aruba
a Hewlett Packard Enterprise company

Validated Solution Guide

ESP DATA CENTER VOLUME 1

Design Guide

About This Guide

This document is from a family of technology guides called Aruba Validated Solution Guides (VSG). VSGs are cross-portfolio solution guides that cover multiple technology areas, including wired, wireless, data center, SD-WAN and security. They are validated by Aruba's Solution TME and Solution Quality Assurance teams on an ongoing basis using a rigorous process. A VSG provides prescriptive guidance focused on the Aruba recommended best practices specific to the solution being covered.

The goal is to describe a solution implementation which addresses the primary use cases for customer networks, while avoiding the corner cases. The intent is to enable partners and customers to efficiently install end-to-end solutions using Aruba technology in a consistent and repeatable manner. The result will be improved stability and supportability by limiting the number of deployment variations.

VSGs are categorized into volumes to differentiate each guide type from the others.

Volumes

- *Design:* Identify products and technologies to meet customer business requirements
- *Deploy:* Step-by-step set of procedures to build the solution
- *Operate:* Recommended procedures to maintain and optimize the solution

Introduction

The Aruba Networks ESP Data Center is built on a technology platform which provides the tools for transforming the data center into a modern, agile, services delivery platform satisfying the requirements of organizations large, small, distributed, and centralized. The Aruba AOS-CX operating system simplifies operations and maintenance with a common switch operating system across the campus, branch, and data center, managed from the cloud or on-premises, and backed by an artificial intelligence capability which provides best practices guidance throughout the operational lifecycle of your network.

Converged ethernet is changing the way compute hosts access storage in the modern data center. Dedicated storage area networks are no longer required. Lossless ethernet and bandwidth management protocols ensure timely reads and writes over a traditional IP LAN. The cost savings and operational simplicity of converged ethernet are major drivers for transformation in the data center today.

At the same time, network topologies have become virtualized. While this virtualization promotes the flexibility required to meet today's transformational data center requirements, it can lead to complexity during implementation and management. The Aruba ESP Data Center mitigates these challenges by leveraging automation in the management plane and capabilities of the Aruba AOS-CX operating system such as automated configuration backups and built-in alerts instrumented on critical network performance metrics.

As you begin the process of designing a new or transformed data center the first step is to understand your organization's cloud applications strategy. This will allow you to determine which applications will remain on-premises and what a right-sized data center looks like for your requirements. When establishing a new data center intended to grow and adapt into the future, plan to implement a spine and leaf underlay supporting software defined overlay networks. The Aruba Networks CX 83xx and 84xx switching platforms provide a best in class suite of products featuring a variety of high throughput port configurations and industry leading operating system modularity providing real-time analytics and always up maintenance.

Purpose of this Guide

This guide covers the Aruba Data Center Network design, including reference architectures along with their associated hardware and software components. It contains an explanation of the requirements that shaped the design and the benefits it provides your organization. This guide will provide an introduction to Aruba data center solutions that support options for both distributed and centralized workloads and will provide best practices recommendations for designing a next generation spine and leaf data center fabric using VXLAN and BGP EVPN.

This guide assumes the reader has an equivalent knowledge of an Aruba Certified Switching Associate.

Design Goals

The overall goal is to create a high-reliability, scalable design that is easy to maintain and adapt to the changing needs of business. The solution components are limited to a specific set of products required for optimal operations and maintenance. The key features addressed by the Aruba Data Center Network include:

- Zero downtime upgrades
- High throughput
- Converged storage networking
- Flexible segmentation
- 3rd Party Integration

You can use this guide to design new networks or to optimize and upgrade existing networks. It is not intended as an exhaustive discussion of all options but rather to present commonly recommended designs, features, and hardware.

Audience

This guide is written for IT professionals who need to design an Aruba Data Center Network. These IT professionals can fill a variety of roles:

- Systems engineers who need a standard set of procedures for implementing solutions
- Project managers who create statements of work for Aruba implementations
- Aruba partners who sell technology or create implementation documentation

Customer Use Cases

Data center networks are changing rapidly. The most pressing challenge is to maintain operational stability and visibility while placing compute and storage resources where they need to be in order to best serve users. In addition, data center teams are being asked to support the rapid pace of DevOps environments including requirements to connect directly with public cloud infrastructure. Given the rapidly changing landscape for data center requirements it's critical that network and system engineers be provided with tools to simplify and automate complex infrastructure configurations.

This guide discusses the following use cases:

- Pay as you grow designs that support network and compute workload elasticity
- Ease of use and agility to quickly deploy and manage workloads using compute, hypervisor and network orchestration
- Improved operations with data center visibility from inside compute host to network infrastructure
- Workload mobility, security and multi-tenancy using standards based overlay technologies
- Network infrastructure automation and management
- Data aggregation and pre-processing

Aruba ESP Data Center Network Design

The Aruba Edge Services Platform (ESP) Data Center provides flexible and highly reliable designs that ensures efficient access to applications and data for all authorized users while simplifying operations and accelerating service delivery.

The Aruba ESP Data Center includes the following key features and capabilities:

- Modern connectivity - Design efficient and scalable networks using the full range of port densities and speed options available in the Aruba CX 8000 switching family.
- Automation - Automated fabric configuration makes building high-performance, scalable data center networks more efficient and less error prone.
- Analytics - On box and cloud analytics ensures alerts are never missed and intermittent failures can be diagnosed quickly.
- Storage networking - Advanced protocols enable lossless ethernet with bandwidth reservation and congestion management.
- Host integration - Virtual network visualization as part of the physical network topology for end-to-end management.

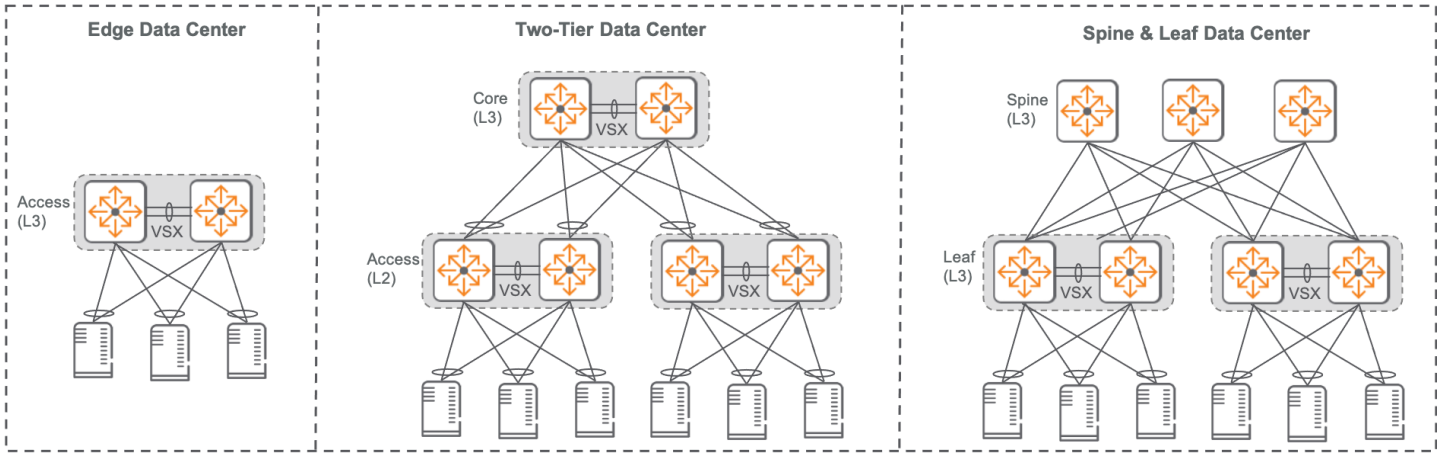
The Aruba ESP Data Center Network design may contain one or more of the following elements:

- Aruba Central
- Aruba Fabric Composer
- Aruba NetEdit
- Aruba CX 8000 Ethernet switches
- Aruba CX 6000 Ethernet switches for out-of-band network management
- Aruba Integration into HPE Solutions

Aruba Fabric Composer	Aruba CX 8300 Series	Aruba Integration into HPE Solutions
		

Aruba ESP Data Center Network Design Options

The Aruba ESP Data Center supports centralized and distributed workloads anywhere within an organization. Each design supports host uplink bundling providing throughput and resiliency for mission critical workloads. Layer 2 domains can be flexibly deployed to suit application requirements and virtual host mobility. Aruba CX switches provide a robust platform for layer 3 services in the data center. When deployed in a spine and leaf topology a layer 3 data center network eliminates the need for loop avoidance protocols and is optimized for high capacity and non-oversubscribed low-latency performance.



Edge Data Center Overview

Enterprises that have migrated most of their workloads to the cloud and do not have a large demand for an on-premises data center can leverage their existing campus network wiring closets, or small server rooms to deploy workloads at the edge. This approach employs a simple and easy to manage design leveraging the same AOS-CX switches that provide wired connectivity to users and IoT devices to also provide server access. The edge data center also supports high bandwidth and low latency access to compute and storage resources for distributed workloads that may not be well suited to cloud deployments.

Two-Tier Data Center Overview

Enterprises with significant, existing, on-premises workloads spanning multiple workgroups will often require a traditional, 2-tier data center design. The design is well understood and common within data centers everywhere. The 2-tier approach ensures sufficient bandwidth and reliability using legacy protocols such as LACP, spanning tree, and OSPF. Hosts are dual-homed to top-of-rack (ToR) switches using VSX LAG. Each ToR switch is dual-homed to the core. Loops are primarily prevented by the use of LACP to aggregate redundant links.

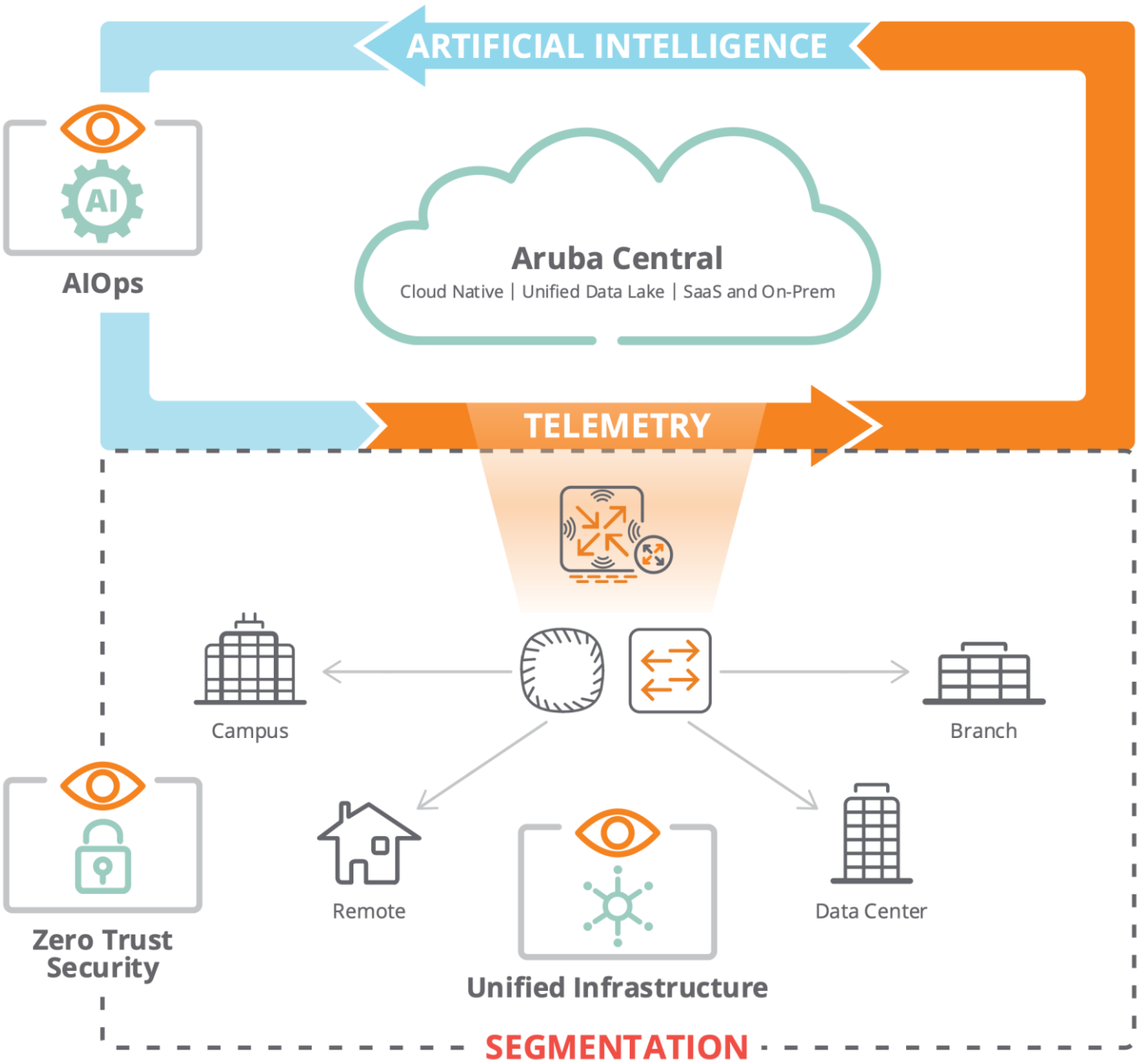
Spine & Leaf Data Center Overview

Enterprises with growing, on-premises workloads and those with workloads spread across data centers should leverage the efficiencies of a CLOS based, spine and leaf architecture. In most cases, a migration to the spine and leaf design should be paired with the implementation of a VXLAN overlay topology. The spine and leaf design ensures high reliability through the use of redundant layer 3 links between leaf nodes and spine switches. Equal cost multi-path (ECMP) routing ensures load balancing and fast fail-over if a link or switch goes down. The fully meshed architecture enables simple, horizontal growth by simply adding another spine switch as needed. VXLAN provides a layer 2 over layer 3 tunneling solution which enables customers to modernize the underlay while preserving legacy service requirements by allowing for physically dispersed L2 segments in the overlay. VXLAN also enables highly segmented designs which can go beyond traditional VLANs when creating secure, discreet groups of resources within the data center.

This guide addresses the most common uses cases of an Aruba Spine and Leaf Data Center Network solution. If you are planning a more complex project that is not covered in this guide, contact an Aruba or partner SE for design verification.

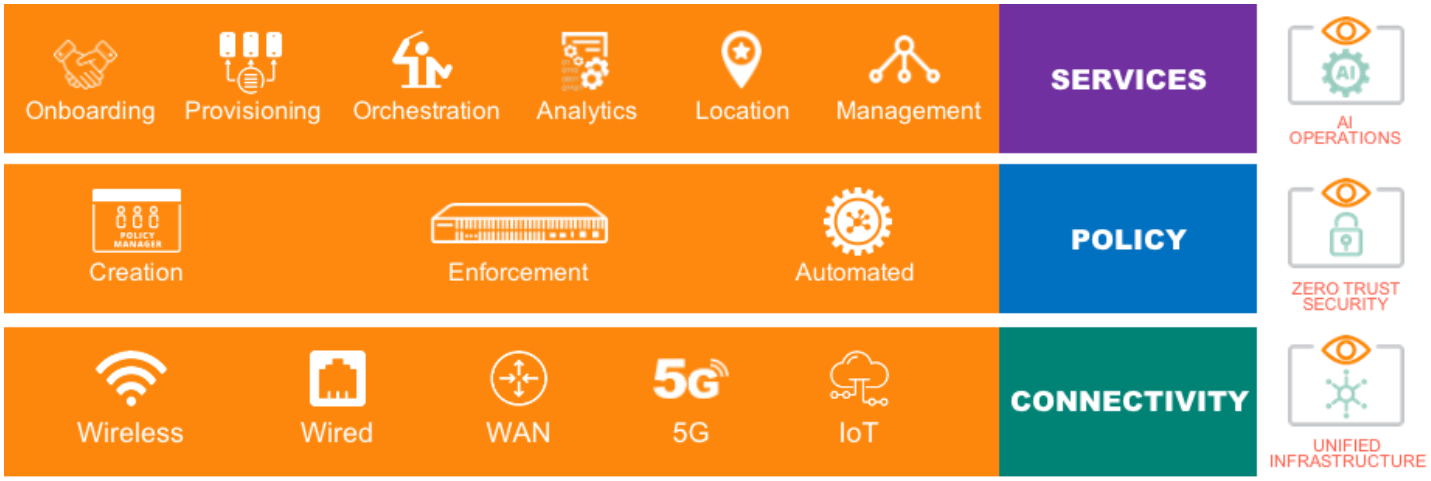
Aruba ESP Data Center Architecture for Spine and Leaf

The Aruba Edge Services Platform (ESP) is an evolution of Aruba's end-to-end architecture, providing a Unified Infrastructure with centralized management leveraging artificial intelligence (AI) operations for improved operational experience that helps enable a Zero Trust security policy. Aruba ESP is the industry's first platform that is purpose-built for the new requirements of the Intelligent Edge.



Aruba ESP Architecture Layers

Aruba ESP offers a breadth of services, including on-boarding, provisioning, orchestration, analytics, location tracking and management. AI Insights reveal issues before they impact users allowing an organization to accomplish tasks quickly and easily with intuitive workflow-centric navigation using views that present multiple dimensions of correlated data. Policies are created centrally and features like Dynamic Segmentation allow the network administrator to implement them over an existing infrastructure. This is possible because the Aruba ESP architecture is built in distinct layers, as shown in the following figure.

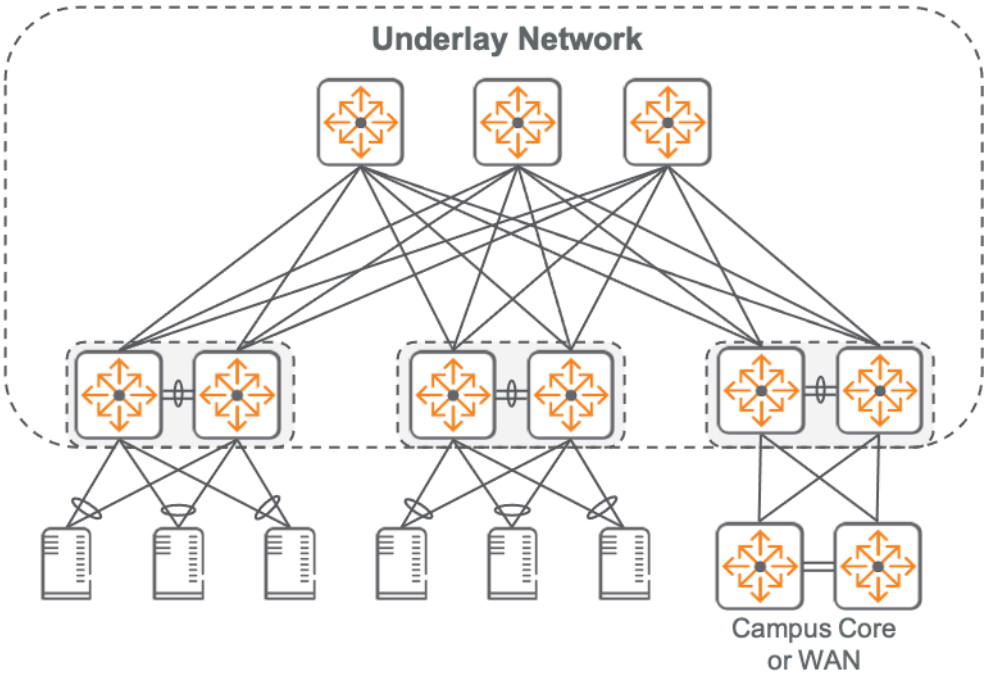


Aruba ESP Data Center Connectivity Layer

The connectivity layer for the Aruba ESP Data Center is implemented on the CX 8000 series Ethernet switches which provide low latency and high bandwidth on a fault tolerant platform designed to carry data center traffic.

Underlay Network

The underlay network is implemented using a spine and leaf fabric topology. It is deployed as a layer-3 routed network. Each leaf is connected to each spine over a routed port and OSPF is the routing protocol. Layer-2 services are not required in the underlay, but can be provided for workloads using virtual overlay networks. The underlay spine and leaf topology optimize performance, increases availability and reduces latency as each leaf is never more than one hop away across multiple load-balanced paths to all other leaf switches.



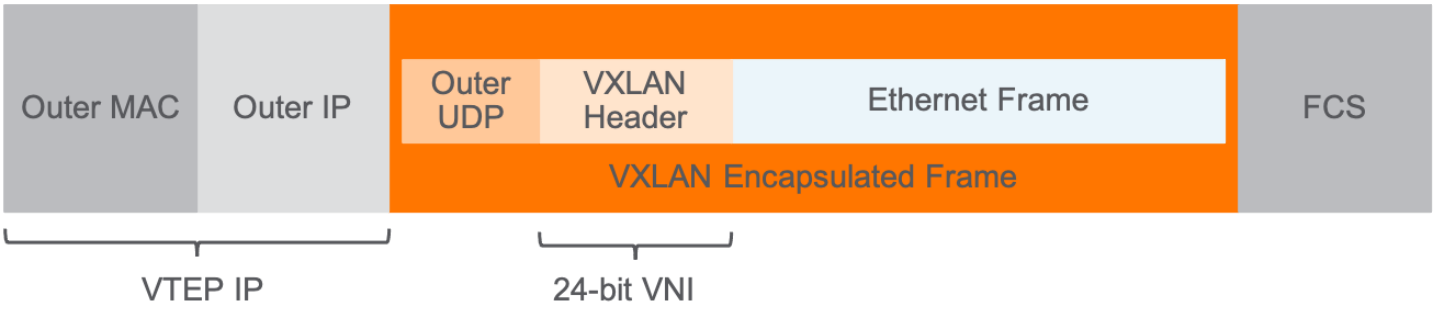
The spine and leaf topology provides a flexible, scalable network design that can expand to accommodate a growing data center without disrupting the existing network. It is easy to begin with a small, one or two rack fabric that can increase capacity without having to replace existing hardware. Top-of-rack ports on leaf switches are used for incrementally adding compute capacity to a rack. Ports on spine switches are used to add additional racks to the fabric. The maximum size of the fabric is determined by the port density on the spine and it is an important consideration for supporting future growth. A minimum of two spine switches is recommended for any size fabric in order to provide high availability and fault tolerance. Additional spine switches increase overall fabric capacity and reduce the fault domain in case a spine must be taken out of service.

Aruba ESP Data Center Policy Layer

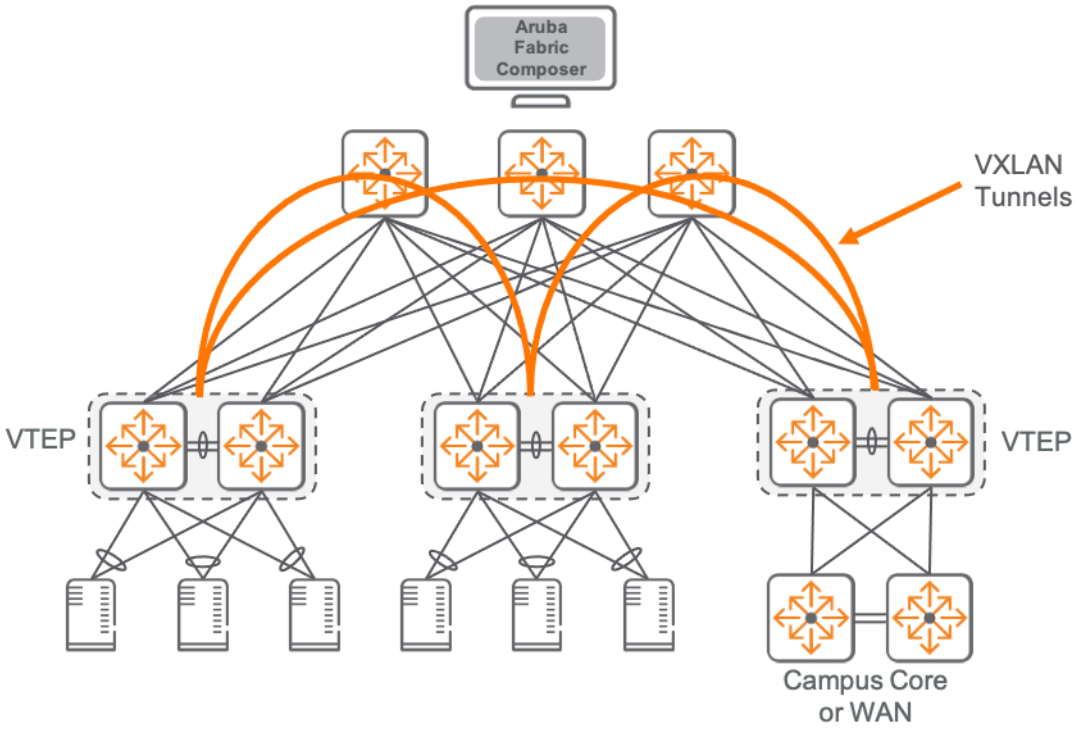
The policy layer for the Aruba ESP Data Center is implemented by the use of overlay technologies and traffic filtering mechanisms for isolating user and application traffic.

Overlay Network

An overlay network is implemented using Virtual Extensible LAN (VXLAN) tunnels that provide both layer-2 and layer-3 virtualized network services to workloads directly attached to the leaf switches. Similar to a traditional VLAN ID, a VXLAN Network Identifier (VNI) identifies an isolated layer-2 segment in a VXLAN overlay topology. Symmetric Integrated Routing and Bridging (IRB) capability allows the overlay networks to support contiguous layer-2 forwarding and layer-3 routing across leaf nodes.



A VXLAN Tunnel End Point (VTEP) is the function within leaf switches that handles the origination and termination of point-to-point tunnels forming an overlay network. A single logical VTEP is implemented when redundant leaf switches are deployed in a rack. Spine switches provide IP transport for the overlay tunnels, but do not participate in the encapsulation/de-encapsulation of VXLAN traffic.

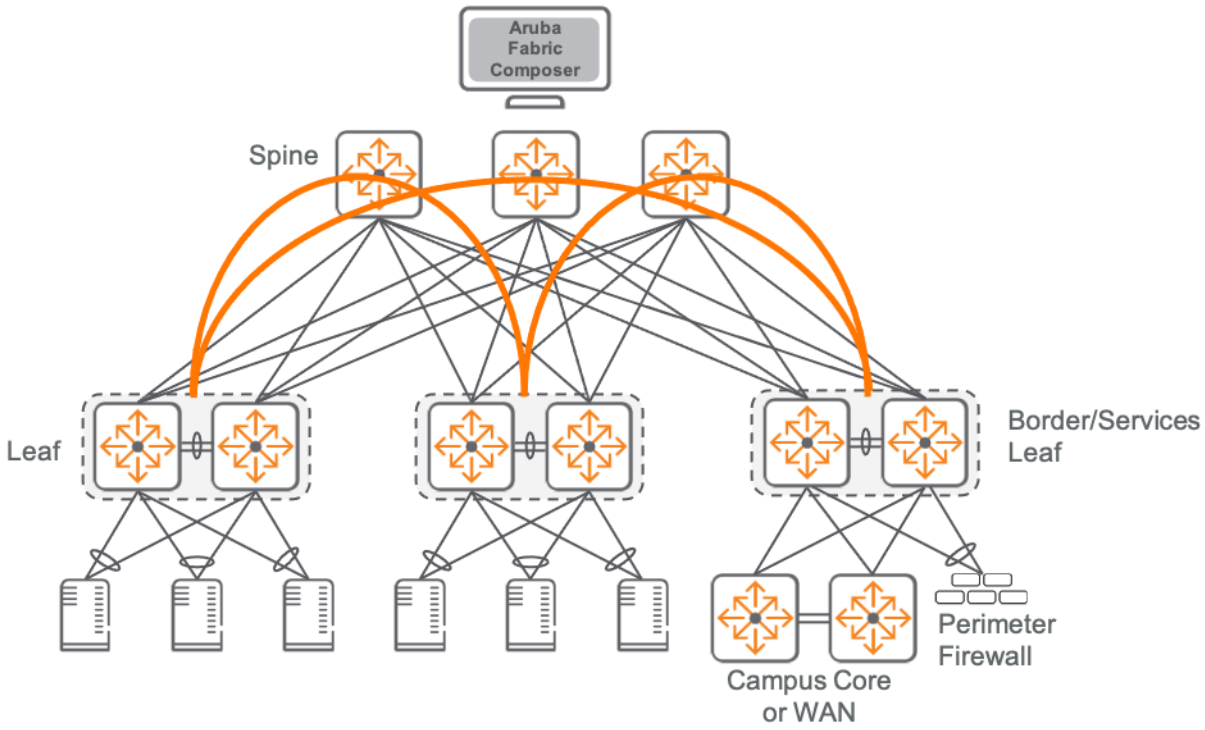


Attached hosts are learned at the leaf switch using Ethernet link layer protocols. Remote learning across the VXLAN fabric is accomplished using Multiprotocol BGP (MP-BGP) as the control plane protocol and a dedicated Ethernet Virtual Private Network (EVPN) address family for advertising host IP and MAC prefixes. This approach minimizes flooding while enabling efficient, dynamic discovery of remote hosts within the fabric.

Security and Segmentation

In a VXLAN spine and leaf design a pair of leaf switches is the single entry and exit point to the data center. This is called the border leaf but it is not required to be dedicated to that function. Compute hosts and firewalls may also be attached. Typically, the border leaf is where a set of policies are implemented to control access into the data center network. These policies are the first layer of security for data center applications. They limit access to only permitted networks and hosts while also monitoring those connections. Protecting the data center perimeter is usually implemented in one or both of the following ways:

- Border Leaf ACLs - When IP subnets inside the data center are designed in a way that can map to security groups or business functions, Access Control Lists (ACL) at the border leaf can provide policy enforcement from user locations into data center applications. If subnets cannot be mapped to security groups, the ACLs can become difficult to manage and scale in larger environments. The primary benefit of perimeter ACLs is that they can be implemented directly on the switching infrastructure to enforce a foundation of policy from which to establish user access into the data center. Policies implemented using switch ACLs specifically target layer 3 and layer 4 constructs. ACLs are not stateful or application aware.
- Perimeter Firewalls - Dedicated security systems at the perimeter can offer advanced monitoring, application aware policy enforcement and threat detection. Perimeter firewalls are typically deployed in transparent or routed mode. In transparent mode the firewalls behave like inline devices. All user and network control traffic will pass transparently through them. In routed mode a firewall will participate in the routing control plane and can be deployed in a configuration which limits the amount of traffic subject to deep inspection in order to maximize the value of that capability. It is important to note that firewalls providing stateful inspection require symmetric forwarding. A stateful firewall must process the connection establishment in order for it to correctly apply policy to the subsequent flow.



Policy inside a VXLAN spine and leaf data center can also be implemented using firewall appliances that are commonly deployed in what is called a services leaf. The firewalls connected at the services leaf are used as the default gateway for hosts requiring specific services accessible through the firewall. An advantage of this approach is the ease with which a Layer-2 overlay network can be used to transport host traffic to the firewall. The disadvantage of this approach is that it relies on a centralized gateway and prevents the use of an active gateway at every ToR for optimal forwarding.

Some vendors offer virtualized firewall services within a hypervisor environment. This approach can provide granular, service level policy enforcement while also allowing for the use of active gateways. VMware NSX is an example of a product able to integrate in this way. VXLAN overlays may be implemented in both hardware and software in order to achieve optimal network virtualization and distributed firewall services while securing east-west traffic inside the data center.

Aruba ESP Data Center Services Layer

The Aruba Data Center solutions include management plane choices enabling an organization to apply the approach which suites their needs best. Aruba Central provides a cloud management solution for the end-to-end Aruba ESP solution. Aruba NetEdit provides the same multi-device configuration editor and topology mapper now found in Aruba Central in an on-premises offering. Aruba Fabric Composer is a fabric automation tool which provides a simplified, work-flow based method of fabric configuration also offered as an on-premises solution.

Aruba Central

Aruba Central is designed to simplify the deployment, management and optimization of network infrastructure. The use of integrated AI-based machine learning, and unified infrastructure management provides a all encompassing platform for digital transformation in the enterprise.

Aruba Central provides advanced services to facilitate transformational data center roll-outs. With NetEdit MultiEditor capability now integrated into Central it's possible deploy complex, multi-device, multi-layer configurations from the cloud to your data center. The Network Analytics Engine provides real-time alerts on the state of your switches and allows for rapid analysis of intermittent problems. Aruba Central is cloud hosted for elasticity and resiliency which also means that end users never have to be concerned with system maintenance or application updates.

Workflow based configurations within Central allow for efficient, error free deployments of Aruba solutions anywhere in the world. The workflows are based on common, best practices approaches to network configuration. They enable new devices to come online quickly using new or existing network configurations.

AIOps

According to Gartner Inc., AIOps (Artificial Intelligence for IT operations) combines big data and machine learning to automate IT operations processes, including event correlation, anomaly detection and causality determination.

Aruba AIOps, driven by Aruba Central eliminates manual troubleshooting tasks, reduces average resolution time as…Aruba's next generation AI uniquely combines network and user-centric analytics to not only identify and inform staff of anomalies, but also applies decades of networking expertise to analyze and provide prescriptive actions

AI Insights are available to monitor connectivity performance, RF management, client roaming, airtime utilization, and wired and SD-WAN performance. Each insight is designed to reduce trouble tickets and ensure SLAs by addressing network connectivity, performance, and availability challenges

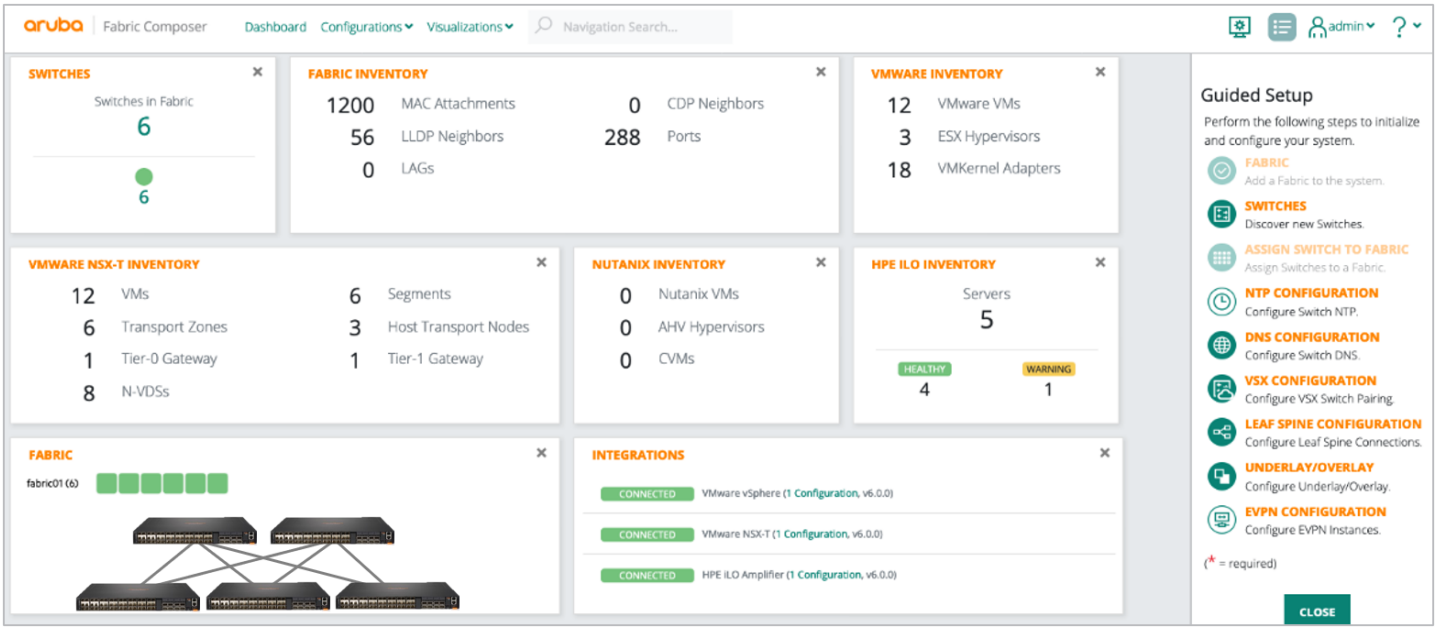
AI Assist uses event-driven automation to trigger the collection of troubleshooting information, to identify issues before they impact the business, and virtually eliminate the time-consuming process of log file collection and analysis. Once log information is automatically collected, IT staff are alerted with relevant logs that can be viewed and even shared with Aruba TAC, who can more quickly assist with root cause determination and remediation.

Aruba Fabric Composer

Aruba Fabric Composer provides API driven automation and orchestration capabilities for the Aruba ESP Data Center. Aruba Fabric Composer discovers and interrogates data center infrastructure in order to automate and accelerate spine and leaf fabric provisioning as well as day-to-day operations across rack-scale compute and storage infrastructure. Aruba Fabric Composer orchestrates a set of switches as a single entity called a fabric and allows the operator to orchestrate data center resources using an application centric approach to visualizing network and host infrastructure.

Visualization of the data center network fabric includes physical and virtual network topologies as well as host infrastructure through integration with Aruba-OS CX, HPE iLO Amplifier, HPE Simplivity, VMWare vSphere, and other leading data center products. In addition to providing a complete view across the fabric, Aruba Fabric Composer makes network provisioning accessible to more than just network staff. It provides a platform for orchestrated deployment of host and networking resources across the fabric through a guided workflow user interface. Aruba Fabric Composer ensures a consistent and accurate configuration of a spine and leaf data center whether or not an overlay network is also deployed.

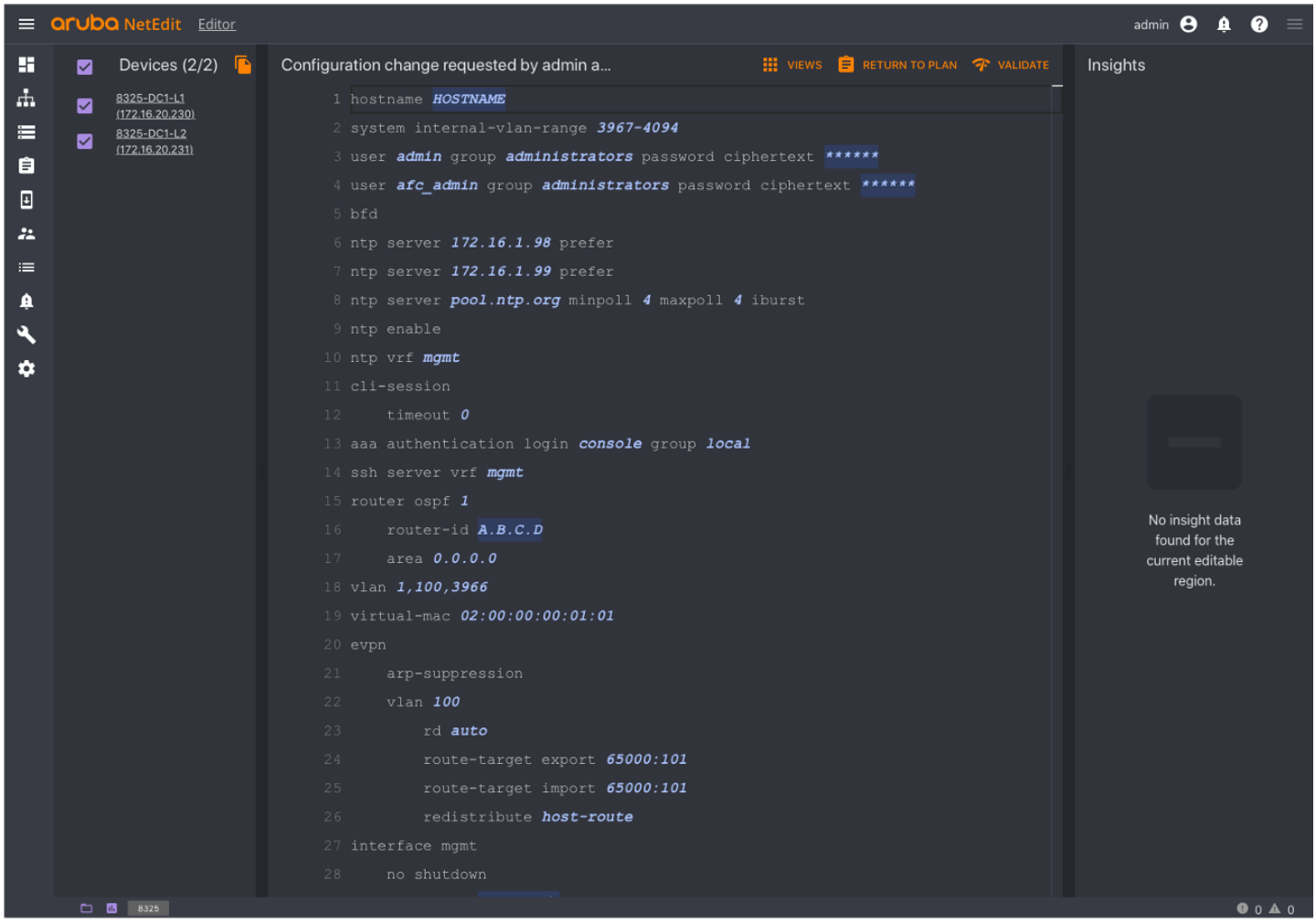
Aruba Fabric Composer is an end-to-end data center network management tool recommended for new data center deployments based on a spine and leaf fabric topology. It is particularly helpful when also deploying an overlay topology using VXLAN-EVPN as the Fabric Composer will configure both the underlay and overlay routing automatically using basic IP information provided by the operator.



Aruba NetEdit

Aruba NetEdit enables IT teams to automate the configuration of multiple switches to ensure deployments are consistent, conformant, and free of errors. It enables automation workflows without the overhead of programming by providing operators with a user-friendly, CLI-like interface. NetEdit also provides a dynamic network topology view to ensure an up-to-date view of the network.

When deploying an Aruba Data Center Network using on-premises tools, plan to deploy NetEdit for detailed configuration management. While Aruba Fabric Composer enables fast, error-free spine and leaf implementations, NetEdit provides the ability to tailor that configuration when necessary. Together, Fabric Composer and NetEdit deliver an automated, integrated and validated network configuration ready to support the needs of any data center network.

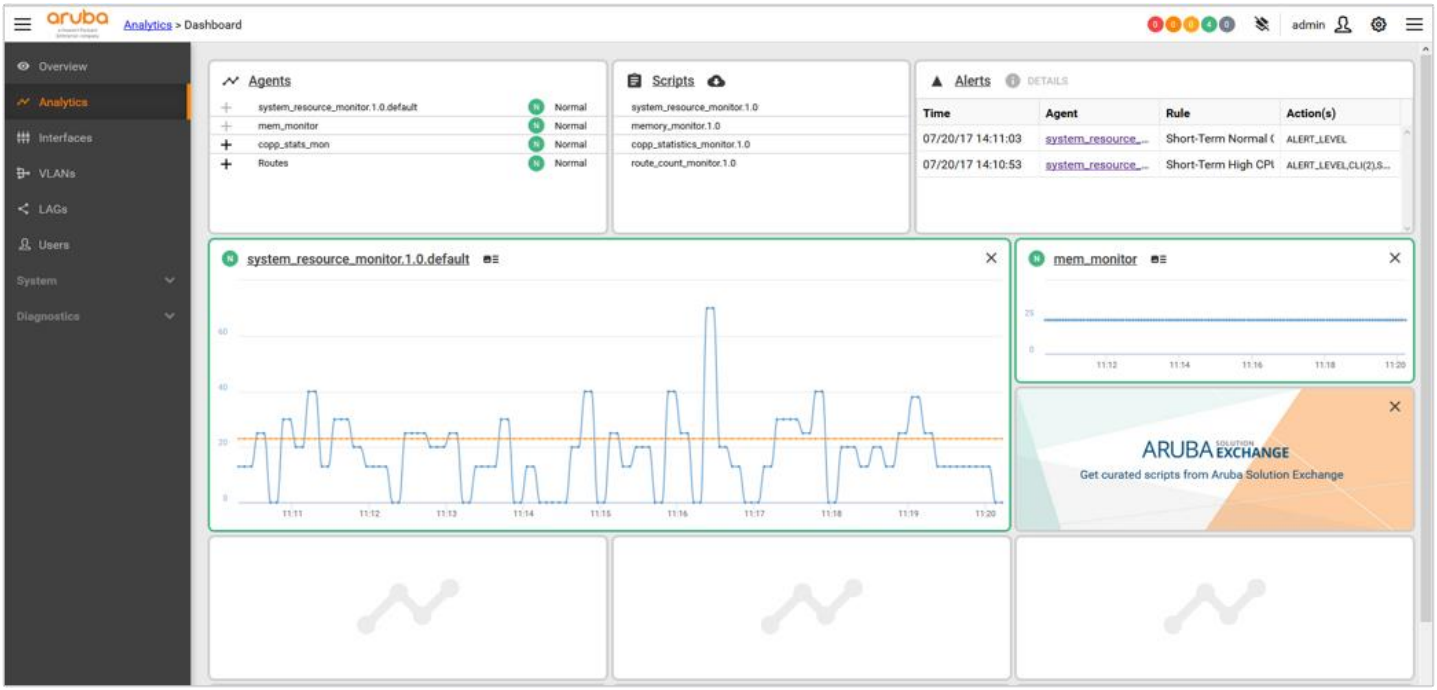


Aruba Network Analytics Engine

NAE provides a built-in framework for monitoring and troubleshooting networks. It automatically interrogates and analyzes network events to provide unprecedented visibility into outages and anomalies. Using these insights, IT can detect problems in real time and analyze trends to predict or even avoid future security and performance issues

A built-in, time series database delivers event and correlation history along with real-time access to network-wide insights to help operators deliver better experiences. Rules-based, real-time monitoring and intelligent notifications automatically correlate to configuration changes. Integrations with Aruba NetEdit and third-party tools such as ServiceNow and Slack provide the ability to generate alerts which trigger actions within an IT service management process.

NAE runs within the AOS-CX operating system in the Aruba CX 6000 and CX 8000 switch series. NAE agents test for conditions on the switch, its neighboring devices, or on traffic that is passing through the network, and then take actions based on the result fo the test.



Choosing an approach

In general, small, Edge Connected data centers are best managed using Aruba Central to ensure consistent configuration anywhere in the world. Larger, centrally located data centers will likely require the use of Aruba NetEdit so that detailed, custom configurations can be written and deployed automatically to multiple network devices. If you're planning to build a spine and leaf topology in your data center, consider using Aruba Fabric Composer as well. When planning to deploy a VXLAN overlay, use of Aruba Fabric Composer is highly recommended in order to simplify the configuration of underlay and overlay services as well as layer 3 segments.

Additional data center services

Planning a data center network involves more than just the network infrastructure itself. It's also necessary to ensure that services are available to bring switches and hosts online and to ensure devices are able to send log messages to a syslog server accessible to people and applications.

It may be useful to leverage the zero touch provisioning (ZTP) capabilities of Aruba switches. In order to do so, the network must provide a dynamic host configuration protocol (DHCP) server on a management LAN with a route to the internet. In addition to the default gateway address, devices will also require at least one domain name service (DNS) server to resolve hostnames required for connectivity to Aruba Central and the Aruba Activate service.

Network time protocol (NTP) ensures that log data from across the network and in the cloud is time stamped correctly for later analysis. NTP is also required for public key infrastructure (PKI) to function correctly. PKI is required for a variety of access security approaches today. A log management or security information and event management (SIEM) solution is a part of most data centers today. Knowing which will be implemented will help with establishing a configuration baseline for all switches in the network.

Aruba ESP Data Center Design for Spine & Leaf

In order to successfully design a spine and leaf data center it is important to consider all layers of the data center network. This section will provide general design considerations across the different layers of the Aruba ESP Data Center. The reference architecture section will include hardware specific recommendations that you can use to finalize your design.

Connectivity Layer Design

This section will provide design recommendations and best practices for physical connectivity of compute and network infrastructure.

Compute Host Connectivity

The first step to design a data center is to identify which types of connectivity are required by the compute hosts. Server hardware will typically have an Ethernet RJ45 port for a lights-out management device such as HPE iLO. Application connectivity is commonly over redundant links using RJ45 or SFP ports.

The lights-out port is typically connected using a Cat5e or Cat6 copper patch cable into a switch on the management LAN. Typically the host to leaf connections will be either 10Gb fiber using SFP+ modules, copper direct attach cables (DAC) or active optical cables (AOC). DAC cables have limited distance support and can be harder to manage due to the thicker wire gauge when compared against optical cables. AOC cables support longer distances than DAC cables and are thinner and easier to manage. Both DAC and AOC cost less than separate fiber patch cables and optical transceivers but only operate at a single speed. It is important to verify that both the host NIC and the ToR switch are compatible with the same DAC or AOC cable. When separate transceivers and optical cables are used it is also important to verify transceiver compatibility with host NIC, ToR switch, and optical cable type. In most cases the supported transceiver on the host will be different than the supported transceiver on the switch. Always consult with a structured cabling professional when planning a new or upgraded data center.

If a converged network for IP storage traffic is being deployed, then also look for NIC cards supporting offload of storage protocols. Similarly, ensure the hardware also supports VXLAN offloading. These features will help to minimize latency of storage traffic by reducing the load on a host CPU.

Applications can be hosted directly on a server using a single operating system. This is commonly referred to as a bare metal server. Multiple hosts can be virtualized on a single physical server using a hypervisor software layer that runs on top of the server operating system. Examples of this would be VMWare ESXi or Microsoft Hyper-V. Hypervisors generally contain some form of virtual switch that provides connectivity to each virtual machine using Layer-2 VLANs or VXLAN tunnels for micro-segmentation. A successful spine and leaf design should support layer-2 and layer-3 connectivity using untagged and VLAN tagged ports to match the connectivity required to the server and/or virtual switch inside the server. Aruba Fabric Composer provides visibility and orchestration of the configuration required to ensure connectivity between the server and Aruba ToR switches is properly established.

Out-of-Band Management

The Aruba ESP Data Center spine and leaf design uses a dedicated management LAN connecting to switch management ports and host lights out management (LOM) ports. Typically a single management switch is deployed for out-of-band management at every rack. A dedicated management switch ensures reliable connectivity to data center infrastructure for automation, orchestration, and traditional management access.

Top of Rack Design

Deploying switches in the top-of-rack position allows for shorter cable runs between hosts and switches. The result is a more modularized solution with host to switch cabling contained inside a rack enclosure and only switch uplinks exiting the enclosure. This approach helps reduce complexity when adding racks to the data center.

In the Aruba ESP Data Center, each rack is serviced by a redundant pair of VSX configured switches. This allows dual-homed hosts to be connected to two physical switches using a link aggregation bundle for fault-tolerance and increased bandwidth.

VSX enables a distributed and redundant architecture that is highly available during upgrades, which is inherent in its architecture. VSX virtualizes the control plane of two switches to function as one device at layer-2 and as independent devices at layer-3. From a data-path perspective, each device does an independent forwarding lookup to decide how to handle traffic. Some of the forwarding databases, such as the MAC and ARP tables, are synchronized between the two devices via the VSX control plane over a dedicated ISL trunk. Each switch builds the layer-3 forwarding databases independently.

A critical requirement to note when deploying a pair of switches in VSX mode is for three ports to connect the switches to each other. Two should be the same speed as the uplinks ports. A third can be any available port.

For backwards compatibility and to also support future growth, choose a ToR switch that supports connectivity rates of 1, 10, or 25 Gb/s. These connection speeds can be implemented using the same fiber optic media types which makes it easy to increase bandwidth by simply upgrading transceivers or DAC/AOC cables. Keep in mind the following when selecting a ToR switch model:

- **Number of and type of server connections:** Typical rack server configurations support 48 host facing ports but lower density ToR options are available in the CX 8360 series.
- **Host connectivity speed:** To simplify management, consolidate hosts connecting at the same speeds to the same racks and switches. Port speed settings between 25Gb and 10Gb may impact a group of adjacent ports in the switch. Consider port group size, which may vary between models, when planning for a rack requiring multiple connection speeds.
- **Number of uplink ports:** ToR switch models support a range of uplink port densities. When using VSX for redundancy, two uplink ports are used for inter-switch links providing data path redundancy and cannot be used for spine connectivity. Take this into consideration when choosing a switch configuration.
- **ToR to spine connectivity:** The amount and port speed of the uplinks will define the oversubscription rate from the hosts to the data center fabric. As an example, in a four spine deployment at 100Gb a non-oversubscribed fabric can be implemented for racks of 40 servers connected at 10Gb rate.
- **Cooling design:** Different ToR models are available for port-to-power and power-to-port cooling. Optional air duct kit can isolate hot air from servers inside the rack in power-to-port configurations. Cabling can absorb heat and restrict airflow, short cable routes and good cable management will improve the airflow efficiency.

Spine Design

The spine layer provides aggregation for the leaf switches. In a spine and leaf design, each ToR switch is connected to each spine switch. Each leaf to spine connection should use the same link speed in order to ensure multiple, equal cost paths within the fabric. This allows ECMP based routing to ensure connectivity if a link goes down.

The port capacity of the spine switches will define the maximum number of racks the data center can connect. In the case of a redundant ToR design, the maximum number of racks will be half the port count on the spine switch. The use of two spines is the recommended minimum for high availability. Additional spines increase overall fabric capacity and reduce the size of the fault domain in case a spine is taken out of service.

- Determine the media and bandwidth requirements for your racks.
- Determine if you will install single or redundant ToR switches
- Determine how many racks you need for current compute and storage requirements
- Determine the spine switches required to support your planned racks
- Design your data center network for no more than 50% capacity to leave room for growth

If your network will have more than two spine switches, pay special attention to the number of uplink ports available on your chosen ToR switch. Each ToR switch must be connected to each spine in order for ECMP to work effectively within the fabric.

In deciding where to place your spine switches, also take into consideration their distance from the leaf switches and what media type you will use to connect them. Leaf to spine connections will be either 40Gb or 100Gb fiber using QSFP transceivers or active optical cables (AOC) which are similar to a DAC in that the cable and transceiver are integrated.

Underlay IP Design

The underlay of a spine and leaf data center network is the layer which provides IP connectivity between leaf and spine switches. The underlay is the part of the network that ensures VXLAN tunneled traffic (the overlay network) can be forwarded across the fabric.

The Aruba ESP Data Center uses OSPF as the underlay routing protocol. OSPF is a widely used and well understood Interior Gateway Protocol (IGP) that provides straightforward configuration and fast convergence. A single OSPF area and point-to-point interfaces are recommended to minimize complexity and the time required for establishing neighbor adjacencies.

Configure the data center switches for a jumbo MTU of 9198 bytes. This accommodates both the storage protocols likely to be deployed as well as the expanded frame header used by VXLAN.

Policy Layer Design

This section will provide design recommendations and best practices for policy layer design of the data center network.

Overlay Control Plane Design

Host mobility refers to the ability to move physical or virtual hosts within a data center network without changing the host network configuration. This flexibility is particularly powerful when paired with virtualized hosts and can be used to ensure optimized compute resources, high availability of applications, and efficient connectivity for distributed workloads.

In order to maintain a data center overlay and successfully forward traffic through it, VTEPs within the fabric require reachability information about fabric connected endpoints. To enable this remote learning of endpoints and VTEPs, a distributed, dynamic control plane is recommended for the following reasons:

- Traditional flood and learn techniques can consume large amounts of bandwidth due to the replication of traffic in a large spine and leaf environment.
- Network configuration is simplified as the ToRs will automatically learn about other ToR switches inside the fabric.
- A distributed control plane provides redundancy and consistent topology state across the data center fabric switches.
- A distributed control plane allows optimal forwarding via the use of distributed gateways at the ToRs. This makes it possible for the default gateway address to remain the same across the fabric.

The use of MP-BGP with EVPN address families between VTEPs provides a standards-based, highly scalable control plane for sharing endpoint reachability information with native support for multi-tenancy. MP-BGP has been used for many years by service providers to offer secure layer-2 and layer-3 VPN services at very large scale. Network operations are simplified by using an iBGP design with route reflectors so that peering is only required between leaf switches and the spine. Some of the BGP control plane constructs that you should be familiar include the following:

- **Route Distinguisher (RD)** - In order to support multi-tenancy and the likelihood of overlapping IP addressing, the use of an RD associated with a BGP prefix allows the unique identification of the virtual network associated with each prefix. In VXLAN a layer-3 VNID maps to a VRF and represents a tenant with a dedicated virtual network and corresponding routing table.
- **Route Target (RT)** - Route targets are used as an attribute to flag the source of specific prefix updates and as a filtering criteria for importing prefixes into a routing table. In a typical data center environment with any to any communication on the same tenant this attribute is not relevant unless route leaking is required between different virtual networks.
- **Route Reflector (RR)** - To optimize the process of sharing reachability information between VTEPs, the use of route reflectors at the spine will allow for simplified iBGP peering. This design allows for all VTEPs to have the same iBGP peering configuration and eliminate the need for full mesh of iBGP neighbors.
- **Address Family (AF)** - Different types of routing tables (IPv4 unicast, IPv6 unicast, L3VPN…) are supported in MP-BGP. The layer 2 VPN address family (AFI=25) and the EVPN address family (SAFI=70) are used to advertise IP and MAC address information between BGP speakers. The EVPN address family routing table contains reachability information for establishing VXLAN tunnels between VTEPs.

The Aruba ESP Data Center design uses two spine layer switches as iBGP route reflectors. The quantity of destination prefixes and overlay networks will consume physical resources in the form of forwarding tables and should be taken into consideration when designing the network. The reference architecture section will provide hardware guidelines for scaling the design of your data center network.

Segmentation Policy Design

Data center applications are deployed in many different ways. Applications can be hosted in bare metal servers or implemented as virtual machines using hypervisors. Containerized apps are highly distributed and usually require connectivity between multiple compute and service nodes. In some cases, a single data center will host applications for multiple tenants while offering a set of shared services across them. Because of how applications are deployed in a modern data center, it would be incorrect to assume that all security threats are external as the majority of traffic is usually contained within the data center.

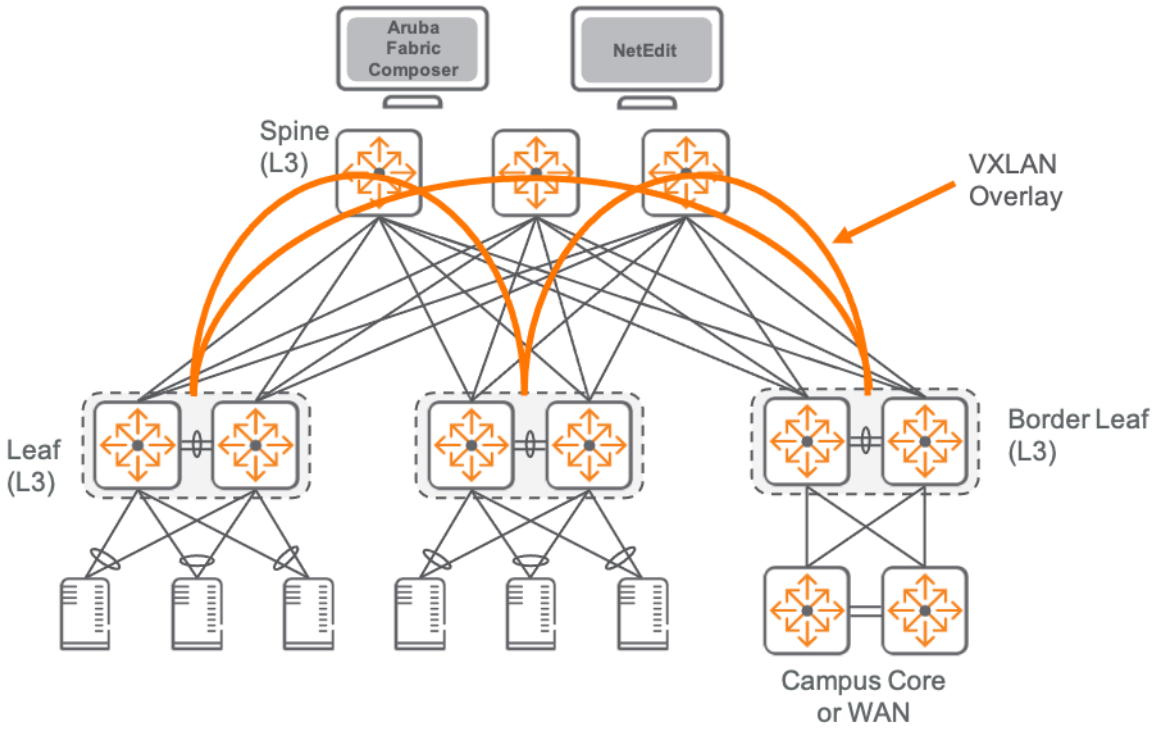
Successful data center policy design begins with understanding the requirements of the applications that will run in the environment. Note that it is often necessary to re-profile legacy applications if sufficient requirements documentation does not exist. From a networking perspective, application profiling should document all of the network connections required for that application to run successfully. These might be to back-end databases or to cloud hosted services. To properly define policy regarding which connections must be permitted and which will be denied, it’s first necessary to know the application profile.

Similarly, a profile of the users accessing the applications and data is typically required. Never leave a data center wide open to a campus, even if it is assumed to be a secure environment. In order to restrict access, it is necessary to understand the various user profiles of the organization associated with the applications and data required for their work. It's particularly important to identify on-campus, remote branch, and mobile field workers so that appropriate data center access profiles can be developed to represent the unique requirements of each.

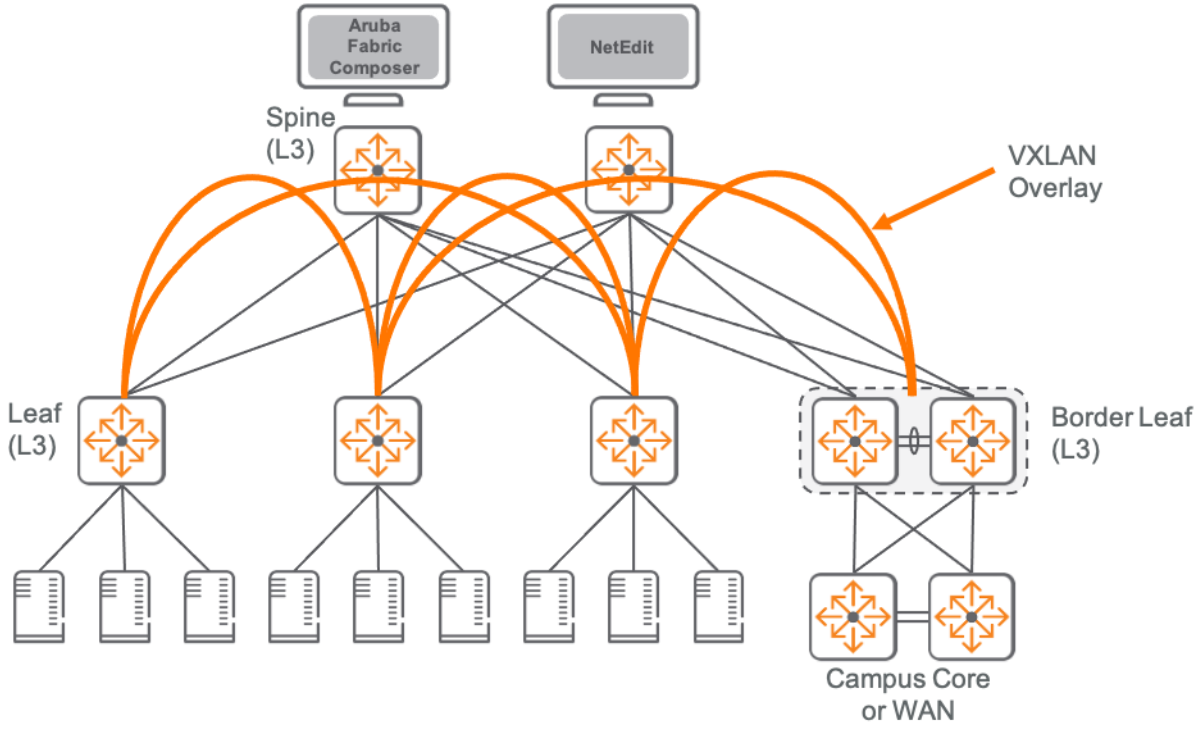
Planning for secure network device management also involves policy considerations. Organizations should plan for building a physically separate management LAN and role-based access control on the network devices. This means that login to a switch requires authentication against an enterprise directory. That would typically be accomplished using the TACACS+ protocol and a policy server such as Aruba ClearPass Policy Manager. Logging facilities, log management, and log analysis should also be considered.

Aruba Reference Architecture for Data Center

The Aruba ESP Data Center reference architecture is designed to support high availability compute racks using redundant top-of-rack switches connected in a layer 3 spine and leaf topology. The spine and leaf topology optimizes performance and also provides a horizontally scalable design that can expand to accommodate a growing data center without disrupting the existing network components. A data center can be initially built with as few as two spine switches. When additional capacity is required up to four spine switches can be deployed in a single fabric. The following figure shows the reference architecture with three spine switches.



Certain application environments do not require high availability at the individual compute host. In this case, a single ToR switch per rack will provide a more cost effective data center network. In this type of implementation, the number of compute hosts deployed per rack should be kept low as a ToR switch under maintenance will impact connectivity to all compute hosts in the rack. The following topology shows a single ToR design with two spine switches.



Reference Architecture Components Selection

The following section provides guidance for hardware selection based on your compute host, availability and bandwidth requirements.

Aruba CX 8300 Data Center Switches

The Aruba CX 8300 portfolio offers two models of fixed configuration data center switches. The CX 8325 model offers the highest port density while the new CX 8360 model offers a variety of port configurations for small and medium spine and leaf topologies. Both models offer the following data center switching capabilities:

- High-speed, fully distributed architecture with line-rate forwarding
- High availability and in-service ToR upgrades with VSX
- Cloud-native and fully programmable modern operating system built on a microservices architecture
- Error free network configuration with software defined orchestration tools
- Distributed analytics and guided troubleshooting provide full visibility and rapid issue resolution
- Hot swappable and redundant load-sharing fans and power supplies
- Power-to-port and port-to-power cooling options for different data center designs
- Jumbo frame support for 9,198 byte frames
- Advanced L2 and L3 features to support VXLAN spine and leaf with MP-BGP / EVPN control plane
- Distributed active gateways for supporting host mobility

Spine Switches

The Aruba ESP Data Center reference architecture is built around two choices of 1RU high density spine switch with QSFP ports capable of 40GbE/100GbE speeds. The 8325 model can support up to 32 compute racks in a single ToR switch topology or up to 16 compute racks in a dual ToR switch topology. The 8360 model can support up to 12 compute racks in a single ToR switch topology or up to 6 compute racks in a dual ToR switch topology.

The primary function of the spine switches is to make routing decisions for the overlay. The primary design considerations when choosing the best switch for your data center spine are:

- Port density
- Ports speeds
- Routing table sizes

SKU	Description	Maximum Rack Capacity
JL626A	8325: 32-port 40GbE/100GbE QSFP+/QSFP28, port-to-power airflow	32 Single ToR / 16 Dual ToR
JL627A	8325: 32-port 40GbE/100GbE QSFP+/QSFP28, power-to-port airflow	32 Single ToR / 16 Dual ToR
JL708A	8360: 12-port 40GbE/100GbE QSFP+/QSFP28, port-to-power airflow	12 Single ToR / 6 Dual ToR
JL709A	8360: 12-port 40GbE/100GbE QSFP+/QSFP28, power-to-port airflow	12 Single ToR / 6 Dual ToR

Leaf Switches

There are two leaf switch models to choose from in the Aruba ESP Data Center reference architecture. Both models are 1RU ToR switches that support high density racks using 10GbE copper or SFP+ ports. SFP ports on the 8360 model also support 10GBASE-T transceivers

For redundant ToR designs, the high and medium density SKUs provide the minimum of four uplink ports that are required for a two spine switch topology. For non-redundant ToR design, medium and low density SKUs provide the minimum of two uplink ports required for a two spine switch topology.

You can mix and match racks of different ToR configurations in a common spine and leaf topology in order to optimize your host aggregation needs. The following table summarize the leaf SKUs available with their corresponding supported designs.

SKU	Description	Rack Design	Spine Design
JL624A	8325: 48-port 1/10/25GbE SFP/SFP+/SFP28, 8-port 40/100GbE QSFP+/QSFP28, port-to-power airflow	High Density - Dual ToR	2 to 4 switches
JL625A	8325: 48-port 1/10/25GbE SFP/SFP+/SFP28, 8-port 40/100GbE QSFP+/QSFP28, power-to-port airflow	High Density - Dual ToR	2 to 4 switches
JL706A	8360: 48-port 100M/1GbE/10GbE 10GBASE-T, 4-port 40/100GbE QSFP+/QSFP28, port-to-power airflow	High Density - Dual ToR	2 switches
JL707A	8360: 48-port 100M/1GbE/10GbE 10GBASE-T, 4-port 40/100GbE QSFP+/QSFP28, power-to-port airflow	High Density - Dual ToR	2 switches
JL700A	8360: 32-port 1/10/25GbE SFP/SFP+/SFP28, 4-port 40/100GbE QSFP+/QSFP28, port-to-power airflow	Medium Density - Dual ToR	2 switches
JL701A	8360: 32-port 1/10/25GbE SFP/SFP+/SFP28, 4-port 40/100GbE QSFP+/QSFP28, power-to-port airflow	Medium Density - Dual ToR	2 switches
JL710A	8360: 24-port 1/10GbE SFP/SFP+, 2-port 40/100GbE QSFP+/QSFP28, port-to-power airflow	Medium Density / Single ToR	2 switches
JL711A	8360: 24-port 1/10GbE SFP/SFP+, 2-port 40/100GbE QSFP+/QSFP28, port-to-power airflow	Medium Density / Single ToR	2 switches
JL702A	8360: 16-port 1/10/25GbE SFP/SFP+/SFP28, 2-port 40/100GbE QSFP+/QSFP28, port-to-power airflow	Low Density / Single ToR	2 switches
JL703A	8360: 16-port 1/10/25GbE SFP/SFP+/SFP28, 2-port 40/100GbE QSFP+/QSFP28, power-to-port airflow	Medium Density / Single ToR	2 switches

Out-of-Band Management Switches

The Aruba ESP Data Center reference architecture uses a management LAN built on dedicated switching infrastructure to ensure reliable connectivity to data center infrastructure for automation, orchestration, and traditional management access. The following table lists the recommended switch models.

SKU	Description	Host ports
JL667A	Aruba CX 6300F 48-port 1GbE and 4-port SFP56 Switch	48
JL668A	Aruba CX 6300F 24-port 1GbE and 4-port SFP56 Switch	24
JL663A	Aruba CX 6300M 48-port 1GbE and 4-port SFP56 Switch	48
JL664A	Aruba CX 6300M 24-port 1GbE and 4-port SFP56 Switch	24
JL724A	Aruba 6200F 24G 4SFP+ Switch	24
JL726A	Aruba 6200F 48G 4SFP+ Switch	48
JL678A	Aruba 6100 24G 4SFP+ Switch	24
JL676A	Aruba 6100 48G 4SFP+ Switch	48

Aruba Fabric Composer

Aruba Fabric Composer is offered as a self-contained ISO or Virtual Machine OVA and can be installed in both virtual and physical host environments as a single instance or as a high-availability, 3 node cluster. Aruba Fabric Composer is available as an annual, per switch software subscription.

Aruba Fabric Composer supports Aruba CX 8325 and 8360 switches.

Ordering information for Aruba Fabric Composer is at the end of the solutions overview.

NetEdit

NetEdit runs as an Virtual Machine OVA on a host. Aruba NetEdit is available from the Aruba Service Portal. Customers must visit the Aruba Airheads Community and create an Airheads account in order to download the NetEdit software.

Ordering information for Aruba NetEdit is at the end of the data sheet.

Reference Architecture Physical Layer Planning

The following section provides guidance for physical layer planning of your data center switches.

Cables and Transceivers

Please refer to the following documents to make sure that you select supported cables and transceivers as you plan for physical connectivity inside your data center:

HPE Server Networking Transceiver and Cable Compatibility Matrix

ArubaOS-Switch and ArubaOS-CX Transceiver Guide

Port Speed Groups

When planning for ToR configurations that require server connectivity at multiple speeds, it is important to note that setting the speed of a port might require adjacent ports to then operate at that same speed.

The CX 8325 series switch has a default speed of 25GbE. To change the speed to 10GbE will impact groups of 12 ports. The CX 8360 series switch allows individual ports to operate at different speeds without impacting adjacent ports unless MACSec is in use. Ports configured to use MACSec must all be configured to operate at the same speed.

Split Ports

Breakout cables can be used to split a 40 Gb/s or 100 Gb/s port, into four lower speed connections (4x10 Gb/s and 4x25 Gb/s). Please refer to the ArubaOS-Switch and ArubaOS-CX Transceiver Guide for selecting supported breakout cables and switch support for split ports.

Media Access Control Security

Media Access Control security (MACsec) is a standard defined in IEEE 802.1AE which extends standard Ethernet to provide frame level encryption on point-to-point links. This feature is typically used in environments where additional layers of data confidentiality are required or where it is impossible to physically secure the network links between systems. Please refer to the following table for details of MACsec support in the Aruba switching portfolio:

SKU	Description	Supported Ports
JL700A	8360: 32-port 1/10/25GbE SFP/SFP+/SFP28, 4-port 40/100GbE QSFP+/QSFP28, port-to-power airflow	1-4 SFP+/SFP28
JL701A	8360: 32-port 1/10/25GbE SFP/SFP+/SFP28, 4-port 40/100GbE QSFP+/QSFP28, power-to-port airflow	1-4 SFP+/SFP28

Reference Architecture Capacity Planning

The following section provides capacity planning guidance for the Aruba ESP Data Center spine and leaf reference architecture.

Bandwidth Calculations

A spine and leaf network design provides maximum flexibility and throughput for your Aruba ESP Data Center implementation. To achieve the greatest level of performance, a spine and leaf topology can be designed for zero oversubscription of bandwidth. This results in a data center network that will never be congested because the bandwidth available to hosts is equal to the bandwidth between leaf and spine switches.

A significant advantage of a spine and leaf design is that additional capacity can be added as needed by adding additional spine switches and/or increasing the speed of the uplinks between leaf and spine switches. A rack with 40 dual homed servers with 10GbE NICs could theoretically generate a total load of 800G of traffic. For that server density configuration a 1:1 (non-oversubscribed) fabric could be built with four spine switches using 4x100GbE links on each. In practice most spine and leaf topologies are built between 2.4:1 and 6:1 server to fabric oversubscription ratio.

Network and Compute Scaling

The Aruba ESP Data Center reference architecture provides enough capacity for most deployments. With the use of distributed gateways and symmetric IRB forwarding the MAC and ARP tables are localized to directly attached compute nodes and are not impacted by the amount of racks deployed. The amount of IP prefixes will be a function of the total number of nodes, fabric links and also the number of physical and/or virtualized servers. The border leaf is typically the node with the highest control plane load as it handles both internal and external connections. Route summarization is a good practice to reduce the redistribution of IP prefixes between domains.

The Aruba data center reference architecture was thoroughly tested in an end-to-end solution environment that incorporates best practices deployment recommendations, applications and load profiles that represent production environments.

Please refer to the product data sheets on Aruba Campus Core and Aggregation Switches for detailed specifications not included in this guide. The following table provides validated multi-dimensional profiles that you can use for capacity planning of your spine and leaf design.

Feature	8325 Leaf	8360 Leaf	8325 Spine	8360 Spine
Host Scale - IPv4/ARP	30,000	50,000	N/A	N/A
Host Scale - IPv6/ND	15,000	50,000	N/A	N/A
Routing - IPv4 Routes	10,000	16,000	72,000	100,000
Routing - IPv6 Routes	1,000	8,000	20,000	100,000
Routing - OSPF Neighbors	4	4	128	64
VXLAN - Overlay VRFs (L3 VNI)	32	32	N/A	N/A
VXLAN - Host VLANs (L2 VNI)	1024	512	N/A	N/A
Active Gateway SVIs	1000	512	N/A	N/A

Summary

Data center networks are changing rapidly. The most pressing challenge is to maintain operational stability and visibility while placing compute and storage resources where they need to be in order to best serve users. In addition, data center teams are being asked to support the rapid pace of DevOps environments including requirements to connect directly with public cloud infrastructure. Given the rapidly changing landscape for data center requirements it's critical that network and system engineers be provided with tools to simplify and automate complex infrastructure configurations.

The Aruba Networks ESP Data Center is built on a technology platform which provides the tools for transforming the data center into a modern, agile, services delivery platform satisfying the requirements of organizations large, small, distributed, and centralized. The Aruba CX operating system simplifies operations and maintenance with a common switch operating system across the campus, branch, and data center, managed from the cloud or on-premises, and backed by an artificial intelligence capability which provides best practices guidance throughout the operational lifecycle of your network.