Designing The Next Generation Campus with Arista's Modern Stacking Approach (EOS SWAG + CloudVision Leaf-Spine Stack Management)

Stacking the Odds: Why Legacy Switch Stacking Falls Short in Modern Networks

Introduction

Switch stacking emerged in the late 1990s and early 2000s as a solution to simplify the management of multiple network switches. By linking switches together into a "stack," administrators could **manage them as a single entity** and provide a **single CLI interface**, reducing complexity in configuration and maintenance. Another key benefit was that stacking conserved IP addresses by allowing multiple switches to operate under a **single management IP**, which was especially valuable in environments where IP address availability was limited. This approach gained popularity because it allowed for easier scalability and centralized control.

However, legacy switch stacking architectures have notable design drawbacks. Traditional stack topologies like Chain and Ring limit the number of switches that can be effectively added to the stack. As the stack grows, the complexity of inter-switch communication increases, and the stacking backplane may not provide adequate bandwidth to handle the traffic. Even in a ring topology, while redundancy exists, the topology's reliance on a master switch or specific links means failures can still cause significant issues and the traffic might take multiple hops through the members of the stack causing severe backplane congestion. Traditional stacking is rigid and less adaptable to changing network requirements. Expanding or modifying the topology often requires downtime or complex reconfiguration. This makes traditional stacking unsuitable for networks requiring, hitless, large-scale growth. Troubleshooting and upgrading stacked switches can also be cumbersome, as issues within one switch can affect the whole stack. As networks grow more complex, designs that focus on distributed, resilient architectures like spine-leaf topologies are often preferred over traditional switch stacking.

Arista's Flexible Approach to Modernize Campus Network Design

Arista's flexible and incremental approach to campus network design provides customers with unparalleled choice for designing their next gen campus network. Before we describe our network design in detail, let's familiarize with Arista's foundational building blocks of campus networks designs.

Arista Campus Design Building Blocks

The campus network is abstracted into topologies and control plane architectures that are repeatable and conform to Arista Validated Designs. This diagram below illustrates the campus building blocks with "pods" representing the L2 and L3 architectures available (Figure-1).

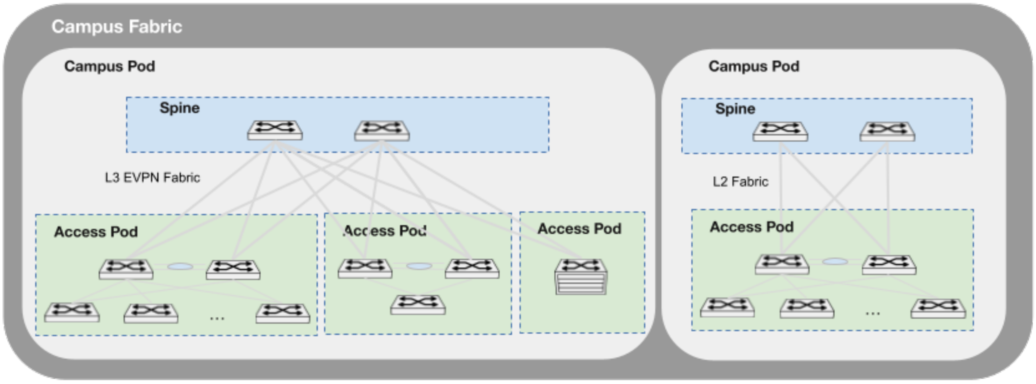

Figure 1: Arista Campus Design with L2 & L3 Architecture

At the topmost level, there is the Campus. Within each Campus, there is at least one Campus Pod. A Campus Pod is an area of the network consisting of a distribution layer of spines that provides uplink connectivity for one or multiple downstream Access Pods. These Access Pods provide endpoint users with access to the wider network.

Campus Fabric: The campus fabric contains campus-pods

Campus Pod: A campus pod contains a set of spines and a layer of access pods. Each spine in a pod provides uplink connectivity for leaf switches in an access pod. Depending on the campus pod topology, spine switches may be configured in an MLAG pair, providing anycast gateway redundancy.

Access Pods: Access pods contain a group of switches that provide users access to the network. These access pods are comprised of leaf switches and member leaf switches

Leaf switches: provide both user access and uplink connectivity for member leaf switches. Leaf switches in a single access pod can be either a standalone device or an MLAG pair of devices

Member Leaf switches: extend an access pod. They provide the access pod with additional port density for user access by uplinking to leaf switches in the access pod. Member leaf switches can form an MLAG pair, which can provide redundancy for connected devices, such as access points.

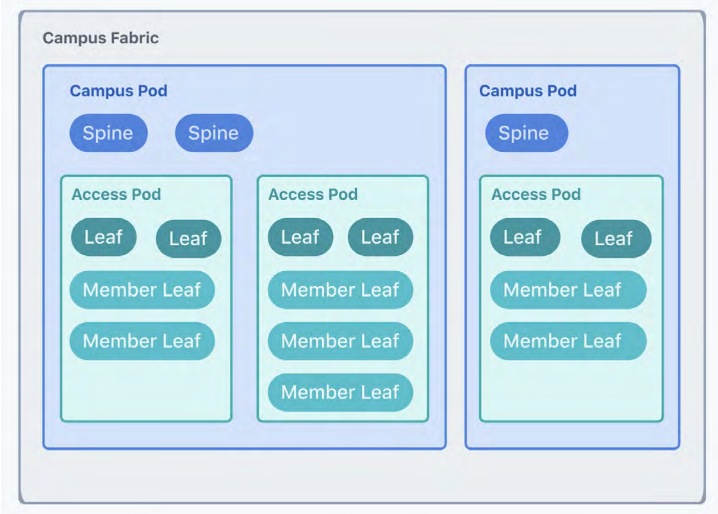

Figure 2: Arista Campus Design Building Blocks

Campus Network Design Flexibility

Arista offers a range of network design choices that provide unparalleled flexibility and high availability, all powered by a common Extensible Operating System (EOS) for a consistent operating model. The **MLAG Leaf-Spine Design** delivers a distributed architecture with active-active redundancy, enabling scalable and fault-tolerant networks optimized for modern campus environments. For environments requiring high-density and centralized control, the **Modular Chassis** option combines scalability with advanced redundancy features like dual supervisors and fabric modules, ensuring performance and resilience. **Standalone switches** provide cost-effective solutions for smaller deployments or edge applications, offering simplicity without compromising performance or reliability.

Lastly, **SWAG (Switch Aggregation Group)** extends the same distributed software in EOS built for Modular Chassis to a group of standalone switches, that can now be configured in a SWAG cluster, providing all the benefits of legacy stacking but done with a modern approach.

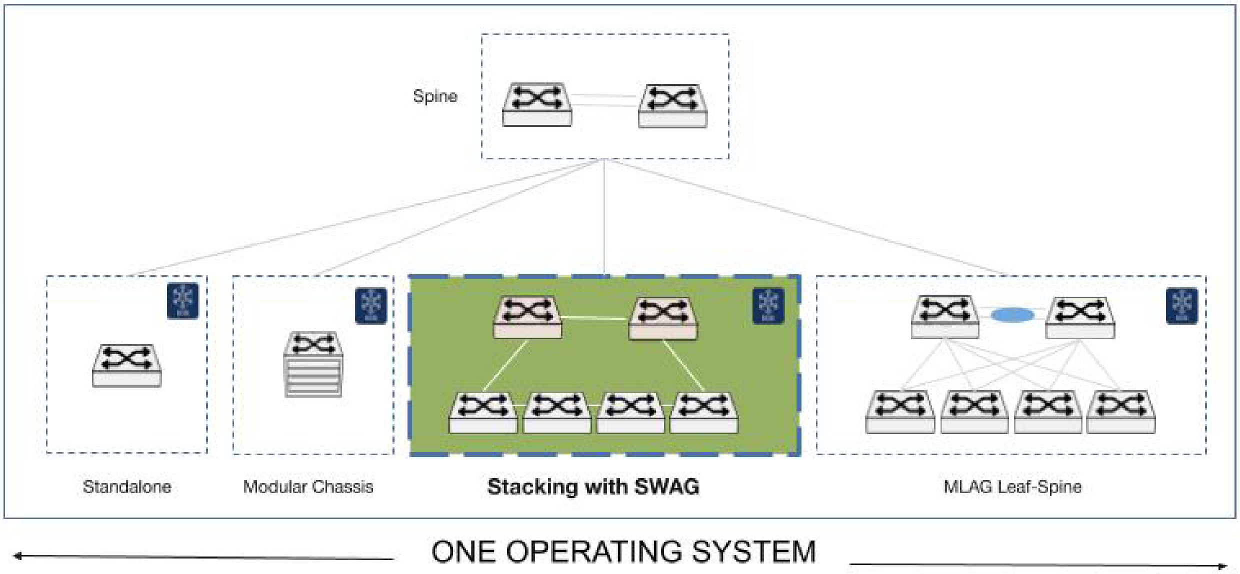

Figure 3: Arista Campus Network Design Choices

Each design leverages Arista's EOS, ensuring consistent functionality, rich telemetry, and simplified operations across all architectures, empowering organizations to tailor their network to meet their specific requirements.

At the management layer, CloudVision Leaf-Spine Stack (LSS) Management provides the ability to manage a group of switches collectively, independent of network design, for common operational tasks, such as: onboarding; provisioning; monitoring; upgrading; or adding/removing devices.

Let's discuss each of these network design choices and management capabilities in detail.

SWAG (Switch Aggregation Group)

A SWAG (Switch Aggregation Group) consists of one or more forwarding switches (workers) and one or two supervisors (active & standby) functioning as a single virtual chassis. This allows EOS to treat these switches as a single unified instance. The management & control plane model of a SWAG feels like a single virtual chassis, but is composed of individual switched instances.

Key Benefits of SWAG

1. Single IP Address to manage all the members of the SWAG cluster
2. Single CLI session to the SWAG cluster via the primary supervisor
3. The SWAG architecture can scale to up to a theoretical 48 switches in a single cluster allowing for increased IP address conservation
4. Can extend the scope of the cluster beyond just 'IDF closet' to a larger domain in the network
5. Supports the highly resilient and the high capacity Arista validated Leaf-Spine topology in addition to Ring and Chain topologies
6. Is built on the same field tested EOS software features designed for Modular Chassis

SWAG Architecture

The figure below depicts a logical switch aggregation group (SWAG) topology along with the control network. In the SWAG design, the EOS software runs on one or two supervisor nodes and it leverages the existing chassis software.

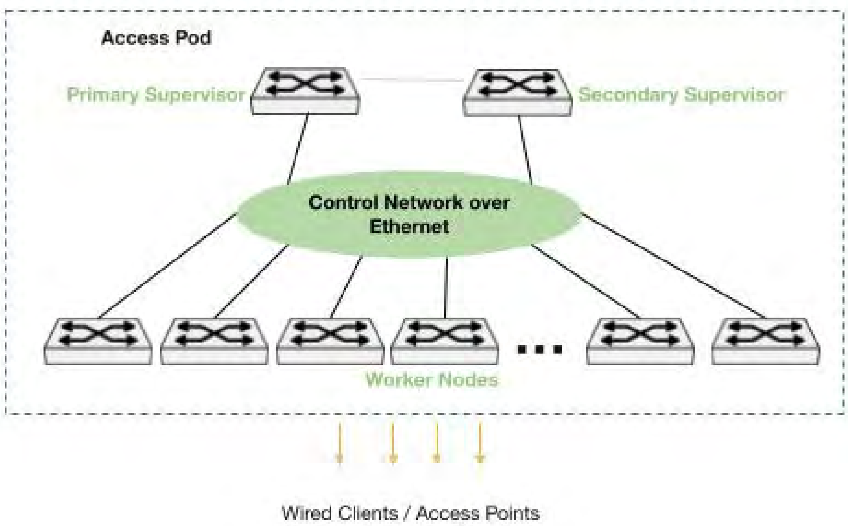

Figure 4: Arista SWAG Architecture

SWAG is booted up, like other Arista switches, with a version of EOS. The Supervisor will ensure that the same EOS runs in every member. The EOS takes the personality of the supervisor on nodes designated as such providing centralized control, management, and monitoring functions. The rest of the switches take the role of workers. E

OS running in the supervisor nodes does the following functions:

- The role of the EOS in the supervisor is almost identical to that of a supervisor in a modular system. The supervisor maintains all the control-plane states, keeps track of the configuration, image management, and all monitoring functions. We'll illustrate this with an example.

- The supervisor also exposes a single SNMP, logging, syslog, alerting, streaming end-point to both Arista CV (CloudVision) and 3rd party applications.
- The supervisor presents a unified CLI and management interface of the entire stack whereby any config is done only in the supervisor. Again, the entire SWAG feels like a modular system from a configuration, management, provisioning, and monitoring perspective.

The Control Network

The Control Network refers to the set of interconnections between nodes of the SWAG that allows any node to communicate with any other node on top of standard ethernet. Arista SWAG allows for the shortest path between any two nodes to be dynamically determined.

The Control Network allows for three types of data traffic flows.

- Node-to-node control, monitoring, and management data traffic
- Control packets from worker nodes to the supervisor and responses to the worker node
- Data packets sourced by end-hosts off a worker so that the packet is sent to that node that has the uplink to the higher layers

Members are admitted into the SWAG, through the control network, either using static definition or through ZTP definitions in the supervisor. The SWAG topology discovery mechanism dynamically determines peer connectivity and the topology also reconverges quickly in the event of link and device failures.

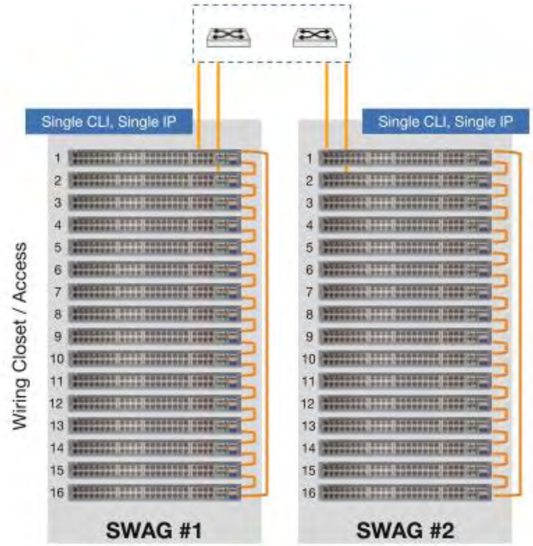

Figure 5: Multiple Members in a SWAG Cluster

Supported Topologies

SWAG supports Ring, Chain and Leaf-Spine topology and the architecture can theoretically scale up to 48 switches in a SWAG cluster.

Leaf-Spine Topology

SWAG also supports the Arista validated Leaf-Spine topology in addition to Ring and Chain. This brings all the benefits of Leaf-Spine Design described earlier to SWAG.

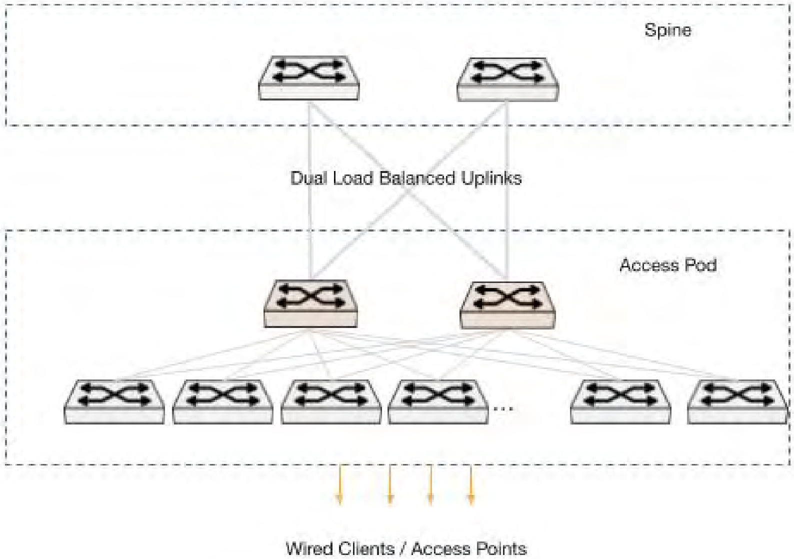

Figure 6: SWAG with Leaf-Spine Topology

In this design, the two switches that are designated as supervisors (active and standby) typically would have the uplink to the Spine (where the Spine itself is deployed in MLAG fashion). and This ensures that the traffic is load balanced between the two uplinks of both the switches.

Each member leaf switch (workers in SWAG) have dual, active/active connections to both the supervisors and traffic can be distributed to both the supervisors across both the links. Though only one supervisor is primary (from a control plane standpoint), the uplink from both the supervisor switches can be active at the same time thus increasing the capacity of the Access Pod.

Ring Topology

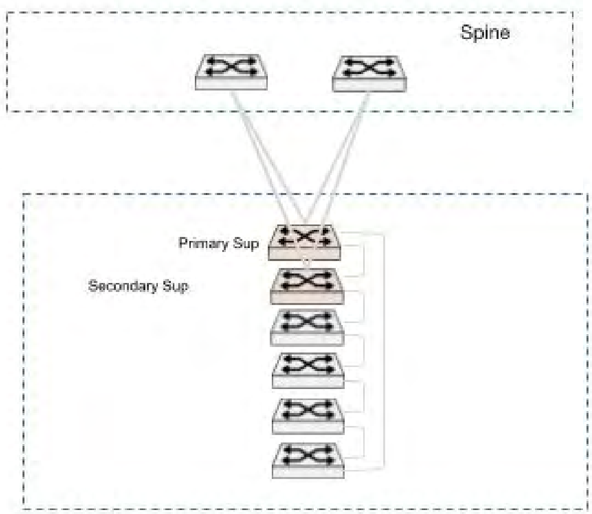

Figure 7: SWAG with Ring Topology

Ring topology is a common design used in stacked switches where the switches are connected in a circular fashion. Each switch in the stack is directly connected to two adjacent switches—one on either side—forming a closed loop. This topology provides redundancy and fault tolerance while enabling high availability for network operations.

Traffic can flow in both directions around the ring, providing a backup path in case of a failure. If one link or switch fails, the stack reconfigures itself to bypass the failed segment, ensuring continued connectivity.

The biggest disadvantage of Ring Topology is that as the stack grows, latency can increase due to the need to traverse more switches within the ring. This becomes even more pronounced when one of the switch in the Ring fails, forcing data to take the longest path in the chain creating performance bottlenecks.

Chain Topology

In the Chain topology multiple switches are connected in a linear sequence, forming a single path for data flow between the switches. Each switch is connected to its immediate neighbor, starting from the first switch and ending at the last, without forming a closed loop as in ring topology.

As you can imagine, this topology suffers from many disadvantages and is really not recommended for Enterprise deployment. This topology lacks redundancy as it provides no backup path. A failure in any stacking cable or switch will break connectivity for all switches downstream of the failure. This makes the topology less resilient compared to ring or modular designs.

As traffic must pass through intermediate switches, congestion can occur, especially if high volumes of traffic are routed through the chain. And as the chain grows longer, latency increases because traffic must traverse multiple switches to reach its destination.

MLAG Leaf-Spine Design

Leaf-spine design with MLAG (Multi-Chassis Link Aggregation Group) is a modern network architecture that combines the high-performance, scalable, and low-latency benefits of the leaf-spine topology with enhanced redundancy and load balancing provided by MLAG. The Leaf-Spine design ensures that traffic from any member in the access pod can reach the spine in a single hop, which minimizes latency and creates a consistent and predictable performance profile.

Below shows an example of how switches can be connected to form an Arista wiring closet with dual load balanced uplinks traversing an IDF riser.

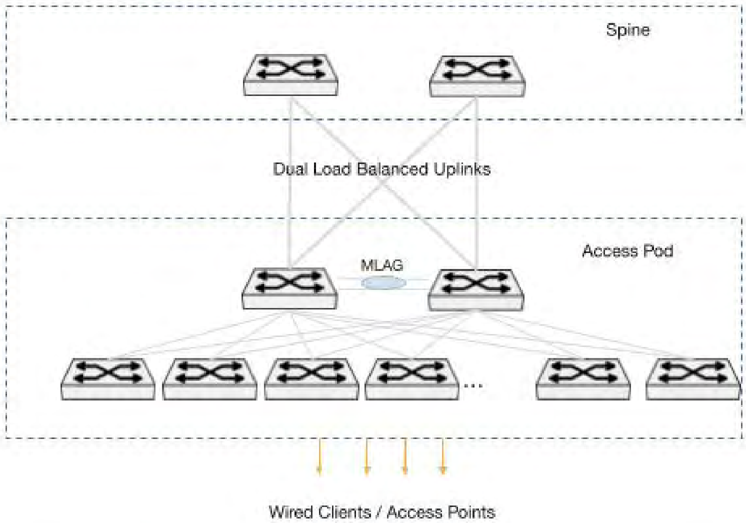

Figure 8: Arista MLAG Leaf-Spine Design

The MLAG Leaf-Spine Design has the following advantages

1. High Availability:
    - If a single leaf switch in the MLAG pair fails, the other switch continues to operate seamlessly, maintaining connectivity for devices.
    - Redundancy at both the link and switch levels reduces single points of failure.
2. Load Balancing:
    - MLAG enables efficient utilization of all available links between devices and leaf switches, improving bandwidth usage.
3. Scalability:
    - The leaf-spine topology allows for linear scalability by adding more leaf switches without disrupting the existing network.
    - MLAG provides redundancy without compromising this scalability.

4. Low Latency:

    - The flat nature of the leaf-spine topology ensures minimal hops between devices, while MLAG ensures fast failover and reliable traffic forwarding.

Modular Chassis

The Arista 750 Series is a line of modular, high-performance switches designed to meet the evolving demands of modern campus networks. These switches offer scalability, flexibility, and advanced features to support a wide range of applications, including IoT devices, Wired Desktops, Medical Devices, Smart Building Sensors, Wi-Fi 6E / Wi-Fi 7 access points, etc. The 750 Series includes two chassis options:

1. **CCS-755:** A 5-slot chassis supporting up to 240 access ports in a 7U form factor.
2. **CCS-758:** An 8-slot chassis accommodating up to 384 access ports in a 10U form factor.

The modular design allows organizations to start with a smaller configuration and add more line cards as the network grows, reducing the need for significant upfront investments.

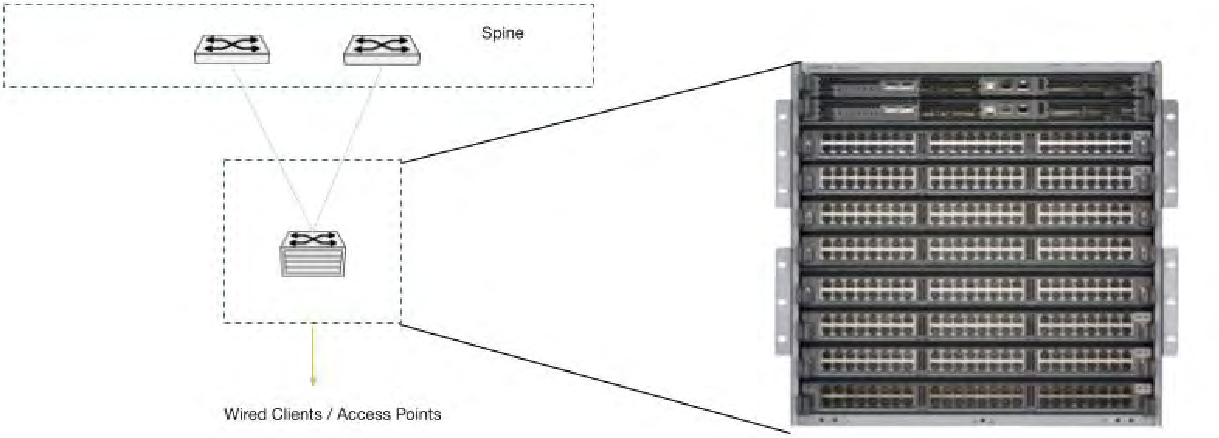

Figure 9: Arista Modular Chassis Design

The 750X series leverages a dedicated centralized switch architecture that provides high throughput, consistent low latency with low jitter and high scale. The switch card runs the dataplane for all forwarding with the management plane and control plane on the separate Supervisor module. With redundant switch cards failover from active to standby occurs without a control plane change, increasing system reliability. Switch cards are hot-swappable and accessed from the system rear, avoiding any impact to network cabling during maintenance

The modular chassis offers the following benefits:

1. Flexibility:

    - Different line cards can support a variety of interface types (e.g., copper, fiber, 10M-1G, 2.5G-10G, 25G, or even 100G ports), making the system adaptable to changing network requirements.

2. Redundancy and High Availability:

    - Components like supervisor modules, power supplies, and fabric modules can be configured redundantly to prevent single points of failure
    - Faster SSS (Stateful SwitchOver) convergence because of dedicated hardware pins

3. Investment Protection:

    - Upgrading or replacing individual line cards is more cost-effective than replacing an entire switch.

4. High Performance:
    - The architecture supports large amounts of data traffic, with the fabric module providing high-speed connectivity between line cards.
5. Single IP Address:
    - Entire chassis can be exposed to the network via a single IP address for management thus preserving IP address space.
6. Single CLI Session:
    - Modular chassis switches can be managed through a single CLI session, which is one of their key operational
7. Ease of Maintenance:
    - Faulty components can be replaced without taking the entire chassis offline, ensuring minimal disruption.

CloudVision Leaf-Spine Stack Management

CloudVision Lead Spine Stack (LSS) Management allows administrators to operate a group of switches collectively to efficiently onboard, provision, configure, monitor, troubleshoot, upgrade and maintain a group of switches as a single logical unit. This streamlines & automates operations by abstracting the complexity of managing multiple switches individually.

Traditional management architecture forces customers to deploy stacking to solve network management problems, but Arista's approach is to address network management problems at the management layer and not at the network layer. CloudVision LSS Management is fully transparent to underlying network design extending the management benefits to all network designs including standalone switches, MLAG Leaf-Spine Design, Modular Chassis and SWAG.

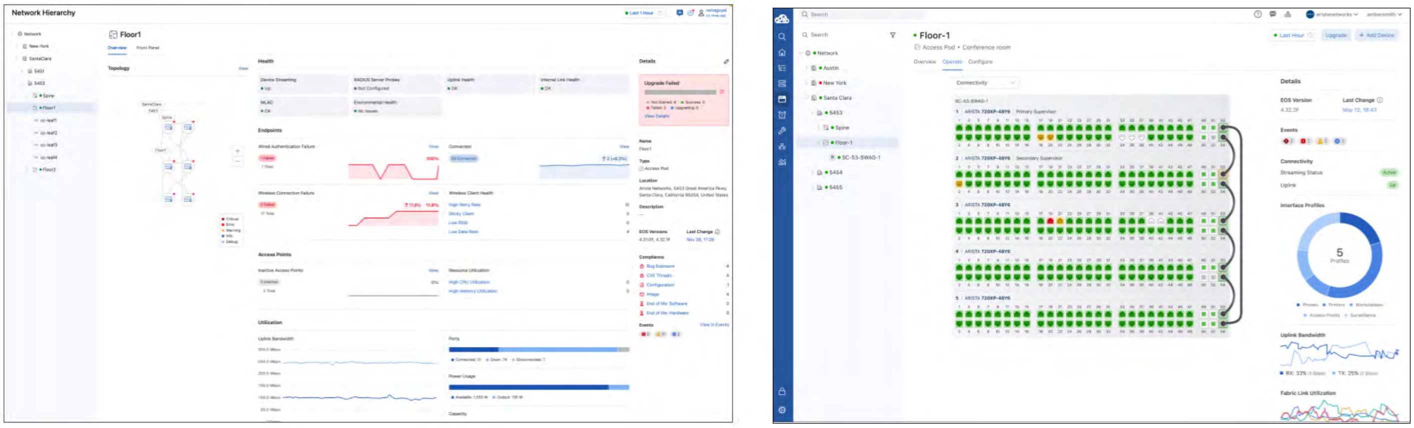

Figure 10: CloudVision Leaf Spine Stack Management Features

Arista's approach to solving this problem is unique in that we are not restricted to grouping switches only in an IDF to be managed as a single logical unit. The concept of hierarchy allows CloudVision to logically group switches across multiple IDF closets, floors, buildings, campus thus providing a very flexible mechanism to manage the network.

CloudVision LSS combines the visibility of real-time streaming telemetry, with common provisioning and diagnostic tasks, providing the user with an information rich, modern operating model. An example of this is a simple use case of cycling an interface – CloudVision provides the user with a live view of the state streaming in from the device, while executing a simple point n' click workflow to cycle the administrative state of the interface.

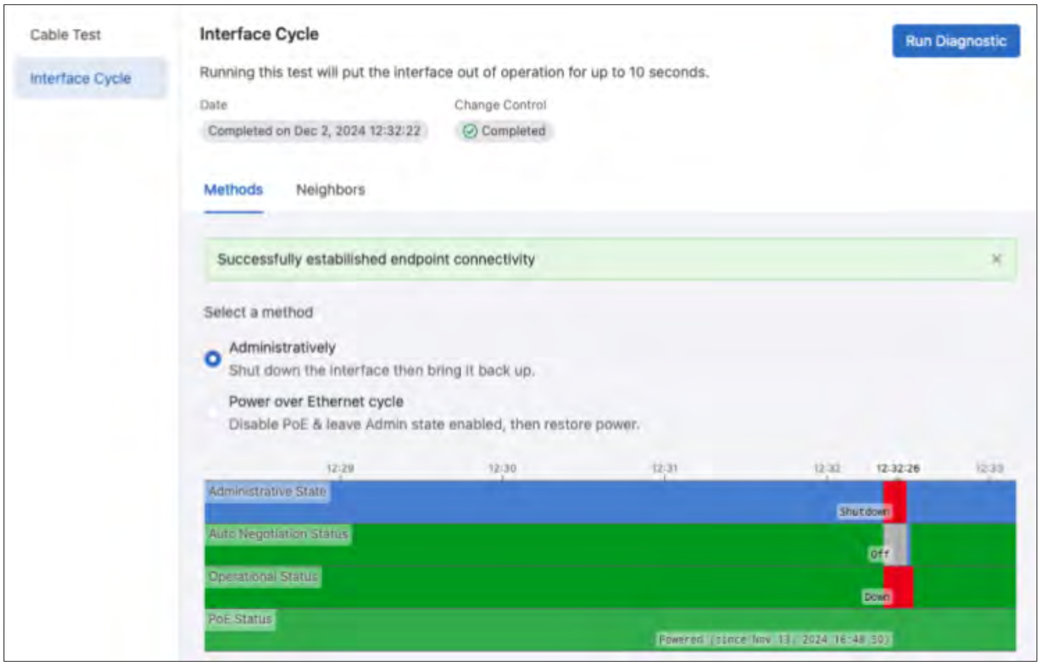

Figure 11: Operations with LSS - Interface Diagnostics

Another common operation is adding a device to the network. CloudVision LSS has introduced guided workflows to quickly onboard a new device and have it operational for end user connections in only a few clicks. A device registers with CloudVision through the usual ZTP process, commonly inband in campus networks, and gives the user full visibility of how the device is connected to the network – something the user would typically be blind to. The workflow uses the information CloudVision has learned about where the device is connected and automates the creation of attributes like hostname and IP addressing.

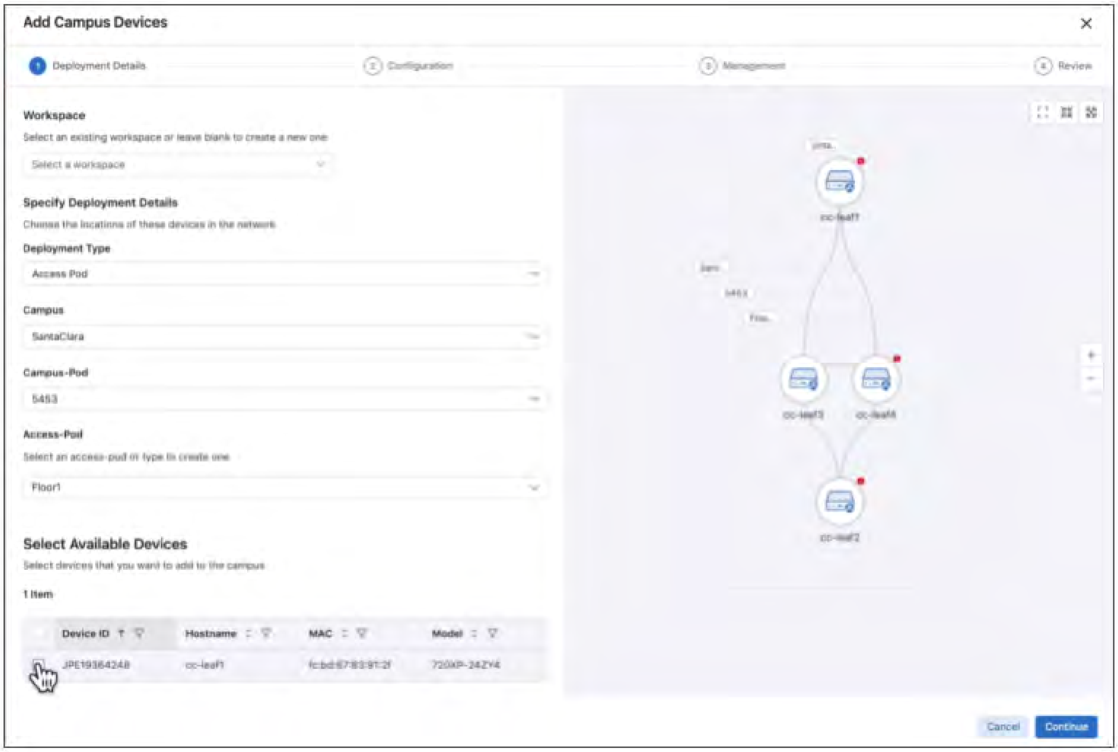

Figure 12: Operations with LSS - Add New Device to Access Pod

Summary

Switch stacking simplifies management by combining multiple switches into a single logical unit, offering benefits like centralized management, IP address conservation, and single CLI interface. However, legacy stacking suffers from limitations in scalability, performance, and flexibility, with constraints like limited switch count, bottlenecks from stacking cables, and downtime during reconfiguration. These limitations inhibit customers from deploying stacking for their modern network designs.

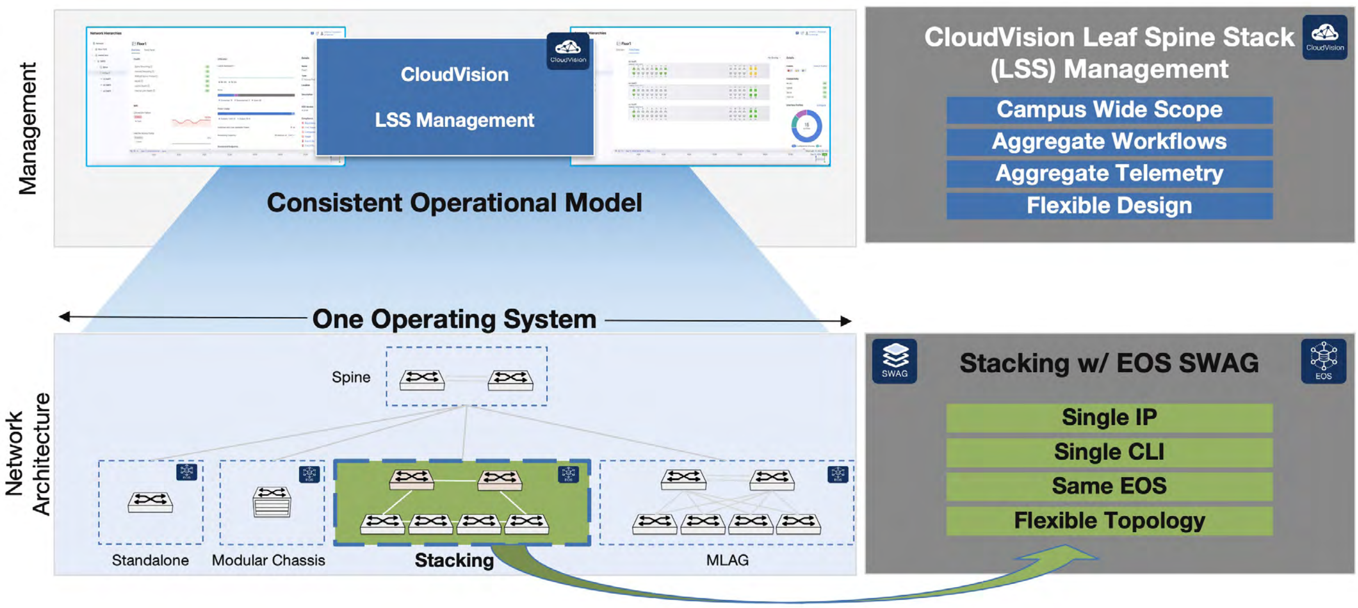

Figure 13: SWAG & Leaf Spine Stak Management Architecture

Using Arista's **Extensible Operating System (EOS)** as the foundation, Arista provides customers with the most flexible choice to design highly scalable campus networks. Multi-Chassis Link Aggregation (MLAG) Leaf-Spine Design facilitates scalable and efficient network designs ensuring consistent low-latency paths and eliminating the bottlenecks associated with traditional stacking topologies. SWAG (Switch Aggregation Group) **SWAG (Switch Aggregation Group)** extends the same distributed software in EOS built for Modular Chassis to a group of standalone switches, that can now be configured in a SWAG cluster providing all the benefits of legacy stacking but done in a more modernized way. SWAG provides single IP address, Single CLI management benefits of stacking, but unlike legacy stacking implementations, uses Leaf-Spine to build a highly scalable stacking architecture

Arista's CloudVision Leaf-Spine Stack (LSS) Management provides the ability to collectively manage a group of switches for single point administrative control and operation extending the management benefits to all network designs.

Santa Clara—Corporate Headquarters
5453 Great America Parkway,
Santa Clara, CA 95054

Phone: +1-408-547-5500
Fax: +1-408-538-8920
Email: info@arista.com

Ireland—International Headquarters
3130 Atlantic Avenue
Westpark Business Campus
Shannon, Co. Clare
Ireland

Vancouver—R&D Office
9200 Glenlyon Pkwy, Unit 300
Burnaby, British Columbia
Canada V5J 5J8

San Francisco—R&D and Sales Office 1390
Market Street, Suite 800
San Francisco, CA 94102

India—R&D Office
Global Tech Park, Tower A, 11th Floor
Marathahalli Outer Ring Road
Devarabeesanahalli Village, Varthur Hobli
Bangalore, India 560103

Singapore—APAC Administrative Office
9 Temasek Boulevard
#29-01, Suntec Tower Two
Singapore 038989

Nashua—R&D Office
10 Tara Boulevard
Nashua, NH 03062