantage of either virtual or switch-embedded Catalyst 9800 WLC as discussed in the Embedded Wireless section.

Tech tip

If additional services are deployed locally such as an ISE PSN, AD, DHCP, or other compute resources, a services block will provide flexibility and scale while providing the necessary Layer 2 adjacency and high availability. A services block is the

recommended design, even with a single service such as a WLC.

Shared Services Routing Table

Once the services block physical design is determined, its logical design should be considered next. Shared services are commonly deployed in the global routing table (GRT) though they are also supported in a VRF. If deployed in a VRF, this routing table should be dedicated only to these shared services.

Discussed in detail later in the External Connectivity section, the endpoint prefix-space in the fabric site will be present on the border nodes for advertisement to the external world. However, these prefixes will be in a VRF table, not the global routing table. This later section discussion options on connecting the border node to shared services, Internet, and outside the fabric.

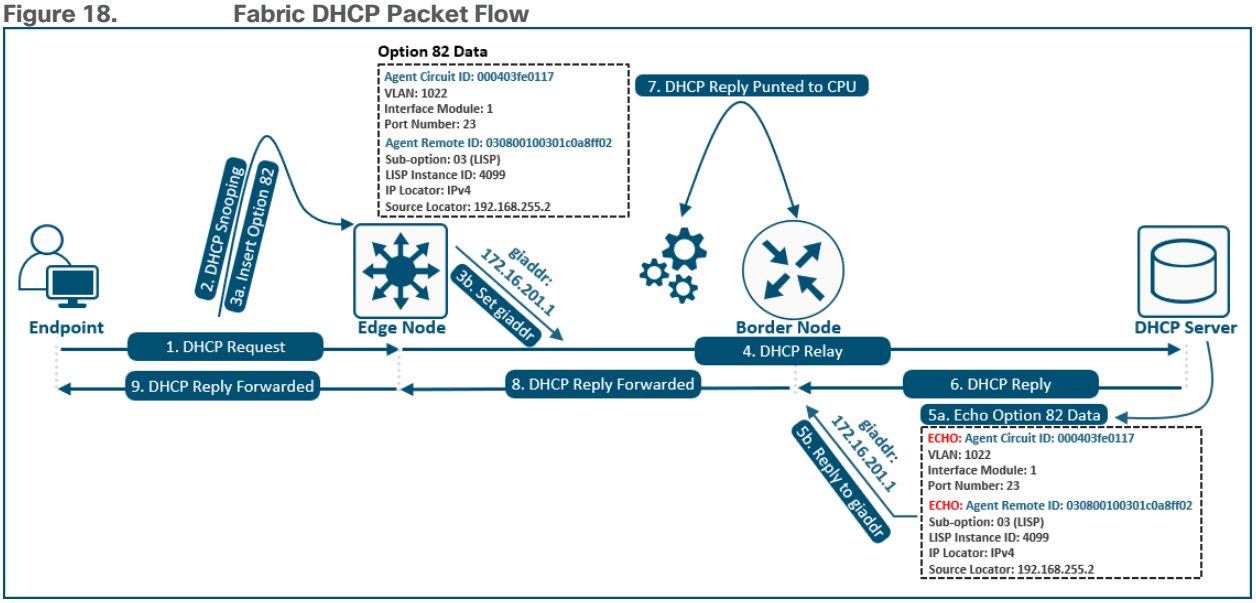
With shared services in a dedicated VRF, route leaking (VRF to VRF leaking) is administratively straightforward as it uses route-targets under the VRF configuration, although it is at the expense of creating another VRF to manage. The alternative approach, shared services in the GRT, requires a different approach to leak routes for access to shared services. The process still requires the same handoff components to the external entity to the border node, though with slightly more touch points. These begin with IP prefix-list for each VN in the fabric that references each of the associated subnets. A route-map is created to match on each prefix-list. Finally, the VRF configuration imports and exports routes that are filtered based on these route-maps.

While the second approach, shared services in GRT, may have more configuration elements, it also provides the highest degree of granularity. Specific routes can be selectively and systematically leaked from the global routing table to the fabric VNs without having to maintain a dedicated VRF for shared services. Both approaches are supported, although the underlying decision for the routing table used by shared services should be based on the entire network, not just the SD-Access fabric sites.

Fabric DHCP Overview and Design

SD-Access does not require any specific changes to existing infrastructure services, because the fabric nodes have capabilities to handle the DHCP relay functionality differences that are present in fabric deployments. In a typical DHCP relay design, the unique gateway IP address determines the subnet address assignment for an endpoint in addition to the location to which the DHCP server should direct the offered address. In a fabric overlay network, that gateway is not unique—the same Anycast IP address exists across all fabric edge nodes within the fabric site. Without special handling either at the fabric nodes or by the DHCP server itself, the DHCP offer returning from the server may not be relayed to the correct edge node where the DHCP request originated.

Figure 18. Fabric DHCP Packet Flow



Fabric DHCP Packet Flow

For simplicity, the DHCP Discover and Request packets are referred to as a **DHCP REQUEST**, and the DHCP Offer and Acknowledgement (ACK) are referred to as the **DHCP REPLY**.

- **Step 1**—Endpoint sends a **DHCP REQUEST** to the edge node.
- **Step 2**—The packet is inspected by DHCP Snooping.
- **Step 3a**—Option 82 data (DHCP Relay Agent Information) is inserted into the **DHCP REQUEST**.
- **Step 3b**—The Gateway IP address (giaddr) is set to the edge node's Anycast IPv4 address (example: 172.16.201.1).
- **Step 4**—Packet is encapsulated and sent to the border node where it is relayed to the DHCP server.
- **Step 5a**—DHCP server receives the **DHCP REQUEST** and offers an IP address within the applicable scope. The original Option 82 information is echoed back in the **DHCP REPLY**.
- **Step 5b**—DHCP server uses the Gateway IP address (giaddr) from **DHCP REQUEST** packet as the destination.
- **Step 6**—The **DHCP REPLY** sent back toward the border, as it also has the same Anycast IPv4 address assigned to a Loopback interface.
- **Step 7**—The **DHCP REPLY** is inspected, and the border node uses the option 82 information to determine the source RLOC (example: 192.168.255.2).
- **Step 8**—**DHCP REPLY** packet is encapsulated and sent back to the original source edge node.
- **Step 9**—Edge node receives the **DHCP REPLY**, de-encapsulates, and forwards to the endpoint which is identified via its MAC address.

To identify the specific DHCP relay source, Cisco DNA Center automates the configuration of the Relay Agent at the fabric edge with DHCP option 82. When a fabric edge node receives a DHCP Discovery message, it adds the DHCP Relay Agent Information using option 82 to the DHCP packet and forwards it across the overlay. Relay Agent Information is a standards-based (RFC 3046) DHCP option. It is a *container* option which contains two parts (two sub-options):

- **Agent Circuit ID**—Identifies the VLAN, the interface module, and interface port number.
- **Agent Remote ID**—Identifies the LISP Instance-ID (the VN), the IP Protocol (IPv4 or IPv6), and the source RLOC.

The relay agent sets the gateway address (*giaddr* field of the DHCP packet) as the IP address of the SVI the DHCP packet was received on. Once the DHCP option 82 information is inserted into the original packet, it is encapsulated in fabric VXLAN and forwarded across the overlay to the fabric border node who then forwards the packet to the DHCP server. The DHCP server, by referring to the relay agent IP address (giaddr) in a DHCP Discover message, allocates an address to the DHCP client from the address pool scope.

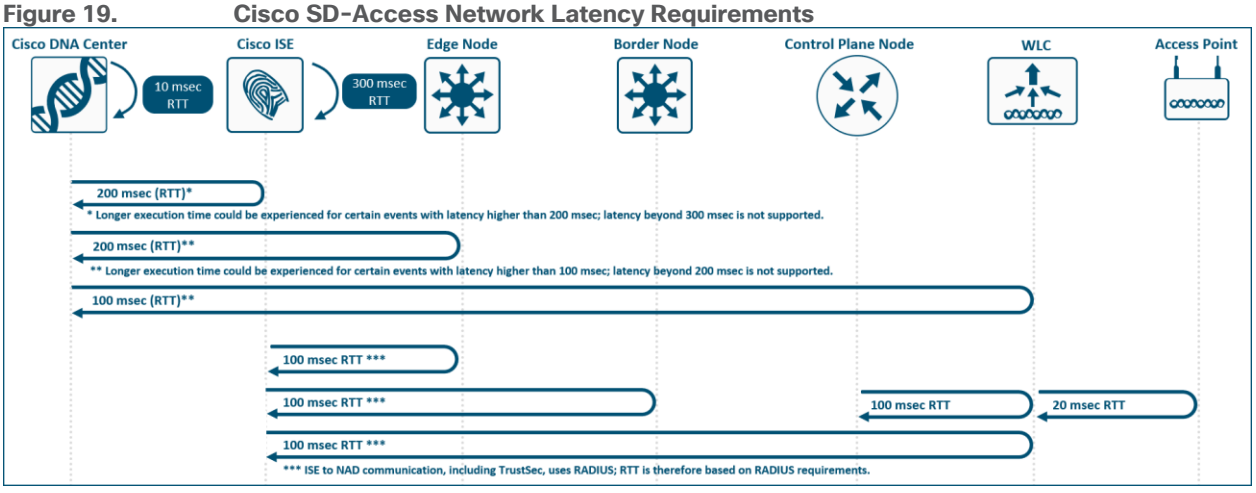
The border node has advanced DHCP relay capabilities which allows DHCP server configuration to remain unchanged for scopes covering fabric endpoints. Border nodes inspect the DHCP offer returning from the DHCP server. The offer includes the RLOC (edge node's loopback) from fabric edge switch which relayed the original DHCP request. The border node references the embedded option 82 information and directs the DHCP offer back to the correct fabric edge destination. This reply is encapsulated in Fabric VXLAN and sent across the overlay.

Tech tip
The DHCP server used in the deployment must conform the RFC standard and echo back the Option 82 information. Modern Microsoft Windows Servers such as 2012 R2 and beyond generally adhere to this standard. Other DHCP server providers such as Infoblox and BlueCat also adhered to this standard, though support may vary by release. Please check the applicable manufacture's release notes and user guides for the DHCP server in used in the deployment.

Latency

Latency in the network is an important consideration for performance, and the RTT between Cisco DNA Center and any network device it manages must be taken into strict account. The RTT should be equal to or less than 100 milliseconds to achieve optimal performance for all solutions provided by Cisco DNA Center including SD-Access. The maximum supported latency is 200ms RTT. Latency between 100ms and 200ms is supported, although longer execution times could be experienced for certain functions including Inventory Collection, Fabric Provisioning, SWIM, and other processes that involve interactions with the managed devices.

Figure 19. Cisco SD-Access Network Latency Requirements



Device Role Design Principles

This section is organized into the following subsections:

Section	Subsection
Device Role Design Principles	Edge Node Control Plane Node Border Node Fabric in a Box Extended Node Roles and Capabilities Cisco DNA Center High Availability

This section discusses design principles for specific SD-Access devices roles including edge nodes, control plane nodes, border nodes, Fabric in a Box, and extended nodes. This section concludes with device platform role and capabilities discussion and Cisco DNA Center High Availability design considerations.

Edge Node Design

In SD-Access, fabric edge nodes represent the access layer in a two or three-tier hierarchy. The access layer is the edge of the campus. It is the place where end devices attach to the wired portion of the campus network. The edge nodes also represent the place where devices that extend the network connectivity out one more layer connect. These include devices such as IP phones, access points, and extended nodes.

The access layer provides the intelligent demarcation between the network infrastructure and the devices that leverage that infrastructure. As such it provides a trust boundary for QoS, security, and policy. It is the first layer of defense in the network security architecture, and the first point of negotiation between end devices and the network infrastructure.

To meet network application and end-user demands, Cisco Catalyst switching platforms operating as a fabric edge node do not simply switch packets but provide intelligent services to various types of endpoints at the network edge. By building intelligence into these access layer switches, it allows them to operate more efficiently, optimally, and securely.

The edge node design is intended to address the network scalability and availability for the IT-managed voice, video, and wireless communication devices along with the wide variety of possible wired endpoint device types. Edge nodes should maintain a maximum 20:1 oversubscription ratio to the distribution or collapsed core layers. The higher the oversubscription ratio, the higher the probability that temporary or transient congestion of the uplink may occur if multiple devices transmit or receive simultaneously. Uplinks should be minimum of 10 Gigabit Ethernet and should be connected to multiple upstream peers.

As new devices are deployed with higher power requirements, such as lighting, surveillance cameras, virtual desktop terminals, remote access switches, and APs, the design should have the ability to support power over Ethernet to at least 60W per port, offered with Cisco Universal Power Over Ethernet (UPOE), and the access layer should also provide PoE perpetual power during switch upgrade and reboot events. New endpoints and building systems may require even more power, and IEEE 802.3bt and Cisco UPOE-Plus (UPOE+) can provide power up to 90W per port. Both fixed configuration and modular switches will need multiple power supplies to support 60–90W of power across all PoE-capable ports.

Control Plane Node Design

The fabric control plane node contains the database used to identify an endpoint's location in the network. This is a central and critical function for the fabric to operate. A control plane node that is overloaded and slow to

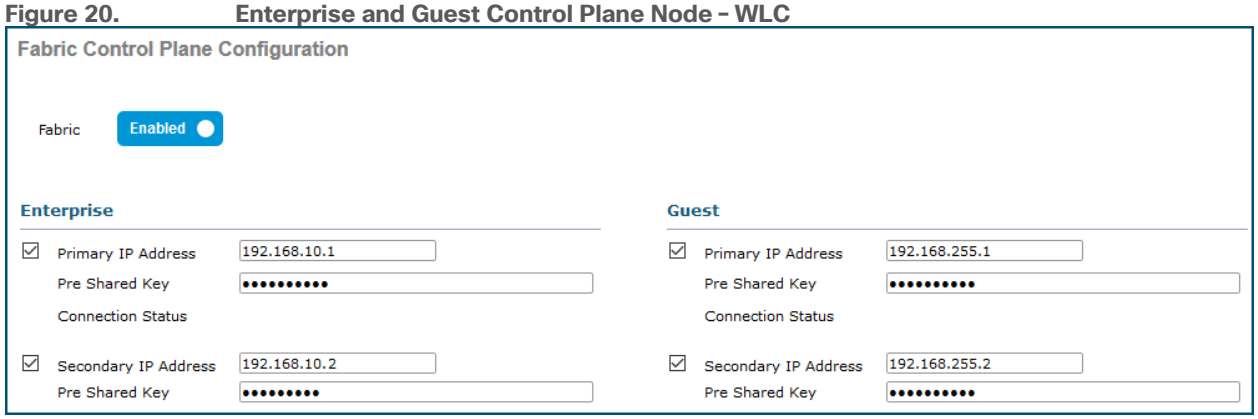
respond results in application traffic loss on initial packets. If the fabric control plane is down, endpoints inside the fabric fail to establish communication to remote endpoints that are not cached in the local database.

For redundancy, it is recommended to deploy two control plane nodes to ensure high availability of the fabric site, as each node contains a copy of control plane information acting in an Active/Active state. The devices supporting the control plane should be chosen to support the HTDB (EID-to-RLOC bindings), CPU, and memory needs for an organization based on the number of endpoints. Border nodes and edge nodes register with and use all control plane nodes, so redundant nodes chosen should be of the same type for consistent performance.

Cisco AireOS and Catalyst WLCs can communicate with a total of four control plane nodes in a site: two control plane nodes are dedicated to the guest and the other two for non-guest (enterprise) traffic. If the dedicated Guest Border/Control plane node feature (discussed later in the guide) is not used, fabric WLCs can only communicate with two control plane nodes per fabric site.

In Figure 20, the WLC is configured to communicate with two control plane nodes for Enterprise (***192.168.10.1*** and ***192.168.10.2***) and two control plane nodes for Guest (***192.168.255.1*** and ***192.168.255.2***).

Figure 20. Enterprise and Guest Control Plane Node – WLC



Distributed Control Plane Node and Border Node

A fabric control plane node operates similarly to a BGP Route Reflector (RFC 4456). The control plane node advertises the fabric site prefixes learned from the LISP protocol to certain fabric peers, I.e. the border nodes. Like route reflector (RR) designs, control plane nodes provide operational simplicity, easy transitions during change windows, and resiliency when deployed in pairs. When the control plane nodes are deployed as dedicated devices, not colocated with other fabric roles, they provide the highest degrees of performance, reliability, and availability. This method also retains an original goal of a Software-Defined Network (SDN) which is to separate the control function from the forwarding functions.

Control plane nodes may be deployed as either dedicated (distributed) or non-dedicated (colocated) devices from the fabric border nodes. In a Fabric in a Box deployment, fabric roles must be colocated on the same device. In Small and Very Small deployment, as discussed in the Reference Models section below, it is not uncommon to deploy a colocated control plane node solution, utilizing the border node and control plane node on the same device. Deploying a dedicated control plane node has advantages in Medium and Large deployments as it can provide improved network stability both during fabric site change management and in the event that a fabric device becomes unavailable in the network.

Dedicated control plane nodes, or *off-path control plane nodes,* which are not in the data forwarding path, can be conceptualized using the similar *DNS Server* model. The control plane node is used for LISP control plane queries, although it is not in the direct data forwarding path between devices. The physical design result is similar to a *Router on a Stick* topology.

The dedicated control plane node should have ample available memory to store all the registered prefixes. Bandwidth is a key factor for communication prefixes to the border node, although throughput is not as key since the control plane nodes are not in the forwarding path. If the dedicated control plane node is in the data forwarding path, such as at the distribution layer of a three-tier hierarchy, throughput should be considered along with ensuring the node is capable of CPU-intensive registrations along with the other services and connectivity it is providing.

Tech tip
To achieve optimal performance in a fabric role, routing platforms should have a minimum of 8 GB DRAM.

One other consideration for separating control plane functionality onto dedicated devices is to support frequent roaming of endpoints across fabric edge nodes. Roaming across fabric edge nodes causes control plane events in which the WLC updates the control plane nodes on the mobility (EID-to-RLOC mapping) of these roamed endpoints. Although colocated control plane is the simplest design, adding the control plane node function on border nodes in a high-frequency roam environments can lead to high CPU on colocated devices. For high-frequency roam environments, a dedicated control plane node should be used.

CSR 1000v as Control Plane Node

The dedicated control plane node can be deployed completely *out of band* (*off-path*) through virtualization. Virtualization technologies have been widely used in enterprise data centers as a reliable technology that can be extended and deployed onto critical and highly available network infrastructure. A virtualized control plane node also follows the NFV (Network Function Virtualization) concepts of Software-Defined Networking (SDN) which calls for separating network functions from specialized hardware through virtualization.

The Cisco Cloud Services Router (CSR) 1000V Series, is an excellent solution for the dedicated off-path control plane node application. In order to meet the intensive CPU and memory demand to handle large site scale, CPU and memory resources can easily be carved out and provisioned according to the requirements. A virtual control plane node also positions the device within the highly-available data center while allowing logical placement at those locations deemed most useful for the fabric site architecture. The CSR 1000v is supported as both a site-local control plane node and a transit control plane node.

Colocated Control Plane Node and Border Node

If the chosen border nodes support the anticipated endpoint, throughput, and scale requirements for a fabric site, then the fabric control plane functionality can be colocated with the border node functionality. While it does provide operational simplicity in that it is two less pieces of equipment to manage, it also reduces the potential for resiliency in the event of software upgrade, device reboots, common upgrades, or updates to configuration.

Border nodes connecting to external resources such as the Internet should always be deployed in pairs to avoid single failure points. As discussed in the next section, border nodes may be used to connect to internal resources such as the data center or used as a migration strategy with the Layer 2 handoff functionality. When considering colocating the control plane node and border node, understand that the lowest common denominator is the Fabric WLCs which can only communicate with two control plane nodes per fabric site. If a fabric site is deployed with external border nodes, internal border nodes, and border nodes with Layer 2 handoff, it is not possible to colocate the control plane node and border node function on all devices deployed as a border. Distributing the border and control plane node will alleviate this and will provide role consistency across the devices deployed as a border node.

Border Node Design

A border node is an entry and exit point to the fabric site. In effect, it speaks two languages: SD-Access fabric on one link and traditional routing and switching on another. The fabric border design is dependent on how the fabric site is connected to networks outside of the fabric site.

Border node functionality is supported on both routing and switching platforms. The correct platform should be selected for the desired outcome. Both routing and switching platform support 1-, 10-, 40-, and 100-Gigabit Ethernet ports. Switching platforms generally have a higher port density than routing platforms and support 25-Gigabit Ethernet (25GBASE / SFP28). Routing platforms generally have a higher performance and scaling numbers for SGT and control plane node related functions, allow for a higher number of BGP peerings, and support advanced WAN technologies such as IPSec. Routing platforms are also supported for SD-WAN infrastructure.

If Cisco DNA Center Assurance is used in the deployment, switching platforms can be used to show *quantitative* application health. Quantitative metrics show how much application traffic is on the network. Routing platforms can be used to show *quantitative* and *qualitative* application health. These metrics go beyond simply showing the amount of application of traffic on the network by displaying how the traffic is being serviced using latency and loss information.

A border node does not have a direct mapping to a layer in the network hierarchy. However, the border node is not necessarily a distribution layer switch or core switch in the network. Border nodes may also be a routing infrastructure, WAN edge, or other network edge devices.

Tech tip
For supported Wide-Area technologies when the border node is a WAN edge router, please see the End-to-End Macro Segmentation section. Border nodes cannot be the termination point for an MPLS circuit.

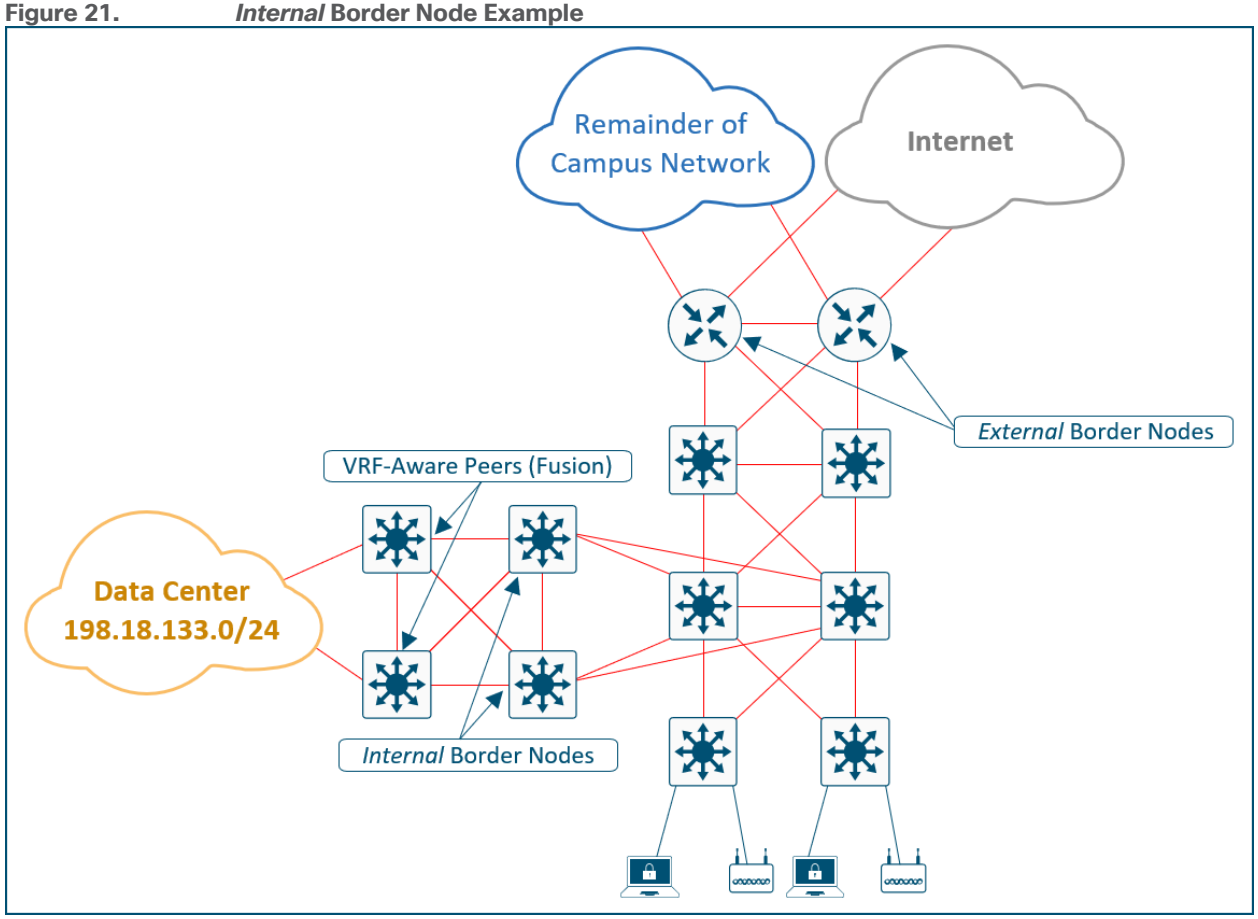
Border Nodes and External Networks

When provisioning a border node in Cisco DNA Center, there are three different options to indicate the type of external network(s) to which the device is connected. Older collateral and previous UI refer to these as *Internal*, *External*, and *Anywhere*. While this nomenclature is no longer used in user interface, these names can still be helpful in describing the external network to the border nodes and designing the fabric for that network connection.

A border may be connected to *internal*, or known, networks such as data center, shared services, and private WAN. Routes that are learned from the data center domain are registered with the control plane node, similarly to how an edge node registers an endpoint. In this way, LISP, rather than native routing, is used to direct traffic to these destinations outside of the fabric.

In Figure 21 below, there are two sets of border nodes. The external border nodes connect to the Internet and to the rest of the Campus network. The internal border nodes connect to the Data Center by way of VRF-Aware peers (fusion devices). If traditional, default forwarding logic is used to reach the Data Center prefixes, the fabric edge nodes would send the traffic to the external border nodes who would then hairpin the traffic to the internal border nodes resulting in an inefficient traffic forwarding. By importing, or registering, the Data Center prefixes with the control plane node using the internal border functionality, edge nodes can send traffic destined for ***198.18.133.0/24*** directly to the internal border nodes. Traffic destined for the Internet and remainder of the campus network to the *external* border nodes.

Figure 21. ***Internal*** **Border Node Example**

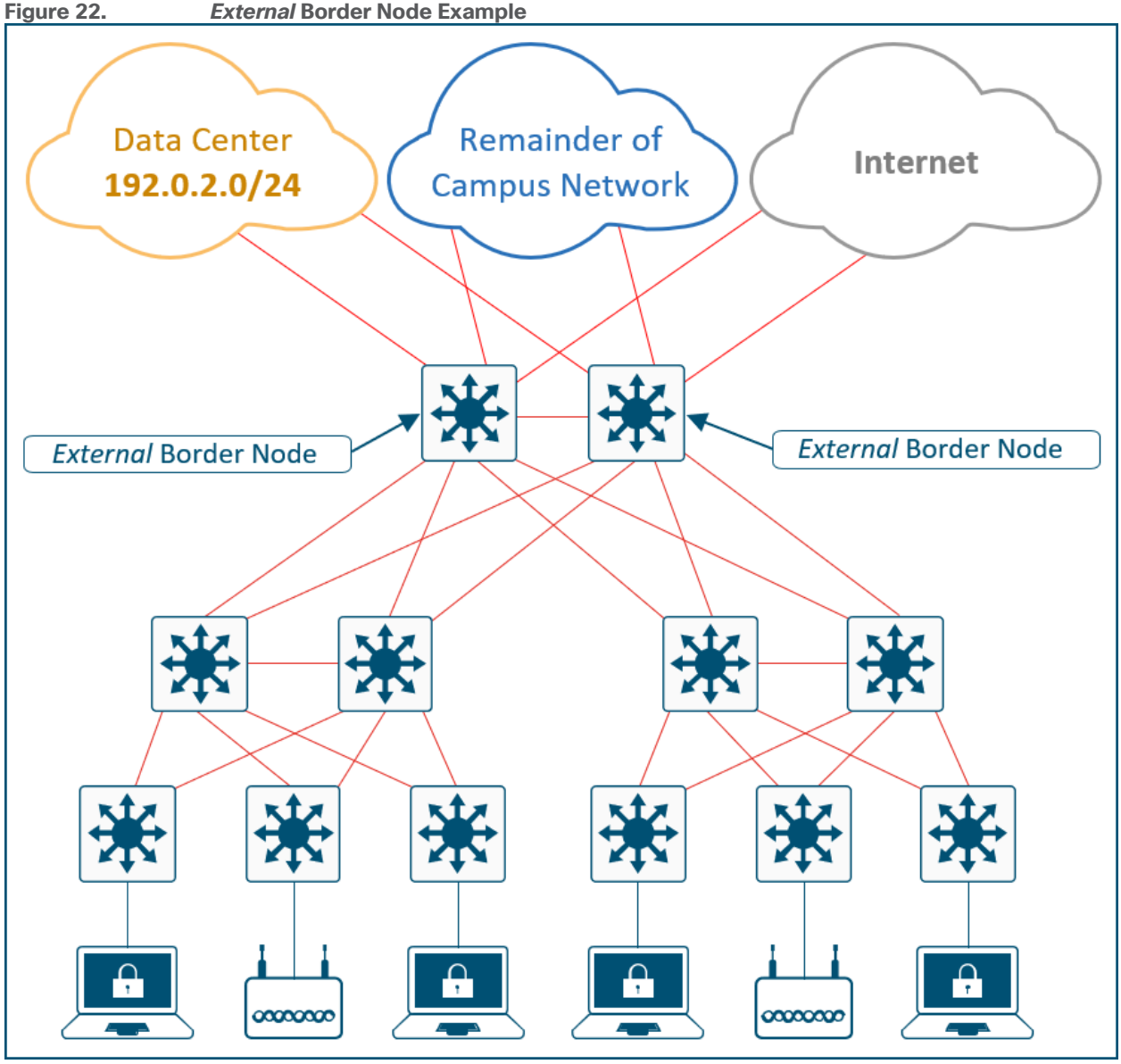


A border may be connected to *external*, or unknown, networks such as Internet, WAN, or MAN. The routes learned from the external domain are not registered (imported) to the control plane node. This border is the default exit point, or gateway of last resort, for the virtual networks in the fabric site.

The control plane node has a mechanism that notifies the fabric devices that a destination prefix is not registered with it. This triggers the device requesting this mapping to simply send traffic to the *external* border node. The majority of SD-Access deployments should provision border nodes as *external* which provisions the device as the *fabric site gateway of last resort*. This deployment type uses default routing (traditional forwarding logic), rather than LISP, to reach all external prefixes.

In Figure 22 below, there are a single pair of borders nodes that represent the common egress point from the fabric site. The border nodes are connected to the Data Center, to the remainder of the campus network, and to the Internet. When the edge nodes forward traffic to any of these external destinations, the same border nodes will be used. Traditional, default forwarding logic can be used to reach these prefixes, and it is not necessary to register the Data Center prefixes with the control plane node.

Figure 22. *External* Border Node Example



A border node may also be connected to both *known* and *unknown* networks such as being a common egress point for the rest of an enterprise network along with the Internet. What distinguishes this border is that ***known*** routes such as shared services and data center, are registered with the control plane node rather than using the default forwarding logic described above. This type of border node is sometimes referred to as an *Anywhere* border node.

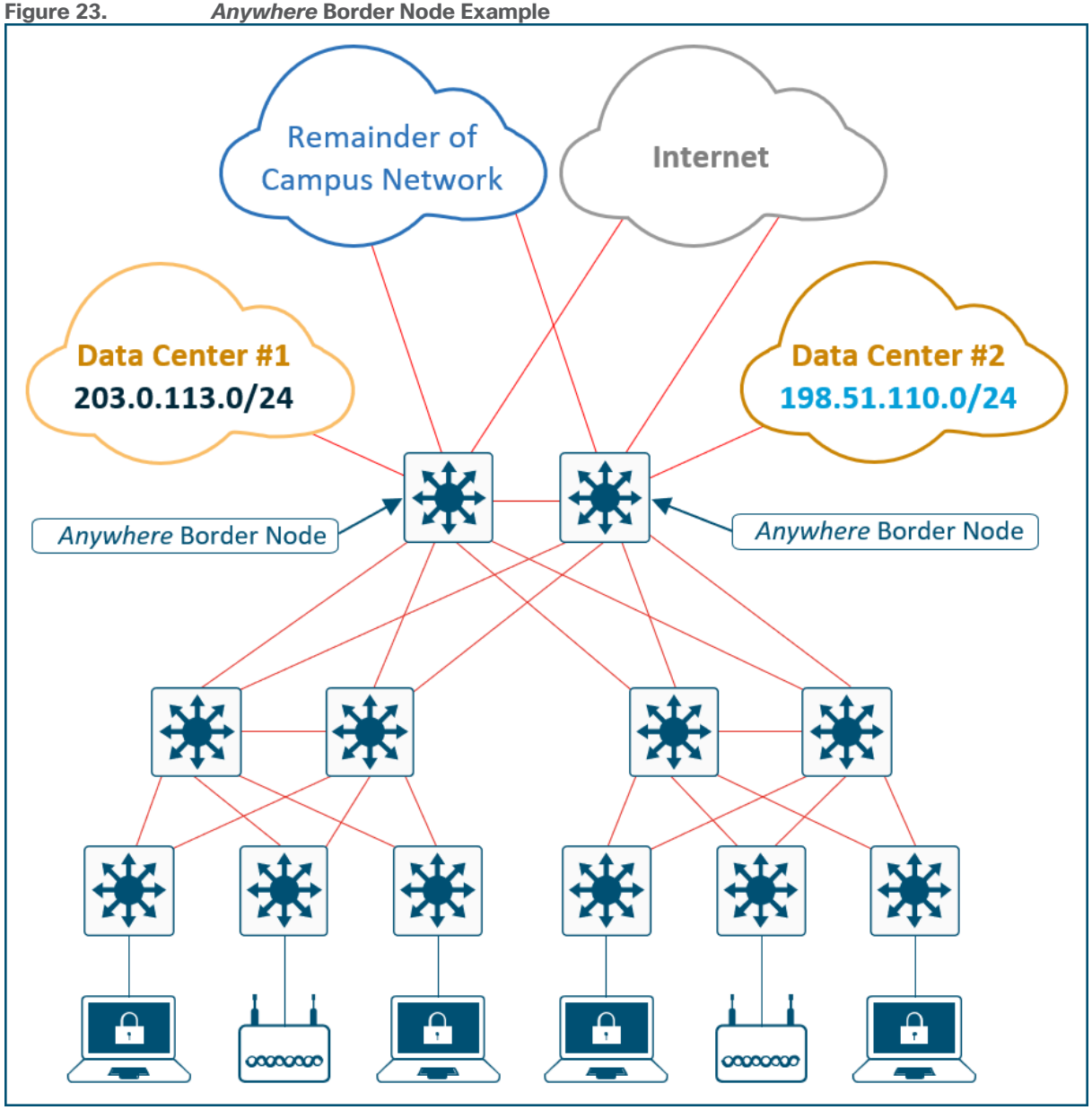
Tech tip
SD-Access for Distributed Campus deployments are the most common use case for a border than connects to both known and unknown routes (*Anywhere*) and also needs to register these known routes with the control plane node. For more information on border node provisioning options and Distributed Campus deployments, please see: Software-Defined Access for Distributed Campus Deployment Guide. Further design considerations for Distributed Campus deployments are discussed below.

The key distinction between these border types is the underlying routing logic that is used to reach ***known*** prefixes. Networks deployed similarly to Figure 8 – SD-Access Fabric Roles (Example) do not commonly import (register) routes with the control plane node. Because there is a common egress point to the fabric site, the

border nodes are the destination for both known and unknown external routes. Registering the known external prefixes in this type of design is not needed, as the same forwarding result is achieved for both known and unknown prefixes. Most deployments should provision a border node using the *external* border node type.

In Figure 23 below, both border nodes are connected to the Internet and to the remainder of the campus network. Each border node is also connected to a separate Data Center with different prefixes. If traditional, default forwarding logic is used to reach these prefixes, the fabric edge nodes may send the traffic to a border not directly connect to the applicable data center. Traffic will have to inefficiently traverse the crosslink between border nodes. By importing the data center prefixes into LISP, the edge nodes can send to the traffic to the border node on the left to reach ***203.0.113.0/24*** and the border node on the right to reach ***198.51.100.0/24***. Either border can be used as the default path to the Internet.

Figure 23. ***Anywhere*** **Border Node Example**



Tech tip
For further descriptions and discussions regarding how the Cisco DNA Center UI represents these three border node types,

please see Guide to SD-Access Border Node Roles on Cisco DNA Center ≥1.3.x on Cisco Community.

A border node may also connect to a traditional Layer 2 switched access network. This allows the same IP subnet to exist in both the traditional network and SD-Access network with the border node performing the translation between these two networks and allowing them to communicate. This feature is called the Layer 2 border handoff and is discussed in depth in later sections.

Because this border node is connected to the traditional network, it is subject to broadcast storms, Layer 2 loops, and spanning-tree problems that can occur in Layer 2 switched access networks. To prevent disruption of control plane node services or border node services connecting to other external or external networks, a border node should be dedicated to the Layer 2 handoff feature and not colocated with other fabric roles or services.

Fabric in a Box Design

Some physical locations may use unique wiring plans such that the MDF and IDF do not conform to the common two-tier and three-tier hierarchical network structure. The result is that the available fiber and copper wiring may require access switches to be daisy-chained or configured in a ring. Any number of wiring variations may exist in a deployment. Due to the unique nature of supporting all three fabric roles on a node, Fabric in a Box has specific topologies that are supported if additional fabric edge nodes or extended nodes are connected to it (downstream from it). The topologies supported differ based on if SD-Access Embedded wireless (now a fourth fabric role on the device) is also implemented.

Like other devices operating as edge node, extended nodes and access points can be directly connected to the Fabric in a Box. In locations where physical stacking is not possible due to the wiring structure, Fabric in a Box can support up to two daisy-chained edge nodes creating a three-tier topology. In this daisy-chained topology, access points and extended nodes can be connected to any of the devices operating in the edge node role, including the Fabric in a Box itself. Embedded wireless is also supported in this scenario. Dual Fabric in a Box is also supported, though should only be used if mandated by the existing wiring structures. When Fabric in a Box is deployed on a Stackwise Virtual pair, an external WLC should be utilized.

Tech tip

Fabric in a Box is supported using a single switch, a switch with hardware stacking, or with StackWise Virtual deployment. Support for StackWise Virtual in fabric role was first introduced in Cisco DNA Center 1.3.3.x for the Catalyst 9500 Series Switches. For specific platforms supported with StackWise Virtual in SD-Access networks, please see the Cisco DNA Center Release Notes.

Extended Node Design

Extended nodes are connected to a single Fabric Edge switch through an 802.1Q trunk port. This trunk port is deployed as an EtherChannel with one or more links aggregated to the upstream fabric edge. Cisco DNA Center automates both the trunk and the creation of the port-channel. Once onboarded through the workflow, switch ports on the extended node support the same dynamic methods of port assignments as an edge node in order to provide macro-segmentation for connected endpoints.

Cisco DNA Center has two different support options for extended nodes: classic extended nodes and policy extended nodes.

Classic Extended Nodes

SD-Access Extended Nodes capabilities are supported on the Cisco Catalyst IE-3300, Catalyst IE-3400, Catalyst IE-3400H, IE-4000 Series, IE-5000, Catalyst Digital Building, and Catalyst 3560-CX Compact Series

switches. In current versions of Cisco DNA Center, Extended Nodes support AAA configuration on their host-connected ports which allows endpoints to be authenticated and authorized with ISE. Dynamic VLAN assignment places the endpoints into specific VLANs based on the credentials supplied by the user. This VLAN is being forwarded for a VRF instance on the upstream edge node creating the first layer of segmentation. SGT assignment, the second layer of segmentation, is provided within Cisco DNA Center through VLAN to SGT mappings.

When deploying extended nodes, consideration should be taken for east-west traffic in the same VLAN on a given extended node. This east-west traffic is forwarded using traditional Layer-2 forwarding logic. Inter-VLAN traffic is attracted to the edge node because the AnyCast gateway for the end hosts resides there. When a host connected to extended node sends traffic to destinations in the same VN connected to or through other fabric edge nodes, segmentation and policy is enforced through VLAN to SGT mappings on the fabric edge node.

Tech tip
For enhanced security and segmentation scalability, consider using the Policy Extended Node because scalable group enforcement can be executed at the ingress point in the network.

Policy Extended Nodes

Additional enhancements are available to devices operating as Policy Extended Nodes. This capability is supported on the Cisco Catalyst IE-3400 and IE-3400H Series Switches. In addition to the operation and management provide by a classic extended node, policy extended nodes directly support SGTs. This provides direct east-west traffic enforcement on the extended node.

Both VLAN and SGT assignment can be received dynamically as a result of the endpoint authentication and authorization process. This allows traffic between sources in the same VLAN and in different VLANs to be enforced on the policy extended node itself. Segmentation to other sources in the fabric are provided through inline tagging on the 802.1Q trunk connected to the upstream fabric edge node.

Design Considerations

Extended nodes and Policy Extended Nodes can only be connected to a single fabric edge switch. They should not be dual-homed to different upstream edge nodes. Daisy chaining is not supported by the zero-touch Plug and Play process used to onboard these switches. The devices must have the appropriate interface type and quantity to support connectivity to both the upstream fabric edge node and the downstream endpoints.

Access points and other Power over Ethernet (PoE) devices can be connected directly to both variants of extended node switches. When connecting PoE devices, ensure that there is enough available PoE power available. This is especially true with Industrial Ethernet Series switches which have significant variety of differing powering options for both AC and DC circuits.

SGT value 8000 is leveraged on the ports between the policy extended node and the edge node. It should not be reused elsewhere in the deployment. Because these ports use inline tagging, this scalable group identifier is used to build the trust between the two peer devices on both ends of the link.

Tech tip
For additional details the behavior of inline tagging described above, please see the Overview of TrustSec Guide, Configuring Native SGT Propagation (Tagging) section.

Platform Roles and Capabilities Considerations

The SD-Access network platform should be chosen based on the capacity and capabilities required by the network, considering the recommended functional roles. Roles tested during the development of this guide are noted in the companion deployment guides at Cisco Design Zone for Campus Wired and Wireless LAN.

Refer to the SD-Access Hardware and Software Compatibility Matrix for the most up-to-date details about which platforms and software are supported for each version of Cisco SD-Access. The physical network design requirements drive the platform selection. Platform capabilities to consider in an SD-Access deployment:

- A wide range of Cisco Catalyst 9000, Catalyst 3850, and Catalyst 3650 Series switches are supported; however, only certain devices are supported for the edge node, border node, and control plane node roles.
- Additional devices such as the Cisco Catalyst 4500, 6500, and 6800 Series and Cisco Nexus 7700 Series are also supported, but there may be specific supervisor module, line card module, and fabric-facing interface requirements.

  Additionally, the roles and features support may be reduced. For example, Catalyst 6000 series switches are not supported as border nodes connected to SD-Access transits and do not support SD-Access Embedded Wireless.
- A variety of routing platforms are supported as control plane nodes and border nodes, such as the Cisco ISR 4400 and 4300 Series Integrated Services routers, Cisco ASR 1000-X and 1000-HX Series Aggregation Services Routers. The Cisco Cloud Services Router 1000V Series is also supported, but only as a control plane node.
- Cisco Catalyst 9800 Series, Aironet 8540, 5520, and 3504 Series Wireless LAN Controllers are supported as Fabric WLCs. Similarly, the Cisco Catalyst 9100 and Cisco Aironet Wave 2 and Wave 1 APs are supported as fabric-mode access points.
- Cisco ISE must be deployed with a version compatible with Cisco DNA Center. ISE can be deployed virtually or on a Cisco SNS (Secure Network Server) appliance.

Tech tip
The Nexus 7700 Series switch is only supported as an *external* border. It does not support colocating the control plane node functionality. It is not supported as a border node connected to SD Access Transit for Distributed Campus deployments nor does it support the Layer 2 handoff functionality and Layer 2 flooding features. It does not support SD-Access embedded wireless. Greenfield deployments should consider Catalyst 9000 Series switches rather than the N7700 Series switch for use in the fabric.

Cisco DNA Center High Availability

Cisco DNA Center is supported in single-node and three-node clusters. Scaling does not change based on the number of nodes in a cluster; three-node clusters simply provide high availability (HA). If the Cisco DNA Center node is deployed as a single-node cluster, wiring, IP addresses, and connectivity should be planned and configured with future three-node clustering in mind.

In a single-node cluster, if the Cisco DNA Center appliance becomes unavailable, an SD-Access network provisioned by the node still functions. However, automated provisioning capabilities and Assurance insights are lost until the single node availability is restored.

For high-availability purposes, a three-node cluster can be formed by using appliances with the same core count. This includes the ability to cluster a first-generation 44-core appliance with a second-generation 44-

core appliance. A three-node Cisco DNA Center cluster operates as a single logical unit with a GUI accessed using a virtual IP, which is serviced by the resilient nodes within the cluster.

Within a three-node cluster, service distribution provides distributed processing, database replication, security replication, and file synchronization. Software upgrades are automatically replicated across the nodes in a three-node cluster. A three-node cluster will survive the loss of a single node, though requires at least two nodes to remain operational. Some maintenance operations, such as software upgrades and file restoration from backup, are restricted until the three-node cluster is fully restored. Additionally, not all Assurance data may be protected while in the degraded two-node state.

For Assurance communication and provisioning efficiency, a Cisco DNA Center cluster should be installed in close network proximity to the greatest number of devices being managed to minimize communication delay to the devices. Additional latency information is discussed in the Latency section.

Tech tip
For physical topology options and failover scenarios for a three-node cluster, please see Cisco DNA Center 3-Node Cluster High Availability Scenarios technote.

Feature-Specific Designs

This section is organized into the following subsections:

Section	Subsection
Feature-Specific Design Requirements	Multicast Layer 2 Flooding Critical VLAN LAN Automation

The following section discusses design consideration for specific features in SD-Access. It begins with a discussion on multicast design, traditional multicast operations, and Rendezvous Point design and placement. Multicast forwarding in the fabric is discussed along with considerations regarding the Layer 2 flooding feature which relies on a multicast transport in the underlay. Next, Critical VLAN is described along with considerations for how it is deployed in SD-Access. This section ends with LAN Automation, its use-case, general network topology design to support the feature, and considerations when the LAN Automation network is integrated into the remainder of the routing domain.

Fabric Multicast Overview

Multicast is supported both in the overlay virtual networks and the in the physical underlay networks in SD-Access, with each achieving different purposes as discussed further below.

The multicast source can either be outside the fabric site (commonly in the data center) or can be in the fabric overlay, directly connected to an edge node, extended node, or associated with a fabric AP. Multicast receivers are commonly directly connected to edge nodes or extended nodes, although can also be outside of the fabric site if the source is in the overlay.

PIM Any-Source Multicast (PIM-ASM) and PIM Source-Specific Multicast (PIM-SSM) are supported in both the overlay and underlay. The overlay or the underlay can be used as the transport for multicast as described in the Forwarding section.

Rendezvous Point Design

In PIM-ASM routing architecture, the multicast distribution tree is rooted at the Rendezvous Point (RP). This is referred to as shared tree or RP-Tree (RPT), as the RP acts as the meeting point for sources and receivers of multicast data. The advantage of using RPs is that multicast receivers do not need to know about every possible source, in advance, for every multicast group. Only the address of the RP, along with enabling PIM, is needed to begin receiving multicast streams from active sources.

A Rendezvous Point is a router (a Layer-3 device) in a multicast network that acts as a shared root for the multicast tree. Rendezvous Points can be configured to cover different multicast groups, or with regards to SD-Access, cover different virtual networks. Active multicast sources are registered with an RP, and network devices with interested multicast receivers will join the multicast distribution tree at the Rendezvous Point.

An RP can be active for multiple multicast groups, or multiple RPs can be deployed to each cover individual groups. The information on which RP is handling which group must be known by all the routers in the multicast domain. For this group-to-RP-mapping to occur, multicast infrastructure devices must be able to locate the Rendezvous Point in the network. In traditional multicast networks, this can be accomplished through static RPs, BSR (Boot Strap Router), Auto-RP, or Anycast-RP.

Anycast-RP allows two or more RPs to share the load for multicast source registration and act as hot-standbys for each other. With multiple, independent RPs in the network, a multicast source may register with one RP and a receiver may register with another, as registration is done with the closest RP (in terms of the IGP metric). Anycast-RP uses MSDP (Multicast Source Discovery Protocol) to exchange source-active (SA) information between redundant RPs. This allows the sources to be known to all the Rendezvous Points, independent of which one received the multicast source registration. Anycast-RP is the preferred method in SD-Access, and the method used during the PIM-ASM automation workflows.

When PIM-ASM is used in the overlay and multiple RPs are defined within the fabric site, Cisco DNA Center automates the MSDP configuration on the RPs and configures the other fabric nodes within a given fabric site to point to these RPs for a given virtual network.

Rendezvous Point Placement

Where an RP is placed in a network does not have to be a complex decision. To aid in this decision process, it can be helpful to compare PIM-ASM and PIM-SSM and understand the multicast tree building.

Protocol independent multicast (PIM) is used to build a path backwards from the receiver to the source, effectively building a tree. This tree has a root with branches leading out to the interested subscribers for a given stream. With PIM-ASM, the root of the tree is the Rendezvous Point. With PIM-SSM, the root of the multicast tree is the source itself.

Source tree models (PIM-SSM) have the advantage of creating the optimal path between the source and the receiver without the need to meet a centralized point (the RP). In a shared tree model (PIM-ASM), the path through the RP may not be the shortest path from receiver back to source. However, PIM-ASM does have an automatic method called *switchover* to help with this. *Switchover* moves from the shared tree, which has a path to the source by way of the rendezvous point, to a source tree, which has a path directly to the source. This capability provides an automatic path optimization capability for applications that use PIM-ASM.

In an environment with fixed multicast sources, RPs can easily be placed to provide the shortest-path tree. In environments with dynamic multicast sources, RPs are commonly placed in the core of a network. In traditional networking, network cores are designed to interconnect all modules of the network together, providing IP reachability, and generally have the resources, capabilities, and scale to support being deployed as a Rendezvous Point.

In SD-Access networks, border nodes act as convergence points between the fabric and non-fabric networks. Border nodes are effectively *the core* of the SD-Access network. Discussed above, border node device selection is based on the resources, scale, and capability to support being this aggregation point between fabric and non-fabric.

Multicast sources are commonly located outside the fabric site–such as with Music on Hold (MOH), streaming video/video conferencing, and live audio paging and alert notifications. For unicast and multicast traffic, the border nodes must be traversed to reach destinations outside of the fabric. The border nodes already represent the shortest path.

Most environments can achieve the balance between optimal RP placement along with having a device with appropriate resources and scale by selecting their border node as the location for their multicast Rendezvous Point.

External RP

The Rendezvous Point does not have to be deployed on a device within the fabric site. External devices can be designated as RPs for the multicast tree in a fabric site. Up to two external RPs can be defined per VN in a fabric site. The External RP address must be reachable in the VN routing table on the border nodes. External RP placement allows existing RPs in the network to be used with the fabric. In this way multicast can be enabled without the need for new MSDP connections. If RPs already exist in the network, using these external RPs is the preferred method to enable multicast.

Multicast Forwarding in SD-Access

SD-Access supports two different transport methods for forwarding multicast. One uses the overlay and is referred to as *head-end replication,* and the other uses the underlay and is called *Native Multicast*. Multicast forwarding is enabled per-VN. However, if native-multicast is enabled, for a VN, head-end replication cannot be used for another VN in the fabric site. These two options are mutually exclusive within the fabric site.

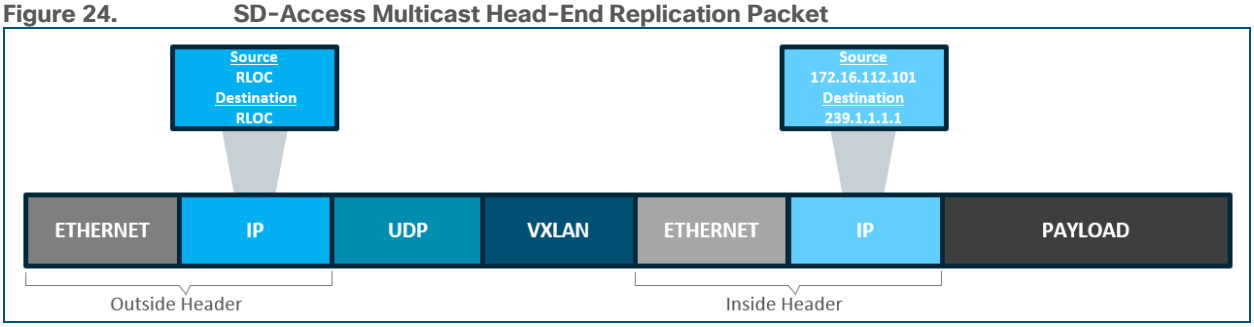
Head-End Replication

Head-end replication (or ingress replication) is performed either by the multicast first-hop router (FHR), when the multicast source is in the fabric overlay, or by the border nodes, when the source is outside of the fabric site.

Head-end replication in fabric operates similarly to Multicast-Unicast mode on a Wireless LAN Controller. The multicast packets from the source are replicated and sent, via unicast, by the FHR to all last-hop routers (LHR) with interested subscribers.

For example, consider a fabric site that has twenty-six (26) edge nodes. Each edge node has receivers for a given multicast group, and the multicast source is connected to one of the edge nodes. The FHR edge node must replicate each multicast packet to all other twenty-five edge nodes. This replication is performed per source, and packets are sent across the overlay. A second source means another twenty-five unicast replications. If the multicast source is outside of the fabric site, the border node acts as the FHR for the fabric site and performs the head-end replication to all fabric devices with interested multicast subscribers.

Figure 24. SD-Access Multicast Head-End Replication Packet



The advantage of head-end replication is that it does not require multicast in the underlay network. This creates a complete decoupling of the virtual and physical networks from a multicast perspective. However, this can create high overhead on the FHRs and result in high bandwidth and CPU utilization. In deployments where multicast cannot be enabled in the underlay networks, head-end replication can be used. Networks should consider Native Multicast due to its efficiency and the reduction of load on the FHR fabric node.

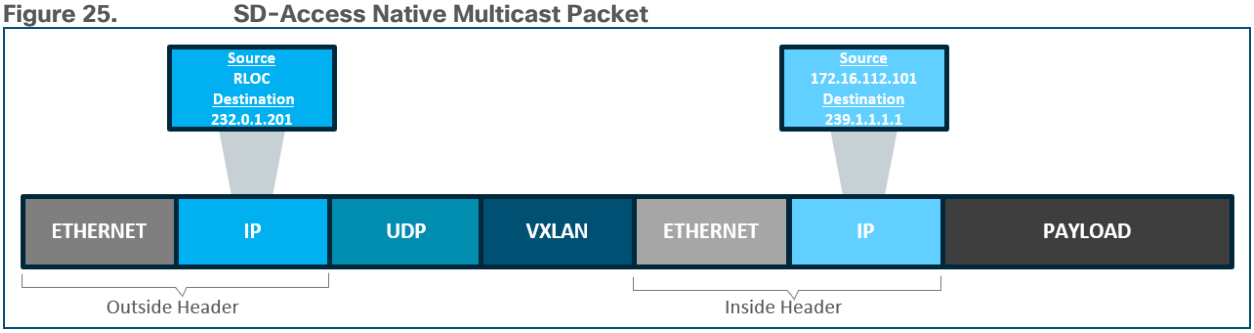
Native Multicast

Native multicast does not require the ingress fabric node to do unicast replication. Rather the whole underlay, including intermediate nodes (nodes not operating in a fabric role) are used to do the replication. To support native multicast, the FHRs, LHRs, and all network infrastructure between them must be enabled for multicast.

Native multicast uses PIM-SSM for the underlay multicast transport. The overlay multicast messages are tunneled inside underlay multicast messages. This behavior also allows overlap in the overlay and underlay multicast groups in the network, if needed. Because the entire underlay network between source and receiver is working to do the packet replication, scale and performance is vastly improved over head-end replication.

Native multicast works by performing multicast-in-multicast encapsulation. Multicast packets from the overlay are encapsulated in multicast in the underlay. With this behavior, both PIM-SSM and PIM-ASM can be used in the overlay.

Figure 25. SD-Access Native Multicast Packet



Layer 2 Flooding

Layer 2 flooding is feature that enables the flooding of broadcast, link-local multicast, and ARP traffic for a given overlay subnet. In traditional networking, broadcasts are flooded out of all ports in the same VLAN. By default, SD-Access transports frames without flooding Layer 2 broadcast and unknown unicast traffic, and other methods are used to address ARP requirements and ensure standard IP communication gets from one endpoint to another.

However, some networks need to utilize broadcast, particularly to support silent hosts which generally require reception of an ARP broadcast to come out of silence. This is commonly seen in some building management systems (BMS) that have endpoints that need to be able to ARP for one other and receive a direct response at

Layer 2. Another common use case for broadcast frames is Wake on LAN (WoL) Ethernet broadcasts which occur when the source and destination are in the same subnet.

Because the default behavior, suppression of broadcast, allows for the use of larger IP address pools, pool size of the overlay subnet needs careful consideration when Layer 2 flooding is enabled. Consider using a /24 (24-bit netmask) or smaller address pool to limit the number of broadcasts, as each of these frames must be processed by every device in the segment. Layer 2 flooding should be used selectively, where needed, using small address pool, and it is not enabled by default.

Layer 2 flooding works by mapping the overlay subnet to a dedicated multicast group in the underlay. Broadcast, link-local multicast, and ARP traffic are encapsulated in fabric VXLAN and sent to the destination underlay multicast group. PIM ASM is used as the transport mechanism.

All fabric edge nodes within a fabric site will have the same overlay VNs and overlay IP subnets configured. When Layer 2 flooding is enabled for a given subnet, all edge nodes will send multicast PIM joins for the respective underlay multicast group, effectively pre-building a multicast shared tree. A shared tree must be rooted at a Rendezvous Point, and for Layer 2 flooding to work, this RP must be in the underlay. This RP can be configured manually or programmatically through LAN Automation.

If LAN Automation is used, the LAN Automation primary device (seed device) along with its redundant peer (peer seed device) are configured as the underlay Rendezvous Point on all discovered devices. MSDP is automated between the seeds to create the Anycast-RP configuration needed in the underlay for Layer 2 flooding. In addition, PIM sparse-mode is enabled on Loopback 0 and all point-to-point interfaces configured through the LAN Automation process on the devices.

If Layer 2 flooding is needed and LAN Automation was not used to discover all the devices in the fabric site, multicast routing needs to be enabled manually on the devices in the fabric site and MSDP should be configured between the RPs in the underlay. Loopback 0 can be used as the *connect-source* and *originator-ID* for the MSDP peering.

Connect-source uses the primary IP address on the configured interface as the source IP address of the MSDP TCP connection. *Originator-ID* allows the MSDP speaker originating a source-active (SA) message to use the IP address of the defined interface as the RP address of the message. *Originator-ID* is the inherent mechanism by which MSDP works to address the RPF check.

If configuring the underlay manually, in order to echo the same configuration elements performed through LAN Automation, Loopback60000 can be used as the RP address on the MSDP peers in the underlay.

About Critical VLAN

By default, when a network access device (NAD) cannot reach its configured RADIUS servers, new hosts connected to the NAD cannot be authenticated and are not provided access to the network. The inaccessible authentication bypass feature, also referred to as *critical authentication*, *AAA fail policy*, or simply *critical VLAN*, allows network access on a particular VLAN when the RADIUS server is not available (down).

When a NAD tries to authenticate an endpoint connected to a port, it first checks the status of the configured RADIUS servers. If a server is available, the NAD can authenticate the host. If all the configured RADIUS servers are unavailable and the critical VLAN feature is enabled, the NAD grants network access to the endpoint and puts the port in the critical-authentication state which is a special-case authentication state. When the RADIUS servers are available again, clients in the critical-authentication state must reauthenticate to the network.

Similarly, critical voice VLAN support works by putting voice traffic into the configured voice VLAN if the RADIUS server becomes unreachable.

Critical VLAN Design Considerations

Within a fabric site, a single subnet can be assigned to the critical data VLAN. The critical voice VLAN does not need to be explicitly defined, as the same VLAN is used for both voice and critical voice VLAN support. This ensures that phones will have network access whether the RADIUS server is available or not. SD-Access uses VLAN 2046 and VLAN 2047 for the critical voice VLAN and critical (data) VLAN, respectively.

As discussed in the Fabric Overlay Design section, SD-Access creates segmentation in the network using two method: VRFs (Virtual networks) for macro-segmentation and SGTs (Group-Based Access Control) for micro-segmentation. By default, users, devices, and applications in the same VN can communicate with each other. SGTs can permit or deny this communication within a given VN.

When designing the network for the critical VLAN, this default macro-segmentation behavior must be considered. For example, consider if the subnet assigned for development servers is also defined as the critical VLAN. In the event of the RADIUS server being unavailable, new devices connecting to the network will be placed in the same VLAN as the development servers. Because these devices are in the same VN, communication can occur between them. This is potentially highly undesirable.

Creating a dedicated VN with limited network access for the critical VLAN is the recommended and most secure approach. In the event of RADIUS unavailability, new devices connecting to the network will be placed in their own virtual network which automatically segments their traffic from any other, previously authenticated hosts.

The dedicated critical VN approach must look at the lowest common denominator with respect to total number of VN supported by a fabric device. Certain switch models support only one or four user-defined VNs. Using a dedicated virtual network for the critical VLAN may exceed this scale depending on the total number of other user-defined VNs at the fabric site and the platforms used.

Tech tip
Please see the Cisco DNA Center data sheet on Cisco.com for device-specific fabric VN scale.

LAN Automation

LAN Automation is the Plug-n-Play (PnP) zero touch automation of the underlay network in the SD-Access solution. The simplified procedure builds a solid, error-free underlay network foundation using the principles of a Layer 3 routed access design. Using the LAN Automation feature, Cisco DNA Center automatically finds and adds switches to the underlay routing network. These discovered switches are then provisioned with an IS-IS (Intermediate System to Intermediate System) configuration, added to the IS-IS domain to exchange link-state routing information with the rest of the routing domain, and added to the Cisco DNA Center Inventory. Once in Inventory, they are in ready state to be provisioned with AAA configurations and added in a fabric role.

About Plug and Play and LAN Automation

LAN Automation is designed to onboard switches for use in an SD-Access network either in a fabric role or as an intermediate device between fabric nodes. The LAN Automation process is based on and uses components from the Cisco Plug and Play (PnP) solution. While understanding the full Cisco PnP solution is not required for provisioning and automation, understanding the pieces aids in network design.

- **Cisco Network Plug and Play Process**—This pre-installed capability is present on Cisco DNA Center. It receives Plug and Play requests from Cisco devices and then provisions devices based on defined rules, criteria, and templates.
- **Cisco Plug and Play IOS Agent**—This software component is embedded in Cisco devices and communicates to the Cisco Network Plug and Play process using the open plug and play protocol over

HTTPS. By default, this agent runs on VLAN 1. When a switch is powered on without any existing configuration, all interfaces are automatically associated with VLAN 1. With Plug and Play, when a device is first powered on, it will begin requesting a DHCP address through all connected, physical interfaces in the Up/Up state so that an IP address is provided to Interface VLAN 1.

- **Primary and Secondary Devices (LAN Automation Seed and Peer Seed Devices)**—These devices are manually configured with IP reachability to Cisco DNA Center along with SSH and SNMP credentials. Once they have been discovered and added to Inventory, these devices are used to help onboard additional devices using the LAN Automation feature.

  Once the LAN Automation task is started from Cisco DNA Center the primary seed device becomes a temporary DHCP server. It sends DHCP Offers and Acknowledgements, from DHCP's DORA, to the discovered devices running the Agent. These packets include DHCP Option 43 to point the Agent's devices to the Cisco DNA Center Plug and Play Process for additional configuration.

Network Design Considerations for LAN Automation

There are specific considerations for designing a network to support LAN Automation. These include IP reachability, seed peer configuration, hierarchy, device support, IP address pool planning, and multicast. Additional design considerations exist when integrating the LAN Automated network to an existing routing domain or when running multiple LAN automation sessions. Each of these are discussed in detail below.

IP Reachability

Devices operating in SD-Access are managed through their Loopback 0 interface by Cisco DNA Center. For the LAN automation seed devices, this means they should be configured with a Loopback 0 interface, and that Cisco DNA Center must have IP reachability to that interface IP address.

On the seed device, this can be achieved through direct routes (static routing), default routing, or through an IGP peering with upstream routers. IS-IS can be used as the IGP to potentially avoid protocol redistribution later. A floating static route to Cisco DNA Center can be considered, though it should have an administrative distance lower than the IGP. To avoid further, potential redistribution at later points in the deployment, this floating static can either be advertised into the IGP or given an administrative distance lower than the BGP.

Tech tip
Further details on the initial IP reachability and redistribution described above are discussed in the Appendices of SD-Access Fabric Provisioning Guide.

The seed device should have SSH enabled along with SSH credentials and SNMP read credentials configured. SNMPv2 is supported though SNMPv3 is recommended. While a single seed can be defined, two seed devices are recommended. The secondary seed can be discovered and automated, although most deployments should manually configure a redundant pair of core or distribution layer switches as the seed and peer seed devices.

Peer Configuration

The peer device (secondary seed) can be automated and discovered through the LAN Automation process. However, it is recommended to configure the device manually. The two seed devices should be configured with a Layer 3 physical interface link between them. Both devices should be configured with IS-IS, and the link between the two should be configured as a point-to-point interface that is part of the IS-IS routing domain. For consistency with the interface automation of the discovered devices, BFD should be enabled on this cross-link between the seeds, CLNS MTU should be set to 1400, PIM sparse-mode should be enabled, and the system MTU set to 9100.

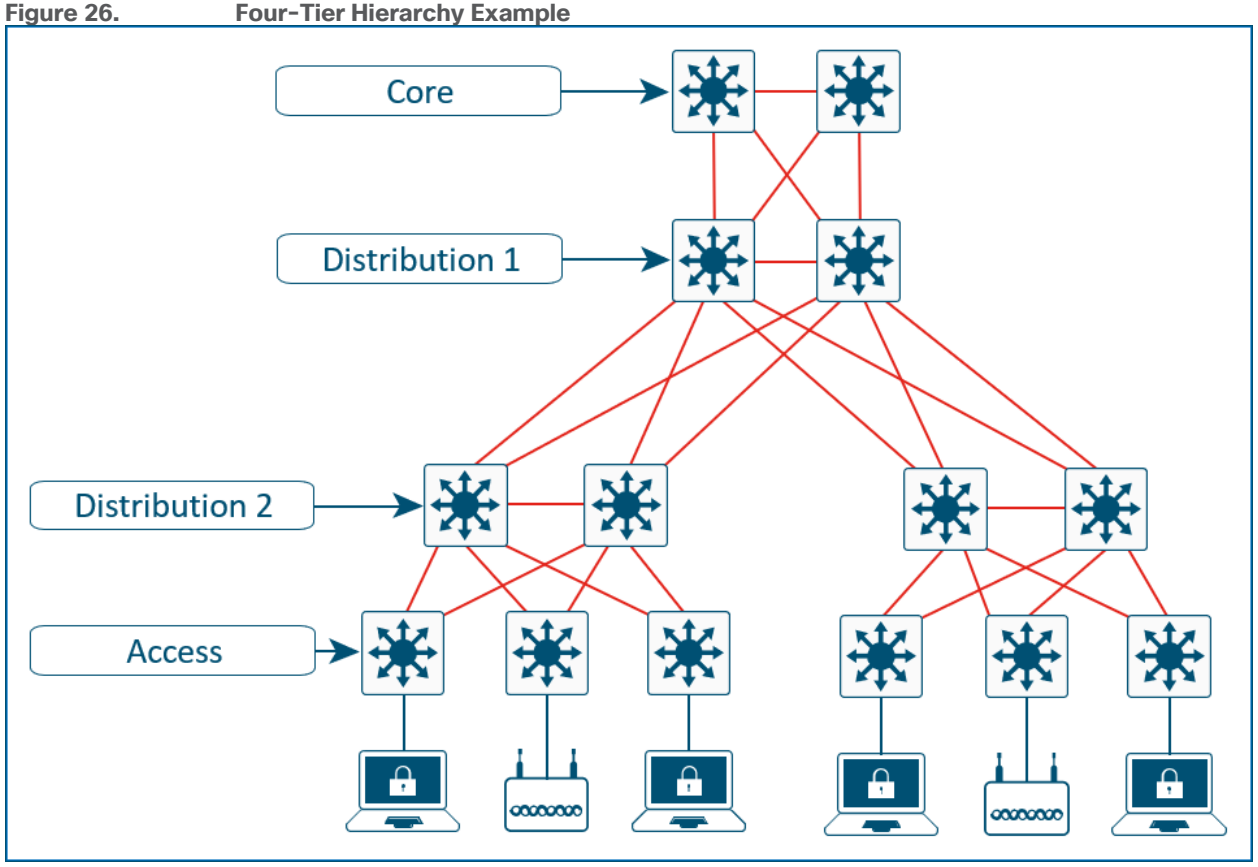
Tech tip
For additional configuration details and BFD parameters, please see SD-Access Fabric Provisioning Guide and Software-Defined Access for Distributed Campus Deployment Guide.

Hierarchy

LAN Automation supports discovering devices up to two CDP hops away from the seed devices. Thus, this feature is supported for both collapsed core/distribution designs and traditional three-tier Campus designs, though the intermediate devices in multitiered network must be Cisco devices.

If the network has more than three-tiers, multiple LAN Automation sessions can be performed sequentially. In Figure 26, if the seed devices are the core layer, then the Distribution 1 and Distribution 2 devices can be discovered and configured through LAN Automation. To discover the devices in the Access layer, a second LAN Automation session can be started after the first one completes. This second session could define Distribution 1 or Distribution 2 as the seed devices for this new LAN Automation workflow.

Figure 26. Four-Tier Hierarchy Example



Device Support

The following chart provides a sample list of common Campus LAN switches supported for LAN Automation.

Table 1. SD-Access LAN Automation Device Support

Primary and Peer Device (Seeds)	Discovered Device
Cisco Catalyst 9000 Series Switches	Cisco Catalyst 9000 Series Switches

Primary and Peer Device (Seeds)	Discovered Device
Cisco Catalyst 6800 Series Switches	Cisco Catalyst 4500E Series Switches[1]
Cisco Catalyst 6500 Series Switches	Cisco Catalyst 3850 Series Switches
Cisco Catalyst 3850 Series Switches	Cisco Catalyst 3650 Series Switches
Cisco Catalyst 3650 Series Switches	

[1]Supervisor Engine 8-E, 9-E only, and using the Supervisor ports only

Tech tip

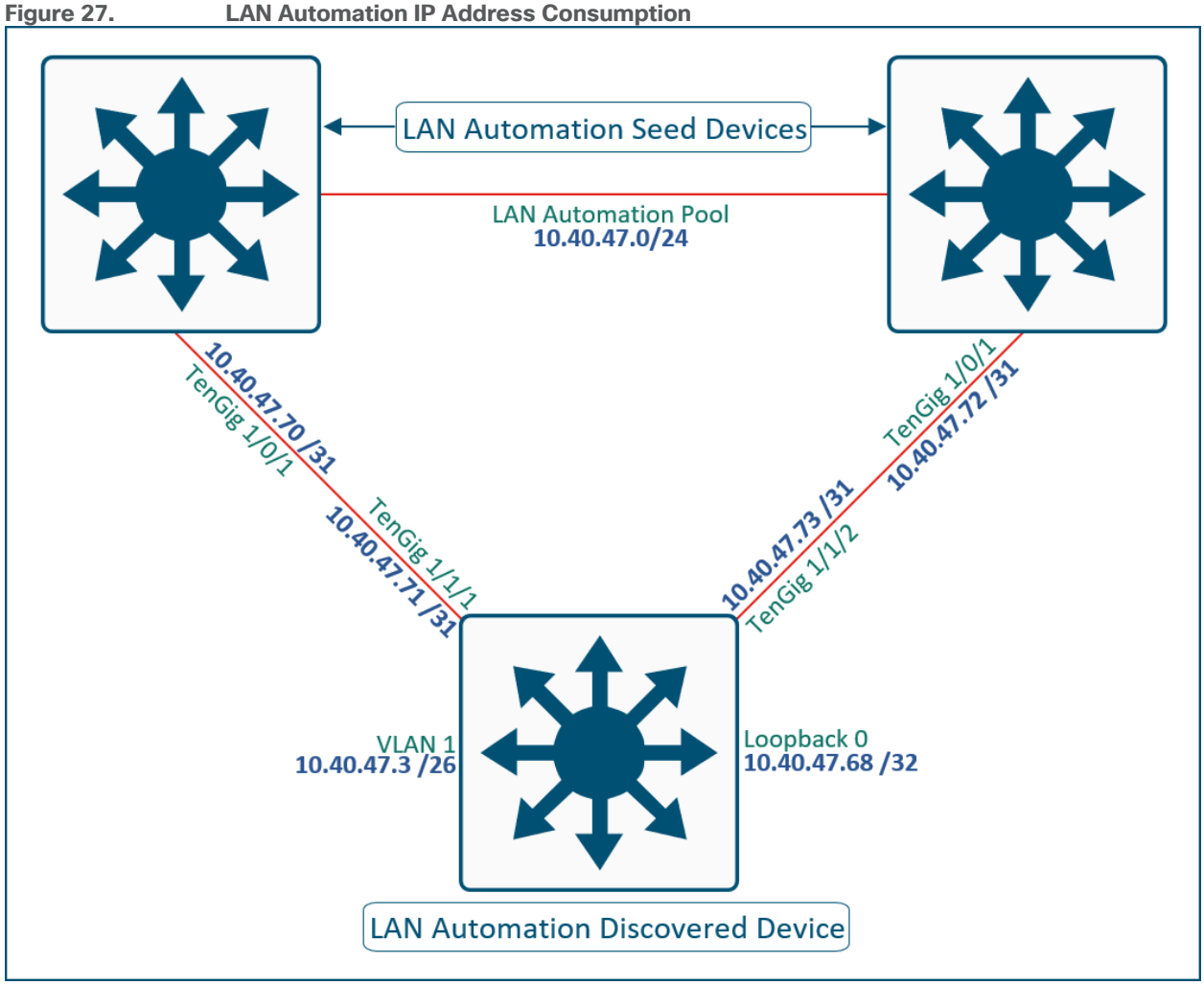
Please consult the Cisco DNA Center Release Notes and Cisco DNA Center SD-Access LAN Automation Deployment Guide for updates, additions, and complete list of devices supported with LAN Automation.

IP Address Pool Planning for LAN Automation

When a device is discovered and provisioned through LAN Automation, Cisco DNA Center automates the Layer 3 routed access configuration on its interfaces. When a device is initially powered on with no configuration, it receives an IP address in VLAN 1 from the DHCP server service temporarily created on the primary device during the initiation of the LAN Automation task.

Cisco DNA Center provisions the discovered device with an IP address on Loopback 0. The interfaces connected to the seed and redundant seed will then each receive an IP address on each end of the link; Cisco DNA Center automates both the seed devices' interfaces and the discovered devices' interfaces. For any given single device onboarded using LAN Automation with uplinks to both seeds, at least six IP addresses are consumed within the address pool.

Figure 27. LAN Automation IP Address Consumption



Tech tip
Interface VLAN 1 used by the PNP Agent on discovered devices to achieve IP reachability to Cisco DNA Center. Once the LAN Automation session is stopped, the IP address on VLAN 1 is removed.

The same IP address pool can be used for multiple LAN Automation discovery sessions. For example, one session can be run to discover the first set of devices. After LAN Automation completes, the same IP address pool can be used a subsequent session provided it has enough available IP addresses.

When Cisco DNA Center assigns IP addresses as part of LAN Automation, it tracks the pool usage within an internal database. When a LAN Automation session starts, a check is run against that internal database to ensure there are at least 128 available IP addresses in the defined address pool. If LAN Automation is run multiple times with the same pool, consider using a minimum /24 address space to ensure enough addresses.

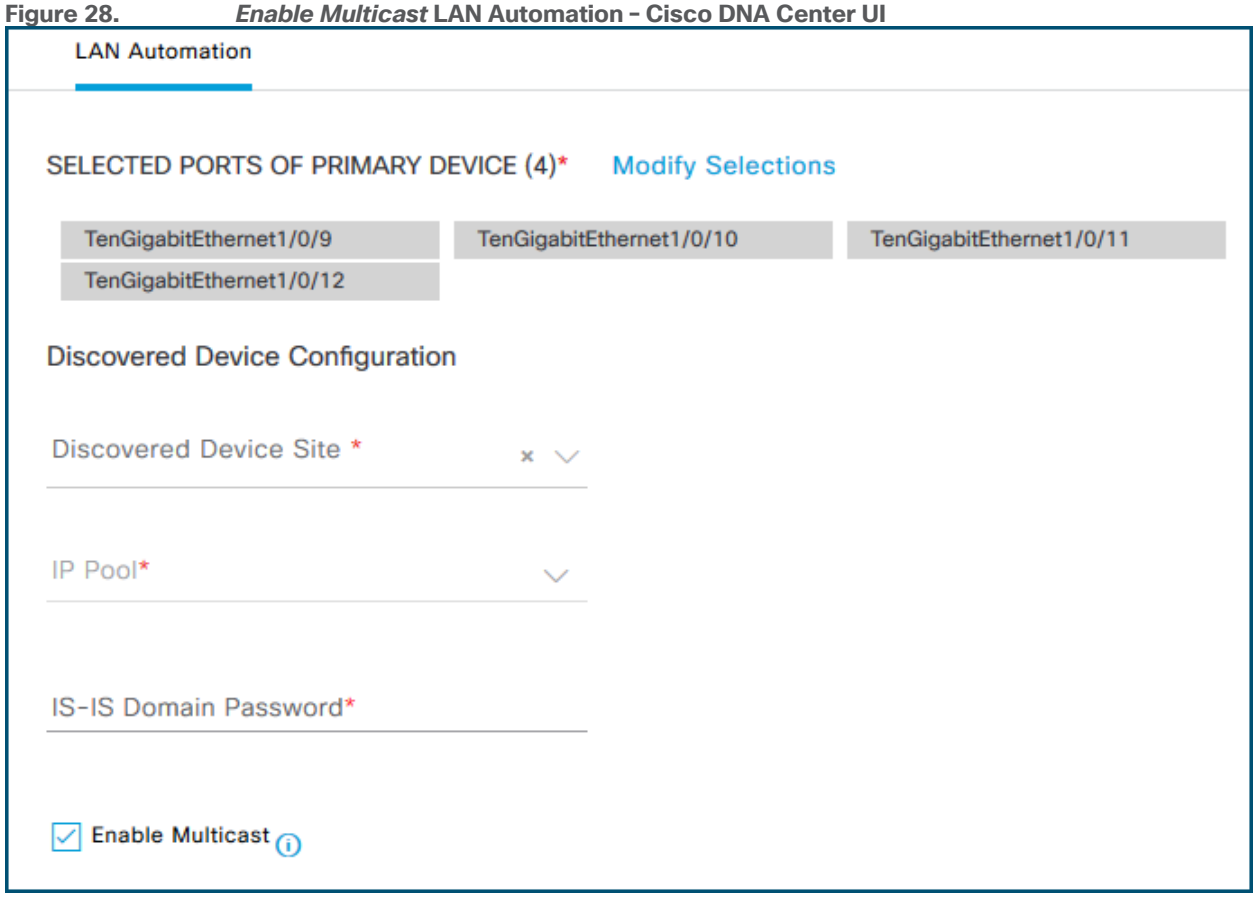
If discovering using the maximum two CDP hops, both the upstream and downstream interfaces on the first-hop device will be configured with routed ports. LAN Automation currently deploys the Loopback 0 interfaces with a /32 subnet mask and the point-to-point routed links with a /31 subnet mask. This provides the highest efficiency of preservation of IP address pool space.

LAN Automation can onboard up to 500 discovered devices during each session. Care should be taken with IP address planning based on the address pool usage described above to ensure that the pool is large enough to support the number of devices onboarded during both single and subsequent sessions.

Multicast and LAN Automation

Enable Multicast is an optional capability of LAN Automation. It is represented by a check box in the LAN Automation workflow as shown the following figure.

Figure 28. ***Enable Multicast*** **LAN Automation – Cisco DNA Center UI**



When this box is checked, PIM sparse-mode will be enabled on the interfaces Cisco DNA Center provisions on the discovered devices and seed devices, including Loopback 0. The seed devices are configured as the Rendezvous Point (RP) for PIM-ASM, and the discovered devices are configured with an RP statement pointing to the seeds. If redundant seeds are defined, Cisco DNA Center will automate the configuration of MSDP between them using Loopback 60000 as the RP interface and Loopback 0 as the unique interface. If subsequent LAN Automation sessions for the same discovery site are done using different seed devices with the *Enable multicast* checkbox selected, the original seed will still be used as the multicast RPs, and newly discovered devices will be configured with the same RP statements pointing to them.

Tech tip
For additional details on multicast RPs, MSDP, and PIM-ASM, please see the Multicast Design section.

Additional IS-IS Routing Considerations

The seed devices are commonly part of a larger, existing deployment that includes a dynamic routing protocol to achieve IP reachability to Cisco DNA Center. When a LAN Automation session is started, IS-IS routing is configured on the seed devices in order to prepare them to provide connectivity for the discovered devices. This IS-IS configuration includes routing authentication, bidirectional forwarding detection , and default route propagation. These provisioned elements should be considered when multiple LAN automation sessions are

completed in the same site, when LAN Automation is used in multiple fabric sites, and when the fabric is part of a larger IS-IS routing domain.

IS-IS Domain-Password

As part of the LAN Automation workflow in Cisco DNA Center, an IS-IS Domain password is required. The IS-IS domain password enables plaintext authentication of IS-IS Level-2 link-state packets (LSP). If the seed devices are joining an existing IS-IS routing domain, the password entered in the GUI workflow should be the same as the existing routing domain to allow the exchange of routing information.

Bidirectional Forwarding Detection

Bidirectional forwarding detection (BFD) is provisioned on seed devices at the router configuration level (*bfd all-interfaces*) and at the interface level connecting to the discovered devices. BFD is also provisioned on the discovered devices at the router configuration level and at interface configuration level connecting to the upstream peers.

When configuring the seed device pair before beginning LAN automation, a Layer 3 routed link should be configured between them and added to the IS-IS routing process. In some platforms, if BFD is enabled at the router configuration level only and not also at the interface level, the IS-IS adjacency will drop. Therefore, BFD should be enabled manually on this cross-link interface to ensure the adjacency remains up once the LAN automation session is started. This also means that when integrating the seed devices into an existing IS-IS network, BFD should be enabled on the interfaces connecting to the remainder of the network.

Default Route Propagation

During LAN Automation, *default-information originate* is provisioned under the IS-IS routing process to advertise the default route to all discovered devices. This command is applied to each seed during the LAN Automation process, including subsequent LAN automation sessions. If integrating with an existing IS-IS network, each seed in a LAN automation session will now generate a default route throughout the routing domain.

Maximum Transmission Unit

MTU defines the largest frame size that an interface can transmit without the need to fragment. From an frame reception perspective, if the received frame is less than or equal to the interface MTU, then the packet can be accepted. It is then sent up the protocol stack to be processed at the higher layers. If the frame is larger than the interface MTU, it is dropped.

Interface MTU should be set consistently across a Layer 2 domain (collision domain/VLAN) to ensure properly communication. Consistent MTU is also required for several other processes and protocols to work properly such as OSPF and IS-IS.

All devices on the physical media must have the same protocol MTU to operate properly. All infrastructure devices in a broadcast domain should have the same MTU. If the broadcast domain is logically extended using an overlay encapsulation protocol, the underlay routers and switches through which this overlay is carried should all be configured with a common jumbo MTU value.

LAN Automation configures a Layer 2 MTU value of 9100 on the seed devices and all discovered devices. Link state routing protocols need matching MTU values for the neighbor relationship to come up, and so the end-to-end MTU value across the routing domain should be the same to accommodate this.

Wireless Design

This section is organized into the following subsections:

Section	Subsection
Wireless Design	Over-the-Top Wireless Mixed-Mode Wireless Guest Wireless

SD-Access supports two options for integrating wireless access into the network. One option is to use traditional Cisco Unified Wireless Network (CUWN) local-mode configurations *over-the-top* as a non-native service. In this mode, the SD-Access fabric is simply a transport network for the wireless traffic, which can be useful during migrations to transport CAPWAP-tunneled endpoint traffic from the APs to the WLCs. The other option is fully integrated SD-Access Wireless, extending the SD-Access beyond wired endpoints to also include wireless endpoints.

Integrating the wireless LAN into the fabric provides the same advantages for the wireless clients as provided to the wired clients in the fabric, including addressing simplification, mobility with stretched subnets, and end-to-end segmentation with policy consistency across the wired and wireless domains. Wireless integration also enables the WLC to shed data plane forwarding duties while continuing to function as the control plane for the wireless domain.

Fabric wireless controllers manage and control the fabric-mode APs using the same general model as the traditional local-mode controllers which offers the same operational advantages such as mobility control and radio resource management. A significant difference is that client traffic from wireless endpoints is not tunneled from the APs to the wireless controller. Instead, communication from wireless clients is encapsulated in VXLAN by the fabric APs which build a tunnel to their first-hop fabric edge node. Wireless traffic it tunneled to the edge nodes as the edge nodes provide fabric services such as the Layer 3 Anycast Gateway, policy, and traffic enforcement.

This difference enables a distributed data plane with integrated SGT capabilities. Traffic forwarding takes the optimum path through the SD-Access fabric to the destination while keeping consistent policy, regardless of wired or wireless endpoint connectivity.

The control plane communication for the APs does use a CAPWAP tunnel to the WLC, which is similar to the traditional CUWN control plane. However, a fabric WLC is integrated into the SD-Access control plane (LISP) communication. When added as a Fabric WLC, the controller builds a two-way communication to the fabric control plane nodes.

Tech tip
Border nodes and edge nodes also build this two-way communication, or *LISP session,* with the control plane nodes.

This communication allows the WLCs to register client Layer 2 MAC addresses, SGT, and Layer 2 segmentation information (Layer 2 VNI). All of this works together to support wireless client roaming between APs across the fabric site. The SD-Access fabric control plane process inherently supports the roaming feature by updating its host-tracking database when an endpoint is associated with a new RLOC (wireless endpoint roams between APs).

Fabric Wireless Integration Design

Fabric-mode APs connect into a pre-defined VN named *INFRA_VN*. The VN is associated with the global routing table (GRT). This design allows the WLC to connect into the fabric site for AP management without needing to leak routes out of a VRF table.

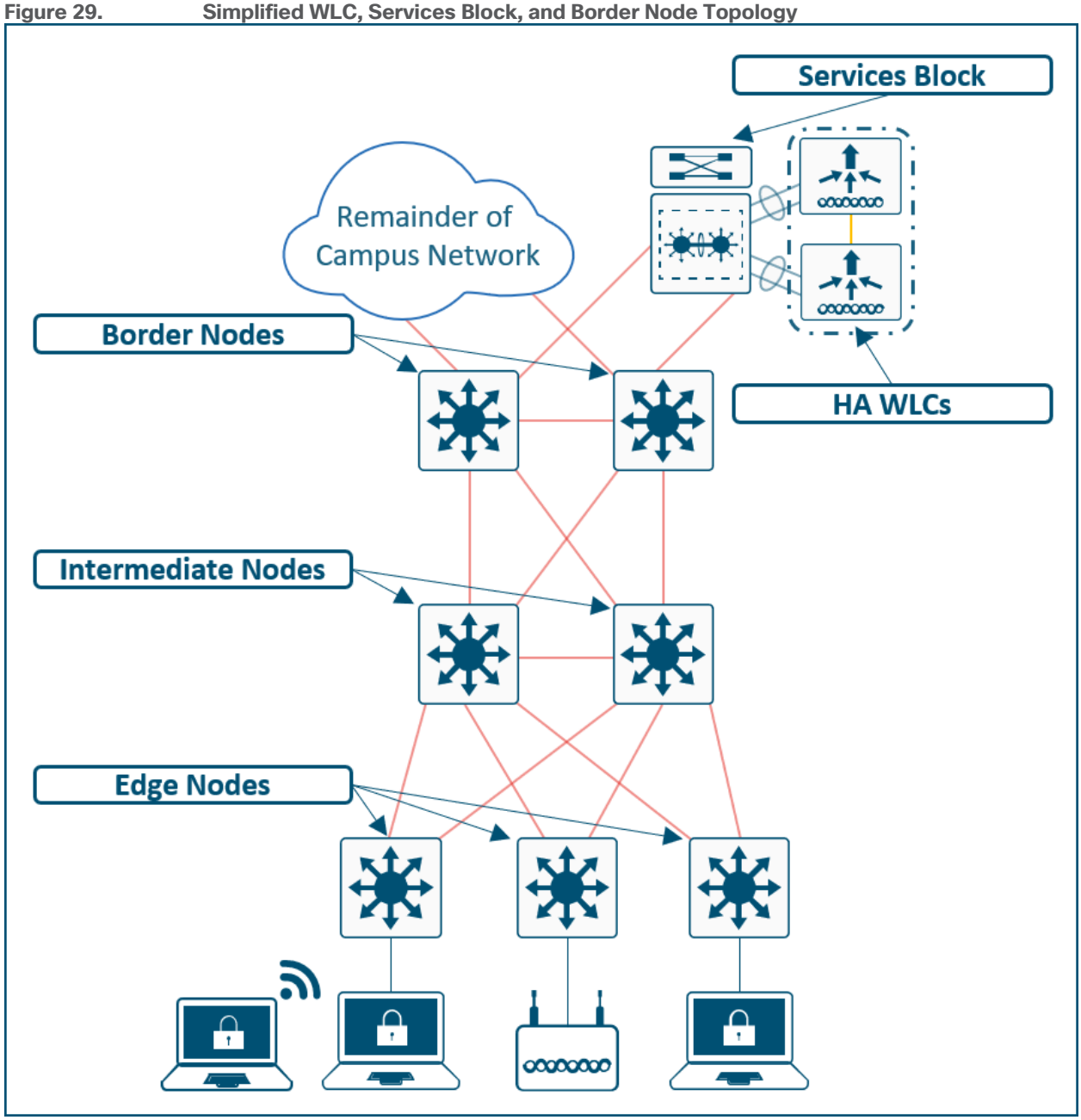
Tech tip
INFRA_VN is also the VN used by classic and policy extended nodes for connectivity.

When integrating fabric-enabled wireless into the SD-Access architecture, the WLC control plane keeps many of the characteristics of a local-mode controller, including the requirement to have a low-latency connection between the WLC and the APs. This latency requirement, 20ms RTT, precludes a fabric WLC from managing fabric-mode APs at a remote site across a typical WAN. As a result, a remote site with SD-Access wireless with a WAN circuit exceeding 20ms RTT will need a WLC local to that site.

Wireless integration with SD-Access should also consider WLC placement and connectivity. WLCs typically connect to a shared services distribution block that is part of the underlay. The preferred services block has chassis redundancy as well as the capability to support Layer 2 multichassis EtherChannel connections for link and platform redundancy to the WLCs. As described in the Services Block section, VSS, StackWise Virtual, switch stacks, and Nexus vPC can be used to accomplish these goals.

In the simplified example diagram below, the border nodes are directly connected to the services block switch with Layer 3 connections. The WLCs are connected to the services block using link aggregation. Each WLC is connected to member switch of the services block logical pair.

Figure 29. Simplified WLC, Services Block, and Border Node Topology



Over-the-Top Centralized Wireless Design

In cases where the WLCs and APs cannot participate in the fabric, a traditional CUWN centralized design model is an option. In Centralized WLC deployment models, WLCs are placed at a central location in the enterprise network. With this deployment model, the CAPWAP tunnels between WLC and APs traverse the campus backbone network. In the *over-the-top* model, this means the wireless infrastructure uses the fabric as a transport but without the benefits of fabric integration.

An *over-the-top* wireless design still provides AP management, simplified configuration and troubleshooting, and roaming at scale. In this centralized *over-the-top* model, the WLAN controller is connected at the data center services block or a dedicated service block adjacent to the campus core. Wireless traffic between WLAN clients and the LAN is tunneled using CAPWAP between APs and the controller. APs can reside inside or

outside the fabric without changing the centralized WLAN design. However, the benefits of fabric and SD-Access are not extended to wireless when it is deployed *over-the-top*.

Tech tip
For additional information about CUWN and traditional campus wireless design, see the Campus LAN and Wireless LAN Design Guide.

Mixed SD-Access Wireless and Centralized Wireless Design

Many organizations may deploy SD-Access with centralized wireless *over-the-top* as a first transition step before integrating SD-Access Wireless into the fabric. For this case, an organization should dedicate a WLC for enabling SD-Access Wireless.

Organizations can deploy both centralized and SD-Access Wireless services as a migration stage. Cisco DNA Center can automate a new installation supporting both services on the existing WLC, though a software WLC software upgrade may be required. In this case, the new installation from Cisco DNA Center on the existing WLC does not take into consideration existing running configurations. Instead, Cisco DNA Center automates the creation of the new replacement services.

Guest Wireless Design

When designing for Guest Wireless, SD-Access supports two different models:

- **Guest as a dedicated VN**—Guest is simply another user-defined VN.
- **Guest Border and Control Plane Node**—Guest traffic is terminated on dedicated Guest border nodes and guests are registered with the HTDB on a dedicated Guest control plane node.

Guest as a VN

Creating a Guest VN is as straightforward as clicking the checkbox when creating a VN in Cisco DNA Center.

Figure 30. Guest Virtual Network Creation – Cisco DNA Center UI

With Guest as VN, guest and enterprise clients share the same control plane node and border node. The Guest SSID is associated to a dedicated Guest VN, and SGTs are used for isolating guest traffic from itself. Guests, by the nature of VRFs and macro segmentation, are automatically isolated from other traffic in different VNs though the same fabric nodes are shared for guest and non-guest.

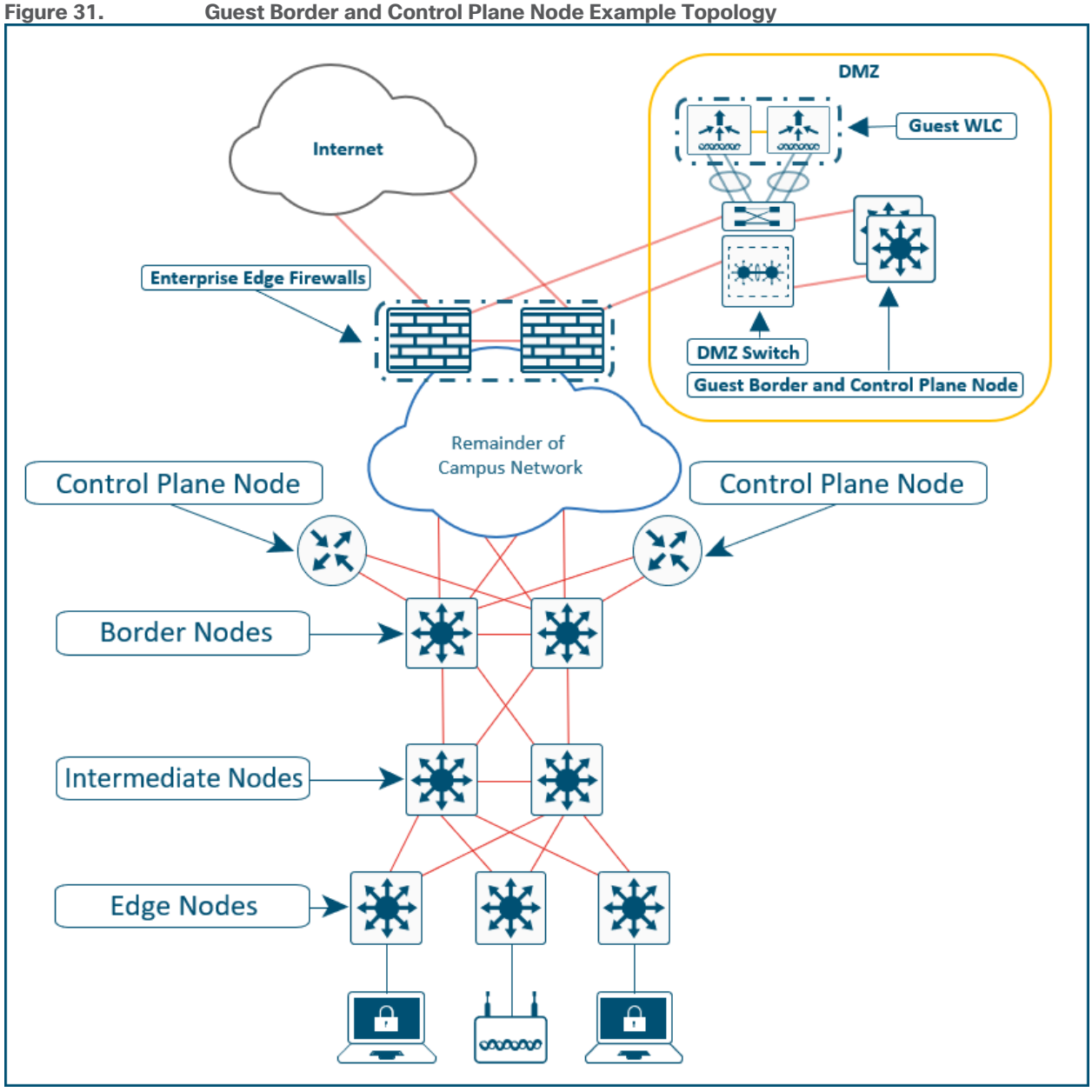
Guest users should be assigned an SGT value upon connecting to the network. This assignment is used to implement an equivalence of a peer-to-peer blocking policy. For a Fabric SSID, all security policy is enforced at the edge node, not at the access point itself. Traditional peer-to-peer blocking, which is enabled on the WLAN in the WLC, would not take effect. To provide consistent policy, an AP will forward traffic to the fabric edge, even if the clients communicating are associated with the same AP. An SGT assigned to Guest users can be leveraged to deny traffic between the same SGTs.

When designing for Guest as a VN, the same design modalities referenced throughout this document for any other virtual network apply to this Guest VN.

Guest Border Node and Guest Control Plane Node

This design leverages a dedicated control plane node and border node for guest traffic. The nodes can be colocated on the same device, for operational simplicity, or on separate devices, for maximum scale and resilience. The guest control plane node and border node feature provides a simplified way to tunnel the Guest traffic to the DMZ which is a common security convention.

Figure 31. Guest Border and Control Plane Node Example Topology



A maximum of two control plane nodes can be deployed for guest traffic. The result is a fabric site can have two control plane nodes for Enterprise traffic and another two for Guest traffic as show in Figure 20.

Like the enterprise traffic, guest traffic is still encapsulated in VXLAN at the AP and sent to the edge node. The edge node is configured to use the guest border node and guest control plane node as well as the enterprise nodes. All guest traffic is encapsulated in fabric VXLAN by the edge node and tunneled to the guest border node. The guest border node commonly resides in the DMZ in order to provide complete isolation from the

enterprise traffic. This solution is similar to the CUWN Guest Anchor solution. Guest users are registered to a guest control plane node, and the guest endpoints receive an IP address in the DHCP scope for the DMZ.

Tech tip
Cisco DNA Center automates and manages the workflow for implementing the wireless guest solution for fabric devices only; wired guest services are not included in the solution.

Dedicated Guest Border and Control Plane Design Considerations

The network infrastructure into the DMZ must follow the MTU requirements for Layer 2 segments: when the broadcast domain is logically extended using an overlay encapsulation protocol, the underlay routers and switches through which this overlay is carried should all be configured with a common jumbo MTU value. A firewall commonly separates the DMZ block from the remainder of the Campus network. The firewall must be configured to allow the larger MTU requirements and to allow the traffic between the fabric edge devices and the guest border and control plane nodes. The Layer 3 IP-based handoff is not automated on the Guest border node and must be configured manually.

Like other RLOCs (Loopback 0 address) of devices operating in a fabric role, the IP address of the guest border node and guest control plane node must be advertised into the fabric site and be available as a /32 route in the global routing table on the edge nodes.

External Connectivity

This section is organized into the following subsections:

Section	Subsection
External Connectivity	Layer 3 Handoff VRF-Aware Peer Non-VRF-Aware Peer Firewall Peer External Connectivity Considerations

External connectivity outside of the fabric site can have several possible variations, and these variations are based on underlying network design. For example, the fabric border node may be connected to an actual Internet edge router, an ISP device, a firewall, a services block switch, or some other routing infrastructure device. Each of these peer devices may be configured with a VRF-aware connection (VRF-lite) or may simply connect to the border node using the global routing table.

Shared services, as discussed in the earlier Routing Table section, may be deployed in a dedicated VRF or the global routing table, and shared services may be connected to a services block or be accessed through data center infrastructure. Internet access itself may be in a VRF, though is most commonly available in the global routing table. While each of these options are viable, though each present a different underlying network design that the fabric site must integrate with.

Layer 3 Handoff

Regardless of the potential variations for the network design and deployment outside of the fabric site, a few things are going to be in common, and the border node will be the device tying these things together:

- **VRF Aware**—A border node will be VRF-aware. All user-defined VNs in the fabric site are instantiated and provisioned as VRFs.

- **Site Prefixes in VRF**—The EID-space prefixes associated with the fabric site will be in VRF routing tables on the border node.
- **Upstream Infrastructure**—The border nodes will be connected to a next-hop device and further routing infrastructure (referenced simply as *next-hop*, for brevity). This upstream infrastructure, while a necessary part of the overall design, is not part of the fabric site and is therefore not automated though SD-Access workflows in Cisco DNA Center.

Cisco DNA Center can automate the configuration on the border nodes, though, and this is done through an *IP-based Layer 3 handoff*. By IP-based, this means native IP forwarding, rather than encapsulation, is used. The fabric packet is de-encapsulated before being forwarded. The configuration is Layer 3 which means it uses subinterfaces, when the border node is a routing platform, or Switched Virtual Interfaces (SVIs), when the border node is a switching platform, to connect to the upstream peers.

This Layer 3 handoff automation provisions VRF-lite by associating each SVI or subinterface with a different fabric VN (VRF). External BGP is used as the routing protocol to advertise the endpoint space (EID-space) prefixes from the fabric site to the external routing domain and to attract traffic back to the EID-space. This BGP peering can also be used to advertise routes into the overlay such as for access to shared services.

While the Layer 3 handoff for external connectivity can be performed manually, automation through Cisco DNA Center is preferred and recommended.

With the Layer 3 IP-based handoff configured, there are several common configuration options for the *next-hop* device. This device may peer (have IP connectivity and routing adjacency) with the border node using VRFs. This *next-hop* device may even continue the VRF segmentation extension to its next hop. This *next-hop* may not be VRF-aware and peer to the border node using the global routing table. Finally, the *next-hop* may be firewall which is special case peering that is not VRF-aware.

VRF-Aware Peer Design

This VRF-Aware peer design begins with VRF-lite automated on the borer node through Cisco DNA Center, and the peer manually configured as VRF-aware. For each VN that is handed off on the border node, a corresponding VN and interface is configured on the peer device. Existing collateral may refer to this deployment option as a *fusion router* or simply *fusion device*.

The generic term *fusion router* comes from MPLS Layer 3 VPN. The basic concept is that the *fusion router* is aware of the prefixes available inside each VPN (VRF), generally through dynamic routing, and can therefore *fuse* these routes together. In MPLS Layer 3 VPN, these generic *fusion routers* are used to route traffic between separate VRFs (VRF leaking). Alternatively, the *fusion router* can also be used to route traffic to and from a VRF to a shared pool of resources in the global routing table (route leaking). Both responsibilities are essentially the same as they involve advertising routes from one routing table into a separate routing table.

This VRF-Aware peer design is commonly used for access to shared services. Shared services are generally deployed using a services block deployed on a switching platform to allow for redundant and highly-available Layer 2 links to the various devices and servers hosting these services. Shared service most commonly exists in the global routing table, though deployments may use a dedicated VRF to simply configuration.

In an SD-Access deployment, the *fusion device* has a single responsibility: to provide access to shared services for the endpoints in the fabric. There are two primary ways to accomplish this task depending on how the shared services are deployed, route leaking and VRF leaking. Both require the fusion device to be deployed as VRF-aware.

- **Route Leaking**—The option is used when the shared services routes are in the GRT. On the fusion device, IP prefix lists are used to match the shared services routes, route-maps reference the IP prefix lists, and

the VRF configurations reference the route-maps to ensure only the specifically matched routes are leaked.

- **VRF Leaking**—The option is used when shared services are deployed in a dedicated VRF on the fusion device. Route-targets under the VRF configuration are used to leak between the fabric VNs and the shared services VRF.

A fusion device can be either a true routing platform, a Layer 3 switching platform, or a firewall must meet several technological requirements. It must support:

- **Multiple VRFs**—Multiple VRFs are needed for the VRF-Aware peer model. For each VN that is handed off on the border node, a corresponding VN and interface is configured on the peer device. The selected platform should support the number of VNs used in the fabric site that will require access to shared services.
- **Subinterfaces (Routers or Firewall)—**A virtual Layer 3 interface that is associated with a VLAN ID on a routed physical interface. It extends IP routing capabilities to support VLAN configurations using the IEEE 802.1Q encapsulation.
- **Switched Virtual Interfaces (Layer 3 switch)—**Represents a logical Layer 3 interface on a switch. This SVI is a Layer 3 interface forwarding for a Layer 3 IEEE 802.1Q VLAN.
- **IEEE 802.1Q**—An internal tagging mechanism which inserts a 4-byte tag field in the original Ethernet frame between the Source Address and Type/Length fields. Devices that support SVIs and subinterfaces will also support 802.1Q tagging.
- **BGP-4**—This is the current version of BGP and was defined in RFC 4271 (2006) with additional update RFCs. Along with BGP-4, the device should also support the Multiprotocol BGP Extensions such as AFI/SAFI and Extended Community Attributes defined in RFC 4760 (2007).

Tech tip
These five technical requirements are supported on a wide range of routers, switches, and firewalls throughout the Cisco portfolio including Catalyst, Nexus, ASA, FTD, Aggregation Services Routers (ASRs), and Integrated Services Routers (ISRs) for both current and even previous generation hardware.

To support this route leaking responsibility, the device should be properly sized according the number of VRFs, bandwidth and throughput requirements, and Layer 1 connectivity needs including port density and type. When the network has been designed with a services block, the services block switch can be used as the fusion device (VRF-aware peer) if it supports the criteria described above. Fusion devices should be deployed in pairs or as a multi-box, single logical box such as VSS, SVL, or vPC. When the fusion device is a logical unit, border nodes should be connected to both members of the logical pair as described in the later external considerations section.

In a *fusion device* environment, the device performing the leaking may not even be the direct next hop from the border. It may be several physical hops away. In this environment, the VRFs must be maintained, commonly using VRF-lite, from the border to the device ultimately performing the route leaking. In this deployment type, the next-hop from the border is VRF-aware along with the devices in the data path towards the *fusion*.

Non-VRF-Aware Peer

This deployment type begins with VRF-lite automated on the border node, and the peer manually configured, though not VRF-aware. For each VN that is handed off on the border node, a corresponding interface is configured on the peer device in the global routing table. This deployment option is commonly used when the fabric site hands off to a WAN circuit, ISP, an MPLS CE or PE device, other upstream routing infrastructure, or even a firewall which is special-case non-VRF peer discussed further in the Firewall section.

This deployment type is common in WAN infrastructure. If this next-hop peer is an MPLS CE, routes are often merged into a single table to reduce the number of VRFs to be carried across the backbone, generally reducing overall operational costs. If the next-hop peer is an MPLS PE or ISP equipment, it is outside of the administrative domain of the fabric network operator. The result is that there is little flexibility in controlling the configuration on the upstream infrastructure. Many times, ISPs have their own peering strategies and themselves are presenting a Layer 3 handoff to connected devices.

Non-VRF aware means that peer router is not performing VRF-lite. It may have the functionality to support VRFs, but it is not configured with corresponding fabric VRFs the way a VRF-Aware peer would be. The non-VRF aware peer is commonly used to advertise a default route to the endpoint-space in the fabric site.

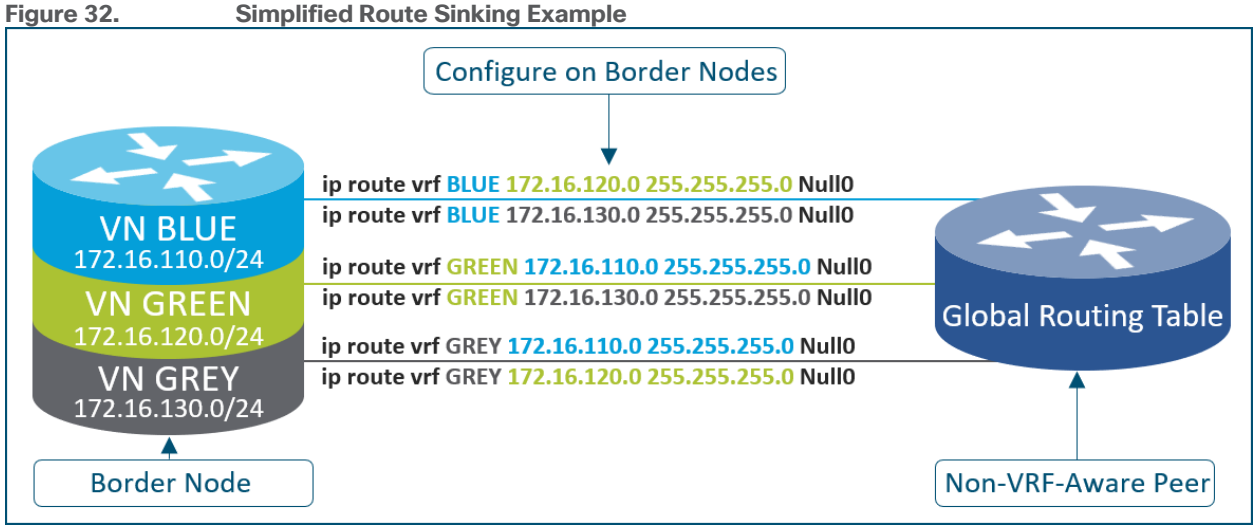
The result is the VNs from the fabric site are merged into a single routing table (GRT) on the next-hop peer. Merging routes into a single table is a different process than route leaking. This deployment type does use the colloquial moniker of *fusion router*.

The challenge with merged tables is the potentiality of East-West communication across the North-South link. Merging the VRFs into a common routing table is best accomplished with a firewall. Firewalls are policy-oriented devices that align well with the segmentation provided through the SD-Access solution.

It is not always possible to use a firewall in environments that use route-table merging such as with WAN circuits listed above. However, degrees of precaution and security can be maintained, even without a firewall. For example, specific scalable group tags (SGTs) or port-based ACLs can limit and prevent East-West communication. Further protection can be added by sinkhole routing. This is done manually on the border node, for each VRF, by pointing the aggregate prefixes for each other VRF to Null0.

In the simplified topology in Figure 32 below, the border node is connected to a non-VRF-aware peer with each fabric VNs and their associated subnet are represented by a color. This type of connection effectively merges the fabric VN routing tables onto a single table (generally GRT) on the peer device. By route sinking as described above, the East-West communication between the VNs can be prevented across the North-South link between the border node and its peer.

Figure 32. Simplified Route Sinking Example



Firewall Peer

If the fabric VNs need to merge to a common routing table, a policy-oriented device such as a firewall should be considered as an upstream peer from the fabric border nodes. Common use cases for a firewall peer include Internet access, access to data center prefixes, WAN connectivity, or Inter-VN communication requirements.

Tech tip
In most deployments, endpoints, users, or devices that need to directly communicate with each other should be placed in the same overlay virtual network. Some networks may have specific requirements for VN to VN communication, though these are less common. VN to VN requirements are often seen during mergers of companies or in some corporate or government structures or similar multi-tenant environment where each agency, tenant, or division is required to have their own VN-space.

A firewall can be used to provide stateful inspection for inter-VN communication along with providing Intrusion Prevent System (IPS) capabilities, advanced malware protection (AMP), granular Application Visibility and Control (AVC), and even URL filtering. Firewalls such as Cisco ASA and Cisco Firepower Threat Defense (FTD) also provide a very rich reporting capability with information on traffic source, destination, username, group, and firewall action with guaranteed logging of permits and drops.

Firewalls can be deployed as a cluster (multiple devices acting as a single logical unit), as an HA pair (commonly Active/Standby), or even as a standalone device. While firewalls do not generally have VRF capabilities, they have other method for providing the same general type of segmentation provided by VRFs. These include contexts, interface-specific ACL, and security-levels (ASA), instances, and security zones (FTD).

Tech tip
Cisco Firepower Threat Defense Release 6.6.0 introduced VRF-lite support. For additional details, please see the Cisco Firepower Release Notes, Version 6.6.0, and Firepower Management Center Configuration Guide, Version 6.6.

Firewall - Security Contexts and Multi-Instance

A single or logical security appliance running ASA software can be partitioned into multiple virtual devices called security contexts. Each context is an independently configured device partition with its own security policy, interfaces, routing tables, and administrators. Multiple contexts logically emulate multiple standalone devices. However, they share the underlying hardware resources such as CPU and memory. Therefore, it is possible for one context to starve one another under load.

Each VN in the fabric can be mapped to a separate security context to provide the most complete separation of traffic. For common egress points such as Internet, a shared context interface can be used.

Tech tip

For devices operating on a Firepower 4100 and 9300 series chassis, the Multi-Instance Capability can be used with the Firepower Threat Defense (FTD) application only. It is similar in construct to security contexts, though allows hard-resource separation, separate configuration management, separate reloads, separate software updates, and full feature support.

FTD does not support multiple security contexts. For additional details on Multi-Instance, please see Cisco Firepower Release Notes, Version 6.3.0, Multi-Instance Capability White Paper, and Using Multi-Instance Capability Configuration Guide.

Firewall - Security Zones

Security zones are a Cisco FTD construct. They are a grouping of one or more matching interfaces that are used to manage and classify traffic flow using various policies and configurations. A given interface can belong to only one zone which provides automatic segmentation between zones.

Like security contexts, each VN in the fabric can be mapped to separate security zone to provide separation of traffic once it leaves the fabric site. Explicit rules can allow for a common egress points such as Internet.

Firewall - Security-Levels

Security-levels are a Cisco ASA construct. A security-level is applied to an interface and defines a relative trust relationship. Security-levels can range from 0 (lowest) to 100 (highest). By default, this relative trust allows traffic to flow from a higher security-level to a lower security-level without explicit use of an access-list. Traffic from a lower security-level cannot flow to a higher security-level without explicit inspection and filtering check such as an ACL.

If interfaces are assigned the same security-level, the default security policy will not allow communicate between these interfaces. Some deployment may require communication between interfaces with the same security-levels, as 0-100 only provides 101 unique values. It is possible to override the default behavior and allow communication between interfaces of the same security-level using a global configuration command on the firewall.

Like contexts and zones, each VN in the fabric can be mapped to different, or even the same, security-level to provide continued separation of traffic outside of the fabric site. On the firewall, a common external interface that faces the public or untrusted network, such as the Internet, can be assigned with a security-level of 0, providing the default traffic flow from high to low.

Additional Firewall Design Considerations

When considering a firewall as the peer device, there are additional considerations. The device must be appropriately licensed and sized for throughput at a particular average packet size in consideration with the enabled features (IPS, AMP, AVC, URL-filtering) and connections per second. It must also have the appropriate interface type and quantity to support connectivity to both its upstream and downstream peers and to itself when deploying a firewall cluster or firewall HA pair. The firewalls must be deployed in routed mode rather than transparent mode.

External Connectivity Design Considerations

The same considerations and conventions apply to external connectivity as they do to connections between layers in Enterprise Campus Architecture: *build triangles, not squares, to take advantage of equal-cost*

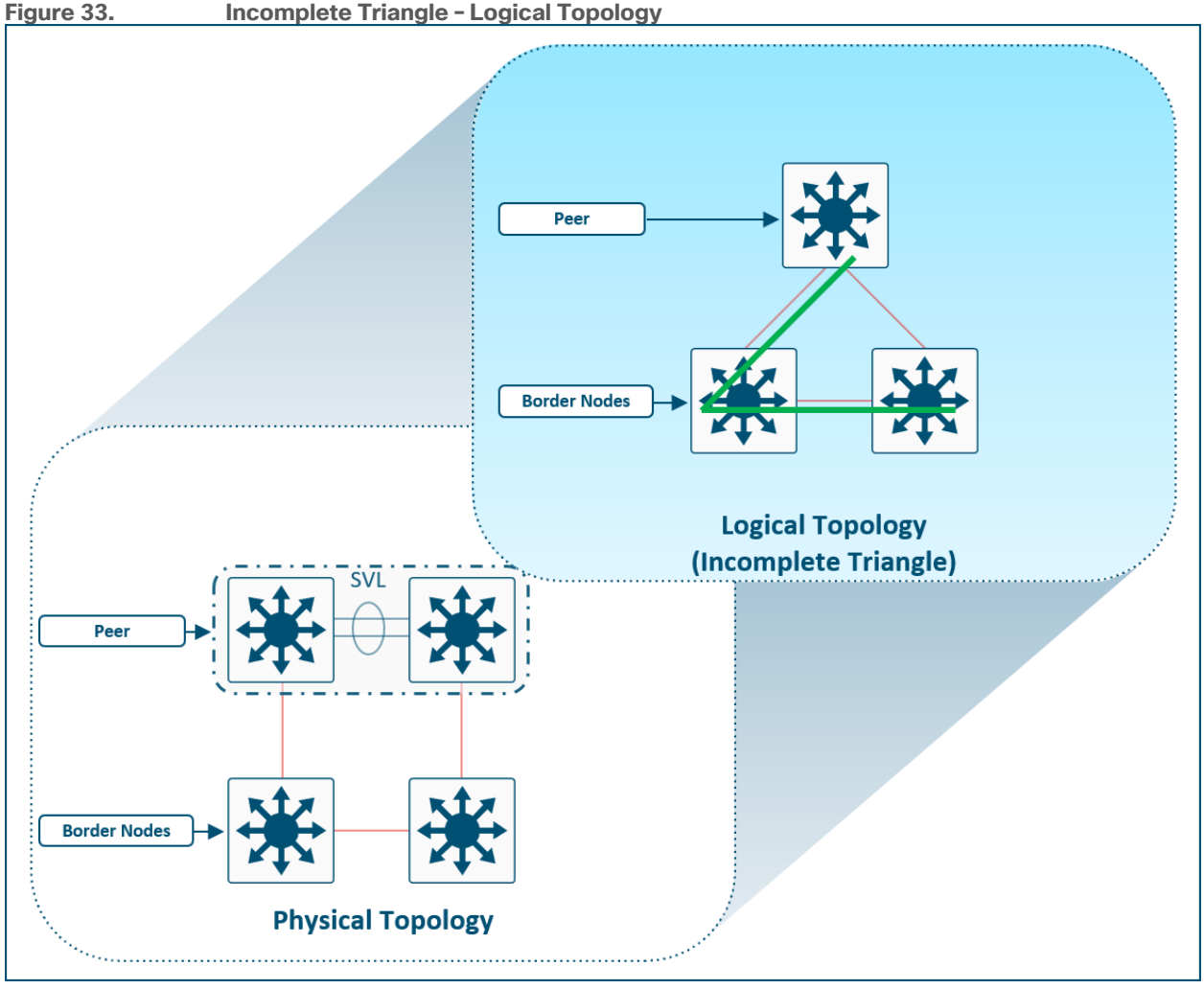
redundant paths for the best deterministic convergence. Border nodes should be deployed in pairs and should each connect to a pair of upstream devices. Border nodes should have a crosslink between each other. If the upstream infrastructure is within the administrative domain of the network operator, these devices should be crosslinked to each other. Border nodes of the same type, such as *internal* and *external*, should be fully meshed.

In some deployments, the upstream device from border nodes may be a single logical unit represented by two or more devices such as VSS, SVL, or even a firewall cluster. To build triangle topologies, the border nodes should be connected to each device in the logical unit.

Tech tip
Figures 33-36 below show the peer device as a StackWise Virtual device, although the failover scenarios represented are also applicable to Active-Standby Firewalls and other HA upstream pairs.

In Figure 33 below, the physical topology use *squares* to connect the devices. Each border node is connected to only one member of upstream logical peer. The border nodes are crosslinked to each other which provides an indirect and non-optimal forwarding path in the event of an upstream link failure. The resulting logical topology is an incomplete triangle.

Figure 33. Incomplete Triangle - Logical Topology



In Figure 34 below, the physical topology uses *triangles* to connect the devices. Each border node is connected to each member of the upstream logical peer. The border nodes are crosslinked to each other. The resulting

logical topology is the same as the physical, and a complete triangle is formed. This is the recommended approach.

Figure 34. Complete Triangle – Logical Topology

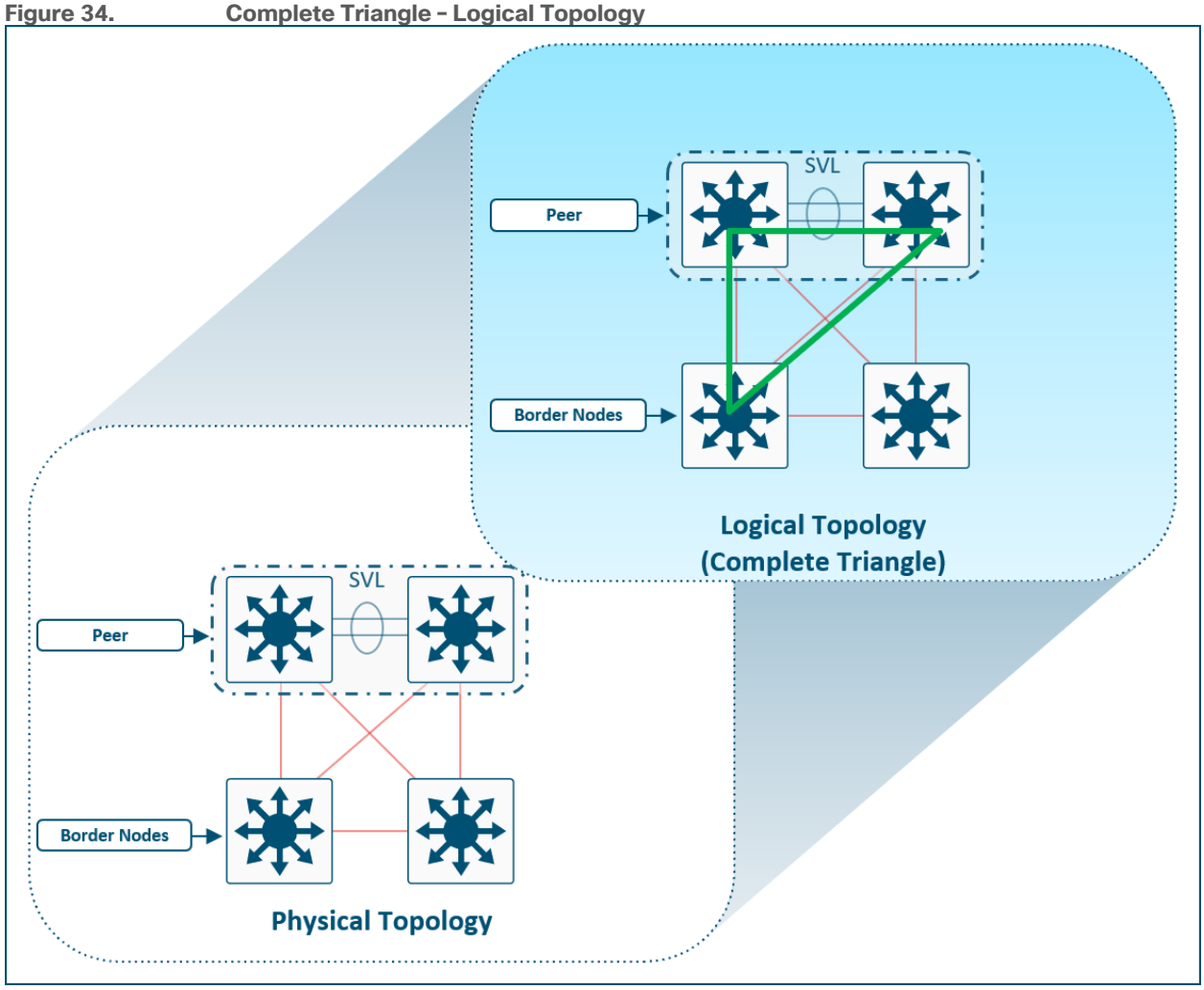
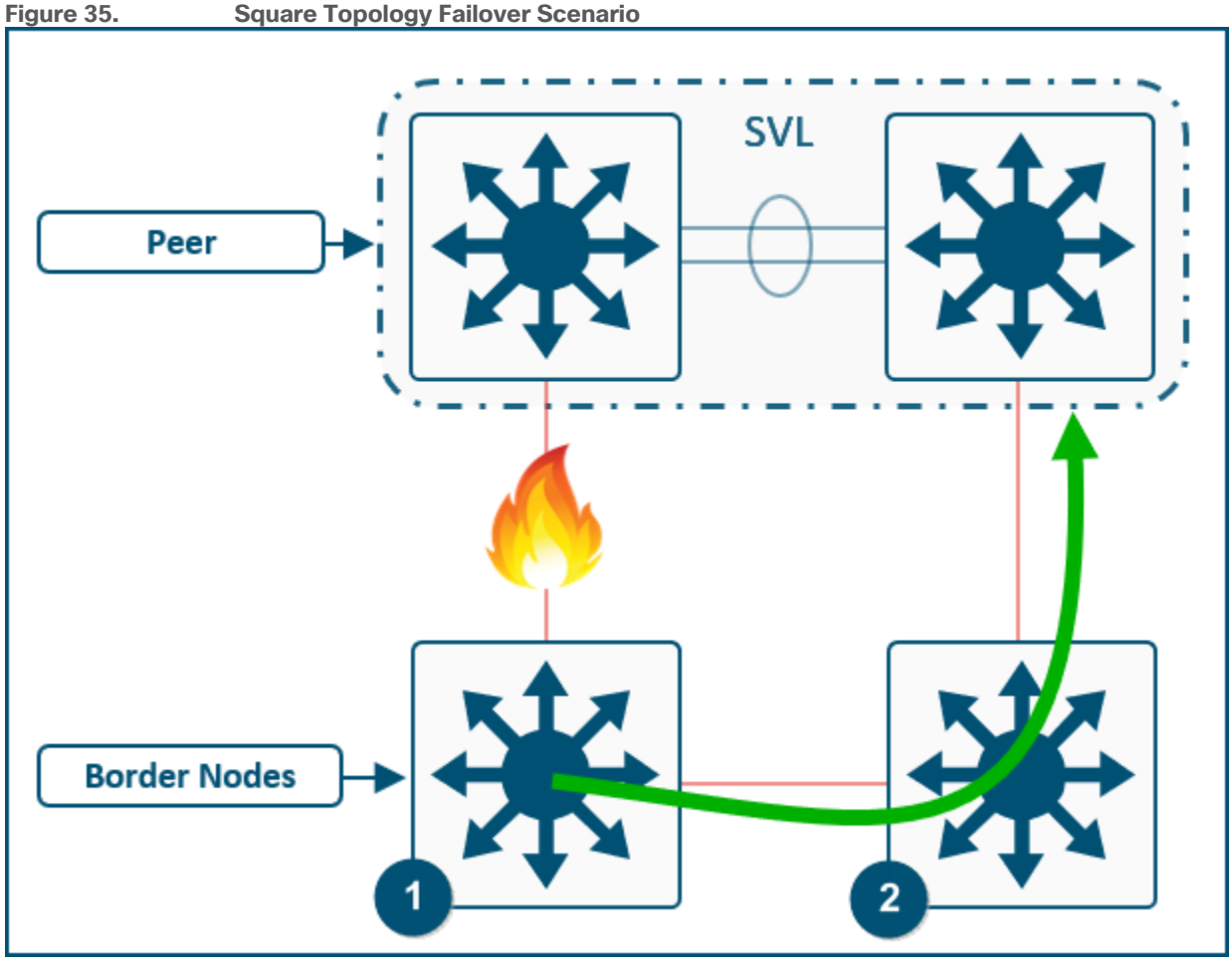


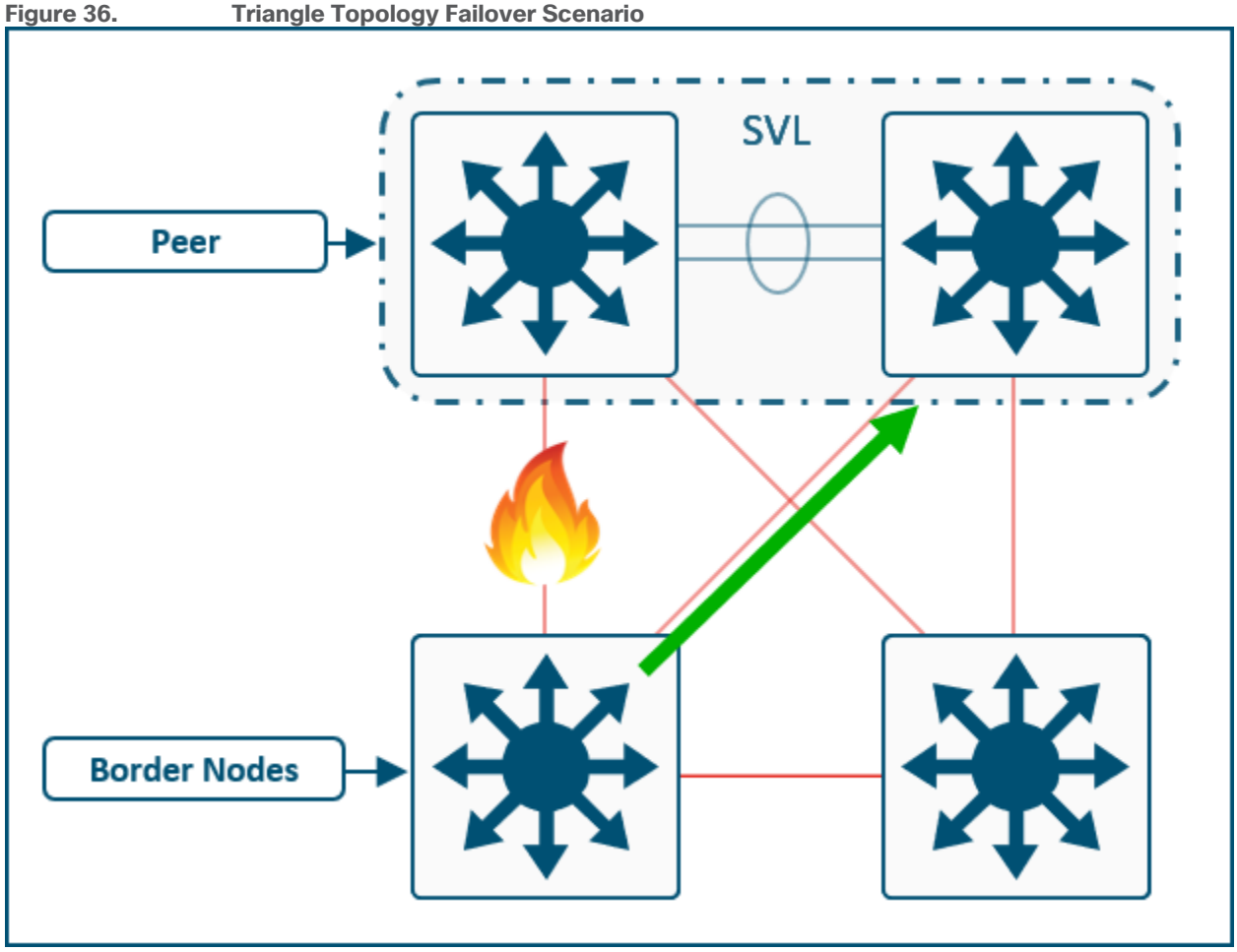
Figure 35 below shows a pair of border node connected to a StackWise Virtual upstream peer. If the link to one StackWise member has a failure scenario, IP reachability still exists, but Border Node #1 must traverse Border Node #2 to reach destinations beyond the upstream peer. And while IP reachability still exists, it is an inefficient forwarding path that requires VRF-awareness (VRF-lite) between the redundant borders to achieve. This topology example represents a single point of failure akin to having a single upstream device from the redundant border nodes.

Figure 35. Square Topology Failover Scenario



In contrast, as shown in Figure 36 below, if the border nodes are connected to both StackWise peers, even in the event of a single member failure, each border node will still have an optimal, redundant forwarding path.

Figure 36. Triangle Topology Failover Scenario



Security Policy Design Considerations

This section is organized into the following subsections:

Section	Subsection
Security Policy Design Considerations	Aspects and Design Criteria

A one-size-fits-all security design is not desirable—security requirements vary by organizations. Security designs are driven by information security policies and legal compliance. The planning phase for a security design is key to ensuring the right balance of security and user experience. The following aspects should be considered when designing security policy for the SD-Access network:

- **Openness of the network**—Some organizations allow only organization-issued devices in the network, and some support a *Bring Your Own Device* (BYOD) approach. Alternatively, user choice can be balanced with allowing easier-to-manage endpoint security by deploying a *Choose Your Own Device* (CYOD) model in which a list of IT-approved endpoints is offered to the users for business use. An identity-based approach is also possible in which the network security policies deployed depend on the device ownership. For example, organization-issued devices may get group-based access, while personal devices may get Internet-only access.
- **Identity management**—In its simplest form, identity management can be a username and password used for authenticating users. Adding embedded security functions and application visibility in the network

provides telemetry for advanced policy definitions that can include additional context such as physical location, device used, type of access network (wired, wireless, VPN), application used, and time of day.

- **Authentication, Authorization, and Accounting (AAA) policies**—Authentication is the process of establishing and confirming the identity of a client requesting access to the network. Authorization is the process of authorizing access to some set of network resources. Accounting is process of recording what was done and accessed by the client.

  For unified experience for wired and wireless endpoints, AAA policies in SD-Access are enforced at the access layer (edge nodes) with the use of SGACLs for segmentation within VNs and dynamic VLAN assignment for mapping endpoints into VNs. Event logs, ACL hit counters, RADIUS accounting, and similar standard accounting tools are available to enhance visibility.

Tech tip
Cisco IOS® Software enhances 802.1X device capabilities with Cisco Identity Based Networking Services (IBNS) 2.0. For example, concurrent authentication methods and interface templates have been added. Likewise, Cisco DNA Center has been enhanced to aid with the transition from IBNS 1.0 to 2.0 configurations, which use Cisco Common Classification Policy Language (commonly called C3PL). See the release notes and updated deployment guides for additional configuration capabilities. For more information about IBNS, see: https://www.cisco.com/go/ibns.

- **Endpoint security**—Endpoints can be infected with malware, compromising data and creating network disruptions. Malware detection, endpoint management, and data exports from the network devices provide insight into endpoint behavior. Tight integration with security appliances such as Cisco Adaptive Security Appliances (ASA) and Cisco Firepower Threat Defense (FTD) and analytics platforms such as Stealthwatch and Cognitive Threat Analytics (CTA) enables the network to have the intelligence to quarantine and help remediate compromised devices.
- **Data integrity and confidentiality**—Network segmentation using VNs can control access to applications such as separating employee transactions from IoT traffic.
- **Network device security**—Hardening security of network devices is essential. The use of the secure device management options, such as enabling device authentication using TACACS+ and disabling unnecessary services, are best practices to ensure the network devices are secured.

Enabling group-based segmentation within each virtual network allows for simplified hierarchical network policies. Network-level policy scopes of isolated control and data planes are possible using VNs, while group-level policy scopes are possible using SGTs within VNs, enabling common policy application across the wired and wireless fabric.

SGTs tag endpoint traffic based on a role or function within the network such that the traffic is subject to role-based policies or SGACLs centrally defined within ISE which references Active Directory, for example, as the identity store for user accounts, credentials, and group membership information. Endpoints can be classified based on that identity store information and can be assigned to an appropriate scalable group. These scalable groups can then be used to create segmentation policies and virtual network assignment rules.

SGT information is carried across the network in several forms:

- **Inside the SD-Access fabric**—The SD-Access fabric header transports SGT information. Fabric edge nodes and border nodes can enforce SGACLs to enforce the security policy.
- **Outside the fabric on a device with Cisco TrustSec capability**—Inline devices with Cisco TrustSec capability carry the SGT information in a CMD header on the Layer 2 frame. This is the recommended mode of transport outside the SD-Access network.

- **Outside the fabric over devices without Cisco TrustSec capability**—SXP allows the control plane communication of SGT to IP mappings over a TCP connection. This can be used to communicate SGTs over network devices that do not support SGT inline tagging.

Tech tip
For additional security policy design considerations, please see the SD-Access Segmentation Design Guide.

Multidimensional Considerations

This section is organized into the following subsections:

Section	Subsection
Multidimensional Considerations	Greenfield and brownfield Number of Users Number of Fabric Devices Geography Shared Services VRF-Aware Peers (Fusion Devices) WAN and Internet Connectivity Unified Policy End-to-End Macro Segmentation End-to-End Micro Segmentation Site Survivability High Availability

An SD-Access network begins with a foundation of the *Cisco Enterprise Architecture Model* with well-designed and planned hierarchical network structures that include modular and extensible network blocks as discussed in the LAN Design Principles section. On this foundation, the network is designing and configured using the Layer 3 routed access model.

While individual sites can have some design and configuration that is independent from other locations, this design and configuration must consider how the site becomes part of the larger campus network including other fabric sites, non-fabric sites, shared services, data center, WAN, and Internet. No element, consideration, or fabric site should be viewed in isolation, and an end-to-end view of the network must be taken into account.

Beyond the business needs, business drivers, and previous listed Design Considerations, additional technical factors must be considered. The results of these technical considerations craft the framework for the topology and equipment used in the network. These factors are multi-dimensional and must be considered holistically. The design strategy for SD-Access is to maximize site size while minimizing site count. Each of the factors below could drive the need to deploy multiple, smaller fabric sites rather than one larger one.

Greenfield and Brownfield

Greenfield networks have the advantage that the network can be designed as new from the ground up. Brownfield networks may have less flexibility due to geography, fiber, or existing configurations. Ultimately, the goal in brownfield environment is to use it in as an SD-Access network, and careful and accurate information, configuration, and topology details for the existing network should be collected in advance to migration.

Migration from a traditional network to an SD-Access network can be accomplished through the following approaches:

- **Layer 2 Handoff**—This feature of connects a traditional network with an SD-Access network. This feature can be used during transitions and migrations in concert with the following approach.
- **Building by building**—Areas of the existing network are converted to SD-Access. This is commonly done closet by closet (IDF by IDF) or building by building. Migration is done, at minimum, one switch at a time. One VLAN at a time is not supported, as the VLAN may span multiple traditional switches.

Tech tip
Layer 2 border handoff considerations are discussed further in Migration section.

Number of Users

The most significant factor in the selection of equipment and topology for a site, apart from existing wiring, is total number of wired and wireless clients in that location. This will determine the number of physical switch ports and access points required which will determine the need for three-tier or two-tier network designs. The number of clients may be small enough that the network is composed of a switch stack or large enough to cover multiple buildings with many thousands of endpoints.

Number of Fabric Devices

The number of fabric devices in a site is a count of all of routers, switches, classic and policy extended nodes, and wireless controllers that are operating in a fabric role. Cisco DNA Center can support a specific number of network devices in total and also a maximum number per fabric site. Each of these scale numbers varies based on the appliance size, and it may also vary by release. The maximum number of devices may be a reason to create several smaller fabric sites rather than one very large site. Please consult Cisco DNA Center Appliance: Scale and Hardware Specifications on the Cisco DNA Center data sheet for the specific maximum number of fabric device per site for the current release.

Geography

Physical geography impacts the network design. It may not have a direct impact on the topology within the fabric site itself, but geography must be considered as it relates to transit types, services locations, survivability, and high availability. Geography impacts the end to end design and the fabric domain.

Locations that are situated within the same metro area (MAN) or campus with multiple buildings in close, physical proximity with interconnect direct fiber can benefit from a SD-Access for Distributed Campus design. A Distributed Campus deployment, by extension, allows for native, unified policy across the locations as well as with the potential to have a single services block location.

Locations connected across WAN or Internet circuits, where the fabric packet is de-encapsulated as it leaves the fabric, must consider shared services location, methods to maintain unified policy constructs across the circuits, and consider the routing infrastructure outside of the fabric.

Shared Services

Services such as DHCP, DNS, ISE, and WLCs are required elements for clients in an SD-Access network. Services are commonly deployed in one of three ways.

- **Fabric Site Local**—For survivability purposes, a services block may be established at each fabric site location. Local services ensure that these critical services are not sent across the WAN/MAN/Internet and ensure the endpoints are able to access them, even in the event of congestion or unavailability of the

external circuit. However, this may drive the need for VRF-aware peering devices to *fuse* routes from the fabric overlay to shared services.

- **Centralized within the Deployment**—In locations distributed across a WAN and in SD-Access for Distributed Campus deployments, services are often deployed at on-premises data centers. These data centers are commonly connected to the core or distribution layers of a centralized location such as a headquarters. Traffic is sent from the remote and branch sites back to the central location, and then directed towards the necessary services.
- **Both Centralized and Fabric-Site Local**—This is a hybrid of the two approaches above. For most fabric sites, services are centralized. Specific fabric sites with a need for services connectivity independent of the status of the WAN circuit use local services.

VRF-Aware Peer (Fusion Devices)

While not a specific reason factor in the decision to deploy multiple fabric sites, shared services must be considered as part of the deployment. A VRF-Aware peer (fusion device) is the most common deployment method to provide access to shared services. For fabric sites needing resiliency, high availability, and site survivability independent of WAN status, local shared services are needed. These locations should plan for the use of a services block and VRF-aware peer to provide the fabric endpoint access to these services.

WAN and Internet Connectivity

External Internet and WAN connectivity for a fabric site has a significant number of possible variations. The key design consideration is to ensure the routing infrastructure has the physical connectivity, routing information, scale, performance, and throughput necessary to connect the fabric sites to the external world.

Unified Policy

Unified policy is a primary driver for the SD-Access solution. With unified policy, access control for wired and wireless traffic is consistently and uniformly enforced at the access layer (fabric edge node). Users, devices, and applications are subject to the same policy wherever and however they are connected in the network.

Within a fabric site, unified policy is both enabled and carried through the Segment ID (Group Policy ID) and Virtual Network Identifier (VNI) fields of the VXLAN-GPO header. This allows for both VRF (macro) and SGT (micro) segmentation information to be carried within the fabric site.

Tech tip
Low-level details on the fabric VXLAN header can be found in Appendix A.

In SD-Access for Distributed Campus, the same encapsulation method used for data packets within the fabric site is used for data packets between sites. This allows unified policy information to be natively carried in the data packets traversing between fabric sites in the larger fabric domain.

When designing for a multi-site fabric that uses an IP-based transit between sites, consideration must be taken if a unified policy is desired between the disparate locations. Using an IP-based transit, the fabric packet is de-encapsulated into native IP. This results in loss of embedded policy information. Carrying the VRF and SGT constructs without using fabric VXLAN, or more accurately, once VXLAN is de-encapsulated, is possible through other technologies, though.

End-to-End Macro Segmentation (VRFs)

Segmentation beyond the fabric site has multiple variations depending on the type of transit. SD-Access transit carries the VRF natively. In IP-based transits, due to de-encapsulation of the fabric packet, VN policy

information can be lost. Several approaches exist to carry VN (VRF) information between fabric sites using an IP-based transit. The most straightforward approach is to configure VRF-lite hop-by-hop between each fabric site. While this is the simplest method, it also has the highest degree of administrative overhead. This method is not commonly utilized, as the IP-based infrastructure between fabric sites is generally under the administrative control of a service provider.

If VRF-lite cannot be used end to end, options still exist to carry VRFs. The supported options depend on if a *one-box method* or *two-box method* is used.

- ***One-Box*** **Method**– The internal and external routing domain are on the same border node.
- ***Two-Box*** **Method**–The internal and external routing domains are on two different boxes. The internal routing domain is on the border node. The external routing domain is on upstreaming routing infrastructure. BGP is used to exchange the reachability information between the two routing domains.

One-Box Method Designs

One-box method designs require the border node to be a routing platform in order to support the applicable protocols. This configuration is done manually or by using templates.

- **Border Node with DMVPN**–On the border node router, a DMVPN cloud is configured per fabric VN.
- **Border Node with IPSec Tunnels**–On the border node router, an IPsec tunnel is configured per fabric VN.

Two-Box Method Designs

All *two-box method* designs begin with a VRF-lite handoff on the border node. The VRF is associated with an 802.1Q VLAN to maintain the segmentation construct. The *two-box* design can support a routing or switching platform as the border node. However, the peer device needs to be a routing platform to support the applicable protocols. The handoff on the border node can be automated through Cisco DNA Center, though the peer router is configured manually or by using templates.

- **Border Node with GRE Peer**–A VRF is handed off via a VLAN to a GRE tunnel endpoint router. On the tunnel endpoint router, one GRE tunnel is configured per fabric VN.
- **Border Node with DMVPN Peer**–A VRF is handed off via a VLAN to a DMVPN router. On the DMVPN router, one DMVPN cloud is configured per fabric VN.
- **Border Node with IPSec Peer**–A VRF is handed off via a VLAN to an IPSec router. On the IPSec router, one IPsec tunnel is configured per fabric VN.
- **Border Node with MP-BGP Peer**– A VRF is handed off via a VLAN to a peer supporting multiprotocol BGP such as MPLS provider. BGP needs a VRF-Aware data plane such as MPLS to have a mechanism to carry the VRF attributes.

Tech tip
For additional details on the supported the *One-Box* and *Two-Box* designs listed above, please see Real World Route/Switch to Cisco SD-Access Migration Tools and Strategies – BRKCRS-3493 (2020, APJC).

End-to-End Micro segmentation (SGTs)

Like VRFs, segmentation beyond the fabric site has multiple variations depending on the type of transit. SD-Access transit carries the SGT natively. In IP-based transit, due to the de-encapsulation of the fabric packet, SGT policy information can be lost. Two approaches exist to carry SGT information between fabric sites using an IP-based transit, inline tagging and SXP.

Inline Tagging

Inline tagging is the process where the SGT is carried within a special field known as CMD (Cisco Meta Data) that can be inserted in the header of the Ethernet frame. This changes the EtherType of the frame to 0x8909. If the next-hop device does not understand the this EtherType, the frame is assumed to be malformed and is discarded. Inline tagging can propagate SGTs end to end in two different ways.

- **Hop by Hop**—Each device in the end to end chain would need to support inline tagging and propagate the SGT.
- **Preserved in Tunnels**—SGTs can be preserved in CMD inside of GRE encapsulation or in CMD inside or IPsec encapsulation.

SGT Exchange Protocol over TCP (SXP)

A second design option is to use SXP to carry the IP-to-SGT bindings between sites. SXP is used to carry SGTs across network devices that do not have support for Inline Tagging or if the tunnel used is not capable of caring the tag.

SXP has both scaling and enforcement location implications that must be considered. Between fabric sites, SXP can be used to enforce the SGTs at either the border nodes or at the routing infrastructure north bound of the border. If enforcement is done at the routing infrastructure, CMD is used to carry the SGT information inline from the border node.

If enforcement is done on the border node, a per-VRF SXP peering must be made with each border node to ISE. A common way to scale SXP more efficiently is to use SXP domains. A second alternative is to peer the border node with a non-VRF-Aware Peer and merge the routing tables. ISE then makes a single SXP connection to each of these peers.

Tech tip
For additional details on deployment scenarios, SGTs over GRE and VPN circuits, and scale information, please see the SD-Access Segmentation Design Guide.

Site Survivability

In the event that the WAN and MAN connections are unavailable, any service accessed across these circuits are unavailable to the endpoints in the fabric. The need for site survivability is determined by balancing the associated costs of the additional equipment and the business drivers behind the deployment while also factoring in the number of impacted users at a given site. Designing an SD-Access network for complete site survivability involves ensuring that shared services are local to every single fabric site. Generally, a balance between centralized and site-local services is used. If a given fabric site has business requirements to always be available, it should have site-local services. Other fabric sites without the requirement can utilize centralized services for the fabric domain.

High Availability

High availability compliments site survivability. A site with single fabric border, control plane node, or wireless controller risks single failure points in the event of a device outage. When designing for high availability in an SD-Access network, it is important to understand that redundant devices do not increase the overall scale. Redundant control plane nodes and redundant border nodes operate in an active-active method, and Fabric WLCs operate as active-standby pairs.

SD-Access Site Reference Models

This chapter is organized into the following sections:

Chapter	Section
SD-Access Site Reference Models	Design Strategy Fabric in a Box Site Reference Model Very Small Site Reference Model Small Site Reference Model Medium Site Reference Mode Large Site Reference Model SD-Access for Distributed Campus Model

Designing Cisco SD-Access fabric site has flexibility to fit many environments, which means it is not a one-design-fits-all proposition. The scale of a fabric can be as small a single switch or switch stack or as big as one or more three-tier campus deployments. SD-Access topologies should follow the same design principles and best practices associated with a hierarchical design, such splitting the network into modular blocks and distribution of function, as described in the Campus LAN and Wireless LAN Design Guide.

Design elements should be created that can be replicated throughout the network by using modular designs. In general, SD-Access topologies should be deployed as spoke networks with the fabric border node as the exit point hub for the spokes which are the access switches operating as edge nodes. As networks grow, varied physical topologies are used to accommodate requirements for specialized network services deployment.

Fabric Site Sizes – Design Strategy

A practical goal for SD-Access designs is to create larger fabric sites rather than multiple, smaller fabric sites. The design strategy is to maximize fabric site size while minimizing total site count. Some business requirements will necessitate splitting locations into multiple sites such as creating a fabric site for an Emergency Room (ER) that is separate from the fabric site that is represented by the remainder of the hospital. The multidimensional factors of survivability, high availability, number of endpoints, services, and geography are all factors that may drive the need for multiple, smaller fabric sites instead of a single large site. To help aid in design of fabric sites of varying sizes, the Reference Models below were created.

Fabric Site Reference Models

In deployments with physical locations, customers use different templates for each of the different site types such as a large branch, a regional hub, headquarters, or small, remote office. The underlying design challenge is to look at existing network, deployment, and wiring, and propose a method to layer SD-Access fabric sites in these areas. This process can be simplified and streamlined by templatizing designs into reference models. The templates drive understanding of common site designs by offering reference categories based on the multidimensional design elements along with endpoint count to provide guidelines for similar site size designs. The numbers are used as guidelines only and do not necessarily match maximum specific scale and performance limits for devices within a reference design.

- **Fabric in a Box site**—Uses Fabric in a Box to cover a single fabric site, with resilience supported by switch stacking or StackWise Virtual; designed for less than 200 endpoints, less than 5 VNs, and less than 40 APs; the border, control plane, edge, and wireless functions are colocated on a single redundant platform.

- **Very small site**– Covers a single office or building; designed to support less than 2,000 endpoints, less than 8 VNs, and less than 100 APs; the border is colocated with the control plane function on one or two devices and uses embedded wireless or optionally hardware WLCs.
- **Small site**–Covers a single office or building; designed to support less than 10,000 endpoints, less than 32 VNs, and less than 200 APs; the border is colocated with the control plane function on one or two devices and a separate wireless controller has an optional HA configuration.
- **Medium site**–Covers a building with multiple wiring closets or multiple buildings; designed to support less than 25,000 endpoints, less than 50 VNs, and less than 1,000 APs; the border is distributed from the control plane function using redundant devices, and a separate wireless controller has an HA configuration.
- **Large site**–Covers a large building with multiple wiring closets or multiple buildings; designed to support less than 50,000 endpoints, less than 64 VNs, and less than 2,000 APs; multiple border exits are distributed from the control plane function on redundant devices, and a separate wireless controller has an HA configuration.

Each fabric site includes a supporting set of control plane nodes, edge nodes, border nodes, and wireless LAN controllers, sized appropriately from the listed categories. ISE Policy Service Nodes are also distributed across the sites to meet survivability requirements.

Tech tip
The guideline numbers for the site reference sizes are based on the design strategy to maximize site size and minimize site count. These guidelines target an approximate ~75% of specific scale numbers as documented on Table 10 and Table 12 of the Cisco DNA Center data sheet, and the specifics are noted in each reference site section. The important concept in fabric site design is to allow for future growth by not approaching any specific scale limit on *Day 1* of the deployment. A fabric site can only support a maximum of four border nodes provisioned as *external* borders. A fabric site with SD-Access Wireless can only support two control plane nodes for non-guest (Enterprise) traffic as discussed in the Wireless Design section and shown in Figure 20.

Fabric in a Box Site Reference Model

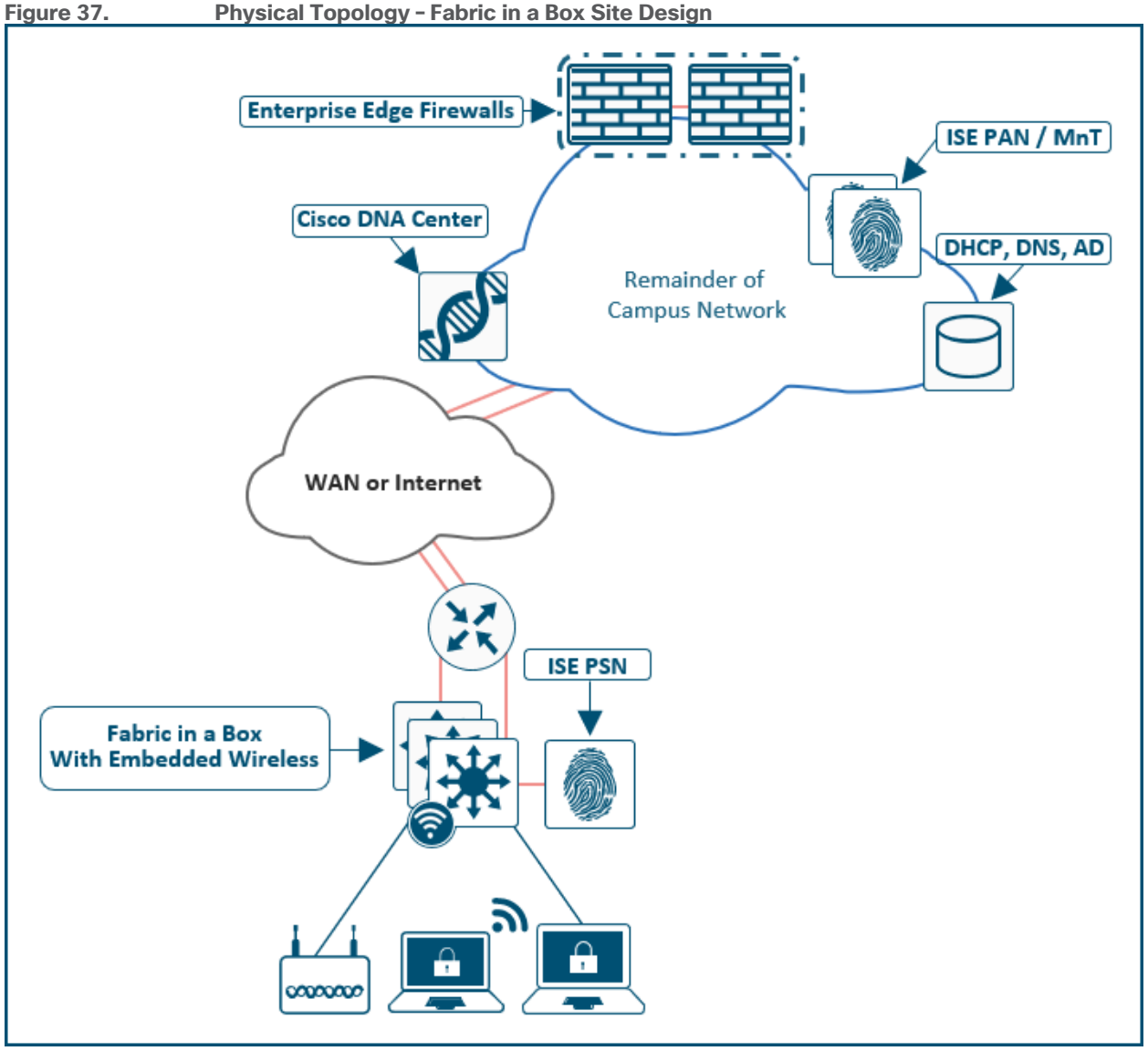
The Fabric in a Box Site Reference Model should target less than 200 endpoints. The central component of this design is a switch stack or StackWise Virtual operating in all three fabric roles: control plane node, border node, and edge node. For switch stack Fabric in a Box deployments, SD-Access Embedded Wireless is used to provide site-local WLC functionality. The site may contain an ISE PSN depending on the WAN/Internet circuit and latency.

Use the table below to understand the guidelines to stay within for similar site design sizes. The numbers are used as guidelines only and do not necessarily match specific limits for devices used in a design of this site size.

Table 2. Fabric in a Box Site Guidelines (Limits may be different)

Endpoints, target fewer than	200
Control plane nodes, colocated	1
Border nodes, colocated	1
Virtual networks, target fewer than	5
IP pools, target fewer than	8
Access points, target fewer than	40

Figure 37. Physical Topology – Fabric in a Box Site Design



Fabric in a Box Site Considerations

Due to the smaller number of endpoints, and so implied lower impact, high availability and site survivability are not common requirements for a Fabric in a Box design. As with all the reference designs, site-local services of

DHCP, DNS, WLCs, and ISE can provide resiliency and survivability although at the expense of increased complexity and equipment such as a services block.

If shared services are deployed locally, the peer device is commonly a switch directly connected to the Fabric in a Box with services deployed as virtual machines on Cisco UCS C-Series Server. An alternative is to deploy a UCS E-series blade servers on the routing infrastructure to virtualize the shared services.

High availability in this design is provided through StackWise-480 or StackWise Virtual which both combine multiple physical switches into a single logical switch. StackPower is used to provide power redundancy between members in a switch stack. StackWise Virtual deployments have power redundancy by using dual power supplies in each switch. If a chassis-based switch is used, high availability is provided through redundant supervisors and redundant power supplies. To support power redundancy, available power supplies would need to be redundant beyond the needs of the switch to support power chassis, supervisor, and line cards.

Wireless LAN controllers can be deployed as physical units directly connected to the Fabric in a Box or deployed as the embedded Catalyst 9800 controller. When using the embedded Catalyst 9800 with a switch stack or redundant supervisor, AP and Client SSO (Stateful Switch Over) are provided automatically. StackWise Virtual deployments of Fabric in a Box need physical WLCs.

Tech tip
Client SSO provides the seamless transition of clients from the active controller to the standby controller. Client information is synced from the Active to the Standby, so client re-association is avoided during a switchover event. For additional information on Client and AP SSO, please see the WLC High Availability (SSO) Technical Reference.

When using stacks, links to the upstream routing infrastructure should be from different stack members. Ideally, the uplinks should be from the member switches rather than the stack master. With chassis switches, links should be connected through different supervisors. To prepare for border node handoff automation along with having initial IP reachability, SVIs and trunk links are commonly deployed between the small site switches and the upstream routing infrastructure.

The Catalyst 9300 Series in a stack configuration with the embedded Catalyst 9800 Series wireless LAN controller capabilities is an optimal platform in this design. Other available platforms such as the Catalyst 9500 Series can be deployed as StackWise Virtual and can provide connectivity options such as SFP+ (10 Gigabit Ethernet) and multi-chassis redundancy capabilities.

Very Small Site Reference Model

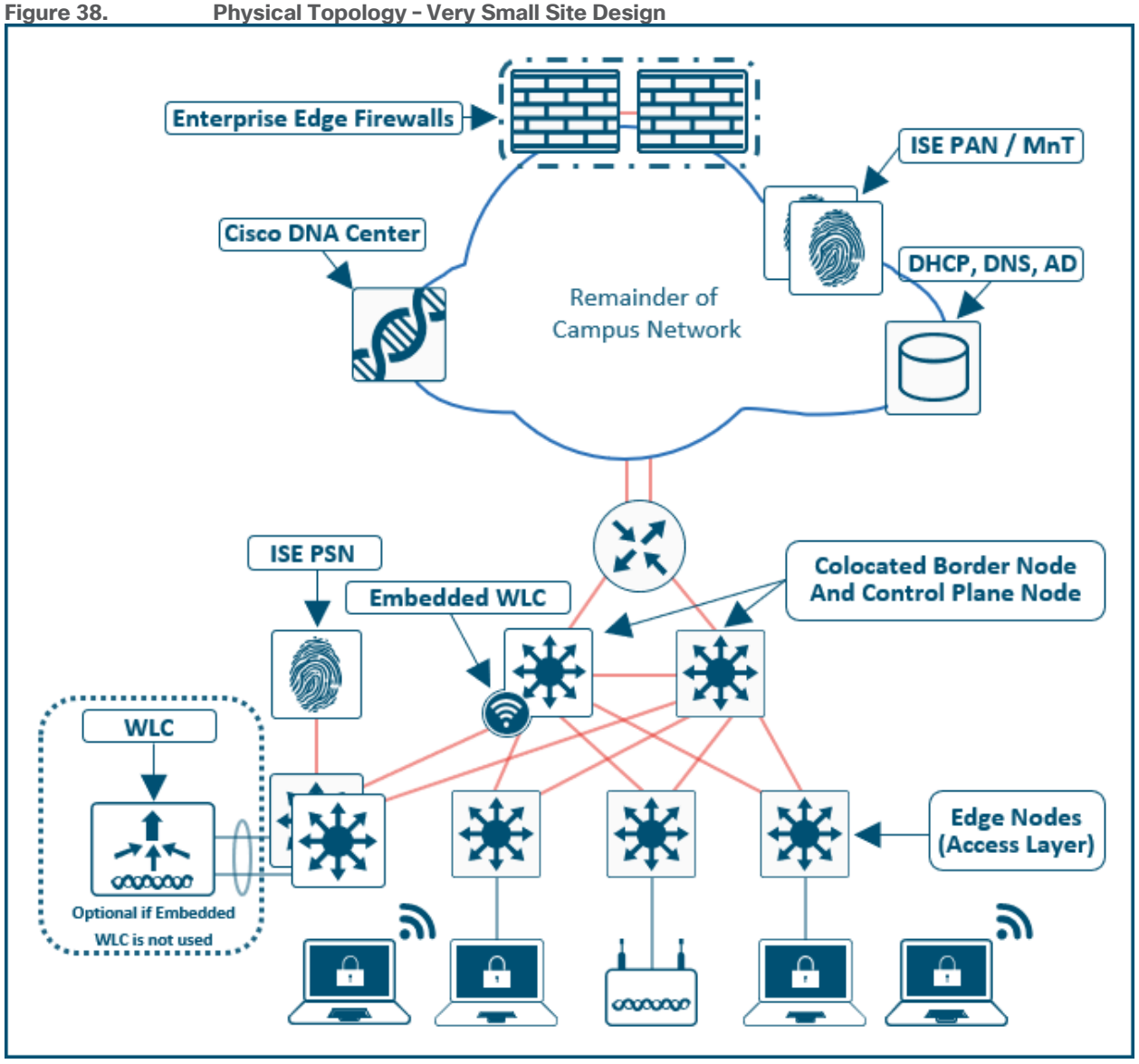
The Very Small Site Reference Model should target less than 2,000 endpoints. The physical network is usually a two-tier collapsed core/distribution with an access layer servicing several wiring closets. Rather than colocating all roles in one device, the Very Small Site Reference Model provides added resiliency and redundancy along with a larger number of endpoints by separating the edge node role onto dedicated devices in the access layer. The border and control plane node are colocated in the collapsed core layer. For SD-Access Wireless, the embedded WLC is provisioned on one of the colocated border and control plane nodes. Optionally, a virtual or hardware-based WLC is used.

Use the table below to understand the guidelines to stay within for similar site design sizes. The numbers are used as guidelines only and do not necessarily match specific limits for devices used in a design of this site size. The target maximum number of endpoints is based on approximately ~50% of the number endpoints supported by the Catalyst 9800 Embedded Wireless controller as documented on the Cisco Access Point and Wireless Controller Selector.

Table 3. Very Small Site Reference Guidelines (Limits may be different)

Endpoints, target fewer than	2,000
Fabric nodes, target fewer than	50
Control plane nodes, colocated	2
Border nodes, colocated	2
Virtual networks, target fewer than	8
IP pools, target fewer than	20
Access points, target fewer than	100

Figure 38. Physical Topology – Very Small Site Design



Very Small Site Considerations

For very small deployments, an SD-Access fabric site is implemented using a two-tier design. The same design principles for a three-tier network are applicable, though there is no need for an distribution layer (intermediate nodes). In a very small site, high availability is provided in the fabric nodes by colocating the border node and control plane node functionality on the collapsed core switches and deploying these as a pair. For both resiliency and alternative forwarding paths in the overlay and underlay, the collapsed core switches should be directly to each other with a crosslink.

Provided there are less than 200 APs and 4,000 clients, SD-Access Embedded wireless can be deployed along with the colocated border node and control plane node functions on a collapsed core switch. For high-availability for wireless, a hardware or virtual WLC should be used.

Small Site Design Reference Model

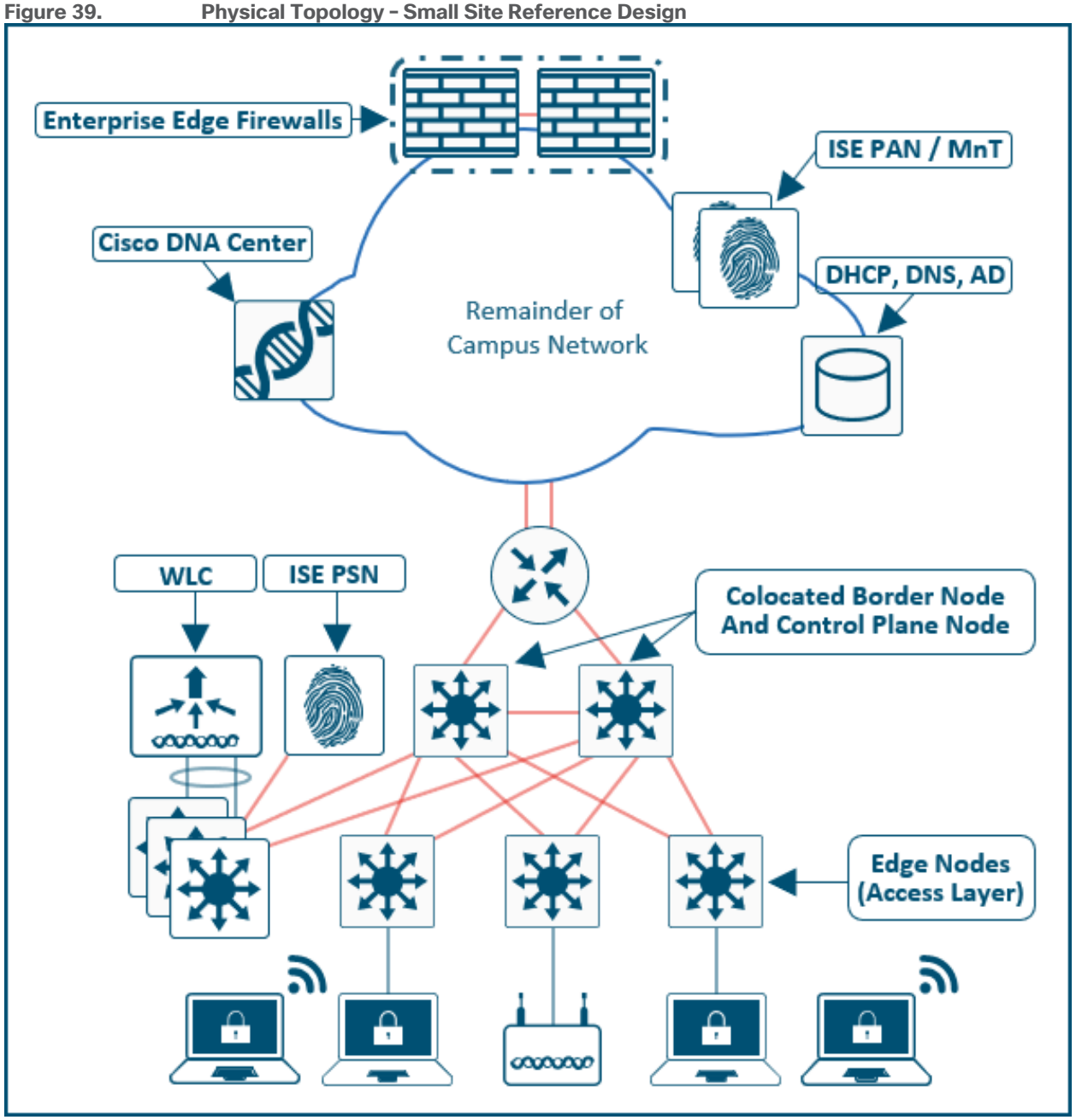
The Small Site Reference Model covers a building with multiple wiring closets or multiple buildings and typically has less than 10,000 endpoints. The physical network is usually a two-tier collapsed core/distribution with an access layer.

Use the table below to understand the guidelines to stay within for similar site design sizes. The numbers are used as guidelines only and do not necessarily match specific limits for devices used in a design of this site size. The target maximum number of Virtual Networks is approximately ~50% of the number of VNs supported by entry and mid-sized Cisco DNA Center appliance as listed on Table 10 of its data sheet.

Table 4. Small Site Guidelines (Limits may be different)

Endpoints, target fewer than	10,000
Fabric nodes, target fewer than	75
Control plane nodes	2
Border nodes	2
Virtual networks, target fewer than	32
IP pools, target fewer than	100
Access points, target fewer than	200

Figure 39. Physical Topology – Small Site Reference Design



Small Site Considerations

For smaller deployments, an SD-Access fabric site is implemented using a two-tier design. The same design principles for a three-tier network applicable, though there is no need for an aggregation layer (intermediate nodes). In a small site, high availability is provided in the fabric nodes by colocating the border node and control plane node functionality on the collapsed core switches and deploying these as a pair. For both resiliency and alternative forwarding paths in the overlay and underlay, the collapsed core switches should be directly to each other with a crosslink.

The client and access point count calls for use of dedicated WLCs either in hardware or virtual machines. To enable highly-available links for WLC through physical connectivity, a services block is deployed. The WLCs are connected to the services block switch through Layer 2 port-channels to provide redundant interfaces. The services block is switch stack or SVL that is connected to both collapsed core switches through Layer 3 routed

links. This services block is deployed as a VRF-aware peer if DHCP/DNS and other shared services are site-local.

Medium Site Design Reference Model

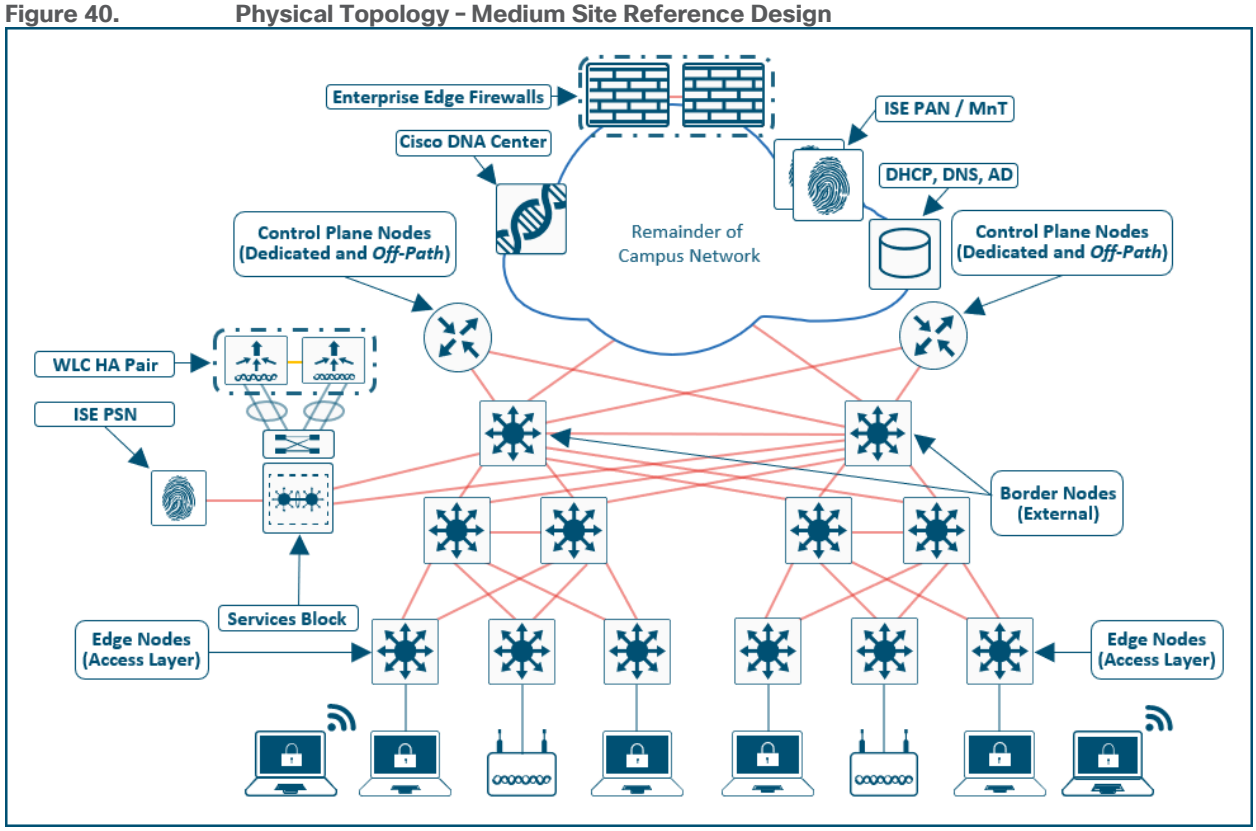
The Medium Site Reference Model covers a building with multiple wiring closets or multiple buildings and is designed to support less than 25,000 endpoints. The physical network is usually a three-tier network with core, distribution, and access layers. It may even contain a routed *super-core* that aggregates multiple buildings and serves as the network egress point to the WAN and Internet. The border and control plane node functionality are provisioned on separate devices rather than colocating.

Use the table below to understand the guidelines to stay within for similar site design sizes. The numbers are used as guidelines only and do not necessarily match specific limits for devices used in a design of this site size. The target maximum endpoint count requires, at minimum, the mid-size Cisco DNA Center appliance to provide for future growth. The maximum fabric nodes and virtual networks are approximately ~75% of the number supported the mid-sized Cisco DNA Center appliance as listed on Table 10 of its data sheet.

Table 5. Medium Site Guidelines (Limits may be different)

Endpoints, target fewer than	25,000
Fabric nodes, target fewer than	450
Control plane nodes (limit to 2 Enterprise + 2 Guest for SD-Access Wireless)	2-4
Border nodes	2
Virtual networks, target fewer than	50
IP pools, target fewer than	200
Access points, target fewer than	1,000

Figure 40. Physical Topology – Medium Site Reference Design



Medium Site Considerations

In a medium site, high availability is provided in the fabric nodes by dedicating devices as border nodes and control plane nodes rather than collocating the functions together. For both resiliency and alternative forwarding paths in the overlay and underlay, the all devices within a given layer, with the exception of the access layer, should be crosslinked to each other. Multiple distribution blocks do not need to be cross-connected to each block, though should cross-connect to all distribution switches within a block. Dedicated control plane nodes are generally connected to the core switches sot that they are highly available for any edge node within the various distribution blocks. For optimal forwarding and redundancy, they should have connectivity through both cores, and if interfaces and fiber is available, crosslink to each other though this is not a requirement.

Physical WLC should be deployed to support the wireless user scale. To enable high availability a WLC HA-SSO pair is deployed with redundant physical connectivity to a services block using Layer 2 port-channels. The WLCs should be connected to each other through their Redundancy Port in accordance with the Tech tip from the Services Block section above. The services block is commonly implemented with fixed configuration switches operating in VSS or StackWise Virtual and connected to the core through Layer 3 routed links. This services block is deployed as a VRF-aware peer if DHCP/DNS and other shared services are site-local.

Large Site Design Reference Model

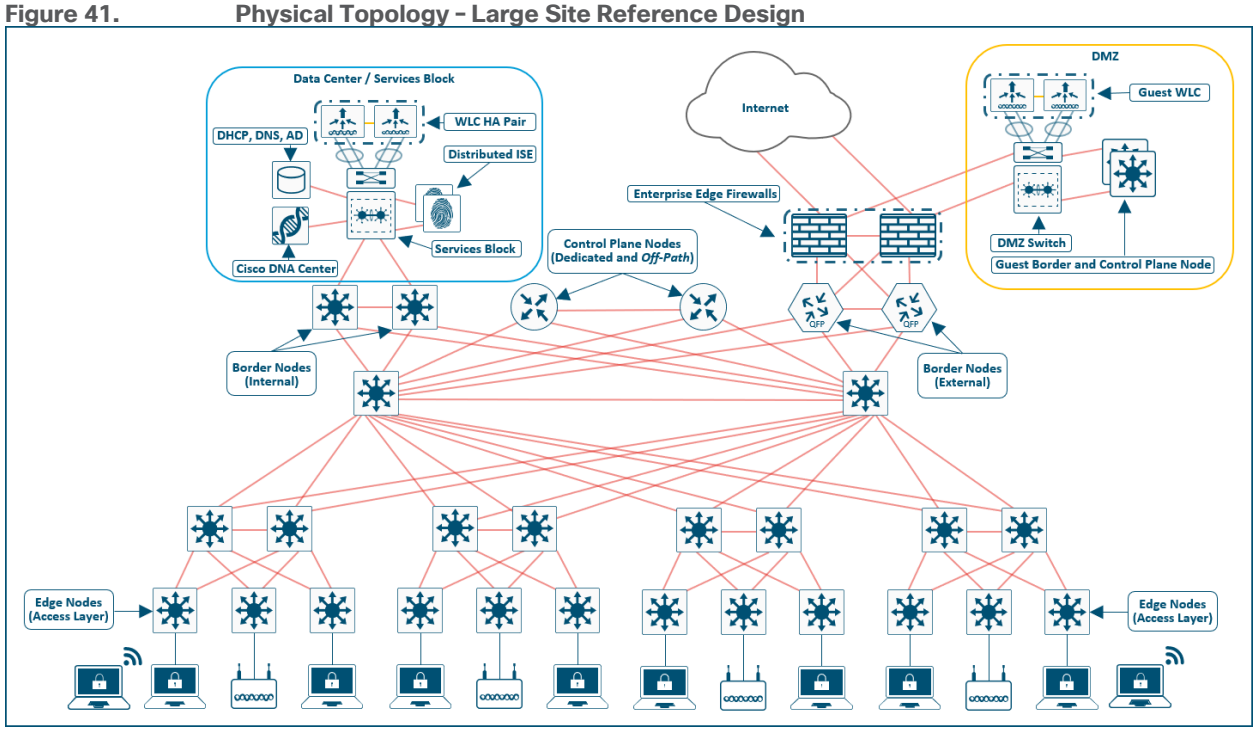
The Large Site Reference Model covers a building with multiple wiring closets or multiple buildings. The physical network is a three-tier network with core, distribution, and access and is designed to support less than 40,000 endpoints. This network is large enough to require dedicated services exit points such as a dedicated data center, shared services block, and Internet services.

Use the table below to understand the guidelines to stay within for similar site design sizes. The numbers are used as guidelines only and do not necessarily match specific limits for devices used in a design of this site size. The target maximum endpoint count requires, at minimum, the large Cisco DNA Center appliance to provide for future growth. The maximum fabric nodes and virtual networks are approximately ~75% of the number supported the large Cisco DNA Center appliance as listed on Table 10 its data sheet.

Table 6. Large Site Guidelines (Limits may be different)

Endpoints, target fewer than	50,000
Fabric nodes, target fewer than	750
Control plane nodes (limit to 2 Enterprise + 2 Guest for SD-Access Wireless)	2-4
Border nodes (2 as Internal and 2 as External)	2-4
Virtual networks, target fewer than	64
IP pools, target fewer than	450
Access points, target fewer than	2,000

Figure 41. Physical Topology - Large Site Reference Design



Large Site Considerations

The large site design is commonly the headquarters (HQ) location in a multiple-fabric site deployment. The enterprise edge firewall (perimeter firewall) is usually deployed at this location, and Internet traffic from remote sites is tunnel back to this site to be processed by the perimeter security stack before being forwarded to the Internet. Cisco DNA Center and the primary ISE PAN are generally deployed at this location.

Control plane nodes and border nodes should be dedicated devices deployed as redundant pairs. Dedicated control plane nodes should be connected to each core switch to provide for resiliency and to have redundant forwarding paths. If interfaces and fiber is available, crosslink the control plane nodes to each other though this is not a requirement; it simply provides another underlay forwarding path.

A wireless LAN controller HA-SSO pair is deployed with redundant physical connectivity to a services block using Layer 2 port-channels. The WLCs should be connected to each other through their Redundancy Ports in accordance with the Tech tip from the Services Block section above. The services block is commonly part of the on-premise data center network. At this headquarters location, the data center core is connected to either the campus core or the distribution switches to provide reachability to services and applications. Cisco Nexus 9000 Series switches with appropriate license level and capabilities are often used in the data center core function.

Dedicated *internal* border nodes are commonly used to connect the fabric site to the data center core while dedicated *external* border nodes are used to connect the site to the MAN, WAN, and Internet. Dedicated redundant routing infrastructure and firewalls are used to connect this site to external resources, and border nodes fully mesh to this infrastructure and to each other.

The Large Site may contain the DMZ where the dedicated Guest fabric border and control plane nodes for Guest Wireless are deployed. This provides complete control plane and data plane separation between Guest and Enterprise traffic and optimizes Guest traffic to be sent directly to the DMZ without the need for an Anchor WLC. Depending on the scale and redundancy needs, these devices are generally deployed with the fabric roles colocated though they may also be distributed.

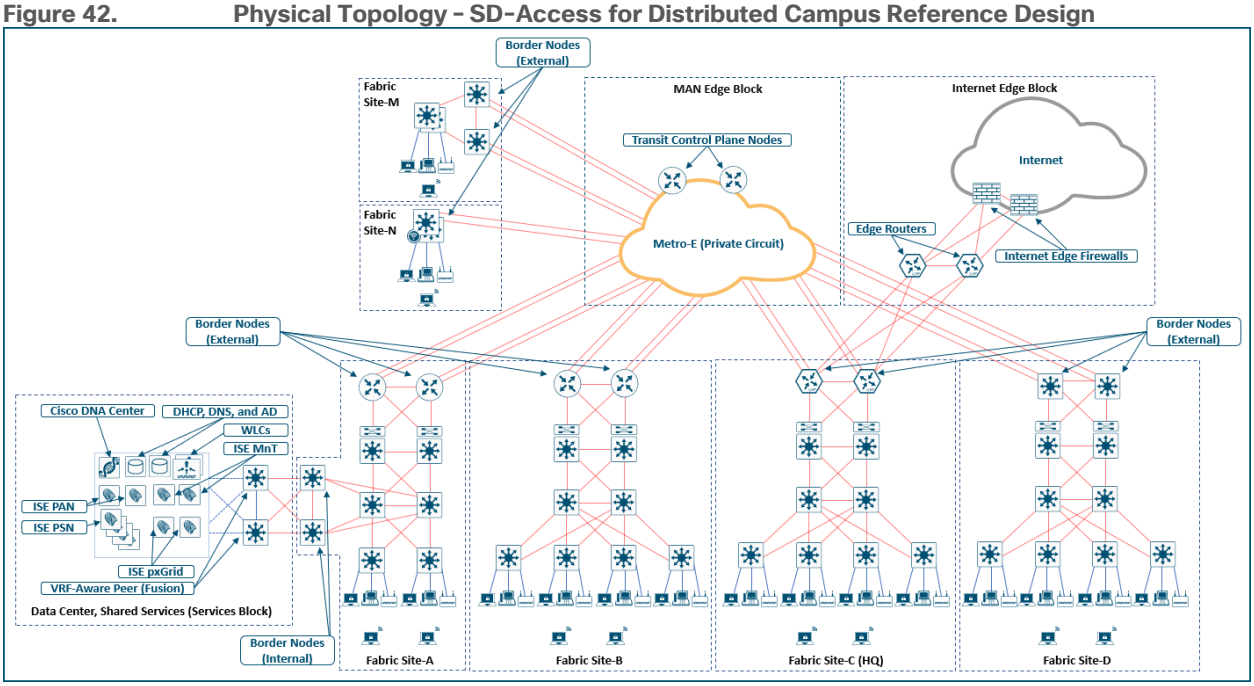
SD-Access for Distributed Campus Design—Reference Model

SD-Access for Distributed Campus is a solution that connects multiple, independent fabric sites together while maintaining the security policy constructs (VRFs and SGTs) across these sites. Control plane signaling from the LISP protocol along with fabric VXLAN encapsulation are used between fabric sites. This maintains the macro- and micro-segmentation policy constructs, VRFs and SGT respectively, between fabric sites. The result is a network that is address-agnostic because end-to-end policy is maintained through group membership.

In the reference topology in Figure 42 below, each fabric site is connected to a metro-Ethernet private circuit. The deployment is a large enterprise campus with dispersed buildings in a similar geographic area with each building operating as an independent fabric site. The border nodes connected to this circuit are configured as *external* borders. IGP peering occurs across the circuit to provide IP reachability between the loopback interface (RLOCs) of the devices. The Metro-Ethernet circuit is the used as the SD-Access transit between the fabric sites.

The headquarters (HQ) location has direct internet access, and one of the fabric sites (Fabric Site-A) has connections to the Data Center where shared services are deployed. *Internal* border nodes at Fabric Site-A import (register) the data center prefixes into the overlay space so the VNs in each fabric site can access these services. For traffic destined for Internet prefixes, traffic is forwarded back to the HQ location so that it can be processed through a common security stack before egressing to the outside world. The transit control plane nodes are deployed in their own area, accessible through the SD-Access transit Metro-E network though not in the direct forwarding path between fabric sites. For diagram simplicity, the site-local control plane nodes are not shown, and edge nodes are not labeled.

Figure 42. Physical Topology – SD-Access for Distributed Campus Reference Design



Distributed Campus Considerations

The core components enabling the Distributed Campus solution are the SD-Access transit and the transit control plane nodes. Both core components are architectural constructs present and used only in Distributed Campus deployments. The SD-Access transit is simply the physical network connection between fabric sites in the same city, metropolitan area, or between buildings in a large enterprise campus.

Reference Model Circuit for SD-Access Transit

The SD-Access transit, the physical network between fabric sites, should be have campus-like connectivity. The same encapsulation method that is used by nodes within a fabric site is used between sites though the SD-Access transit. This physical network should therefore strive for the same latency, throughput, connectivity as the campus itself.

This reference model transit is high-bandwidth (Ethernet full port speed with no sub-rate services), low latency (less than 10ms one-way as a general guideline), and should accommodate the MTU setting used for SD-Access in the campus network (typically 9100 bytes). The physical connectivity can be direct fiber connections, leased dark fiber, Ethernet over wavelengths on a DWDM system, or metro Ethernet systems (VPLS, etc.) supporting similar bandwidth, port rate, delay, and MTU connectivity capabilities.

It is possible to support an SD-Access transit on circuit types with criteria different from the reference model listed above. The primary requirement is to support jumbo frames across the circuit in order to carry the fabric-encapsulated packets without fragmentation. The latency supported by Cisco DNA Center itself as described in the Latency section (100ms RTT recommended, 200ms RTT supported) is the maximum supported latency for these non-Campus-like circuits.

Tech tip
For wide-area deployment using a standard 1500-byte MTU, configurating a smaller ***tcp adjust-mss value*** such as 1250 on the client- and AP-facing SVIs can be performed. If the UDP application uses an MTU value larger than the ***tcp adjust-mss value***, please adjust the MTU value on the UDP application server. It is also recommended that ICMP Type 3, Code 4 is permitted end to end throughout the network to allow requisite application control communication to take place for non-

TCP MTU reduction.

Key Considerations for SD-Access Transits

The SD-Access transit (the physical network) between sites is best represented, and most commonly deployed, as direct or leased fiber over a Metro Ethernet system. While Metro-E has several different varieties (VPLS, VPWS, etc.), the edge routers and switches of each fabric site ultimately exchange underlay routes through an IGP routing protocol. In SD-Access, this is commonly done using the IS-IS routing protocol, although other IGPs are supported as listed in the Underlay Network Design section. WAN circuits with appropriate latency such as MPLS are also supported.

IP reachability must exist between fabric sites. Specifically, there must be a known underlay route between the Loopback 0 interfaces on all fabric nodes. Existing BGP configurations and BGP peering on the transit control plane nodes could have complex interactions with the fabric configuration and should be avoided. BGP private AS 65540 is reserved for use on the transit control plane nodes and automatically provisioned by Cisco DNA Center. It should not be used elsewhere in the deployment. The transit control plane nodes should have IP reachability to the fabric sites through an IGP before being discovered or provisioned into the fabric role.

Traversing the transit control plane nodes in the data forwarding path between sites is not recommended. Transit control plane nodes should always be deployed as a matching pair of devices to provide resiliency and high availability.

Transit Control Plane Node Considerations

The transit control plane nodes do not have to be physically deployed in the transit area (the metro connection between sites) although common topology documentation often represents them in this way. These devices are generally deployed in their own dedicated location accessible through the physical transit network or deployed virtually in the data center as described in the CSR 1000v section above. While this is not a requirement, it is a recommended practice.

Like site-local control plane node design, which itself is based on BGP Route Reflector best practices, transit control plane nodes should not act as a physical-transit hop in the data packet forwarding path. Rather, they function similarly to a DNS server: they are queried for information, though data packets do not traverse through them.

The transit control plane nodes cannot be collocated with any other fabric role. They should be highly available through redundant physical connections. Routing platforms should have at least 8GB and preferably 16 GB or more DRAM to store all the registered prefixes for the entire fabric domain.

Migration to SD-Access

This chapter is organized into the following sections:

Chapter	Section
Migration to SD-Access	Migration Strategies Layer 2 Border Handoff

SD-Access greenfield networks can be created by adding the infrastructure components, interconnecting them, and using Cisco DNA Center with Cisco Plug and Play and LAN Automation features to automate provisioning of the network architecture from the ground up. Migrating an existing network requires some additional planning. Here are some example considerations:

- Does the network require reconfiguration into a Layer 3 Routed Access model?
- Do the SD-Access components in the network support the desired scale for the target topologies, or do the hardware and software platforms need to be augmented with additional platforms?
- Is the organization ready for changes in IP addressing and DHCP scope management?
- What is the strategy for integrating new overlays with common services (for example: Internet, DNS/DHCP, data center applications)?
- Are SGTs or dynamic ACLs already implemented, and where are the policy enforcement points? If SGTs and multiple overlays are used to segment and virtualize within the fabric, what requirements exist for extending them beyond the fabric? Is infrastructure in place to support Cisco TrustSec, VRF-Lite, MPLS, or other technologies necessary to extend and support the segmentation and virtualization?
- Can wireless coverage within a roaming domain be upgraded at a single point in time, or does the network need to rely on over-the-top strategies?

Migration Strategies

There are three primary approaches when migrating an existing network to SD-Access.

- **Parallel** –An SD-Access network is built next to an existing brownfield network. Switches are moved from the brownfield network to the SD-Access network by physically patching cables. This approach makes change management and rollback extremely simple. However, the parallel network requires additional rack space, power, and cabling infrastructure beyond what is currently consumed by the brownfield network.
- **Incremental**–This strategy moves a traditional switch from the brownfield network and converts it to an SD-Access fabric edge node. The Layer 2 Border handoff, discussed in the next section, is used to accomplish this incremental migration. This strategy is appropriate for networks that have equipment capable of supporting SD-Access already in place or where there are environmental constraints such as lack of space and power.
- **Hybrid**–The hybrid approach uses a combination of parallel and incremental approaches. For example, a new pair of core switches are configured as border nodes, control plane nodes are added and configured, and the existing brownfield access switches are converted to SD-Access fabric edge nodes incrementally.

Layer 2 Border Handoff

Layer 2 Border Handoff provides an overlay service between the SD-Access network and the traditional network, allowing hosts in both to communicate, ostensibly, at Layer 2. When a traditional network is migrating to an SD-Access network, the Layer 2 Border Handoff is a key strategic feature. Endpoints can remain in place in the traditional network while communication and interaction are tested with the endpoints in the fabric without needing to re-IP address these hosts. Hosts can then be migrated over to fabric entirely either through a parallel migration which involves physically moving cables or through an incremental migration of converting a traditional access switch to an SD-Access fabric edge node.

The Layer 2 Border Handoff allows the fabric site and the traditional network VLAN segment to operate using the same subnet. Communication between the two is provided across the border bode with this handoff that provides a VLAN translation between fabric and non-fabric. Cisco DNA Center automates the LISP control plane configuration along with the VLAN translation, Switched Virtual Interface (SVI), and the trunk port connected to the traditional network on this border node.

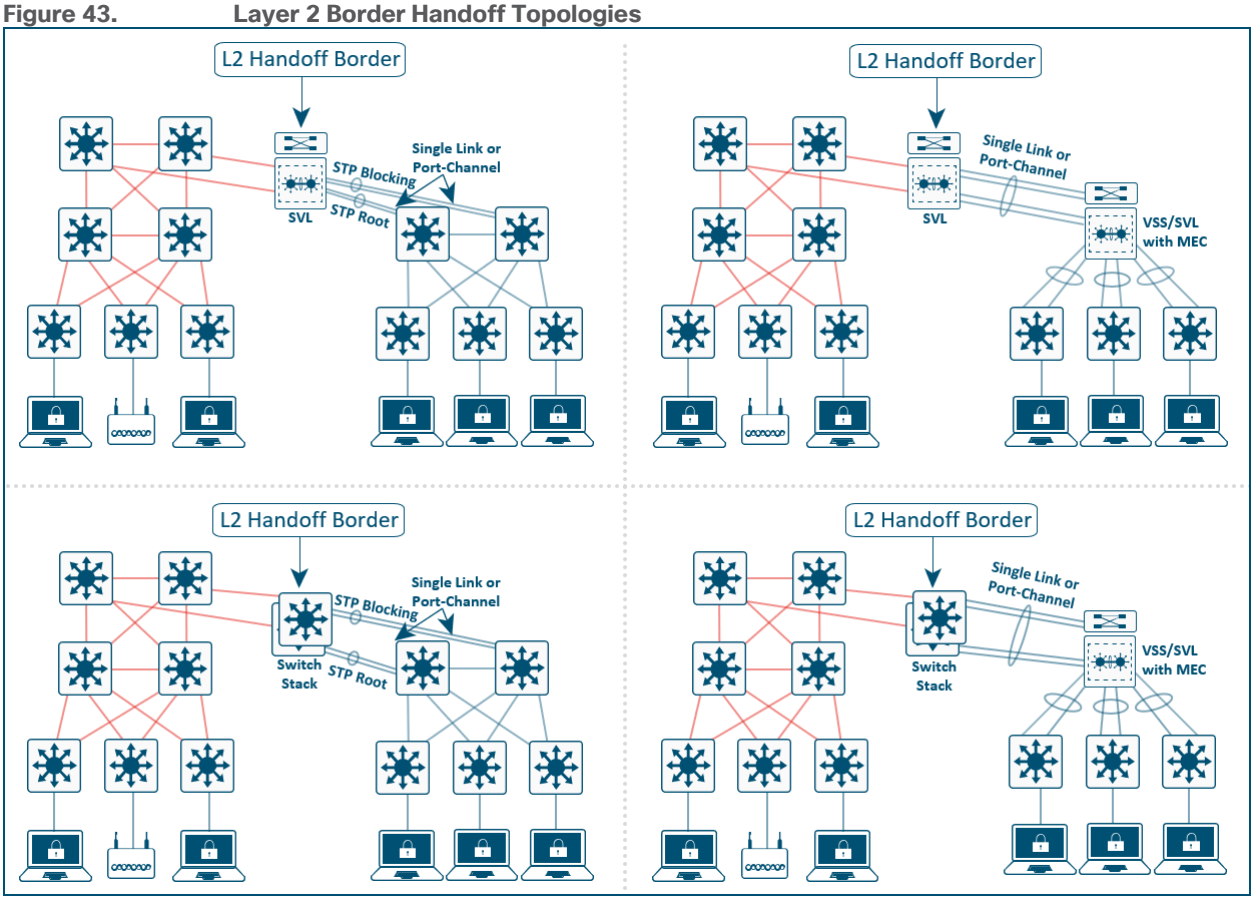
Multicast is supported across the Layer 2 handoff, allowing multicast communication between the traditional network and the SD-Access network. The multicast forwarding logic operates the same across the Layer 2 handoff border node as it does in the fabric, as described in the multicast Forwarding section, and the traditional network will flood multicast packets using common Layer 2 operations.

For OT (Operational Technology), IoT, and BMS (Building Management Systems) migrating to SD-Access, the Layer 2 border handoff can be used in conjunction with Layer 2 Flooding. This enables Ethernet broadcast WoL capabilities between the fabric site and the traditional network and allows OT/BMS systems that traditionally communicate via broadcast to migrate incrementally into the fabric.

Deployment Models and Topology

The traditional network switches can be connected to a single border node with a Layer 2 handoff. A traditional network switch should not be multihomed to multiple border nodes. Dual-homing, however, is support using link aggregation. Multichassis EtherChannel (MEC) is supported to a single border if the traditional network switches are operating in multi-box, single logical-box construct such as a hardware switch stack, Virtual Switching System (VSS), or StackWise Virtual (SVL). Redundancy for the border node itself can be provided through hardware stacking or StackWise Virtual.

Figure 43. Layer 2 Border Handoff Topologies



The Border node with the Layer 2 handoff should be a dedicated role. While it is technically feasible for this device to operate in multiple roles (such as a border node with Layer 3 handoff and control plane node), it is strongly recommended that a dedicated device be used. Because this device is operating at Layer 2, it is subject to the spanning-tree (STP) design impacts and constraints of the brownfield, traditional network, and a potential storm or loop in the traditional network could impact the Layer 2 handoff border node. Dedicating this

border node to the function of connecting to the traditional network separates the impact away from the remainder of the fabric network which can continue to operate normally independent of the traditional network.

Design Considerations

Devices operating with an Edge Node role, including Fabric in a Box, are not supported with Layer 2 Border Handoff. The border node connected to an SDA transit should not be the same device with using the Layer 2 border handoff. It is recommended and a best practice that the Layer 2 border handoff device be dedicated and not colocated with any other function. The device must be operating in transparent mode for VLAN Trunking Protocol (VTP) to avoid unintended modification of the traditional network's VLANs. The traditional network can use any VLAN except ***1, 1002-1005, 2045-2047***, and ***3000-3500*** which are either reserved in Cisco DNA Center or reserved for special use in Cisco software.

The border configured with the Layer 2 handoff becomes the default gateway for the VLAN in the traditional network. This same IP address and SVI will be present in the traditional network and must be placed in administrative down state and/or removed before the handoff automation on the border node. Shutting down and removing this SVI can be performed manually on the traditional network devices or through templates in Cisco DNA Center.

Appendices

This chapter is organized into the following sections:

Chapter	Section
Appendix A	SD-Access Fabric Protocols Deep Dive
Appendix B	References Used in this Guide
Appendix C	Glossary of Terms and Acronyms
Appendix D	Recommended for You and Additional Resources
Feedback	Comments, Suggestions, and Discussion Links

Appendix A - SD-Access Fabric Protocols

This section is organized into the following subsections:

Section	Subsection
SD-Access Fabric Protocols	Fabric Data Plane Fabric Control Plane

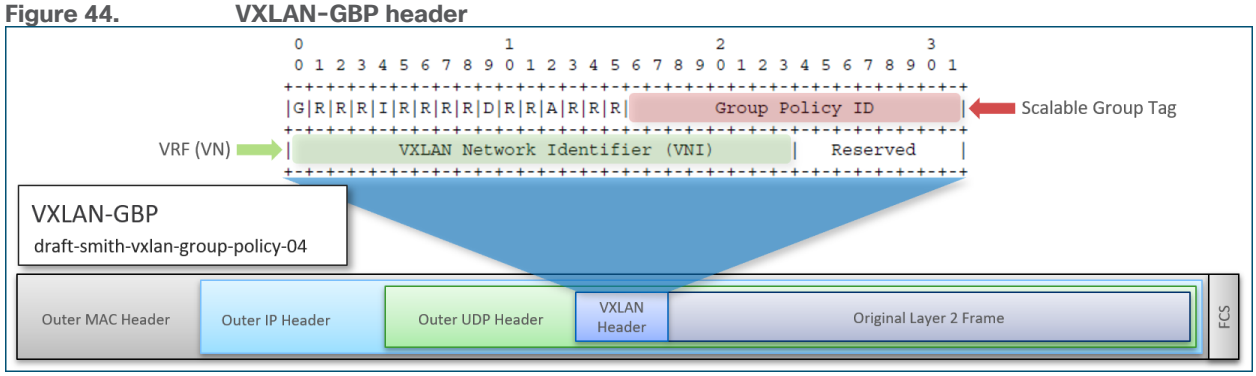
Fabric Data Plane

RFC 7348 defines the use of virtual extensible LAN (VXLAN) as a way to overlay a Layer 2 network on top of a Layer 3 network. Each overlay network is called a VXLAN segment and is identified using a 24-bit VXLAN network identifier, which supports up to 16 million VXLAN segments.

The SD-Access fabric uses the VXLAN data plane to provide transport of the full original Layer 2 frame and additionally uses LISP as the control plane to resolve endpoint-to-location (EID-to-RLOC) mappings. The SD-Access fabric replaces sixteen (16) of the reserved bits in the VXLAN header to transport up to 64,000 SGTs using a modified VXLAN-GPO (sometimes called VXLAN-GBP) format described in https://tools.ietf.org/html/draft-smith-vxlan-group-policy-04.

The Layer 3 VNI maps to a virtual routing and forwarding (VRF) instance for Layer 3 overlays, whereas a Layer 2 VNI maps to a VLAN broadcast domain, both providing the mechanism to isolate data and control plane to each individual virtual network. The SGT carries group membership information of users and provides data-plane segmentation inside the virtualized network.

Figure 44. VXLAN-GBP header



Fabric Control Plane

RFC 6830 through RFC 6836 along with later RFCs define LISP as a network architecture and set of protocols that implement a new semantic for IP addressing and forwarding. In traditional IP networks, the IP address is used to identify both an endpoint and its physical location as part of a subnet assignment on a router. In a LISP-enabled network, an IP address or MAC address is used as the endpoint identifier for an endpoint, and an additional IP address is used as an RLOC to represent the physical network device the endpoint is connected directly to or directly through such as with an access point or extended node. The Loopback 0 address of the network device is used as the RLOC address. The EID and RLOC combination provides the necessary information for traffic forwarding. The RLOC address is part of the underlay routing domain, and the EID can be assigned independently of the location.

The LISP architecture requires a mapping system that stores and resolves EIDs to RLOCs. This is analogous to using DNS to resolve IP addresses for host names. EID prefixes (either IPv4 addresses with /32 mask, MAC Address, or IPv6 Addresses with /128 masks) are registered with the map server along with their associated RLOCs. When sending traffic to an EID, a source RLOC queries the mapping system to identify the destination RLOC for traffic encapsulation. As with DNS, a local node probably does not have the information about everything in a network but instead asks for the information only when local hosts need it to communicate (pull model). This information is then cached for efficiency.

Although a full understanding of LISP and VXLAN is not required to deploy a fabric in SD-Access, it is helpful to understand how these technologies support the deployment goals. Included benefits provided by the LISP architecture are:

- **Network virtualization**—A LISP Instance ID (IID) is used to maintain independent VRF and VLAN topologies. From a data-plane perspective, the LISP Instance ID maps to either a Layer 2 or Layer 3 VNI.
- **Subnet stretching**—A single subnet can be extended to exist at multiple RLOCs. The separation of EID from RLOC enables the capability to extend subnets across different RLOCs. As a result of the availability of the Anycast Gateway across multiple RLOCs, the client configuration (IP address, subnet, and gateway) can remain unchanged, even as the client moves across the stretched subnet to different physical attachment points.
- **Smaller routing tables**—Only RLOCs need to be reachable in the global routing table for communication within a fabric site. Local EIDs (connected endpoints) are cached at the local node while remote EIDs (endpoints connected to or through other fabric devices) are learned through conversational learning. Conversational learning is the process of populating forwarding tables with only endpoints that are communicating through the node. This allows for efficient use of forwarding tables.

Appendix B – References Used in Guide

Anycast RP Technology White Paper:
https://www.cisco.com/c/en/us/td/docs/ios/solutions_docs/ip_multicast/White_papers/anycast.html

Campus Network for High Availability Design Guide, Tuning for Optimized Convergence:
https://www.cisco.com/c/en/us/td/docs/solutions/Enterprise/Campus/HA_campus_DG/hacampusdg.html#wp1107578

Campus Network for High Availability Design Guide:
https://www.cisco.com/c/en/us/td/docs/solutions/Enterprise/Campus/HA_campus_DG/hacampusdg.html

Cisco Catalyst 9800-CL Wireless Controller for Cloud Data Sheet:
https://www.cisco.com/c/en/us/products/collateral/wireless/catalyst-9800-cl-wireless-controller-cloud/nb-06-cat9800-cl-cloud-wirel-data-sheet-ctp-en.html

Connected Communities Infrastructure Solution Design Guide:
https://www.cisco.com/c/en/us/td/docs/solutions/Verticals/CCI/CCI/DG/cci-dg/cci-dg.html

Cisco DNA Center & ISE Management Infrastructure Deployment Guide: https://cs.co/sda-infra-pdg

Cisco DNA Center and SD-Access 1.3 Scale Metrics – Cisco Communities:
https://community.cisco.com/t5/networking-documents/cisco-dna-center-and-sd-access-1-3-scale-metrics/ta-p/3897030

Cisco DNA Center 1.3.3.0 Data Sheet, Fabric VN Scale:
https://www.cisco.com/c/en/us/products/collateral/cloud-systems-management/dna-center/nb-06-dna-center-data-sheet-cte-en.html#CiscoDNACenter1330DeviceAwareFabricVNLimitFabricVNScale

Cisco DNA Center 3-Node Cluster High Availability Scenarios and Network Connectivity Details:
https://www.cisco.com/c/en/us/support/docs/cloud-systems-management/dna-center/214471-cisco-dna-center-3-node-cluster-high-ava.html

Cisco DNA Center Latency Design Guidance – Cisco Community:
https://community.cisco.com/t5/networking-documents/cisco-sda-design-guidance-and-best-practices/ta-p/3865954

Cisco DNA Center Release Notes: https://www.cisco.com/c/en/us/support/cloud-systems-management/dna-center/products-release-notes-list.html

Cisco DNA Center SD-Access LAN Automation Deployment Guide:
https://www.cisco.com/c/en/us/td/docs/cloud-systems-management/network-automation-and-management/dna-center/tech_notes/b_dnac_sda_lan_automation_deployment.html

Cisco Enterprise Architecture Model – Cisco Networking Academy:
https://www.ciscopress.com/articles/article.asp?p=2202410&seqNum=6

Cisco Enterprise Internet Edge Design Guide:
https://www.cisco.com/c/en/us/td/docs/solutions/Enterprise/Security/IE_DG.html

Cisco Enterprise Mobility 8.5 Design Guide: https://www.cisco.com/c/en/us/td/docs/wireless/controller/8-5/Enterprise-Mobility-8-5-Design-Guide/Enterprise_Mobility_8-5_Deployment_Guide/WirelessNetwork_GuestAccessService.html

Cisco Extended Enterprise Non-Fabric and SD-Access Fabric Design Guide:
https://www.cisco.com/c/en/us/td/docs/solutions/Verticals/EE/DG/ee-dg/ee-dg.html

Cisco Firepower Release Notes, Version 6.3, New Features: https://www.cisco.com/c/en/us/td/docs/security/firepower/630/relnotes/firepower-release-notes-630/new_features.html#id_70703

Cisco Firepower Release Notes, Version 6.6, New Features: https://www.cisco.com/c/en/us/td/docs/security/firepower/660/relnotes/firepower-release-notes-660/features.html#id_122921

Cisco Firepower Threat Defense Multi-Instance Capability on Cisco Firepower 4100 and 9300 Series Appliances White Paper: https://www.cisco.com/c/en/us/products/collateral/security/firepower-ngfw/white-paper-c11-742018.html

Cisco IOS Software Configuration Guide, Release 15.0SY, Chapter: Stateful Switchover (SSO): https://www.cisco.com/c/en/us/td/docs/switches/lan/catalyst6500/ios/15-0SY/configuration/guide/15_0_sy_swcg/stateful_switchover.pdf

Cisco Identity Services Engine Administrator Guide, Release 2.6, Chapter: Set Up Cisco ISE in a Distributed Environment: https://www.cisco.com/c/en/us/td/docs/security/ise/2-6/admin_guide/b_ise_admin_guide_26/b_ise_admin_guide_26_chapter_011.html

Cisco ISE Performance and Scale - Cisco Community: https://community.cisco.com/t5/security-documents/ise-performance-amp-scale/ta-p/3642148

Cisco Live - Cisco SD-Access - Under the Hood - BRKCRS-2810 (2019, Cancun): https://www.ciscolive.com/global/on-demand-library.html?search=brkcrs-2810#/session/15615652543060019s7C

Cisco Live - Real World Route/Switch to Cisco SD-Access Migration Tools and Strategies - BRKCRS-3493 (2020, APJC): https://www.ciscolive.com/global/on-demand-library.html?search=brkcrs-3493#/session/1571888607137001yDeW

Cisco Live Enterprise Campus Design: Multilayer Architectures and Design Principles - BRKCRS-2031 (2019, San Diego): https://www.ciscolive.com/global/on-demand-library.html?search.event=ciscoliveus2019&search=BRKCRS-2031#/session/1542224309065001rZIQ

Cisco on Cisco Best Practices, Cisco High Availability LAN: https://www.cisco.com/c/dam/en_us/about/ciscoitatwork/downloads/ciscoitatwork/pdf/How_Cisco_IT_Achieved_Highly_Available_Local_Area_Network.pdf

Cisco Service Ready Architecture for Schools Design Guide, 9 July 2010: https://www.cisco.com/c/en/us/td/docs/solutions/Enterprise/Education/SchoolsSRA_DG/SchoolsSRA-DG.pdf

Cisco Software-Defined Access 1.0 White Paper: https://www.cisco.com/c/dam/en/us/solutions/collateral/enterprise-networks/software-defined-access/white-paper-c11-740585.pdf

Cisco UCS C-Series Rack Servers: https://www.cisco.com/c/en/us/products/servers-unified-computing/ucs-c-series-rack-servers/index.html

Cisco UCS E-Series Servers: https://www.cisco.com/c/en/us/products/servers-unified-computing/ucs-e-series-servers/index.html

Cisco Unified Access Design Guide, 18 October 2011: https://www.cisco.com/c/en/us/td/docs/solutions/Enterprise/Borderless_Networks/Unified_Access/Unified_Access_Book.pdf

Configuring a Rendezvous Point Technology White Paper:
https://www.cisco.com/c/en/us/td/docs/ios/solutions_docs/ip_multicast/White_papers/rps.html

Enterprise Campus 3.0 Architecture: Overview and Framework:
https://www.cisco.com/c/en/us/td/docs/solutions/Enterprise/Campus/campover.html

Enterprise Mobility 4.1 Design Guide, Chapter: Cisco Unified Wireless Technology and Architecture, Centralized WLC Deployment:
https://www.cisco.com/c/en/us/td/docs/solutions/Enterprise/Mobility/emob41dg/emob41dg-wrapper/ch2_Arch.html

Firepower Management Center Configuration Guide, Version 6.6, Chapter: Virtual Routing for Firepower Threat Defense: https://www.cisco.com/c/en/us/td/docs/security/firepower/660/configuration/guide/fpmc-config-guide-v66/virtual-routing-for-firepower-threat-defense.html

Graceful Restart, Non Stop Routing and IGP Routing Protocol Timer Manipulation Solution Overview:
https://www.cisco.com/c/en/us/products/collateral/ios-nx-os-software/high-availability/solution_overview_c22-487228.html

Guide to SD-Access Border Node Roles on Cisco DNA Center ≥1.3.x - Cisco Community:
https://community.cisco.com/t5/networking-documents/guide-to-choosing-sd-access-sda-border-roles-in-cisco-dnac-1-3/ta-p/3889472

Hierarchical Network Design Overview - Cisco Networking Academy:
https://www.ciscopress.com/articles/article.asp?p=2202410&seqNum=4

High Availability Campus Network Design - Routed Access Layer using EIGRP or OSPF System Assurance Guide: https://www.cisco.com/c/dam/en/us/td/docs/nsite/campus/ha_campus_routed_access_cvd_ag.pdf

High Availability Campus Network Design--Routed Access Layer using EIGRP or OSPF:
https://www.cisco.com/c/en/us/td/docs/solutions/Enterprise/Campus/routed-ex.html

High Availability SSO Deployment Guide for Cisco Catalyst 9800 Series Wireless Controllers, Cisco IOS XE Amsterdam 17.1 (12 March 2020): https://www.cisco.com/c/dam/en/us/td/docs/wireless/controller/9800/17-1/deployment-guide/c9800-ha-sso-deployment-guide-rel-17-1.pdf

Medium Enterprise Design Profile Reference Guide, Chapter 2 - Design Zone for Security:
https://www.cisco.com/c/en/us/td/docs/solutions/Enterprise/Medium_Enterprise_Design_Profile/MEDP/chap2.html

NX-OS VXLAN Configuration Guide, Chapter: VXLAN Overview:
https://www.cisco.com/c/en/us/td/docs/switches/datacenter/sw/nx-os/vxlan/configuration/guide/b_NX-OS_VXLAN_Configuration_Guide/overview.pdf

Overview of TrustSec, 2014 January: https://www.cisco.com/c/dam/en/us/solutions/collateral/borderless-networks/trustsec/C07-730151-00_overview_of_trustSec_og.pdf

Rendezvous Point Engineering White Paper: https://www.cisco.com/c/en/us/products/collateral/ios-nx-os-software/ip-multicast/whitepaper_c11-508498.html

Small Enterprise Design Profile Reference Guide, Chapter 2 - Design Zone for Security:
https://www.cisco.com/c/en/us/td/docs/solutions/Enterprise/Small_Enterprise_Design_Profile/SEDP/chap2.html

Using Multi-Instance Capability on the Firepower 4100/9300 Configuration Guide:
https://www.cisco.com/c/en/us/td/docs/security/firepower/fxos/multi-instance/multi-instance_solution.html

WLC High Availability (SSO) Deployment Guide Technical Reference, 6 August 2014:
https://www.cisco.com/c/en/us/td/docs/wireless/controller/technotes/7-5/High_Availability_DG.html

Appendix C – Acronym Glossary

AAA–Authentication, Authorization, and Accounting

ACP–Access-Control Policy

ACI–Cisco Application Centric Infrastructure

ACK–Acknowledge or Acknowledgement

ACL–Access-Control List

AD–Microsoft Active Directory

AFI–Address Family Identifier

AMP–Cisco Advanced Malware Protection

AP–Access Point

API–Application Programming Interface

APIC– Cisco Application Policy Infrastructure Controller (ACI)

ASA–Cisco Adaptative Security Appliance

ASM–Any-Source Multicast (PIM)

ASR–Aggregation Services Router

Auto-RP–Cisco Automatic Rendezvous Point protocol (multicast)

AVC–Application Visibility and Control

BFD–Bidirectional Forwarding Detection

BGP–Border Gateway Protocol

BMS–Building Management System

BSR–Bootstrap Router (multicast)

BYOD–Bring Your Own Device

CAPWAP–Control and Provisioning of Wireless Access Points Protocol

CDP–Cisco Discovery Protocol

CEF–Cisco Express Forwarding

CMD–Cisco Meta Data

CPU–Central Processing Unit

CSR–Cloud Services Routers

CTA–Cognitive Threat Analytics

CUWN–Cisco Unified Wireless Network

CVD–Cisco Validated Design

CYOD–Choose Your Own Device

DC–Data Center

DHCP–Dynamic Host Configuration Protocol

DM–Dense-Mode (multicast)

DMVPN–Dynamic Multipoint Virtual Private Network

DMZ–Demilitarized Zone (firewall/networking construct)

DNA–Cisco Digital Network Architecture

DNS–Domain Name System

DORA–Discover, Offer, Request, ACK (DHCP Process)

DWDM–Dense Wavelength Division Multiplexing

ECMP–Equal Cost Multi Path

EID–Endpoint Identifier

EIGRP–Enhanced Interior Gateway Routing Protocol

EMI–Electromagnetic Interference

ETR–Egress Tunnel Router (LISP)

EVPN–Ethernet Virtual Private Network (BGP EVPN with VXLAN data plane)

FHR–First-Hop Router (multicast)

FHRP–First-Hop Redundancy Protocol

FMC–Cisco Firepower Management Center

FTD–Cisco Firepower Threat Defense

GBAC–Group-Based Access Control

GbE–Gigabit Ethernet

Gbit/s–Gigabits Per Second (interface/port speed reference)

GRE–Generic Routing Encapsulation

GRT–Global Routing Table

HA–High-Availability

HQ–Headquarters

HSRP–Cisco Hot-Standby Routing Protocol

HTDB–Host-tracking Database (SD-Access control plane node construct)

IBNS–Identity-Based Networking Services (IBNS 2.0 is the current version)

ICMP– Internet Control Message Protocol

IDF–Intermediate Distribution Frame; essentially a wiring closet.

IEEE–Institute of Electrical and Electronics Engineers

IETF—Internet Engineering Task Force

IGP—Interior Gateway Protocol

IID—Instance-ID (LISP)

IOE—Internet of Everything

IoT—Internet of Things

IP—Internet Protocol

IPAM—IP Address Management

IPS—Intrusion Prevention System

IPSec—Internet Protocol Security

ISE—Cisco Identity Services Engine

ISR—Integrated Services Router

IS-IS—Intermediate System to Intermediate System routing protocol

ITR—Ingress Tunnel Router (LISP)

LACP—Link Aggregation Control Protocol

LAG—Link Aggregation Group

LAN—Local Area Network

L2 VNI—Layer 2 Virtual Network Identifier; as used in SD-Access Fabric, a VLAN.

L3 VNI— Layer 3 Virtual Network Identifier; as used in SD-Access Fabric, a VRF.

LHR—Last-Hop Router (multicast)

LISP—Location Identifier Separation Protocol

MAC—Media Access Control Address (OSI Layer 2 Address)

MAN—Metro Area Network

MEC—Multichassis EtherChannel, sometimes referenced as ***MCEC***

MDF—Main Distribution Frame; essentially the central wiring point of the network.

MnT—Monitoring and Troubleshooting Node (Cisco ISE persona)

MOH—Music on Hold

MPLS—Multiprotocol Label Switching

MR—Map-resolver (LISP)

MS—Map-server (LISP)

MSDP—Multicast Source Discovery Protocol (multicast)

MTU—Maximum Transmission Unit

NAC—Network Access Control

NAD—Network Access Device

NAT—Network Address Translation

NBAR—Cisco Network-Based Application Recognition (NBAR2 is the current version).

NFV—Network Functions Virtualization

NSF—Non-Stop Forwarding

OSI—Open Systems Interconnection model

OSPF—Open Shortest Path First routing protocol

OT—Operational Technology

PAgP—Port Aggregation Protocol

PAN—Primary Administration Node (Cisco ISE persona)

PCI DSS—Payment Card Industry Data Security Standard

PD—Powered Devices (PoE)

PETR—Proxy-Egress Tunnel Router (LISP)

PIM—Protocol-Independent Multicast

PITR—Proxy-Ingress Tunnel Router (LISP)

PnP—Plug-n-Play

PoE—Power over Ethernet (Generic term, may also refer to IEEE 802.3af, 15.4W at PSE)

PoE+—Power over Ethernet Plus (IEEE 802.3at, 30W at PSE)

PSE—Power Sourcing Equipment (PoE)

PSN—Policy Service Node (Cisco ISE persona)

pxGrid—Platform Exchange Grid (Cisco ISE persona and publisher/subscriber service)

PxTR—Proxy-Tunnel Router (LISP – device operating as both a PETR and PITR)

QoS—Quality of Service

RADIUS—Remote Authentication Dial-In User Service

REST—Representational State Transfer

RFC—Request for Comments Document (IETF)

RIB—Routing Information Base

RLOC—Routing Locator (LISP)

RP—Rendezvous Point (multicast)

RP—Redundancy Port (WLC)

RP—Route Processer

RPF—Reverse Path Forwarding

RR—Route Reflector (BGP)

RTT—Round-Trip Time

SA—Source Active (multicast)

SAFI—Subsequent Address Family Identifiers (BGP)

SD—Software-Defined

SDA—Cisco Software Defined-Access

SDN—Software-Defined Networking

SFP—Small Form-Factor Pluggable (1 GbE transceiver)

SFP+— Small Form-Factor Pluggable (10 GbE transceiver)

SGACL—Security-Group ACL

SGT–Scalable Group Tag, sometimes reference as Security Group Tag

SM—Spare-mode (multicast)

SNMP—Simple Network Management Protocol

SSID—Service Set Identifier (wireless)

SSM—Source-Specific Multicast (PIM)

SSO—Stateful Switchover

STP—Spanning-tree protocol

SVI—Switched Virtual Interface

SVL—Cisco StackWise Virtual

SWIM—Software Image Management

SXP—Scalable Group Tag Exchange Protocol

Syslog—System Logging Protocol

TACACS+—Terminal Access Controller Access-Control System Plus

TCP—Transmission Control Protocol (OSI Layer 4)

UCS— Cisco Unified Computing System

UDP—User Datagram Protocol (OSI Layer 4)

UPoE—Cisco Universal Power Over Ethernet (60W at PSE)

UPoE+— Cisco Universal Power Over Ethernet Plus (90W at PSE)

URL—Uniform Resource Locator

VLAN—Virtual Local Area Network

VN—Virtual Network, analogous to a VRF in SD-Access

VNI—Virtual Network Identifier (VXLAN)

vPC—virtual PortChannel (Cisco Nexus)

VPLS—Virtual Private LAN Service

VPN—Virtual Private Network

VPNv4—BGP address family that consists of a Route-Distinguisher (RD) prepended to an IPv4 prefix

VPWS—Virtual Private Wire Service

VRF—Virtual Routing and Forwarding

VSL—Virtual Switch Link (Cisco VSS component)

VSS—Cisco Virtual Switching System

VXLAN—Virtual Extensible LAN

WAN—Wide-Area Network

WLAN—Wireless Local Area Network (generally synonymous with IEEE 802.11-based networks)

WoL—Wake-on-LAN

xTR—Tunnel Router (LISP – device operating as both an ETR and ITR)

Appendix D – Recommended for You

Cisco Enterprise Networks Validated Design and Deployment Guide: https://cs.co/en-cvds

Cisco SD-Access Resource – Cisco Community: https://cs.co/sda-resources

Cisco SD-Access Segmentation Design Guide: https://cs.co/sda-segment-sdg

Cisco Software-Defined Access for Distributed Campus Prescriptive Deployment Guide: https://cs.co/sda-distrib-pdg

Cisco Software-Defined Fabric Provisioning Prescriptive Deployment Guide: https://cs.co/sda-fabric-pdg

Design Zone for Branch, WAN, and Internet Edge: https://www.cisco.com/c/en/us/solutions/design-zone/networking-design-guides/branch-wan-edge.html

Design Zone for Branch, WAN, and the Internet Edge – Guide Archive: https://www.cisco.com/c/en/us/solutions/design-zone/networking-design-guides/branch-wan-edge/cvd-archive-wan.html

Design Zone for Campus Wired and Wireless LAN – Guide Archive: https://www.cisco.com/c/en/us/solutions/design-zone/networking-design-guides/campus-wired-wireless/cvd-archive-campus.html

Design Zone for Campus Wired and Wireless LAN: https://www.cisco.com/c/en/us/solutions/design-zone/networking-design-guides/campus-wired-wireless.html

Design Zone for Campus, Design Guides: https://www.cisco.com/c/en/us/solutions/enterprise/design-zone-campus/design-guide-listing.html

Design Zone for WAN and Branch/Remote Sites: https://www.cisco.com/c/en/us/solutions/enterprise/design-zone-branch-wan/design-guide-listing.html

Design Zone for Wireless and Mobility: https://www.cisco.com/c/en/us/solutions/design-zone/mobility-design-guides.html

YouTube Playlists, Cisco DNA Center: https://www.youtube.com/channel/UCSyV6GuKZvmAHndJkDg8_8A/playlists

YouTube Playlist, Cisco DNA Center, Release 1.3.3.x: https://www.youtube.com/playlist?list=PLGJ69loi36wJ3GBDg_-Ybk1vRrZDWKab2

YouTube Playlist, Cisco ISE – Identity Services Engine: https://www.youtube.com/user/CiscoISE

YouTube Playlist, Cisco SD-Access: https://www.youtube.com/channel/UCHDBMhhEDalzaiFGXOhHJDg/featured

Feedback

For comments and suggestions about this guide and related guides, join the discussion on Cisco Community at https://cs.co/en-cvds.

Americas Headquarters
Cisco Systems, Inc.
San Jose, CA

Asia Pacific Headquarters
Cisco Systems (USA) Pte. Ltd.
Singapore

Europe Headquarters
Cisco Systems International BV Amsterdam,
The Netherlands

Cisco has more than 200 offices worldwide. Addresses, phone numbers, and fax numbers are listed on the Cisco Website at **https://www.cisco.com/go/offices**.