Cisco Catalyst 9000

A new era of intent-based networking

Cisco Catalyst 9000

A new era of intent-based networking

Preface

Authors

This book represents a collaborative effort between Technical Marketing, Product Management, Engineering, and Sales teams during a week-long intensive session at Cisco Headquarters in San Jose, CA.

- Bob Sayle - Sales
- Dave Zacks - Technical Marketing
- Dimitar Hristov - Technical Marketing
- Fabrizio Maccioni - Technical Marketing
- Ivor Diedricks - Product Management
- Jay Yoo - Engineering
- Kenny Lei - Technical Marketing
- Mahesh Nagireddy - Technical Marketing
- Minhaj Uddin - Technical Marketing
- Muhammad Imam - Technical Marketing
- Sai Zeya - Technical Marketing
- Shawn Wargo - Technical Marketing

Acknowledgements

A special thanks to Cisco's Enterprise Networking Business Product Management, Engineering, and Sales teams who supported the realization of this book. Thanks to Carl Solder and Muninder Sambi for supporting this effort. We would also like to thank Cynthia Resendez for her exceptional resource organization and support throughout our journey, and Sehjung Hah for his help with planning and logistics.

We are also genuinely appreciative to our Book Sprints (www. booksprints.net) team:

- Adam Hyde (Founder)
- Barbara Rühling (CEO and Facilitator)
- Henrik van Leeuwen (Illustrator)
- Juan Carlos Gutiérrez Barquero (Technical Support)
- Agathe Baëz (Book Producer)
- Raewyn Whyte and Susan Tearne (Proofreaders)
- Barbara and the team created an enabling environment that allowed us to exercise our collaborative and technical skills to produce this technical publication to meet a growing demand.

Organization of this Book

This book is best read in the order presented. However, based on the roles of the reader and their interests, some chapters can be reviewed out of sequence. This book is organized into sections, with each section having multiple chapters.

This book introduces the Catalyst 9000 family, reviews the business drivers for enterprises, and illustrates how the Catalyst 9000 addresses the challenges faced by enterprise IT. Following this, the architectural foundations of the Catalyst 9000 platform, both from a hardware perspective with Cisco's innovative Unified Access Data Plane (UADP) ASIC, as well as the cutting-edge capabilities provided by Cisco IOS XE software, are explored. These foundational elements enable the Catalyst 9000 family to address the many demands placed on enterprise networks today.

How the Catalyst 9000 platforms meet these demands is outlined in the next sections covering High Availability, Security, Quality of Service, Application Visibility and Control, IoT and User-Centric Platform Design. Cisco IOS XE software brings an open, standard and model-based approach to network management interfaces, and these capabilities are reviewed in the Programmability and Automation section. Then an examination of Application Hosting on the Catalyst 9000 is provided. Finally, this book examines the present state and future evolution of network design, and how Catalyst 9000 leads the way towards the ongoing transformation of enterprise network architectures.

Intended Audience

Network administrators, engineers, and architects are always under pressure to meet the business needs of their organizations. This book focuses on Cisco's new and innovative Catalyst 9000 family of switches, and how they help to solve the many challenges that networking professionals face today. The Catalyst 9000 provides state-of-the-art technologies driven by open, flexible, and powerful hardware and software. Networking professionals will be able to utilize this book to understand the Catalyst 9000 family, delve deep into its architecture, and understand how it provides a strong foundation for next-generation networks.

This book assists network professionals, IT managers, executives, and anyone with an interest in the latest and greatest networking technologies to understand and embrace the new era of intent-based networking that the Catalyst 9000 enables.

Book Writing Methodology

Fix your eyes on perfection and you make almost everything speed towards it
- W.E. Channing

Simplicity, consistency, and performance have been overriding themes in designing the Catalyst 9000 family. The idea of this book is to present readers with the current challenges in enterprise networking and explore how the Catalyst 9000 platform solves those challenges. The Catalyst 9000 provides cutting-edge hardware and software capabilities, easily adapting to future protocols and network architectures without losing sight of simplicity. This book explores this powerful new networking platform - the basis for the new era of networking.

A group of twelve Cisco Engineers from diverse backgrounds accepted the challenge of writing a book about a platform that changes the paradigm of enterprise networking. At the end of day one, the task seemed even more daunting, given the breadth of capabilities that Catalyst 9000 brings to networks. However the team persisted, and after hundreds of hours of diligent penmanship, this book was born! The Book Sprints (*www.booksprints.net*) methodology captured each of our unique strengths, fostered a team-oriented environment, and accelerated the overall time to completion.

#NetworkIntuitive
#NewEraOfNetworking

Introduction

Executive Summary

The world is rapidly changing due to the demand of evolving IoT, ubiquitous mobility, cloud adoption, and rapidly advancing security threats. Enterprises of all sizes around the world are replacing legacy systems with digital technologies for competitive advantage, higher productivity, and lower operating costs. As more businesses embrace this change, networks have to adapt. Business cannot build networks the same way they have for the past 30 years. Organizations need to create dynamic networks that can constantly learn, adapt, protect, and evolve.

The Network. Intuitive.

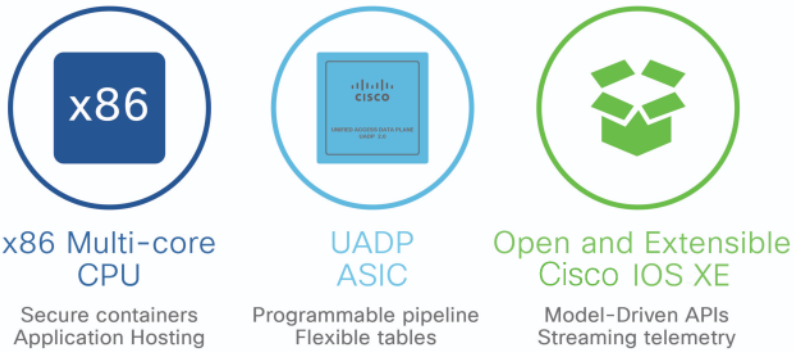
Cisco's Digital Network Architecture (DNA) and Software-Defined Access (SD-Access) help organizations unlock opportunities, enhance security, be more agile, and operate more efficiently. Designed to be intuitive, the network recognizes intent, mitigates threats through segmentation and encryption, learns and adapts over time. Cisco's Catalyst 9000 switches are the next generation in the legendary Catalyst family of enterprise LAN access, distribution and core switches. These are the first purpose-built platforms designed to take advantage of DNA and SD-Access.

The new Cisco Catalyst 9000 family of switches has been designed as the foundation for an entirely new era of networking - "The Network. Intuitive." This book explores the Catalyst 9000 family of switches and examines how these platforms meet the ever-changing needs of the enterprise network, today and well into the future.

For the first time in the industry, a single family of fixed, stackable and modular switches can run a single software image with a common ASIC across every platform in campus and branch networks. Design considerations can now be focused entirely on the scale requirements for different places in the network. This provides significant reduction in Total Cost of Ownership (TCO) for enterprise networks.

The Catalyst 9000 series is based on two foundational aspects:

- **Common hardware** - built with a flexible, programmable ASIC and CPU architecture.
- **Common software** - built with an open, Linux-based, modular operating system, with simple feature licenses.



The Catalyst 9000 family is built on a common ASIC architecture powered by **Unified Access Data Plane (UADP) ASIC**. This serves as an innovative, programmable and flexible silicon foundation for the platform. UADP enables network infrastructures to adapt to new technologies, trends, and business needs over time. Catalyst 9000 platforms are also built on a standard multi-core 64-bit **x86 CPU** architecture. A common CPU architecture provides predictable software processing and control-plane management, providing the horsepower to tackle next-generation network architectures and making it easier to diagnose and resolve issues while providing a platform for application hosting.

Every Catalyst 9000 platform runs on the open and modular **Internet Operating System Cisco IOS XE**. This improves portability across Cisco enterprise platforms (including Catalyst switches, ISR/ASR routers, and Wireless LAN Controllers). It increases feature development velocity, improves High Availability, and makes it easier to consistently

deploy features across the Campus network. Cisco IOS XE provides a well-defined set of APIs, improving management and simplifying automation and programmability.

└the bottom line

Catalyst 9000, built on common hardware and software powered by Cisco's innovative UADP ASIC, x86 CPU and Cisco IOS XE software, is the foundation for the new era of networking.

Industry Trends

The common trends seen in the industry today fall into four main categories - IoT, mobility, cloud and security.

IoT trends and considerations

The digital transformation of business processes and operations includes connecting new devices, sensors, and machines in an effort to improve productivity, reduce risk, and increase security. Billions of machine-to-machine connections will emerge over the next several years that require machine learning intelligence based on analytics and business policy. Enterprise campus networks will be required to support this influx of machine connectivity.

Cisco's “Internet of Things: Workloads and Key Projects 2017” survey predicts organizations will undertake IoT data aggregation, filtering, and analysis at the network edge. The primary drivers for processing IoT data at the network edge are to improve security and speed up data analysis. The network needs to evolve to support the current and future demands of IoT.

Mobility trends and considerations

Wireless and mobility are driving the enterprise network infrastructure market. BYOD and mobile applications make it possible for workers to access corporate data from smartphones, tablets, and personal laptops, creating challenges for back-end IT infrastructures and greater demand for enterprise mobility among workers. Mobility is now a strategic asset. It is the predominant way workers and visitors access the corporate network and the Internet. Mobility must be an integral part of the future enterprise network.

Cloud trends and considerations

Enterprises are augmenting internal IT with cloud services, be it on-premises or collocated private cloud, or public cloud services. Cloud infrastructures mix virtual and

physical elements, with workloads moving between on-premises and off-premises resources. Campus networks have to not only interface with off-premise and public clouds but ensure the same application performance, security, and policy adherence for those workloads as if they were still on-premises.

Security trends and considerations

All of these new connections open up profound security implications. Each new connection is a potential attack vector. Attacks are becoming more and more sophisticated, and worse, they are often obscured via encryption. Campus networks must be able to secure these new connections by detecting anomalies and recognizing potentially malicious behaviors and patterns in real-time at scale.

Business Use Cases

Catalyst 9000 switches extend Cisco's networking leadership with breakthrough innovations in IoT, mobility, cloud and security.

Enabling the IoT Use Case

There are many new devices being connected to the network such as sensors, alarm systems, HVAC systems, badge readers, and so on that have not traditionally been connected or have been using proprietary protocols. The Catalyst 9000 platforms, together with Cisco Identity Services Engine (ISE), are able to automatically profile these devices, provide security and segmentation, and apply policies to them.

Devices are starting to advertise their services using the Bonjour (mDNS) protocol. DNA Service for Bonjour delivers visibility to these services across locations and segments of the network, assigns policy based on these services, and orchestrates all of this from a centralized point with DNA Center.

Some IoT devices, such as LED lighting, require always-on power. The Catalyst 9000 supports Perpetual PoE and Fast PoE to keep the lights on while the switch reloads.

In order to support professional media and audio applications, Catalyst 9000 supports Audio Video Bridging (AVB) and IEEE 1588 timing.

└the bottom line
The Catalyst 9000 is the ideal platform for connecting the Internet of Things.

Enabling the Mobility Use Case

Wired and wireless networks have historically been built and operated by different teams. Catalyst 9000 with SD-Access delivers central orchestration and the assurance

of a single, integrated wired and wireless network. This allows the network to scale seamlessly without bottlenecks as more wireless clients are added to the network. The APs connect directly to the Catalyst 9000 switches for data plane forwarding directly in hardware.

The policies for wired and wireless are the same in this network architecture. Network segmentation and group-based policies are consistent between wired and wireless traffic, making operations simple.

The SD-Access architecture delivers simplified and seamless roaming for devices across the network.

The Catalyst 9000 delivers the industry's highest mGig and PoE capacity allowing customers to build the densest wireless environments, leveraging 802.11ac Wave 2 and future wireless innovations.

└the bottom line
Catalyst 9000 offers the optimal foundation for converging wired and wireless access.

Enabling the Cloud Use Case

In order to make deployment and operation of the network more agile, Cisco has added a programmatic framework and tools to drive use of automation through NETCONF, RESTCONF, and gNMI APIs with YANG models.

Streaming telemetry facilitates near-real-time visibility to operational data.

The Catalyst 9000 supports application hosting with local storage enabling fog computing and network function virtualization. This supports distributed intelligent agents infused into the network for analytics, assurance, security, and cloud-connected

applications. Customers are able to host third-party applications on the Catalyst 9000 platforms, making this the most flexible platform in the industry.

└the bottom line
Catalyst 9000 is an open and fully programmable platform for enabling the move to cloud.

Enabling the Security Use Case

There are a diverse and growing set of devices connecting to enterprise networks. Network segmentation may be used to constrain devices and users so that communication is only possible once allowed. Catalyst 9000 platforms support numerous segmentation capabilities at a macro (network segment) and micro (user or device group) level with support in hardware for SD-Access, MPLS, VRF-Lite, and TrustSec.

As more network traffic is becoming encrypted, it is critical that these threats are detected and mitigated at the point where it connects to the network. The Catalyst 9000 detects and mitigates malware hiding in encrypted traffic using Encrypted Traffic Analytics (ETA). Even better, ETA detects anomalies in encrypted traffic without decrypting it.

The platforms collect metadata in hardware about all the flows traversing the network, using full Flexible Netflow. Combining this with Cisco Security solutions, such as Cisco Stealthwatch, provides detection of denial-of-service attacks and other malicious activity.

With the Catalyst 9000, the links between switches can be encrypted using up to 256-bit AES MACsec, operating at line rate. This encryption can also be used for connections between the switch and endpoints.

Finally, Cisco Trustworthy Systems security solution protects the network switches themselves. An holistic approach provides comprehensive verification of hardware and software integrity by securing the device, network communications, and hosted applications.

└the bottom line

Catalyst 9000 provides the most secure switching environment in the network industry.

Catalyst 9000 Family

Overview

The new Cisco Catalyst 9000 switching platform is the next generation in the Cisco Catalyst family of enterprise LAN access, distribution, and core switches. It is the first purpose-built platform designed to take advantage of Cisco DNA and SD-Access. This new switching system extends Cisco's networking leadership with breakthrough innovations in mobility, IoT, the cloud, and security. Leveraging the UADP ASIC, the Cisco Catalyst 9000 platform delivers much higher performance and adds a host of new features and functionality.

DIAGRAM Catalyst 9000 Switches

The Catalyst 9000 switching platforms are built on a common and strong hardware and software foundation. The commonality and consistency bring simplicity and ease of operations for network engineers and administrators, reducing total operational cost, and creating a better experience.

Common Hardware

The hardware has a common design, both internally and externally. Internally the hardware uses a common ASIC, the Unified Access Data Plane (UADP) ASIC, providing flexibility for packet handling. The hardware also has another common component, namely, the switch CPU. For the first time in the history of Catalyst switches, there is an x86-based CPU onboard, allowing it to host additional applications beyond those normally possible on a network switch.

Externally the hardware is designed by one of the best designers of the world - Pininfarina, designer of the famous Ferrari. This level of industrial design brings an enhanced user experience for the Catalyst 9000 family. It provides ergonomic design and common attributes that simplify device operations. More details are provided in Chapter 11 User-Centric Platform Design.

Common Software

The Catalyst 9000 family of switches run the **exact same binary image of Cisco IOS XE**. Cisco IOS XE is an enhanced, open and programmable OS. With a 30 year history behind it and thousands of features, Cisco IOS XE is arguably the most feature-rich OS in the networking industry. Having a single binary image shared across Catalyst 9000 platforms enables end-to-end feature support and allows feature parity at any point in the network. This also helps with qualifying software releases as only a single image is needed to be tested for the entire campus network.

The strong hardware and software foundation of Catalyst 9000 enables it to face the challenges of enterprise networks today. At the same time, it brings consistency and simplicity for customers. The Catalyst 9000 family has three members - Catalyst 9300 stackable, Catalyst 9400 modular chassis, and Catalyst 9500 fixed-configuration core. These platforms are discussed in further detail in the following chapters.

Catalyst 9300

The Cisco Catalyst 9300 Series is Cisco's lead stackable enterprise fixed switching platform. At 480 Gbps of stacking bandwidth and with up to eight devices in a stack, it is the industry's highest-density stacking bandwidth solution.

DIAGRAM Catalyst 9300

Platform Overview

All models of Catalyst 9300 are 1RU high with dual power supplies and redundant fans. Different models offer a variety of connectivity and scale. These models can be organized into four sub-families. Every sub-family has 24-port and 48-port copper models:

1. Data-only models - Optimized for devices such as desktops and printers that just need data connectivity from 10 Mbps to 1 Gbps.
2. PoE/PoE+ models - Provide the same capability as the data models plus added support for 30W of Power over Ethernet (PoE+). All the ports support PoE / PoE+ and all ports can be active simultaneously with PoE+.

3 UPoE models - These models provide the same capability as the PoE+ models with the added support of 60W of PoE. Any of the ports can be configured with UPoE, but the maximum available total PoE power per switch is 1800W.

4 mGig models - Provide connectivity at multiple speeds up to 10 Gbps on mGig ports. Wireless access points supporting 802.11ac Wave 2 are the most common devices requiring mGig connectivity, but wired connections to desktops can also benefit. All ports on these models support UPoE, but the total available PoE per switch is 1800W. There are two different models in this sub-family:

   - 24 port mGig: All 24 ports support 100 Mbps, 1 Gbps, 2.5 Gbps, 5 Gbps, and 10 Gbps.
   - 48 port mixed mGig: The first 36 ports support 100 Mbps, 1 Gbps, and 2.5 Gbps. The last 12 ports support the full range of mGig speeds.

Network Modules

All Catalyst 9300 switches have an optional slot for uplink network modules. There are four variants of uplink modules. In addition, the ports on these modules are not limited to uplink connectivity; they can be used to connect to hosts as well.

1 (4) 1G RJ-45 ports (supports 10 Mbps, 100 Mbps, and 1 Gbps).

2 (4) mGig ports (no PoE).

3 (8) 10G SFP+ / SFP ports.

4 (2) 40G QSFP+ ports.

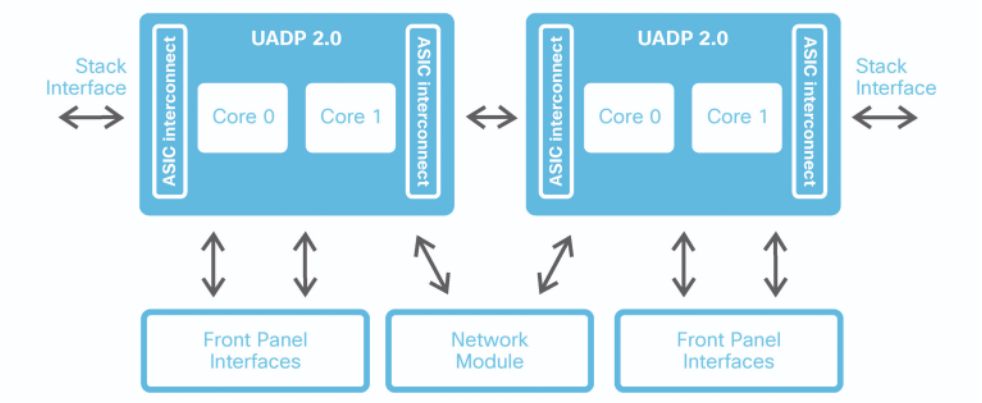
5 (2) 25G SFP28 ports.

Catalyst 9300 switches are compatible with Catalyst 3850 uplinks modules. However, Catalyst 9300 uplinks modules are not compatible with Catalyst 3850.

Architecture

The non-mGig models of the Catalyst 9300 are powered by a single UADP 2.0 ASIC. The mGig models are equipped with two UADP 2.0 ASICs. All ports on the Catalyst 9300 operate at line rate for all packet sizes.

DIAGRAM Catalyst 9300 Architecture



Stackwise-480

Catalyst 9300 provides the ability to stack up to eight switches, combining them together to operate as a single, logical switch. This allows network engineers to manage, configure and troubleshoot the stack of switches as one. Chapter 6 High Availability provides additional details on the operation of StackWise-480.

Power Supply and Fan

Catalyst 9300 switches support dual redundant power supplies. These power supplies are available in 350W AC, 715W AC, 1100W AC, and 715W DC options. The power supplies can be mixed in any combination, for example, AC and DC.

Catalyst 9300 switches are equipped with three field-replaceable fans. These fans are operated in an N+1 redundant mode.

StackPower

The Catalyst 9300 provides the ability to create a shared pool of power using dedicated stack power cables. In the event of power supply failure or more PoE power draw, the switch can utilize the power from the shared pool to support the extra load. Stack power can be deployed in two modes: power-sharing and redundant mode. Additional details are provided in Chapter 6 High Availability.

Catalyst 9400

The Cisco Catalyst 9400 Series is Cisco's leading modular enterprise switching access platform. It provides unparalleled investment protection with a chassis architecture capable of supporting up to 9 Tbps of system bandwidth. It also offers unmatched power delivery for high-density PoE deployments, delivering 60W Power over Ethernet to endpoints. The 9400 Series delivers state-of-the-art High Availability with capabilities such as dual supervisors and N+1/N+N power supply redundancy. The platform is enterprise-optimized with an innovative dual-serviceable fan tray design and side-to-side airflow and is closet-friendly with a ~16-inch depth. A single system can scale up to 384 access ports.

DIAGRAM Catalyst 9400 Family

Platform Overview

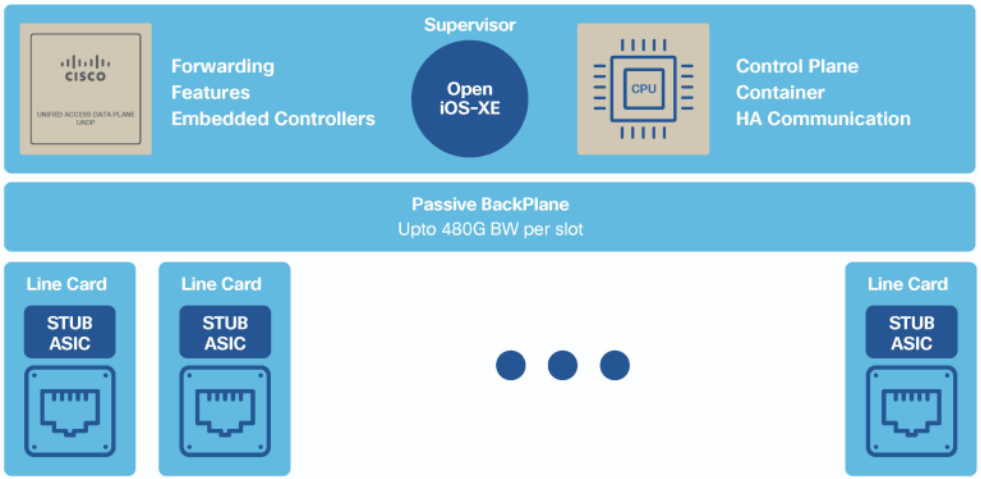
Catalyst 9400 switches provide up to 480G per slot bandwidth. There are three models that offer different densities to fit different size requirements: 4 slot, 7 slot, and 10 slot chassis. All three chassis options provide dual supervisor slots for maximum availability. The chassis is designed to support more than 720G of bandwidth between the two

supervisor slots, which will enable a future supervisor to support multiple 100G ports. With the growing need for increased Power over Ethernet, the chassis has the capability of providing more than 4,800W of PoE power per slot.

Architecture

The Catalyst 9400 is based on a centralized architecture, which means all forwarding, services, and queuing are done on the supervisor while the line cards are considered transparent, containing only stub ASICs and PHYs. The simplicity of this centralized design allows easy upgrade of features by just upgrading the supervisor while keeping the existing line cards. This provides significant investment protection.

DIAGRAM Catalyst 9400 Architecture



Supervisors

There are currently two versions of supervisor available for the Catalyst 9400: Sup-1 and Sup-1XL. Both supervisors are powered by (3) UADP 2.0 XL ASICs. The three ASICs

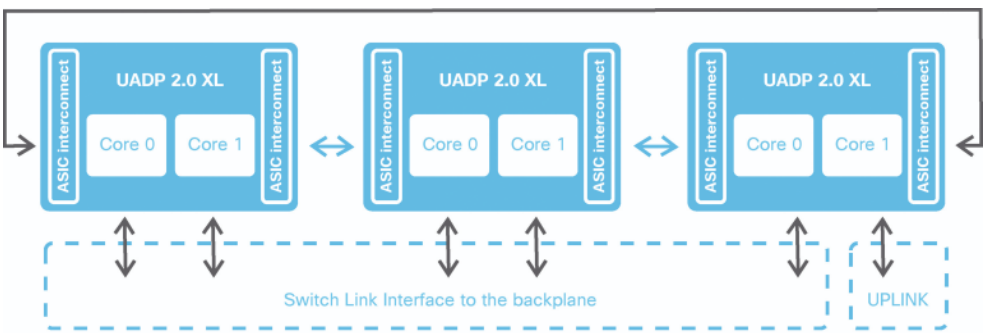
are interconnected through a 720G ASIC interconnect for packets passing between the ASICs.

The Sup-1 provides 80 Gbps of bandwidth per slot for all chassis models and is optimized for access deployments.

Sup-1XL provides 80 Gbps of bandwidth per slot in the 10-slot chassis, 120 Gbps of bandwidth per slot for the 7-slot chassis, and 240 Gbps per slot for the 4-slot chassis. Sup-1XL also adds support for different flexible templates to accommodate various deployment models such as access, distribution, core, SD-Access, or NAT.

UADP has Switch Link Interfaces (SLIs) connecting line card stub devices through the backplane. Each SLI, running at 10G rate with Sup-1/Sup-1XL, aggregates a group of front panel ports, known as an SLI port group. Future supervisors can run the SLIs at a higher speed and provide more bandwidth for the existing line cards. This provides additional investment protection for the existing line cards.

DIAGRAM Catalyst 9400 Supervisor-1XL Architecture



Supervisor Uplinks

Both Sup-1 and Sup-1XL have (8) SFP / SFP+ ports and (2) QSPF+ ports on the front. The architecture of Sup-1 and Sup-1XL provides 80G total uplink bandwidth and supports 1G / 10G / 40G interfaces. Interfaces 1 - 8 support 1G / 10G SFP / SFP+, and interfaces 9 and 10 support 40G QSFPs.

Line Cards

Catalyst 9400 switches offer mGig, Cisco UPoE, data and 10G line cards for different connectivity requirements.

- Copper RJ-45 modules:

    1. 48-port data line card: All 48 ports on this module support 10 Mbps, 100 Mbps, and 1 Gbps.
    2. 48-port PoE+/PoE line card: All features supported on the data line card with added support for PoE+ (30W) and PoE (15.4W).
    3. 48-port UPoE line card: All features supported on PoE+/PoE line card with added support of UPoE (60W). All 48 ports within the slot can provide UPoE simultaneously.
    4. 48-port mGig line card: The first 24 ports are the traditional 10/100/1000 copper RJ-45 ports and the last 24 ports are mGig ports that support 100 Mbps, 1 Gbps, 2.5 Gbps, 5 Gbps and 10 Gbps. All 48 ports on this module support UPoE (60W), PoE+ (30W) and PoE (15W). All 48 ports within the slot can provide UPoE simultaneously.

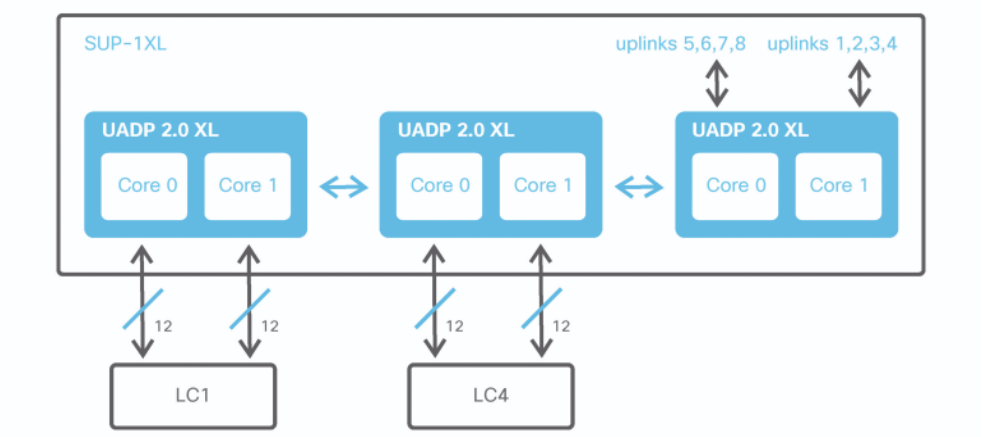
- Fiber SFP/SFP+ modules:

    1. 24-port SFP line card: Supports 100 Mbps and 1 Gbps speeds.
    2. 48-port SFP line card: Double the density compared to the 24-port SFP line card.
    3. 24-port SFP+/SFP line card: Supports 100 Mbps, 1 Gbps, and 10 Gbps. These ports provide connectivity to 10G hosts as well as to uplink devices.

Line Card Slot Bandwidth

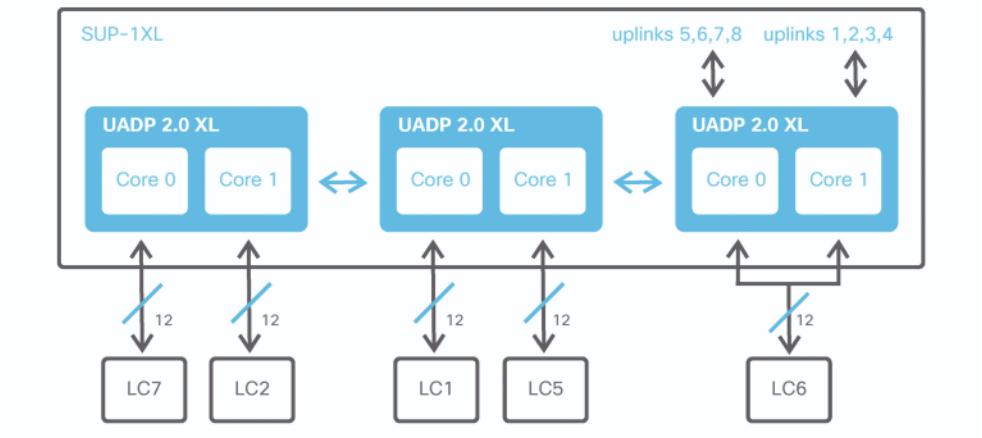
The three diagrams below illustrate how the Sup-1XL bandwidth is being utilized for line card support in the 4-slot, 7-slot, and 10-slot chassis and also shows the distribution of the number of SLIs to each line card.

DIAGRAM 4 Slot Chassis



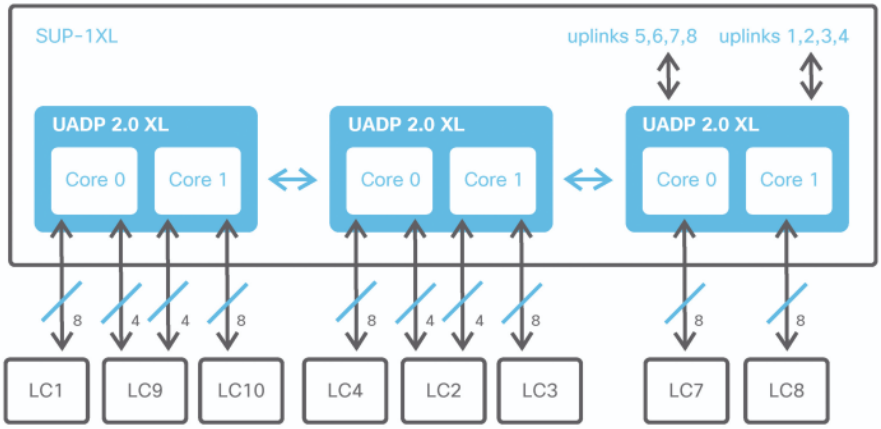
4 slot chassis: 24 SLIs are active for each line card slot. Each UADP services 1 line card.

DIAGRAM 7 Slot Chassis



7 slot chassis: 12 SLIs are active for each line card slot. Each UADP services 2 line cards.

DIAGRAM 10 Slot Chassis



10 slot chassis: 8 SLIs are active for each line card slot. Each UADP services 3 line cards.

Line Card Oversubscription

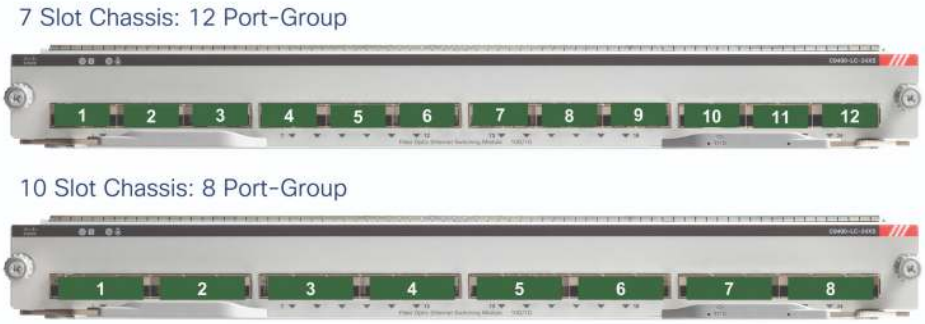
All variants of 1G line cards operate at line rate for all packet sizes. The 10G fiber line card and mGig line cards are oversubscribed with Supervisor-1 and -1XL. Also important to note is that line cards are designed to take advantage of higher per-slot bandwidth with future supervisors by running more number of SLIs at a higher speed.

Line Card Performance mode

With the oversubscribed 10G modules, if 10G line rate performance is needed, both the 24-port SFP / SFP+ and mGig line card can be enabled for maximum performance. When performance mode is enabled, the system uses only one 10G port in each SLI port group.

The following diagram shows the SLI port group mappings for the 10G fiber line card with the 7-slot and 10-slot chassis.

DIAGRAM 7-slot and 10-slot SLI Port Groups



Stateful Switchover (SSO)

Catalyst 9400 supports redundant supervisors with SSO. In SSO mode, the redundant supervisor engine is fully synchronized with configuration on the active supervisor. It subsequently maintains states for different protocols and minimizes the time the switch is unavailable during a supervisor failure or switchover. Additional details are provided in Chapter 6 High Availability.

In-Service Software Upgrade (ISSU)

ISSU is an upgrade process available on the Catalyst 9400 to allow upgrading supervisor software while traffic forwarding continues. This process is built on top of Stateful Switchover. ISSU increases network availability and reduces downtime caused by planned software upgrades. Additional details are provided in Chapter 6 High Availability.

Power Supply

The power supplies for the Catalyst 9400 come in small form factor while providing high capacity and efficient output. The 7-slot and 10-slot chassis provide eight power

supply bays while the 4-slot chassis provides four power supply bays. The Catalyst 9400 combines N+1, and N+N redundant options for power supplies. Additional details are provided in Chapter 6 High Availability.

Fan Tray

The fan tray of the Catalyst 9400 contains multiple individual fans operating in an N+1 redundant mode. Fans operate at variable speeds based on the system temperature and altitudes. This makes efficient use of the power and provides lower noise levels. The fan tray on the Catalyst 9400 can be replaced from the front or the rear of the chassis. This is a tremendous help with operations and reduces downtime since the cable management for wiring in a typical wiring closet could make it unwieldy to remove the cables from the front of the chassis to service the fan tray.

DIAGRAM Catalyst 9400 Fan Tray

Catalyst 9500

Catalyst 9500 Series switches are the industry's first purpose-built fixed 1-RU core and distribution layer switches. These switches deliver exceptional table scales (MAC / route / ACL) and buffering capabilities. This platform delivers up to 3.2 terabits per second of switching capacity and up to 2 billion packets per second of forwarding performance. The platform offers non-blocking 100 Gigabit Ethernet (QSFP28), 40 Gigabit Ethernet (QSFP+), 25 Gigabit Ethernet (SFP28) and 10 Gigabit Ethernet (SFP+) switches with granular port densities.

DIAGRAM Catalyst 9500

Platform Overview

The Cisco Catalyst 9500 platform consists of fixed configuration switches based on the Cisco Unified Access Data Plane (UADP) ASIC architecture. The platform runs on the Cisco IOS XE operating system that supports model-driven programmability, has the capacity to host containers, and run 3rd-party applications and scripts natively within the switch. The platform also supports all the foundational high-availability capabilities including dual redundant power supplies and variable-speed highly efficient redundant fans.

100 Gigabit Ethernet Switches

- C9500-32C - Cisco Catalyst 9500 Series high-performance switch with (32) 100GE ports and (2) UADP 3.0 ASICs.

- C9500-32QC - Cisco Catalyst 9500 Series high-performance switch with (32) 40GE or (16) 100GE ports and (1) UADP 3.0 ASICs.

40 Gigabit Ethernet Switches

- C9500-24Q - Cisco Catalyst 9500 Series switch with (24) 40GE ports and (4) UADP 2.0 ASICs.
- C9500-12Q - Cisco Catalyst 9500 Series switch with (12) 40GE ports and (2) UADP 2.0 ASICs.

25 Gigabit Ethernet Switches

- C9500-48Y4C - Cisco Catalyst 9500 Series high-performance switch with (48) 25GE + 4x100/40 GE ports and (1) UADP 3.0 ASIC.
- C9500-24Y4C - Cisco Catalyst 9500 Series high-performance switch with (24) 25GE + 4x100/40 GE ports and (1) UADP 3.0 ASIC.

10 Gigabit Ethernet Switches

- C9500-40X - Cisco Catalyst 9500 Series switch with (40) 1/10GE ports and (2) UADP 2.0 ASICs.
- C9500-16X - Cisco Catalyst 9500 Series switch with (16) 1/10GE ports and (1) UADP 2.0 ASIC.

Architecture

The Catalyst 9500 switches operate at line rate and offer configurable system resources to optimize support for specific features. The switch architecture consists of three main components:

- UADP ASIC,
- x86 CPU complex, and
- ASIC interconnect.

UADP ASIC

The Catalyst 9500 family of switches are built on two variants of UADP ASIC: UADP 2.0 and UADP 3.0. The architecture of both ASICs are similar, but they differ in switching capacity, port density, port speeds, buffering capability and forwarding scalability.

UADP 2.0 ASIC is built using 28-nanometer technology with two cores, with each core capable of supporting up to 120 Gbps of bandwidth for a total of 240 Gbps supporting a maximum forwarding capacity of 375M packets per second. Switches equipped with the UADP 2.0 ASIC support a total of up to 224K IPv4 / 112K IPv6 hardware tables, up to 54K of security ACL TCAM, and 2 x 16MB of shared buffer.

UADP 3.0 ASIC is built on 16-nanometer technology using two cores, with each core capable of supporting up 800 Gbps of bandwidth for a total of 1.6 Tbps supporting a maximum forwarding capacity of 1B packets per second. Switches equipped with the UADP 3.0 ASIC support a total of up to 416K for IPv4 / IPv6 hardware tables entries, up to 54K of security ACL TCAM, and 36MB of unified buffer.

X86 CPU Complex

The Catalyst 9500 series switches are all equipped with the same CPU, system memory, and flash storage. Catalyst 9500 series switches come with a 2.4 Ghz x86 quad core CPU, 16GB DDR4 RAM, and 16GB of internal flash storage. For application hosting or general purpose storage, these switches support USB 3.0 SSD storage, and models equipped with UADP 3.0 support up to 960GB M2 SATA SSD storage options.

ASIC Interconnect

Catalyst 9500 switches use high-speed ASIC interconnect links for inter-ASIC communication. UADP 2.0 has up to 720Gbps of interconnect bandwidth and UADP 3.0 has up to 800Gbps of interconnect bandwidth between two ASICs. Packets destined to local ports within the ASIC do not use ASIC interconnect links.

DIAGRAM Cisco Catalyst 9500 Switch Block Diagram - UADP 2.0

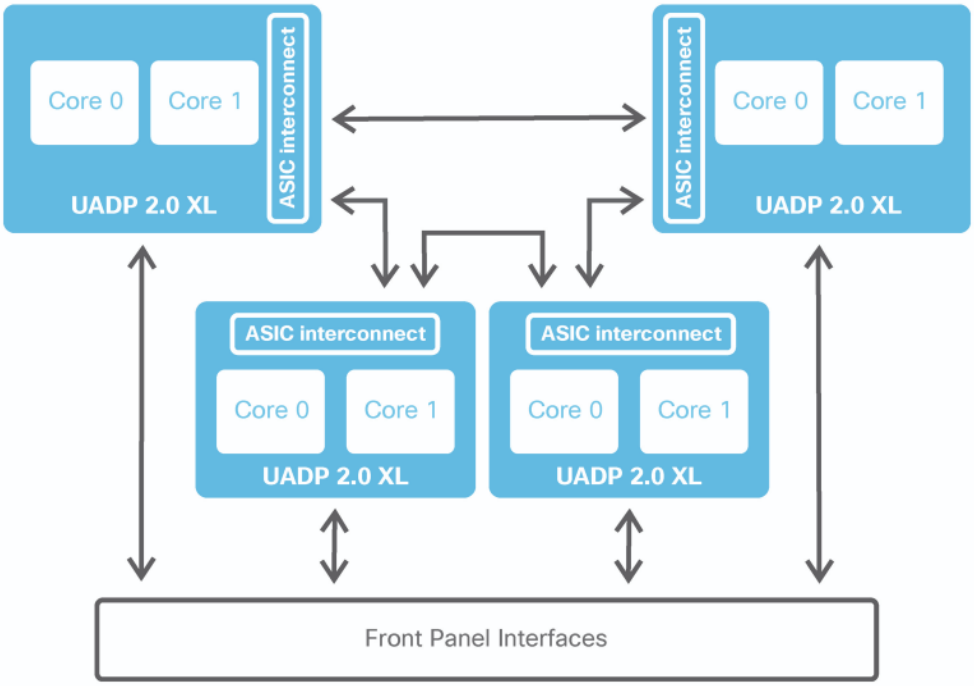
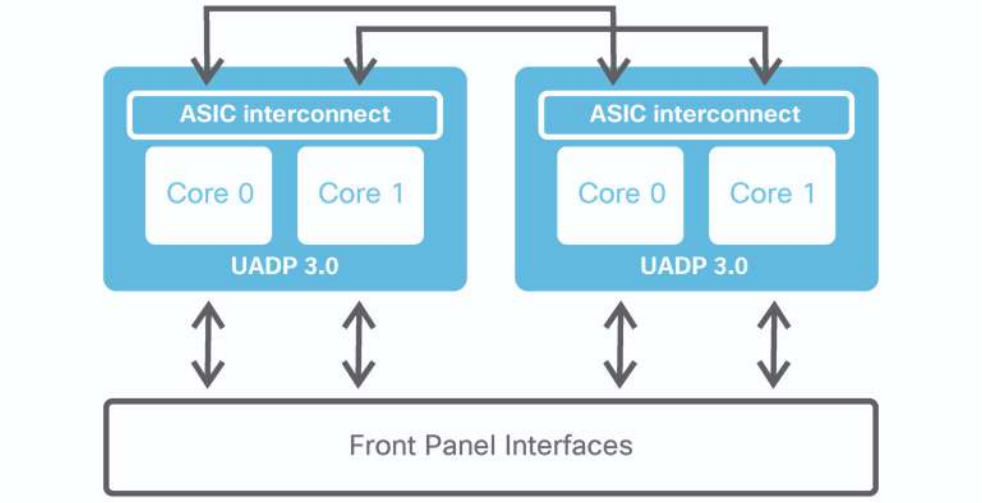


DIAGRAM Cisco Catalyst 9500 High-Performance Switch Block Diagram - UADP 3.0



Network Modules

The Cisco Catalyst 9500 Series supports optional network modules for uplink ports on the C9500-40X and C9500-16X switch models. The default switch configuration does not include the network modules. All ports on the network module are line rate and all software features supported on switch downlink ports are also supported on network module ports. These switches support Online Insertion and Removal (OIR) of the network modules.

DIAGRAM Catalyst 9500 Network Modules

Power Supply

The Catalyst 9500 switch supports up to two AC or DC small form factor platinum rated power supply units for a total system capacity up to 650W, 950W & 1600W. Power supplies can be installed in the following combinations: two AC, two DC or a mix of AC and DC power supplies. The power supplies work together in redundant load-sharing mode, in which each power supply operates at approximately 50 percent of its capacity. If one power supply fails, the other power supply can provide power for the entire system. These switches support OIR for power supplies.

Fan and Fan-Tray

Cisco Catalyst 9500 switches have a total of five independent fans or dual fan trays depending on the SKU. Each individual fan operates at variable speeds. The hardware is capable of accommodating a failure of up to one individual fan or fan tray. The remaining fans will automatically increase their RPM to compensate and maintain sufficient cooling. These switches support OIR of the fans or fan trays for up to 120 seconds.

25G and 100G – Enabling Higher Speeds in Enterprise

Cisco has been pioneering several initiatives to bring new Ethernet technologies to market. These include 100GBASE Quad Small Form Factor Pluggable QSFP28 and 25GBASE Small Form Factor Pluggable SFP28. The 25GBASE options include dual-rate optics (25G and 10G) supporting enterprise campus distances to facilitate next-generation network speed and architecture transformations. These innovations enable flexible options and backward compatibility to drive network speeds beyond the current 10G and 40G capabilities while minimizing cost and real estate changes.

Cisco Catalyst 9500 25G & 100G Switch Portfolio

Catalyst 9500 switches are the foundation for the next-generation 25G and 100G-enabled high-speed enterprise campus network. To support these newer speeds in core and distribution, Cisco offers a full suite of switch options with industry-leading port density and flexibility.

Cisco 25G Optics Portfolio

The Cisco 25GBASE SFP28 portfolio offers customers a wide variety of high-density and low-power 25 Gigabit Ethernet connectivity options for the evolving campus network.

Features and Benefits of Cisco 25G Optics

- Support for dual rate optics in enterprise networking to provide unsurpassed **investment protection**
- **Interoperable** with other IEEE-compliant 25G interfaces where applicable
- **Certified and tested** for superior performance, quality, and reliability
- High-speed electrical interface **compliant to IEEE 802.3by**

For more information, refer to *Cisco 25GBASE SFP28 optics and copper modules*

Cisco 100G Optics Portfolio

The Cisco 100GBASE QSFP28 portfolio offers customers a wide variety of high-density and low-power 100 Gigabit Ethernet connectivity options for enterprise core and distribution layers. In a 3-tier campus network, as the distribution layer moves to 25G or 40G, it is desirable to have a 100G-enabled core. It is important to note that form factors of 40G and 100G optics are compatible.

Features and Benefits of Cisco QSFP28 Optics

- **Hot-swappable** input/output device that plugs into a 100G Gigabit Ethernet Cisco QSFP port.
- **Interoperable** with other IEEE-compliant 100GBASE interfaces where applicable.
- **Certified and tested** for superior performance, quality, and reliability.
- High-speed electrical interface **compliant to IEEE 802.3bm.**

For more information, refer to *Cisco 100GBASE QSFP optics and copper modules*

Packaging, Licensing and Support

Catalyst 9000 has a new simple licensing model. The previous generation of switches had different license types (LAN Base, IP Base, IP Services, and Enterprise Services) with the added complexity that they were not the same across the multiple Catalyst switch families. The Catalyst 9000 family uses the same software packaging and licensing model across all platforms.

The new model provides the following benefits:

- simplifies the packaging of features,
- delivers a more cost-effective solution for consuming features, and
- lowers up-front costs by adding more features and support

Software Packages and Licenses

The Catalyst 9000 offers two software packages:

- **Essentials** - provides the baseline network functionality used to operate a network.
- **Advantage** - includes all the functionality in the Essentials package and adds advanced capabilities such as advanced security, availability, automation, and assurance.

Each software package comes with two licenses: a perpetual license that stays with the hardware for the life of the device and a subscription license that is renewable at the end of its term. The two software license types are:

- **Network** - enables a base set of network functions perpetually.
- **DNA** - enables a set of unique platform and solution capabilities for a term of three, five or seven years.

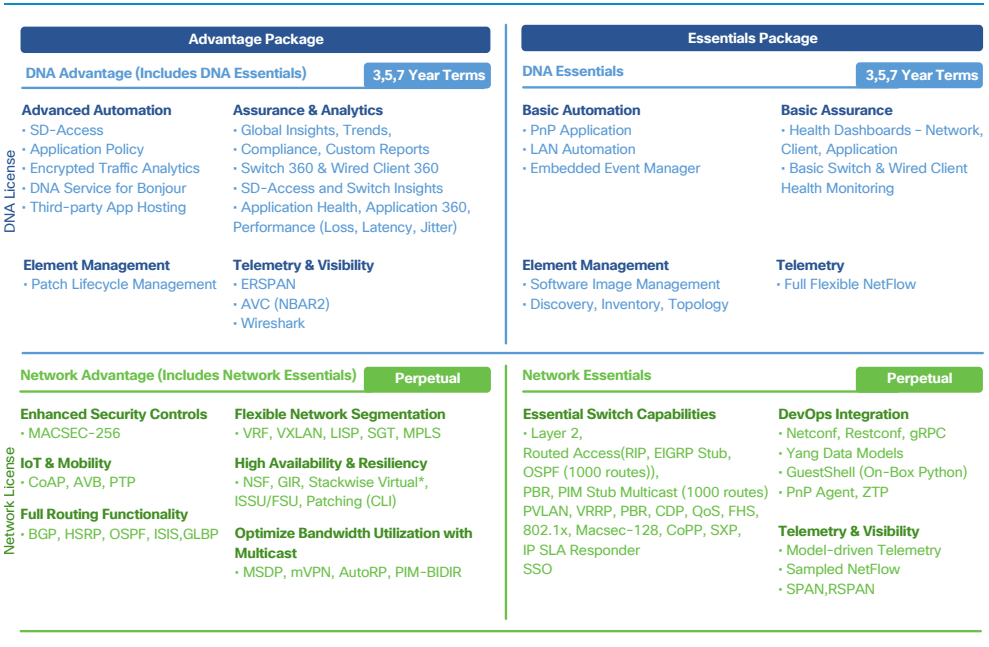
A package is the combination of (2) licenses, with a feature level:

- **Essentials package** = (1) **network-essentials** license **+** (1) **dna-essentials** license
- **Advantage package** = (1) **network-advantage** license **+** (1) **dna-advantage** license

When purchasing a Catalyst 9000 switch, one of these two software packages must be selected. As noted, the packages tightly couple the two licenses. The licenses may not be bought individually.

A high-level grouping of functionality in each package is shown in the following diagram:

DIAGRAM Advantage and Essentials Packages



Note Not all functionality is available across all platforms.

For a complete list of the features of each package, use the *Cisco Feature Navigator*.

Technical Support

Technical support for the Catalyst 9000 family covers both the hardware and the feature packaging just discussed. For technical assistance troubleshooting hardware problems and providing replacement components or chassis, Cisco provides the following general options:

- An Enhanced Limited-Lifetime Warranty covering:
  - 90 days of Technical Support (beginning on the date of initial purchase) from Cisco's Technical Assistance Center (TAC).
  - Hardware troubleshooting and replacement honored for the life of the switch

For terms and conditions, please refer to Enhanced Limited-Lifetime Warranty on www.cisco.com.

- Premium services available through Cisco or a Cisco partner include both technical support and hardware replacement.
  - Consult your Cisco or Partner sales teams for available offers.

Software support comes in the following forms:

- The switch base functionality enabled by its network license is valid for the device's lifetime. Software updates for network licensed components are perpetual.
- 90 days of Technical Support for the switch base functionality (beginning on the date of initial purchase) from Cisco's Technical Assistance Center (TAC).

For terms and conditions, please refer to Enhanced Limited-Lifetime Warranty on www.cisco.com.

- Software support for those features enabled through a switch's DNA Essentials or DNA Advantage license is included while the subscription is valid. Support for these features ends once the subscription expires.

Note that even though DNA license subscriptions are term-based, once they expire, any features unlocked by the license will continue to function. However, once the license term expires, all technical support for features supported by the expired DNA licenses also end.

ASICs - The Power of Programmable Silicon

What is an ASIC?

An Application Specific Integrated Circuit (or ASIC) is a silicon microchip designed for a specific task (e.g. bridging or routing packets), rather than being used for general-purpose processing such as a CPU. ASICs are fundamental to how an Ethernet switch works.

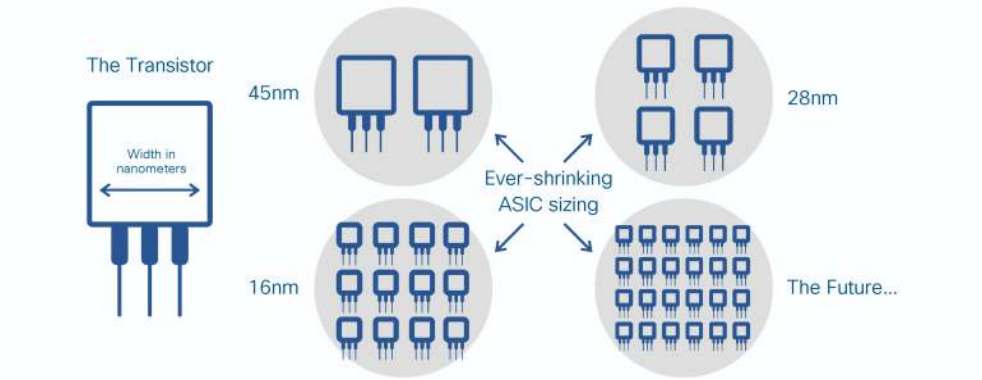
ASICs are custom-designed for the products they are part of and the solutions they support. In a network switch, an ASIC handles packet recognition, manipulation and L2/L3 forwarding at extremely high speeds (tens or hundreds of gigabits per second, trending towards terabits per second), while also allowing a rich set of services for the traffic, such as prioritization (e.g. QoS), accounting (e.g. NetFlow), segmentation (e.g. VRFs and SGTs), traffic filtering and enforcement (e.g. ACLs), path selection (e.g. PBR), and much more.

ASIC microchips are measured in nanometers (billionths of a meter). This is the size of the various components, such as transistors, that the ASIC is built from. The three main advantages of smaller ASICs are:

- increased speed (electrons have shorter distances to travel).
- lower power consumption and less energy wasted as heat.
- lower cost (improved chip yield by decreasing the chance of hitting a silicon defect).

Modern ASICs are generally manufactured at sizes ranging from 45 to 28 nanometers with some newer models as small as 16 nanometers.

DIAGRAM Microchip Packaging

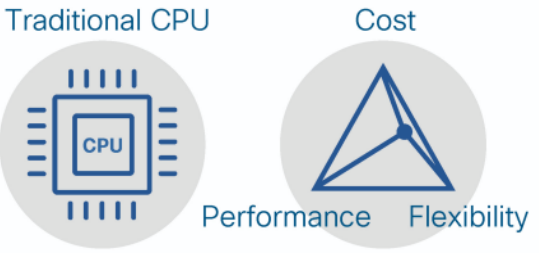


Why Do We Need ASICs?

A generic CPU is too slow for forwarding traffic in a switch. While a general-purpose CPU may be fast at running random access applications on a laptop or server, manipulating and forwarding network traffic is a different matter. Traffic handling requires constant lookups against large memory tables (e.g. L2 tables for MAC addresses, L3 tables for IP routes, L4 ACLs for Security and QoS, etc)

In a generic CPU, all of these tables are held in off-chip memories (not located on the CPU itself) and incur significant performance penalties for frequent memory access. There are also limited data paths and buffers to handle incoming packets (remember, this is millions or even billions of packets per second). Once packets have been received and queued, the CPU must perform the actual processing functions, finding destination lookups, rewriting packet formats, etc. For these reasons, a CPU is not well-suited for this purpose.

DIAGRAM Traditional CPU - More Flexibility, Less Performance, Cost Neutral



└the bottom line

CPUs are flexible but slow. ASICs are necessary to meet the requirements of enterprise networks.

The following chapters examine both traditional types of network ASICs and the latest state-of-the-art programmable ASICs. Administrators will discover not only why ASICs are central to how a switch operates but also how modern ASICs form the foundation of the enterprise network, now and in the future.

Why Programmable ASICs?

Traditional ASICs

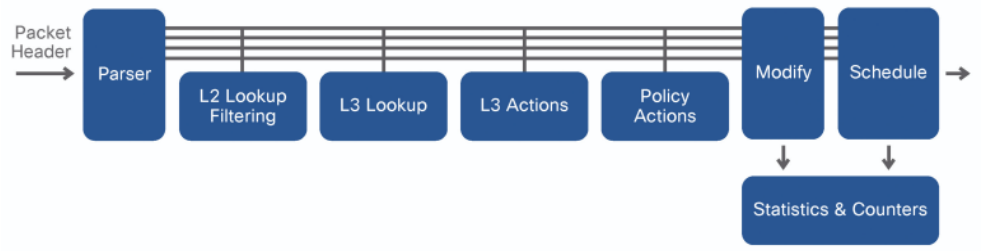
Many ASICs have been used in Cisco switches and routers over the years. Each of these ASICs was designed and developed for the specific features and scale needed for different roles in the campus network. Each also has different capabilities, speeds, and scaling properties suitable for their roles in the network.

However, they all have something in common: this class of networking ASICs is known as fixed ASICs. All aspects of these ASICs (behavior, speed, scale, etc.) are 'baked into' them as part of the manufacturing process and cannot be changed without creating a whole new version of the ASIC.

Another reason they are called fixed ASICs is their processing behavior. As the name suggests, all incoming packets are subject to a fixed series of steps known as a processing pipeline. The typical fixed ASIC processing pipeline stages are mostly similar to the following:

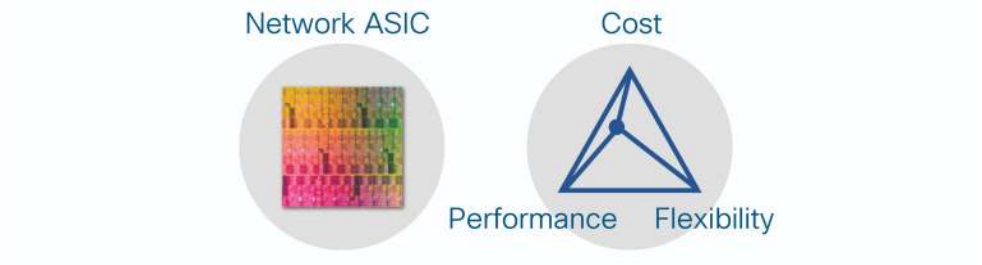
1. Parse incoming packets (examine headers).
2. Layer 2 processing (e.g. MAC lookup).
3. Layer 3 processing (e.g. IP lookup).
4. Policy processing (e.g. ACL lookup).
5. Packet rewrite and traffic counters.
6. Queue scheduling and transmission.

DIAGRAM Traditional ASIC - Processing Pipeline



It is also worth noting that due to the way fixed ASICs are designed and manufactured, along with the time to integrate the ASIC into a network switch, can often require many years before delivering the final product. Fixed ASICs are very cost-effective and efficient but are not flexible nor adaptable. They are only able to handle the types of packets that the chip is hard-wired to process.

DIAGRAM Traditional ASIC - More Performance, Less Flexibility, Cost Neutral



Network and Protocol Evolution

Why do ASICs need to change? To provide an example, the ASIC in Catalyst 3750 can only forward IPv4 and IPv6 packets in hardware. It was designed before VXLAN was developed, so it cannot handle VXLAN in hardware. Since it is a fixed ASIC, it is not

possible to change the processing pipeline to handle new protocols such as VXLAN. An entirely new ASIC would need to be created.

This lack of flexibility may have been acceptable when networks, and the related protocols, did not change much. In the new era of networking, however, everything is "Software-Defined", with ever-evolving protocols and scale requirements. This requires ASICs to support new packet formats and encapsulations such as VXLAN-GPO, GPE, and NSH.

└the bottom line

Traditional fixed ASIC architecture, and thus the network itself, requires hardware replacement to adopt advanced and innovative capabilities that our new software-defined world demands.

Programmable ASICs

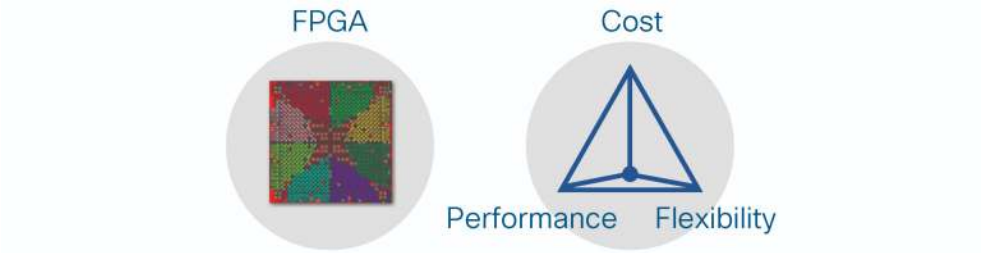
How to get the best of both worlds? How to get the speed we need for multi-gigabit or multi-terabit network devices and also the flexibility to keep pace with new network innovations? These questions led to the concept of programmable ASICs - flexible network microchips designed to adapt to new capabilities as the need emerges, yet still offer the performance networks demand.

Early attempts led to the development of the Field Programmable Gate Array (FPGA). These are essentially simplified ASICs, with reprogrammable logic gates, that can change the original behavior after manufacturing. Although FPGAs do provide a level of flexibility, they are actually very expensive to develop and support. They are not built for any particular task and have little or no onboard memory requiring other chips to provide memory access.

These limitations typically relegate FPGAs to a special-purpose role in most network devices. An FPGA may be used to augment the packet forwarding capabilities of a fixed ASIC for that 'one special feature' the fixed chip does not have. For example, the Catalyst 4500 Supervisor-8 uses an FPGA to provide VXLAN encapsulation, which the switch ASIC does not support. It is usually too expensive to use FPGAs as the primary forwarding engine for a switch (design and manufacturing costs, board space, heat, and

power). Ultimately, this raises the total cost of a switch using combined FPGA and ASIC designs to achieve flexibility.

DIAGRAM FPGA - More Cost, Moderate Flexibility, Moderate Performance



In summary, CPUs are flexible but do not scale for high-speed forwarding; fixed ASICs are fast and scalable but inflexible, and FPGAs are flexible and scalable but very expensive. What is the answer?

Cisco saw this need coming several years ago, and as a result of that foresight, designed and developed the flexible, programmable **Unified Access Data Plane (UADP)** ASIC.

The UADP ASIC combines the flexibility needed to address new and emerging networking protocols and encapsulations, with the speed of a fixed ASIC, and the appropriate cost and scalability to address multiple different areas of the campus network: core, distribution, and access. With UADP, Cisco has truly begun an entirely new era of networking.

The following chapters explore the Unified Access Data Plane ASIC, which is at the heart of the Catalyst 9000 family of switches.

UADP - Programmable ASIC Silicon

UADP - Cisco's Flexible, Programmable Switching ASIC

The Unified Access Data Plane (UADP) ASIC forms the heart of the Catalyst 9000 family of switches. More specifically, the Catalyst 9000 switches are based on the latest generation of UADP ASIC which are evolved from earlier versions used in the Catalyst 3850/3650.

Flexibility is the key attribute that makes UADP the ideal foundation for the world's most advanced switches. This enables the Catalyst 9000 family to:

- handle new frame encapsulations, allowing new features and protocols.
- reprogram their memory tables, allowing switches to adapt to changing needs.
- support multiple interface types and chassis configurations, to address evolving network designs.
- maintain consistent high performance, to address a growing diversity of applications.
- provide a rich, integrated set of flexible traffic handling and accounting services.

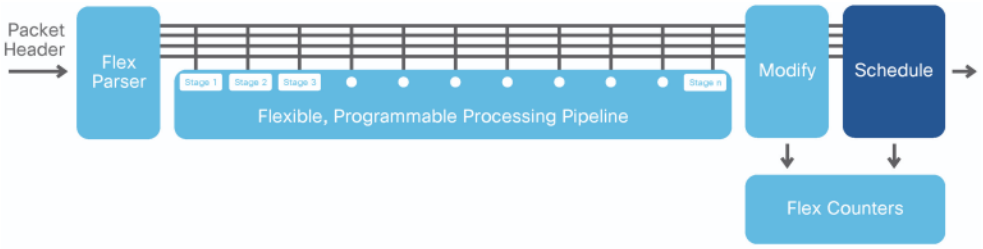
Flexibility At Every Stage

Flexibility has been designed into every aspect of UADP from the beginning. In UADP, almost every processing stage that would be present in a fixed-configuration chip has replaced with a flexible counterpart.

The job of the parser stage is to recognize packet types and headers and analyze them for further processing in the ASIC pipeline. In traditional ASICs, the parser stage is fixed, making it impossible to upgrade the fixed ASIC to recognize or process new packet types and headers in hardware. The UADP contains a reprogrammable FlexParser that can parse a packet for different types of headers.

Unlike the traditional fixed-processing pipeline, the UADP multi-stage flexible pipeline (L2/L3 Forwarding, Policy, Rewrite, Queuing, etc) is also completely reprogrammable (via firmware microcode). There is an ingress pipeline and an egress pipeline, which is not available in most fixed ASICs.

DIAGRAM Programmable UADP - Processing Pipeline

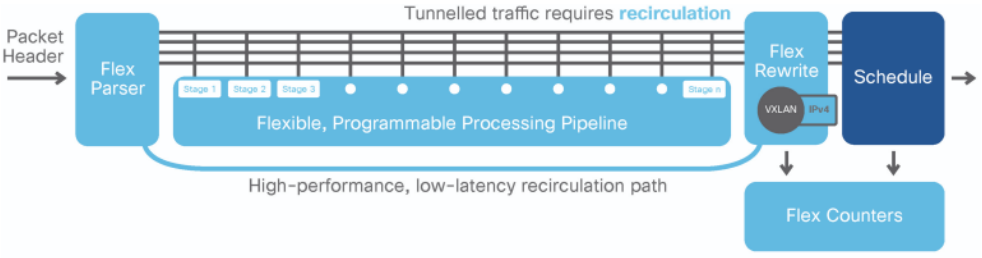


Packet Recirculation

Traffic tunneling is a common design in modern networks. GRE, MPLS, and VXLAN are considered tunnels because they add an additional header to the outer portion of a packet when sending (known as encapsulation), and remove the header when the packet is received (known as decapsulation). Any time packets need to be tunneled in an ASIC, the original packet needs to be processed more than once (known as recirculation) to add the additional header(s).

A quick review of what happens during tunneling reveals why. When a packet arrives and the ASIC decides (based on the switch configuration) that the packet needs to be sent through a tunnel (e.g. VXLAN), a new tunnel header needs to be added in front of the original packet with the source IP of the local side of the tunnel, and the destination IP of the remote side of the tunnel. Since the destination IP address has now changed to that of the remote side of the tunnel, the packet needs to be recirculated through the processing pipeline to forward to this new destination, along with all of the policies that may apply to the tunnel.

DIAGRAM Programmable UADP - Packet Recirculation



In UADP, a packet can be recirculated in approximately 500 nanoseconds (or half a microsecond). In the event that tunneling is required, the impact to forwarding performance is minimal. A packet can be recirculated up to 16 times, while only 2 to 3 passes are normally required. The bandwidth available for recirculation is flexible, meaning packet recirculation can also use the spare bandwidth not currently being used by the front-panel switch ports.

This ability to recirculate packets many times, if necessary, enables even complex use-cases to be accommodated via the UADP flexible pipeline architecture. Now that traffic tunneling is commonplace, it is apparent that UADP was built and optimized for tunneling.

Integrated Micro-Engines

Certain advanced functions executed by UADP may be very processing-intensive. Several tasks, such as fragmentation and encryption, are based on well-known fixed algorithms, and it may not make sense to waste cycles within the UADP pipeline. In such cases, an on-chip micro-engine is available that can process these well-known functions in parallel, thus saving the valuable UADP pipeline performance for other functions.

Some examples of micro-engine functions built into UADP include:

Encryption and Decryption - Security and data confidentiality are paramount in modern networks. Every UADP comes with two levels of hardware encryption built in. The first, operating at the port level at line rate on all UADP ports, is MACsec (802.1AE). MACsec provides link-level encryption to guarantee data confidentiality, meaning packets are encrypted during transmission to, and decrypted when received from, a MACsec-enabled link.

UADP also provides a Datagram Transport Layer Security (DTLS) micro-engine which can encrypt traffic based on packet formats such as CAPWAP and VXLAN. This can serve as the basis for encryption of tunnel overlay traffic. The UADP hardware also allows these two functions (MACsec and DTLS) to be combined on transit if desired. Both use the AES encryption algorithm with up to 256-bit keys, using Galois Counter Mode (AES-256-GCM).

Fragmentation - Any time the Maximum Transmission Unit (MTU) size of a link is exceeded in the network, the original packet may need to be fragmented, and then reassembled at the other side. For example, when traffic is tunneled and the output interface MTU is too small to accommodate the tunnel header plus the original packet. UADP can handle fragmentation actions in hardware, unlike many other ASICs. This ability is important in environments where MTUs may not be easily adjusted end-to-end, and with the growing use of traffic tunneling.

Integrated NetFlow - Accounting for all traffic flowing through the network is important for multiple use cases. The most obvious is for network baselining and capacity planning. Using NetFlow, the entire state of the end-to-end session (TCP or UDP) is tracked by the switch, allowing important information about the entire packet flow to be extracted and analyzed. UADP implements full Flexible NetFlow (FNF) collection capability in hardware. This Catalyst 9000 series is capable of collecting NetFlow statistics for every packet transiting the switch, as an inherent part of overall packet handling.

Cisco Encrypted Traffic Analytics (ETA) utilizes the NetFlow capability and inspects flows to extract vital information about them such as the Initial Data Packet (IDP)

exchange, as well as information about the Sequence of Packet Lengths and Times (SPLT) for encrypted transactions. By integrating this with Cisco Stealthwatch and cloud-based machine learning capabilities using Cognitive Threat Analytics, a high-accuracy network 'fingerprint' analysis can be performed to determine if the encrypted flow represents 'normal' Internet-bound traffic, or whether it may represent a threat posed by encrypted malware.

Policy and ACL - Using integrated Ternary Content Addressable Memory (TCAM) blocks located on-chip for maximum performance, the UADP ASIC provides multiple options for traffic classification and policy enforcement. TCAM matching provides the ability to match traffic flows using IPv4 or IPv6 addresses, special tags such as Virtual Network (VN) ID and Scalable Group Tag (SGT), QoS, CoS, or DSCP values, or other packet markings. UADP can apply the appropriate policies configured by the network administrator. Examples include permit/deny, QoS remarking, path selection, packet copy, and other actions. The UADP flexible pipeline can reference up to two packet matches for multiple parallel actions, without degrading performance.

Packet Replication - Certain application traffic types may require packet replication (creation of multiple copies). For example, an ingress multicast stream may require replication to multiple receivers on the switch. The UADP architecture is optimized for replication, because each packet is held in a central buffer memory during processing, and then a single or multiple copies can be transmitted to all receivers. This is a very efficient replication approach that minimizes latency and space-consuming memory-to-memory packet copies.

Integrated Stacking Capability

UADP is a very powerful and flexible ASIC. In many cases, based on port types and densities, an entire switch may be built around a single UADP ASIC.

However, it may be necessary to connect multiple UADP ASICs together into an integrated system. UADP supports the following inter-ASIC connection options:

- stacking multiple switches together via external cables, to build a single integrated stack;
- linking multiple ASICs together on the baseboard, in a fixed switch; or

- linking multiple ASICs together on a Supervisor module, in a modular switch.

UADP was designed with a dedicated high-speed ASIC Interconnect interface, in addition to the front-panel switch ports, to provide these flexible design options.

Programming UADP with Microcode

Cisco IOS XE uses multiple layers of software. Some of these software layers are more closely associated with the hardware than others. For example, features that a user interacts with (e.g. via the CLI) are typically at a higher layer and less dependent on the hardware. Hardware drivers and other infrastructure pieces of the software (known as microcode) directly interact with the hardware. This microcode layer of the software actually programs the ASIC. Please refer to *Chapter 5 Cisco IOS XE* for more details on the architecture of Cisco IOS XE.

The UADP Family

The history of UADP ASIC began in 2013 when Cisco introduced the Catalyst 3850 switch. As discussed, the ASIC design and manufacturing process is very complex and can take several years for any individual component or product. Several years of innovative work went into developing UADP.

UADP 1.0 took longer to design than most other fixed ASICs at the time as many components were entirely new and designed to be flexible. UADP 1.0 was the first version of a family of UADP ASICs which all share a common architecture. UADP 1.0 was built on a 65 nanometer (nm) process, while the latest UADP 3.0 was built on 16 nm. UADP has progressed significantly in terms of ASIC technology and has incorporated more transistors with each generation. Each additional transistor means additional performance, scalability, features, and functionalities can be built into the ASIC.

UADP 1.0/1.1

UADP 1.0 is a single core ASIC with 1.3 billion transistors, capable of 56Gbps of aggregate bandwidth. It was the first to deliver many of the programmable features discussed in previous sections. Due to its flexible nature, UADP 1.0 was one of the first ASICs to enable support for different, flexible types of packet encapsulations. The first generation of Catalyst 3850 and 3650 used UADP 1.0.

By 2015, a newer version of the same ASIC design (version 1.1) was introduced. The core element and the architecture of the ASIC remained essentially the same, but with some important new additions. The key difference between UADP 1.0 and 1.1 is the use of a dual-core architecture inside the ASIC.

Unlike UADP 1.0, the UADP 1.1 has two ASIC cores with 3 billion transistors. The result is similar to using two UADP 1.0 chips in a single ASIC package. UADP 1.1 also provides higher aggregate bandwidth and performance of up to 160Gbps (80Gbps per core), as well as some new and updated micro-engines. Some of the new features that UADP 1.1

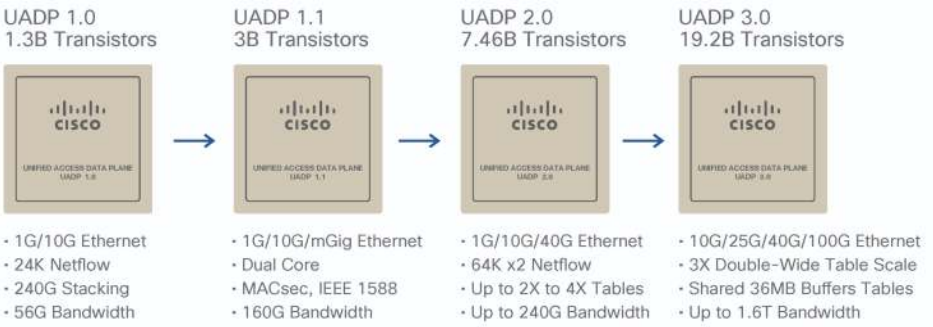
supports include IEEE 1588 timestamps and MACsec 256-bit encryption (AES-256-GCM). The second generation of MultiGigabit and SFP+ versions of Catalyst 3850 and 3650 use UADP 1.1.

UADP 2.0

Catalyst 9000s are built on the next generations of UADP - the UADP 2.0 and 3.0 ASICs.

The UADP 2.0 is a dual-core ASIC (similar to UADP 1.1) with 7.46 billion transistors to provide even higher aggregate bandwidth up to 240Gbps. UADP 2.0 also has larger, more flexible memory tables that can be reprogrammed, giving the option to deploy the same device in multiple network areas, as discussed in Chapter 14 Campus Network Design.

DIAGRAM Comparison of UADP 1.0, 1.1, 2.0, and 3.0



UADP 2.0 ASICs have two variants: UADP 2.0 and UADP 2.0 XL. Both have the same architecture, but the aggregate bandwidth, table scale and overall performance of UADP 2.0 has been optimized for Access layer devices. Catalyst 9300 platforms utilize UADP 2.0.

UADP 2.0 XL has been optimized for modular access and/or distribution layer switches. It has larger memory table sizes (hence the XL designation) with greater aggregate

bandwidth and overall performance to support the port speeds and density of these roles. UADP 2.0 XL also has dual data paths of 720Gbps inter-ASIC connectivity, making it more suitable for platforms where multiple ASICs may be required. The first Catalyst 9500 platforms and the Catalyst 9400 Supervisor-1 and Supervisor-1XL use UADP 2.0 XL.

TABLE UADP 2.0 and 2.0 XL Comparison

	UADP 2.0	UADP 2.0 XL
Total Bandwidth	Up to 160G	Up to 240G
Table Sizes	Standard	XL Tables
TCAM Entries	20K	54K
Buffers	16MB	32MB
Stack Bandwidth	240G	720G
Stack Ring	1	2

UADP 3.0

The need for network speed is never-ending, driven by new wireless speeds, the increasing number of attached devices (IoT) and high-definition video conferencing. New technologies are appearing every day and driving new requirements for network performance and scale.

The UADP 3.0 is a dual-core ASIC with 19.2 billion transistors, to provide an aggregate bandwidth up to 1.6Tbps. UADP 3.0 is the most recent version of UADP, designed to address the challenges brought on by new interface speeds (e.g. 25G and 100G) and new network designs and solutions. A single UADP 3.0 ASIC is capable of 16 ports of 100G line rate.

In addition to increased bandwidth, UADP 3.0 also incorporates several new improvements to make it the ideal ASIC for campus core and distribution. UADP 3.0 has larger shared packet buffers (36MB) to support the interface speed increases. It has

larger double-wide memory table sizes to store both IPv4 (32bit) and IPv6 (128bit) addresses in a single entry. Many other ASICs and previous generations of UADP only support single-width tables, requiring an additional lookup cycle to support IPv6. The first products available with the new UADP 3.0 ASIC are the new 10/25/40/100G Catalyst 9500 platforms.

Cisco IOS XE

IOS Evolution

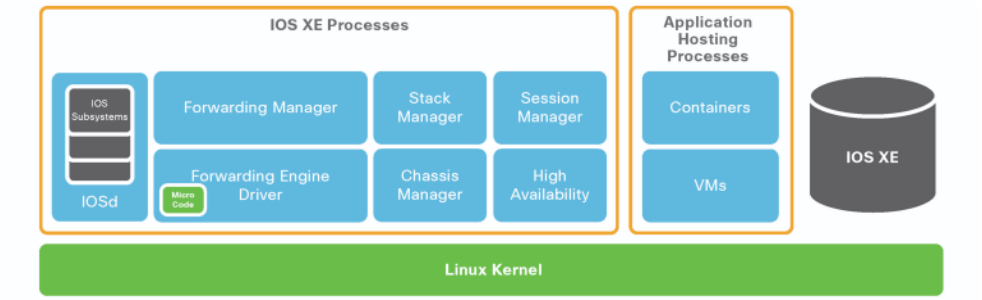
The history of the Internetwork Operating System (IOS) goes back to Cisco's first product, the AGS Multi-Protocol Routers launched in 1986. At the time IOS was a pretty rudimentary, monolithic Operating System (OS). It was one of the very first network operating systems in the industry. Thousands of features have since been added to IOS and over the last 30 years and as the industry has gone through different transitions, IOS has evolved into a more feature rich OS.

Over time, IOS has also branched out into many different versions for different products, with a wide variety of variations. At the same time, Cisco's product portfolio has expanded into switches and routers of various kinds. Purpose-built network areas have evolved, such as data centers and service providers and new operating systems were introduced for these areas, such as NX OS and IOS XR.

A few years ago, Cisco introduced IOS XE, designed to restructure the monolithic code of IOS into a more modular and modern software architecture. With Cisco IOS XE, the OS was subdivided into multiple components to achieve modularity and portability of the features. A low-level Linux Kernel was introduced to provide CPU load-balancing, memory management, and enhanced hardware resource management. IOS now runs as a modular process on top of the Linux kernel (known as IOSd). This approach allows other modular functions to be introduced, such as Wireshark and a Wireless LAN Controller (WLC). More applications will be embedded on Cisco IOS XE in the future, following a similar approach.

Cisco IOS XE is continually evolving. With new applications continually appearing, the established models for configuration and monitoring, such as CLI and SNMP, are beginning to be replaced by standardized APIs for configuration and monitoring data models.

DIAGRAM Classic IOS to IOS XE comparison



The latest Cisco IOS XE software can address key customer needs:

- providing a common OS for enterprise networks,
- rapid introduction of new features and technologies,
- a secure OS to protect the network,
- modularity and high availability,
- programmability and automation.

Considering these requirements, Cisco IOS XE 16.1 was introduced on the Catalyst 3850/3650 switches which use the programmable UADP 1.0 ASIC. Since then, other enterprise network platforms have also adopted Cisco IOS XE as it has the desired software flexibility and scalability for customer needs. This unified OS software release brings multiple advantages:

- fewer software images to manage,
- faster certification of software features,
- unified, consistent experience across platforms,
- ability to run any feature anywhere.

In addition, if there is a need to bring a feature from a core layer platform to an access layer platform, this is much easier due to the use of a unified code release. In most

cases, importing the feature from one platform to another only requires platform dependent code changes.

Catalyst 9000 switches has taken this one step further. The entire Catalyst 9000 product family runs not only the same software release version but also uses the same binary software image. Customers do not have to worry about managing multiple binary files for different platforms (e.g. Catalyst 9300, 9400 and 9500) and can download the single binary file for one of the Catalyst 9000s and run it on all them. This provides added value and significant simplification for software image selection, deployment, and use.

Cisco IOS XE Architecture

Cisco IOS XE is built on top of Linux OS. Various components of Cisco IOS XE run as individual sub-processes and share a common information database that stores the operational state of all the features in a consistent format. This modular OS architecture not only provides key features such as process restartability and patching but also enables the use of Linux Containers or Virtual Machines (VMs) for hosting Cisco and third-party applications.

The Cisco IOS XE architecture has three significant enhancements:

- IOS modularity,
- Cisco IOS XE database,
- VMs and containers.

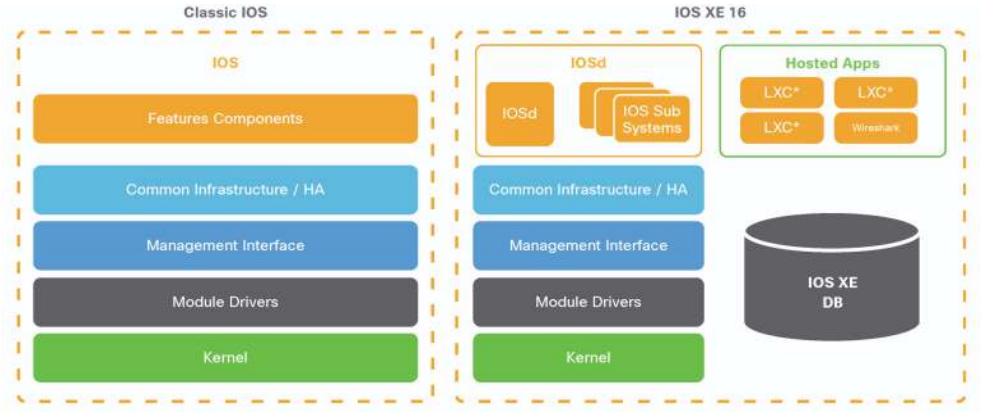
Modular OS

With Cisco IOS XE, the classic IOS code is divided into multiple modules. The majority of the base IOS code is hosted as a daemon (IOSd) which is comprised of traditional IOS features and components such as switching and routing protocols.

IOSd is further divided into multiple IOS subsystems, providing the capability to service one of the sub-systems without affecting the remaining IOSd code. IOSd also provides resiliency in case of individual subsystem failure as it is completely segmented from the remaining IOS code.

This particular OS modularization helps with updating IOS by applying software patches (known as Software Maintenance Upgrades, or SMUs), without affecting the running system.

DIAGRAM Cisco IOS XE



Cisco IOS XE Database

Previously, IOS stored all switching and routing protocol state and related forwarding information in a distributed manner. The process state information was stored in many different parts of memory, in different formats, making it non-optimal nor consumable outside the switch.

The Cisco IOS XE architecture decouples the data from the code. A new feature in the OS is the IOS XE database that stores the configuration and operational state of the system. The stored data is in a standardized format. Major benefits of storing the state information in a centralized database include being able to share information easily between different components of IOS XE.

This standard IOS XE database makes system data easier to express as data models. IOS XE has an interface to convert the database to common data models such as YANG, and provides efficient export using Model Driven Telemetry (MDT). MDT is explained in greater detail in Chapter 12 Programmability and Automation.

Containers & VMs

Cisco IOS XE supports LXC containers and VMs hosting capability on the Catalyst 9000. Please refer to Chapter 13 Application Hosting.

Cisco IOS XE Benefits

With the modern ever-changing software-defined environment, it is imperative that the OS software foundation of Catalyst 9000 be open, easy to use, flexible, and secure. Cisco IOS XE is an open and modular Operating System common across multiple enterprise network products and brings a number of benefits to customers. Cisco IOS XE's modularity, standard database, object-based models, and containers provide key capabilities that help network administrators and engineers with operational tasks and reduce operational costs.

Benefits include a single software image across Catalyst 9000 switches, simplifying network administration and improving software lifecycle management. This provides a consistent format and experience, with consistent provisioning interfaces across all devices. A "run any feature anywhere" approach means that features can be ported very quickly to other platforms. Recent examples of software imported to Catalyst platforms in a short amount of time are MPLS, NAT, and NBAR2.

Some additional key benefits include Cisco IOS XE Install mode, a new WebUI, and Cisco Trustworthy Systems.

Cisco IOS XE Install Mode consumes less memory because the packages are already extracted from the .bin file. With install mode, Catalyst 9000 switches boot IOS faster compared to bundle mode. Install mode is the recommended mode, and advanced high availability features such as ISSU, Patching, and FSU are only supported with install mode.

Cisco IOS XE WebUI was introduced to help customers navigate the device through a standard Web browser. Users can perform simple configurations, troubleshooting, and monitoring high levels of CPU and memory utilization. Users can also configure advanced feature such as AVC to monitor the various applications.

Cisco built the Catalyst 9000 family of switches to be trustworthy to help prevent attacks against a network. As a trustworthy system, the Catalyst 9000 family verifies the authenticity of the platform, prevents malicious code execution, establishes run-

time defenses, and secures communication. For more information, please refer to Chapter 7.3 Trustworthy Systems.

High Availability

Overview

Building networks and network equipment with high-availability (HA) is essential to ensuring business continuity. Catalyst 9000 switches offers several traditional techniques for achieving HA and even introduces some new ones. This section explores these techniques:

- high-availability on the Catalyst 9300 - Stacking, Power, Fast Software Upgrade;
- high-availability on the Catalyst 9400 - Chassis, Supervisors, Power, ISSU;
- high -availability on the Catalyst 9500 - Stackwise Virtual, ISSU;
- Graceful Insertion and Removal,
- Cisco IOS XE patching.

Catalyst 9000 switches utilizes two HA techniques:

Stateful Switchover

Stateful switchover (SSO) offers minimal disruption to Layer 2 sessions for redundant device configuration. SSO replicates forwarding tables and both the running and start-up configuration between an active and a standby component. In the event that the active device fails, the system immediately switches control over to the standby device.

Non-Stop Forwarding

Usually, when a networking device restarts, all routing peers of that device detect that it went down and then came back up. This transition results in what is called a routing flap, which could spread across multiple routing domains. Routing flaps caused by restarts create routing instabilities, which are detrimental to the overall network performance. Non-Stop Forwarding (NSF) helps to suppress routing flaps in SSO-enabled devices. NSF allows for the forwarding of data packets to continue along known

routes while the routing protocol information is being restored following a switchover. With NSF, peer networking devices do not experience routing flaps.

The sub-sections in this chapter explain how each switch family utilizes SSO and NSF in slightly different ways for high-availability.

High Availability on the Catalyst 9300

The Catalyst 9300 delivers access layer redundancy with features such as Stackwise-480, StackPower and Fast Software Upgrades (FSU).

Catalyst 9300 Switch Stacks

The fixed configuration Cisco Catalyst 9300 series switches include stacking to expand port density, switching capacity, and enable redundancy in wiring closets. Moreover, stacking delivers operational simplicity by combining multiple switches together to form a single logical switch.

Cisco IOS XE on Catalyst 9300s supports mixed stacking between any 9300 models for up to eight members. They are physically connected in a ring with special stacking cables connected to the back of each switch using their stacking ports.

Catalyst 9300 stacks deliver deterministic and non-blocking switching performance for up to 448 ports. The switching performance delivers hardware-accelerated, integrated network services such as:

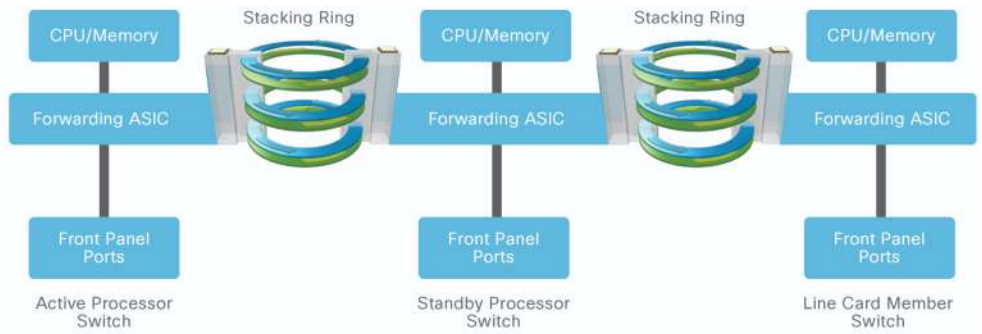
- PoE up to 15.4W,
- PoE+ up to 30W,
- Universal PoE up to 60W,
- Quality of Service,
- Access control lists,
- Flexible Netflow.

Note All switches in a stack must run the same version of Cisco IOS XE and licensing.

Stackwise-480 Architecture

Catalyst 9300 switches enable stacking using a stack-ring fabric known as Stackwise-480. Stackwise-480 refers to the total available stack capacity: 480 Gbps. The fabric consists of six counter-rotating rings (40 Gbps/ring), and the system's throughput is a function of the aggregated throughput of these rings (240 Gbps). Throughput doubles by employing spatial reuse on the stack's rings. Spatial reuse is enabled by destination packet-stripping. Normally, within ring architectures, packet stripping from the ring happens on the source switch where the packet originated and when ring members are processing a packet, no other data may be passed into the ring. Spatial reuse, however, allows multiple flows to co-exist. Spatial reuse frees available bandwidth on the ring as the destination switch strips the packet destined to itself allowing insertion of additional packets onto the ring by other stack members.

DIAGRAM Stackwise-480 Architecture



Stackwise-480 creates a unified control and management plane by electing one switch in the stack as an active switch and another switch as a hot-standby. Remaining switches become stack members. The active switch is responsible for all Layer 2 and Layer 3 network control processing and for synchronizing all state information with the hot-standby. The active switch unifies management for the entire stack; administrators perform all configuration and monitoring via the active switch.

The forwarding architecture is designed to provide distributed switching across all member switches in the stack. Each switch in the stack optimizes data plane

performance by utilizing its local hardware resources. This includes forwarding tasks and network services such as QoS and ACLs. Distributing stack processing delivers wire-speed performance, increases overall system resource capacity, prevents overloading of the active switch processor and optimizes stack-ring bandwidth capacity.

Stackwise-480 SSO/NSF Support

Catalyst 9300 switches support a wide range of Layer 2 and Layer 3 stateful capabilities to provide non-stop network communication. The active switch in a stack synchronizes the protocol state machines, software forwarding tables, and system configuration to the stack's standby switch. Supported protocols are listed in the table below:

TABLE Stackwise-480 Stateful Protocol Support

Layer	HA-Aware Protocols
Layer 2	STP, VLAN, VTP, DTP, CDP, UDLD, SPAN and RSPAN, 802.1x, PAgP and LACP, IGMP Snooping
Layer 3 - IPv4	ARP, EIGRP, OSPF, IS-IS, BGP, MPLS LDP
Layer 3 - IPv6	EIGRPv6, OSPFv3, IS-ISv3, BGPv6, ICMPv6
Services	QoS, ACL, PBR, NetFlow, Port Security

StackPower

Cisco StackPower aggregates all of the available power within a switch stack into one common power pool and shares power among stack members. Up to four switches can be configured in a power stack. It requires the use of Cisco StackPower cables connected to a special port on the back of each switch. Stackwise-480 must first be enabled before StackPower may be used. Thus, if there is an 8-member data stack, then two power stacks of four switches each can be configured to utilize the complete 8-member stack.

Cisco StackPower reduces the number of power supplies required per switch stack and the number of outlets required in the wiring closet. Additional savings accrue from minimizing energy wasted due to inefficient power-supply operation at lower loads and from the reduction in cooling within a wiring closet. The technology also eliminates the need for external power shelves, thus freeing up additional space and power outlets in the wiring closet.

StackPower Operational Modes

Cisco StackPower has two modes of operation: shared and redundant.

In sharing mode, the default, all input power is available for use anywhere in the stack. The total available power is used for power budgeting decisions. If a power supply fails, then the remaining available power from the budget is utilized and there is no impact on either the system components or the POE Devices. If there is not enough power in the budget then POE devices could be shut down, followed by the switches based on the priority. By default, load shedding order is as follows:

1 Low priority ports,

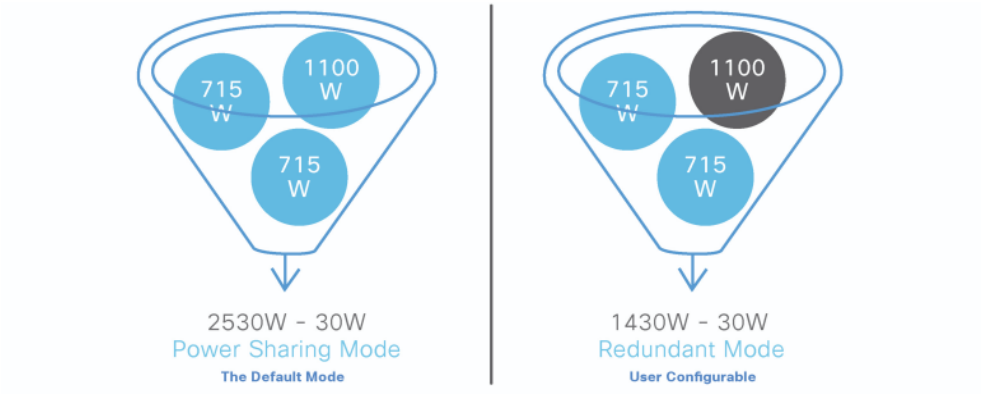
2 High priority ports,

3 Switches.

Power priority is configurable. By default, all ports in the system are considered low priority.

In redundant mode, power from the largest power supply is subtracted from the power budget. This reduces the total available power, but it allows backup power to be available in the event of power supply failure.

DIAGRAM Comparing Sharing and Redundant StackPower Modes



StackPower also reserves 30W in case a new switch is added to the stack.

Cisco StackPower also allows deployment of larger power pools by using a Cisco Expandable Power System (XPS 2200). This system shares power with up to eight switches.

Fast Software Upgrade

During a regular software upgrade on a Catalyst 9300 standalone switch or stacked switch, user traffic is disrupted until the new image is fully loaded. This is due to the control plane and data plane being reset simultaneously while the upgraded software is underway.

The Fast Software Upgrade (FSU) feature decouples updating control plane and data plane functions so that system upgrades have less impact. Via this process, traffic continues to flow unimpeded while the switch updates its control plane. The FSU process only impacts forwarding when the data plane is upgraded. FSU reduces upgrade downtime by almost half when compared to the regular upgrade process.

The FSU feature is supported on both stacking and standalone systems. Further, it can be used to manage switch reloads as well as software upgrades.

High Availability on the Catalyst 9400

The Cisco Catalyst 9400 Series provides several features to minimize outages:

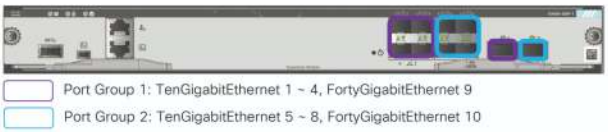
- Dual supervisor stateful switchover with Non-Stop Forwarding,
- Supervisor engine uplink redundancy,
- In-Service Software Upgrade,
- Power supply redundancy,
- Power priority.

Dual Supervisor Stateful Switchover with Non-Stop Forwarding

Supervisor engine redundancy is enabled by default when a second supervisor module is inserted into the chassis. The redundant supervisor engine is automatically synced with the active supervisor's running and startup configuration. The switch's current forwarding state is also replicated to the redundant supervisor and is continuously updated. Stateful switchover (SSO) is triggered if the active supervisor engine fails. If Non-Stop Forwarding (NSF) is configured along with SSO, then routing is not impacted during the switchover; otherwise, only Layer 2 switching is unaffected.

Supervisor Engine Uplink Redundancy

DIAGRAM Supervisor Uplink Port Groups

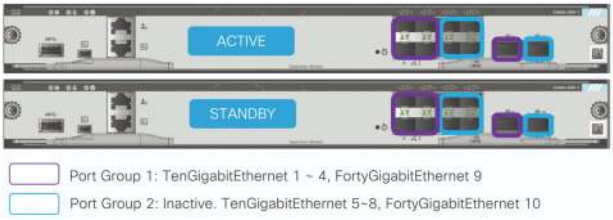


Referring to the picture above, each supervisor has (8) 10 GE interfaces and (2) 40 GE interfaces. Interfaces 1-8 are 10 GE interfaces, and ports 9 and 10 are 40 GE interfaces. The supervisor further combines these interfaces into two port groups: interfaces 1-4 and 9 belong to the first port group and interfaces 5-8 and 10 belong to the second.

Each port group supports a maximum of 40 Gbps of traffic. The supervisor provides flexibility of using those 40 Gbps with either (4) 10 GE interfaces or (1) 40 GE interface.

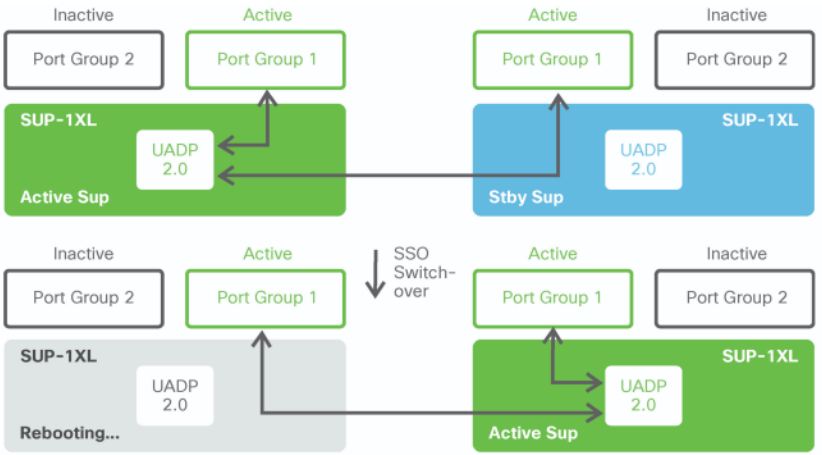
The supervisor is capable of supporting mixed interface types. The network administrator can choose to connect the four 10 GE interfaces in one port group and also connect the 40 GE interface on the other port group.

DIAGRAM Dual Supervisor Uplink Configurations



When a Catalyst 9400 has dual-supervisors installed, the switch automatically disables the second port group on both supervisors. When there is a supervisor switchover, the active uplink ports on the rebooting supervisor continue to forward traffic without interruption.

DIAGRAM Uplinks Stay Active During a Stateful Switchover



In-Service Software Upgrade

In-Service Software Upgrade (ISSU) allows customers to eliminate planned outages for full feature software upgrades. It provides upgrade, downgrade and rollback of the Cisco IOS XE software without incurring an outage. ISSU technology uses SSO and NSF as auxiliary features.

ISSU is an administrative process implemented through a set of exec-level CLI commands issued in a specific order.

ISSU Prerequisites

Before an ISSU can be performed on a Cisco Catalyst 9400 switch, the following must be verified:

- The switch must be using a redundant supervisor engine of the same model.
- The active supervisor must have access to the new Cisco IOS XE image or pre-load it into flash.

- NSF is not required but is recommended if the switch is running routing protocols.
- The system must be running in the install mode.

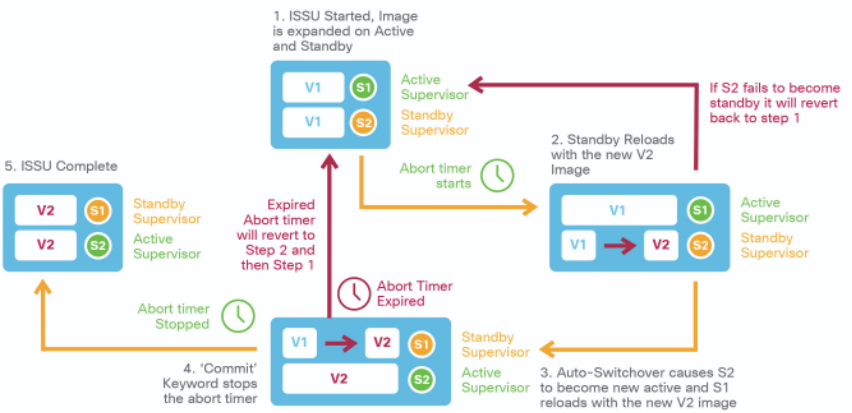
ISSU Process

The ISSU process has five steps:

1. An administrator starts ISSU. The new image is expanded on both the active and standby supervisors.
2. The standby supervisor reloads with the new image.
3. The switch performs an auto-switchover to the standby supervisor, which now becomes the active supervisor. Then the other supervisor reloads the new image, and it transitions into the standby role.
4. Once the ISSU process completes successfully, the upgrade is committed. If not, then the ISSU process automatically rolls back to the previous image.
5. At this point, ISSU is completed.

The diagram below provides further details on how the ISSU process works:

DIAGRAM ISSU Process Details



Note During supervisor switchover, there will be a sub-second traffic reconvergence.

Power Redundancy Mode

The Cisco Catalyst 9400 Series 4-slot chassis has four power supply bays and the 7-slot and 10-slot models have eight bays. The power supplies can operate in a combined or redundant mode.

Combined mode is the default. In this mode, all power supplies are active and sharing the system's load. If a power supply fails, the remaining power supplies pick up the load.

Redundant mode supports two configurations: N+1 and N+N. N+1 provides protection against a single power supply failure. N+N provides protection against multiple supply failures as well as a power input circuit failure.

N+1 Power Redundancy Mode

This is a user-configured mode which allows the user to designate any one of the power supplies as a backup. The designated backup power supply remains in a standby mode. If any one of the active power supplies fail, the backup power supply is activated.

DIAGRAM N+1 Power Redundancy Mode

N+N Power Redundancy Mode

This is also a user-configured mode. Here, an operator divides the power supplies into two groups: active and backup. The power supplies in the active group share the system

load and the backup power supplies remain in standby mode. The two groups can be connected to the same or different input circuits. If the primary input source fails, or any one of the active power supplies fails, all backup power supplies are activated.

DIAGRAM N+N Power Redundancy Mode

Power Priority

The Catalyst 9400 series supports power priority for line card slots. If the system requires more power than the available system power, due to additional PoE draw or sudden failures, the system begins shedding power. Supervisors and fan tray always have the highest priority, and this cannot be modified. By default, the switch turns off line cards starting with the bottom slots and then works its way up to the top. Each line card's power priority, however, can be individually configured.

Stackwise Virtual

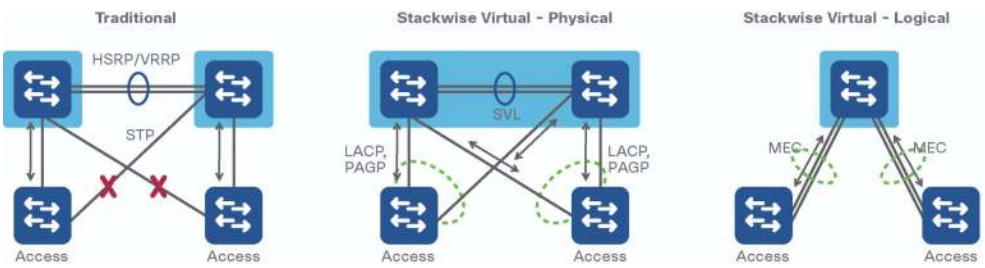
The Cisco Catalyst 9500 series Stackwise Virtual feature allows the merging of two physical switches together into a single, logical switch. The two switches operate as one; they share the same configuration and forwarding state. This is analogous to the virtual switch system (VSS) feature in the previous generation of the Catalyst switching line.

Note Both Catalyst 9500 switches need to use identical hardware, software and license level for Stackwise Virtual to work.

Stackwise Virtual greatly simplifies the design of a campus network. It enables the creation of a loop-free topology because the two switches operate as one. Thus, the spanning tree domain treats a Stackwise Virtual pair as one bridge node instead of two.

Stackwise Virtual also incorporates high-availability features such as stateful switchover, non-stop forwarding and In-Service Software Upgrade, which provides Device-Level Redundancy and eliminates the need for Layer 3 Redundancy Protocols such as HSRP and VRRP. It also supports MultiChassis EtherChannel (MEC), which provides both link redundancy and increased bandwidth.

DIAGRAM Benefits of transitioning from Traditional Architecture to Stackwise Virtual

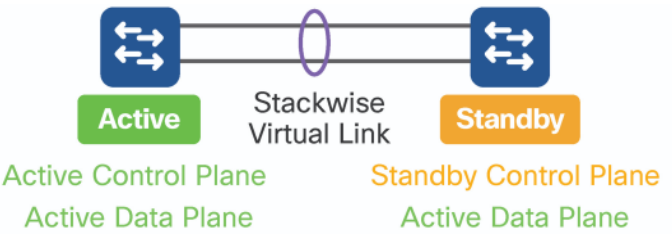


Stackwise Virtual Architecture

Within a Stackwise Virtual pair, one device is designated as the active virtual switch; the other, the standby virtual switch. The active virtual switch manages all control plane functions and distributes the system's current state to its standby peer:

- Management (e.g. running config, startup config, SNMP, Telnet, SSH);
- Layer 2 Protocols (e.g. BPDU, LACP);
- Layer 3 Protocols (e.g. EIGRP, OSPF, BGP, LDP).

DIAGRAM A Stackwise Virtual Domain



From the data plane or traffic forwarding perspective, both switches in a Stackwise Virtual pair actively forward traffic. They each perform local forwarding decisions and, when necessary, forward traffic to neighboring switches via L2/L3 MEC or through Layer 3 equal cost multi-pathing (ECMP).

Stackwise Virtual Components

Stackwise Virtual is formed by leveraging the following components:

1. Stackwise Virtual Link,
2. Dual-Active Detection link,
3. MultiChassis EtherChannel.

Stackwise Virtual Link

The Stackwise Virtual Link (SVL) is a vital part of forming a Stackwise Virtual domain. It provides the signalling path used for synchronizing the two switch control planes, and it serves as the data path for any user data traffic which needs to pass between the two switches. The SVL is an EtherChannel and can be configured using any 10/40G interfaces. Stackwise Virtual supports up to eight interfaces to form an SVL. Cisco recommends using more than one link in an SVL for redundancy between the two switches and to increase cross-switch bandwidth in the event that an uplink or downlink fails on one of the peers.

Dual-Active Detection

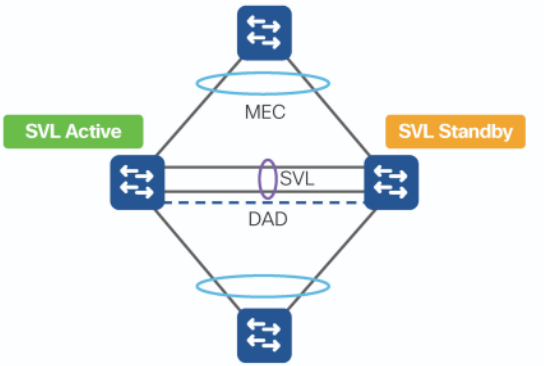
If the SVL link fails for some reason, the communication is broken between the Stackwise Virtual pair. This could cause a dual-active scenario in which both switches assume the active role thus causing adverse effects to both forwarding and network topology. To avoid a dual-active scenario, Cisco recommends configuring Dual-Active Detection (DAD). DAD detects the dual-active scenario and then disables ports on one switch to prevent blackholing traffic. This same switch is rebooted to resolve the dual-active state. Dual-Active Detection can be deployed using either a dedicated link on any 1/10/40G interfaces or it may be configured using ePAgP. Up to four interfaces may be used for a DAD link, and it is recommended to have more than one for redundancy purposes.

Note The management interface may not be used for DAD.

MultiChassis EtherChannel

MultiChassis EtherChannel (MEC) simplifies the connection between Stackwise Virtual switches and neighboring switches by allowing dual-homed connections to be configured as EtherChannel links as opposed to individual links. MEC provides either Layer 2 or Layer 3 multipathing resulting in increased bandwidth and physical link redundancy. Stackwise Virtual supports 128 MECs and it can be configured using mode "ON", LACP, or PAgP.

DIAGRAM The Stackwise Virtual Components



Stackwise Virtual High-Availability

In the event of a failure on the active Stackwise Virtual switch, the standby switch immediately becomes active and continues forwarding traffic. Stackwise Virtual leverages NSF and SSO to achieve a switchover within sub-seconds.

Note NSF should be enabled explicitly for the routing protocols

Stackwise Virtual In-Service Software Upgrades

Stackwise Virtual supports In-Service Software Upgrades. ISSU helps network administrators avoid a network outage when performing a Cisco IOS XE software upgrade on a Stackwise Virtual pair. ISSU supports upgrades, downgrades, and rollbacks.

ISSU Pre-requisites

Before an ISSU can be performed on a Catalyst 9500 Stackwise Virtual switch, the following must be verified:

- The active switch must have access to the new Cisco IOS XE image or pre-load it into flash.
- The switch must be running in install mode.
- NSF should be enabled.

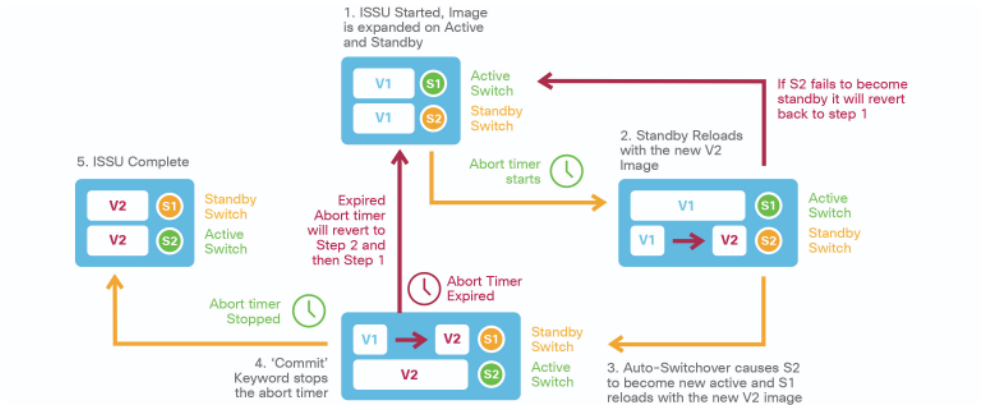
ISSU Process

The ISSU process has five steps:

1. An administrator starts ISSU. The new image is expanded on both the active and standby switches.
2. The standby switch reloads with the new image.
3. The standby switch becomes the new active switch with Auto-switchover. Then the other switch loads the new image, and it transitions into the standby role.
4. Once the ISSU process completes successfully, it commits the upgrade. If not, then the ISSU process automatically rolls back to the previous image.
5. At this point, ISSU is completed.

The diagram below provides further details on how the ISSU process works.

DIAGRAM The ISSU Process



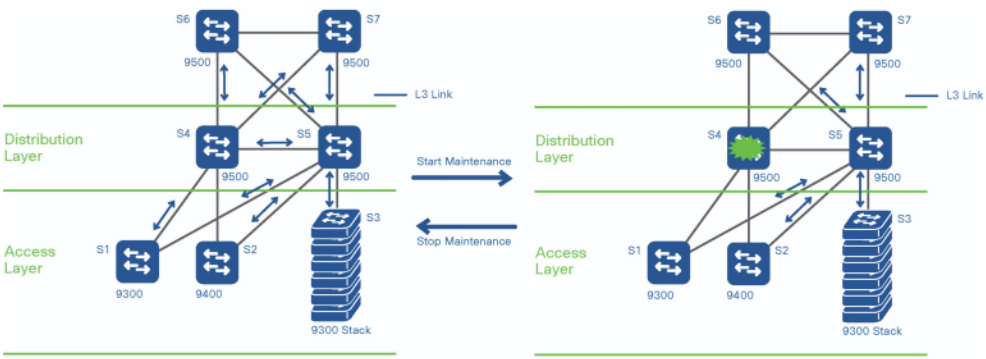
Note During the switchover, there will be a sub-second traffic convergence

Graceful Insertion and Removal

Once a switch is forwarding traffic, there is no simple way to remove it from the network without impacting its active flows. This presents a problem to administrators who need to perform maintenance such as hardware replacement, software upgrades, and troubleshooting. Graceful Insertion and Removal (GIR), solves this problem. GIR leverages redundant paths and existing routing protocols to gracefully isolate a device without impacting active flows. Conversely, GIR also gracefully reinserts the device back into service when the work is complete.

Essentially, GIR allows the network administrator to easily manipulate the routing and first-hop gateway metrics of a network device that is about to undergo maintenance to make it a very unattractive path. It does this by inflating metrics or sending messages to indicate to peers that this device is no longer the best path for traffic. Once traffic moves away from the device, maintenance actions can be undertaken. Once the maintenance is complete, returning these metrics to their former values then smoothly restores normal traffic flow. Below is the sample routed access topology where S4 is put into maintenance mode and all the traffic is diverted gracefully to the redundant path:

DIAGRAM Graceful Insertion and Removal in Action



Note GIR is intended for use in the core and distribution layers.

GIR offers a two-step command structure to what would otherwise require a multi-step, manual and risky process to achieve the same results. The GIR command structure is simple. The `start maintenance` command removes a switch from operation and the `stop maintenance` command returns it to service.

GIR uses two profiles, maintenance-mode profile and normal-mode profile, to manage removal and insertion, respectively. The maintenance-mode profile contains all the commands that are executed to move a switch into maintenance mode. The normal-mode profile, on the other hand, houses all the CLI commands to reinsert the switch into the network.

The following protocols are currently supported by GIR for both IPv4 and IPv6:

TABLE GIR Protocols and Removal and Insertion Methods

Protocol	Graceful Removal	Graceful Insertion
ISIS	Refresh LSPs with Overload bit = 1	Refresh LSPs with Overload bit = 0
OSPF	Send LSAs with max metric	Refresh LSAs with original metric
HSRP	Advertise the resign message	Advertise the Hello Message
VRRP	Advertise Priority 0	Restore the original priority

Catalyst 9000 switches also provide the flexibility to define custom maintenance profiles in order to set site-specific methods for performing removal or insertion. For example, if the device has VRRP or HSRP configured, by default the GIR maintenance shuts down the Layer 2 ports connecting to the access layer while changing the active/standby roles. With a customized template, however, a user can define the isolation of just the HSRP/VRRP and not shut down the Layer 2 ports.

Catalyst 9000 series switches also provide system-generated snapshots to record the state of a switch before and after maintenance. Snapshots are useful for verifying that a switch is operating correctly when it returns to service.

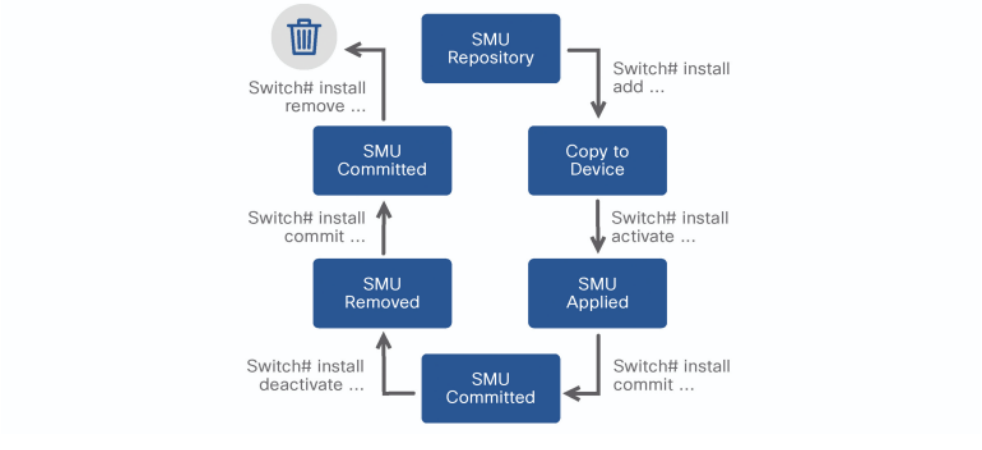
Patching Cisco IOS XE

Software defects happen. It is an unpleasant truth. When a switch encounters an operating system defect, it affects the network's behavior and, consequently, the business operations. Fixing a defect can be just as daunting as encountering them. Often, existing code must be requalified to ensure the problem is truly resolved as well as to prove that the new code does not introduce new issues. In addition, distributing new software across a network infrastructure, and applying it, generally results in needing to coordinate system downtime.

Cisco IOS XE introduces patching to solve this problem. Being a modular operating system, Cisco IOS XE allows point fixes within the software without having to upgrade the entire image. This means bugs and security vulnerabilities can be resolved without having to requalify an entirely new image and potentially without having to reset a switch and incur downtime.

Patching is also referred to as Software Maintenance Upgrade (SMU). The diagram below shows the workflow. When applying a SMU, an administrator first uploads the patch onto the switch. Then, the patch is activated so that the switch understands a new patch is available. Activating the patch applies it. At this point, the patch is active, but if necessary, an administrator can rollback by deactivating it and then remove it to delete the SMU from local storage. If however, the fix is acceptable, the administrator then commits the patch, which makes it persistent across system reloads.

DIAGRAM Sofware Maintenance Upgrade Workflow



Two types of patching are available: cold and hot. Application of a cold patch requires reloading the switch in order to fix the issue. Hot patches, however, do not require a system reload. Instead, the switch only needs to restart the process being patched.

SMU patches can be deployed via multiple methods. The switch command line interface can be utilized on a switch-by-switch basis by an administrator. For distributing patches across a large infrastructure, Cisco recommends using scripts that leverage the Catalyst 9000 APIs or Cisco DNA Center and this includes a software image management utility.

Security and Identity

Overview

This section focuses solely on the new security functionality unique to the Catalyst 9000 platforms. Security features from prior Catalyst switching families have been carried forward into Catalyst 9000 platforms.

Encrypted Traffic Analytics

The rapid rise of encrypted traffic is changing the threat landscape. For example between 2015 and 2016, encrypted traffic nearly doubled, growing from 21% to 40%. By 2019, Gartner predicts that 80% of all web traffic will be encrypted. The bad guys know this, and they are using it to their advantage by making use of encryption to evade detection and hide malevolent activity.

Before the introduction of the Catalyst 9000 series, detecting attacks that hide inside encrypted sessions required unwieldy and expensive measures. In short, it meant installing decryption hardware in the middle of encrypted flows. Such systems can hinder a user's experience by introducing unnecessary latency, and the technique exposes a company to additional legal obligations and privacy issues.

Cisco solves this problem by delivering Encrypted Traffic Analytics (ETA) on Catalyst 9000 switches. ETA identifies malware communications in encrypted traffic via passive monitoring: no extra equipment is required and unnatural traffic redirection need not be performed. ETA achieves this with the extraction of relevant data elements and by employing machine learning techniques that include cloud-based, global security data.

ETA starts from a tried-and-true monitoring technology: Flexible NetFlow (FNF). FNF runs locally on a Catalyst 9000 switch and tracks every conversation, or flow, that passes through it. It collects a range of information about these exchanges in a flow record. Common record values include source and destination addresses, ports, and byte counts.

ETA introduces new flow metadata to help it identify malicious activity hiding within an encrypted flow. These are the Initial Data Packet (IDP) and the Sequence of Packet Length and Times (SPLT).

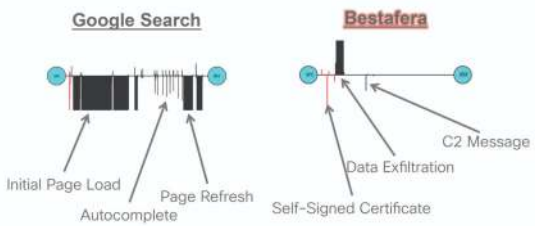
Initial Data Packet

Initial data packets (IDP) are the first packets between two hosts. In the case of ETA, they occur during the handshake used to set up a secure session. Even for an encrypted session, the initial transport layer security (TLS) exchange between two endpoints is passed in clear text. The ETA process can see the TLS handshake and report what it learns, such as which TLS version is being used or which application is carrying the encrypted session. This can be very helpful information when performing a security audit.

Sequence of Packet Lengths and Times

The sequence of packet lengths and times (SPLT) is the length and interarrival time between packets in a flow. Length is measured in bytes and interarrival time in milliseconds. The SPLT provides ETA visibility beyond the first packets of an encrypted flow. ETA matches each flow's SPLT measurements against known malicious behavior in order to identify an attack. For example, consider the picture below.

DIAGRAM Comparison of SPLTs Between Normal and Malicious Behavior



The two graphs shown compare SPLT measurements between a common browser search on Google and a Bestafera attack. Bestafera is malware that acts as a trojan horse on Windows machines. In both examples, the source (src) is a user's PC. In the Google search example, the destination (dst) is one of Google's search engines, and on the right, it is a bad guy's machine. The line between each source and destination pair

represents time. The vertical lines above the timeline show the amount of data sent from the user to either Google or the hacker, and below the timeline is the amount of data downloaded to the user's PC.

The SPLT trace in the Google page search tells the story of a user browsing to the site. The user then types in a search term which triggers Google's autocomplete function. Upon selecting the desired search string, the page reloads with the search's results. The Bestafera graph, by comparison, shows something much different. Here, while running on the user's machine, the Bestafera program reaches out to its command and control server and downloads a digital certificate. With the certificate in hand, the malware opens an encrypted channel and exfiltrates sensitive data from its victim. It then sits idle and periodically checks-in over a command and control (C2) channel for further instructions.

Cisco Cognitive Analytics

ETA integrates with Cisco Stealthwatch and Cisco's Cognitive Analytics, a cloud-based service, to apply machine learning intelligence to ETA's metadata. Cognitive processes IDP and SPLT flow data as previously described and then it compares the results to Cisco's Threat Intelligence Map. The threat intelligence map feeds Cognitive Analytics' engine with security data collected worldwide by Cisco Talos, Cisco's security research division. Cognitive uses the data to model 20 different features across 150 million known or risky endpoints on the Internet. The final result is a more accurate assessment of a particular flow as benign or malicious.

Cryptographic Compliance

ETA also identifies the encryption capabilities used by every network conversation. It reports on the different cryptographic parameters in use such as the TLS version, key exchange technique, and the authentication algorithm used. This allows a security auditor to get a clear picture of which cryptographic algorithms and parameters are in use on the network to verify organizational encryption policies.

ETA on a Catalyst 9000 Switch

Catalyst 9000 switches are the ideal platforms for supporting ETA because they collect full flexible NetFlow information. The collection is performed in hardware directly in the UADP ASIC without any network performance degradation. Additionally, Catalyst 9000 switches leverage their on-board x86 multi-core CPU to deal with the additional overhead needed for collecting IDP and SPLT data.

Trustworthy Systems

There are attacks against the network infrastructure. Networking equipment can either be hijacked through the installation of unauthorized software or by exploiting deficiencies in the running operating system. Even worse, a counterfeit device constructed for easy infiltration by a hacker could unknowingly be installed by an administrator. When these events happen, the network node becomes a point where an adversary can intercept private communications, exfiltrate sensitive data, and launch other attacks against hosts, servers, or the network itself. To help prevent attacks against a network, Cisco built the Catalyst 9000 series of switches to be trustworthy.

Cisco Trust Anchor

All Catalyst 9000 switches employ a local Cisco Trust Anchor (CTA). The CTA is a specially-designed, tamper-resistant chip used to power a device's built-in protections. If this chip is removed, the switch will not operate. The CTA provides a few technologies that drive on-box security.

Random Number Generator

Random number generators (RNG) are fundamental to encryption. The CTA employs a NIST-compliant (NIST SP 800-90A and B certifiable) random number generator that extracts entropy from a true random source from within the chip itself.

Secure Unique Device Identifier

The switch has a secure unique device identifier (SUDI), an X.509v3 certificate. It is generated and installed during manufacturing and it is chained to a publicly identifiable root certificate authority. The SUDI's fields contain the switch's product identifier and its serial number. Including these two fields uniquely binds the SUDI to the switch so that the device can be verified to be authentic Cisco hardware.

The CTA stores the SUDI certificate, its associated key pair, and its entire certificate chain. Furthermore, each SUDI public-private key pair is cryptographically bound to a specific CTA chip. That private key is never exported.

Secure Storage

A Catalyst 9000's CTA additionally provides a highly secure, on-chip storage area. Common items placed here include encryption keys, passwords, LSC and LDevID.

Cisco Trust Anchor Technologies

By building upon the CTA's core components, the Catalyst 9000 series provides hardware authentication, OS integrity and a secure boot process.

Authentic Hardware Check

Every network module or supervisor has its own CTA for hardware authenticity. When a module is inserted, a special library is used to read the module's local CTA and verify its authenticity. Using this makes it impossible to install counterfeit modules into a switch.

Image Integrity

Providing image integrity means a user can be assured that the code they are about to run has not been modified. It is a critical step in establishing trust in a software executable. The integrity process involves creating a unique digital signature for the executable with a hashing algorithm. If the integrity check succeeds, then the code is valid and can be trusted.

Secure Boot

Catalyst 9000 switches follow a secure boot process. It begins by first establishing a root of trust, a secure starting point. The CTA is the root of trust and is used as the basis to establish a trusted chain of valid software during the boot cycle.

Run-Time Defenses

With a trusted operating system loaded, protecting the firmware while it runs is the last step in setting the switch's trustworthiness. Runtime defenses for the Cisco IOS XE have been extended in a number of ways:

Address space layout randomization (ASLR) technology has been added to randomize the locations in memory where different code or data are loaded. That disables the attacking program's ability to know where to jump to inject code or to steal secrets.

When building the Cisco IOS XE code, Cisco used the Safe C Library. Cisco software engineers leveraged some of the following features from this alternative to libc:

- Bounds-checking memory and string functions as well as object size checking.
- Extra compile time warnings to stop developers from injecting security risks.
- Insertion of runtime checks that detect buffer and integer overflows.
- Advanced kernel protection measures.

MACsec

Why was MACsec developed?

An individual with an intention to harm the network could add a tap or Layer 1 / 2 device between two directly connected network devices, and the network administrator might just see a link disconnect and reconnect - but might not ever find the added device. The intruder would now be able to listen to the entire data that is sent over the link and use this data for any harmful purpose. To prevent any possible intrusion on links, Media Access Control Security (MACsec) was developed. MACsec provides value by giving protection against:

- denial of service attacks,
- man in the middle,
- passive wiretapping,
- playback attacks,
- masquerading.

DIAGRAM Intruder Steals Data on Wire

MACsec, which is implemented in Catalyst 9000 switches, is fully compliant with IEEE 802.1AE and 802.1X-2010 standards and operates at Layer 2 in the OSI stack. Layer 2 deployment was used as it involves almost every packet that is transmitted on the link without compromising network performance.

To establish a MACsec session between two directly connected Layer 2 peers, negotiation of keys needs to happen. There are two Session Key Exchange protocols that can form the session - MACsec Key Agreement (MKA) and Security Association Protocol (SAP).

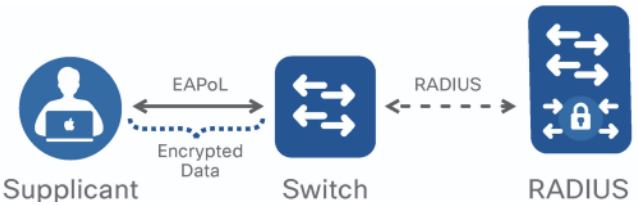
Today, MACsec is enabled between switch-to-switch, user or server host to switches, or router to switches, to ensure data is protected.

Topologies

Host to Switch - usually these are front panel ports.

A host is required to run Cisco Anyconnect Client to do software Encryption/Decryption on the host. Today Anyconnect supports up to AES 128 Bit encryption.

DIAGRAM Host to Switch MACsec

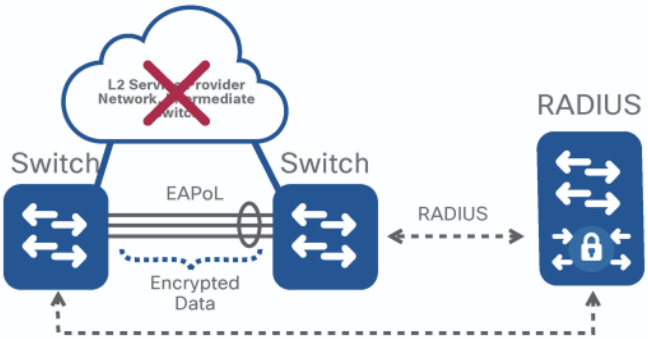


Host to Switch Topology:

- does not support manual Key Exchange.
- supports Dot1X key exchange (ACS/ISE is required and provides large scale).
- supports per-user Authentication/Encryption.

Switch to Switch -usually uses the network module uplink ports or Supervisor uplinks in Catalyst 9400 chassis

DIAGRAM Switch to Switch MACsec



Switch to Switch Topology:

- supports manual Key Exchange config.
- does not support Dot1X key exchange via RADIUS.
- supports EAP-TLS for MKA where dot1x supplicant is enabled on every switch and x.509 certificates are used instead of shared keys.

Supported Platforms

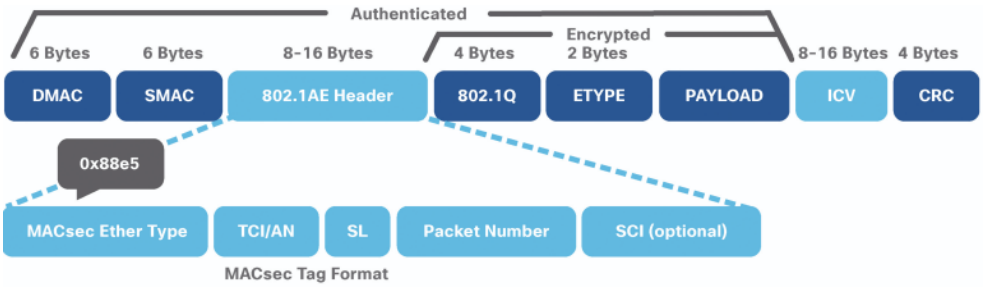
MACsec requires hardware support on the switches to process encryption/decryption at line rate. The Catalyst 9000 family has an integrated block in the UADP ASIC which does line rate MACsec encryption and decryption at any speed.

TABLE MACsec Support per Platform

	MACSec	Cat 9300	Cat 9400	Cat 9500
Switch to Switch	128 Bits SAP	Line rate	HW ready	Line rate
	128 Bits MKA	Line rate	HW ready	Line rate
	256 Bits MKA	Line rate	HW ready	Line rate
Host to Switch	128 Bits MKA	Line rate	HW ready	Line rate
	256 Bits MKA	HW ready	HW ready	HW ready

The reason MACsec support needs to have hardware support comes from the new packet frame format which is used to establish the MACsec session. MACsec uses a new Ethertype (0x88e5) to differentiate these packets.

DIAGRAM MACsec Frame Format



- Authentication – uses only AE Header 16 bytes.
- Encryption – uses AE Header and ICV 16 + 16 bytes.
- No impact to IP MTU/fragmentation.
- All the fields after the destination and source MAC are encrypted, which includes MPLS or dot1Q tags.

- There is no impact on dot1Q Cos or MPLS EXP Marking as they are used after decryption and prior to encryption.
- The Line Rate calculation needs to exclude the added Overhead Header fields for packets. The Bit rate for port is not changed as it does not depend on the packet size.
- "inner shim header" is 20 bytes and is a constant.
- Line Rate per Port = (packet size + inner shim header) / (packet size + inner shim header + **overhead**.
- Line rate in case of Authentication and packet 1500 bytes: (1500 + 20) bytes / (1500 + 20 + 16) bytes = 99% of the bandwidth.
- Line rate in case of Encryption and packet 1500 bytes: (1500 + 20) bytes / (1500 + 20 + 32) bytes = 98% of the bandwidth.

QoS and Queuing

Overview

What is Congestion?

Congestion is a situation where the destination port is unable to forward packets, and as a result, some packets being sent are dropped or delayed more than expected.

Why Do We Care about Congestion?

At times of congestion, packets may be dropped if the hardware buffers are unable to handle additional incoming packets. When packets are dropped, client applications must retransmit. With many application retransmissions, network performance may actually decrease, because the retransmissions will also experience congestion, and the process continues until the congestion is resolved.

Simply adding more hardware buffers will not necessarily alleviate congestion problems since latency-sensitive traffic needs to be forwarded to its destination as quickly as possible.

QoS and Congestion Management in the Catalyst 9000

Catalyst 9000 switches use a modular QoS CLI (MQC) model to configure policers, shapers and traffic remarking.

Modular QoS Model

Catalyst 9000 switches use the MQC model because of following values:

- delivers a consistent QoS configuration model.
- based on policies, classes and types.

- supports two-level hierarchical policies.
- traffic can be classified by class, queue, port or VLAN.
- supports class-based policing and shaping

For more information on *Cisco* MQC *on Catalyst* 9000, please refer to Cisco.com.

QoS Marking

There are three standard types of QoS marking, depending on the type of protocol:

- L2 COS/User Priority – 3 bits (from 0-7).
- MPLS EXP – 3 bits (from 0-7).
- L3 DSCP, ToS for IPv4 (Traffic Class for IPv6) – 6 bits (0 to 63).

For more information on QoS *marking*, please refer to Cisco.com.

Buffers and Queues

The UADP ASIC provides a shared PBC buffer memory. The UADP 2.0 cores each use separate or split buffer memory and an internal data path for packets between cores. The UADP 3.0 supports a new unified packet buffer between its two cores for faster memory access and to increase burst absorption since a single packet buffer is available to all ports on the ASIC.

TABLE Buffer Scale Information

	UADP 2.0	UADP 2.0 XL	UADP 3.0
Buffer MB per Core	8	16	36
Buffer MB per ASIC	16	32	36

Hardware Buffer Allocation

Traditionally, hardware buffers are statically allocated for each queue; however, this can lead to insufficient buffers for all queues in the event of bursting. To remedy this, the Catalyst 9000 platforms use Dynamic Threshold Scaling (DTS).

The hardware buffer is split into multiple segments:

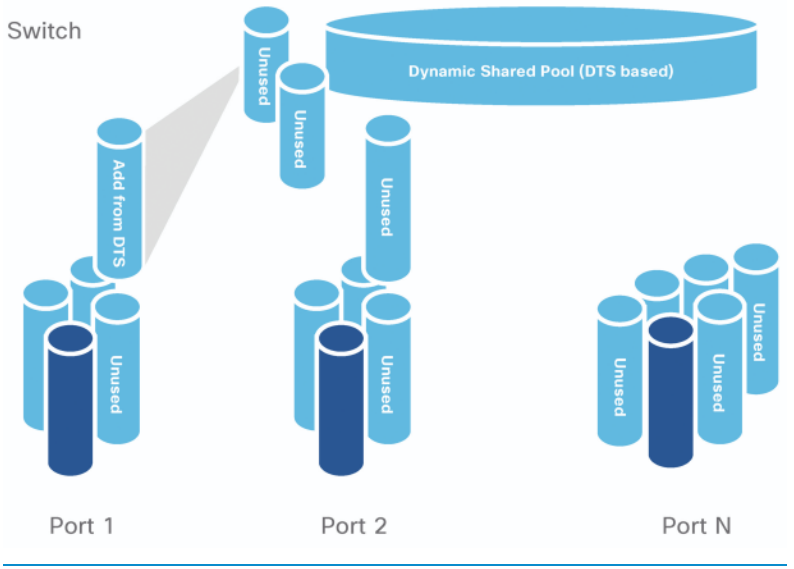
- Ingress buffers (11%) are used for packets scheduled towards the stack and ASIC interfaces.
- Egress stack buffers (25%) are used to receive traffic from the stack ports.
  - The buffer is sized to accommodate up to eight stack members.

- Egress port buffers (64%) are the largest buffers and are used for the port queue structures.
  - These buffers can be shared between different queues and ports using DTS.

DTS Shared Pools

DTS creates a shared, dynamic pool of unused buffers. Buffers from the shared pool can be dynamically assigned to any port that needs more buffer space due to bursting or congestion.

DIAGRAM DTS Shared Pools



DTS provides the following:

- Per-port buffers are split into dedicated (hard) and shared (soft) categories:
  - Shared buffers are good for absorbing packet bursts.
  - Dedicated buffers are good for predictable performance.

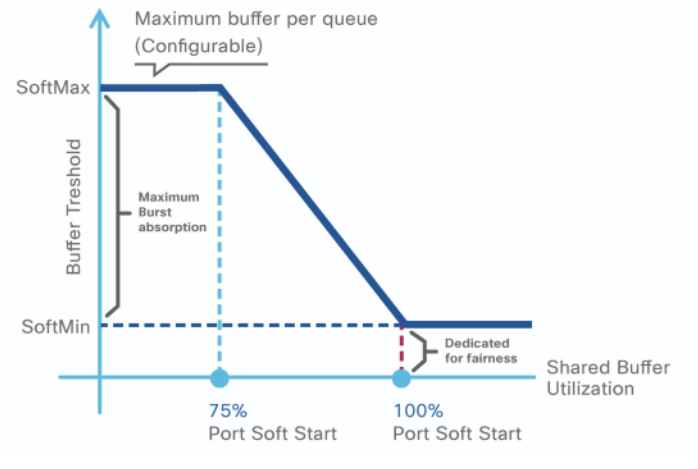
- Dedicated buffers are used first, followed by the shared buffers:
  - Dynamic Threshold Algorithm (DTA) is used to manage the shared buffers.
- The assignment of buffer sizes is flexible (across dedicated and shared):
  - There is a configurable dedicated threshold per port/queue.
  - There is a configurable global maximum shared threshold.
  - The shared pool is automatically adjusted by the DTS algorithm.

DTS Parameters

The parameters used to tune DTS are:

- SoftMin - Minimum shared buffer space given per port.
- SoftMax - Maximum shared buffer that a port can consume from the shared pool.
- Port Soft Start - Time when the SoftMax starts to decrease.
- Port Soft End - Time when the SoftMin and SoftMax are equal.

DIAGRAM DTS Parameters in Use



For more information on *DTS buffer* configuration, please refer to Cisco.com.

Hardware Queues

By default, the Catalyst 9000 series implements a 2-queue structure with (1) Strict priority queue which matches DSCP 16, 24, 32, 48, 56 and (1) Normal queue which matches all other DSCP values.

Users may re-configure hardware queuing up to eight queues with three thresholds per queue. Two queues may be used for differentiated priority queuing. Each port on a Catalyst 9000 may have its own egress queuing policy. The switch uses Weighted Round-Robin (WRR) to schedule egress traffic from its transmit queues.

Congestion Management

The Catalyst 9000 series adds the Weighted Random Early Discard (WRED) algorithm to its queuing process. WRED helps minimize the impact of dropping high priority traffic during congestion. On any port, up to four queues may use WRED.

Note The priority queue cannot use WRED.

QoS and Queuing in the UADP ASIC

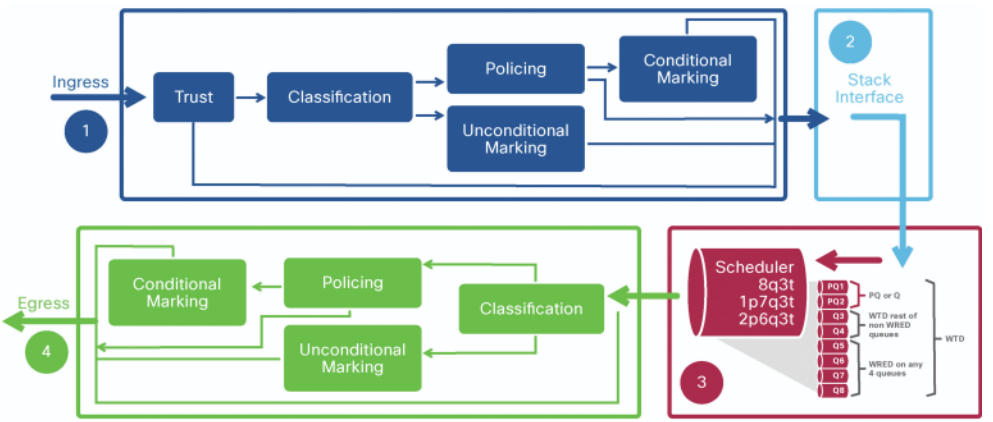
This section describes how the various QoS and queuing processes are applied to a packet as it traverses the UADP.

The packet walk for QoS can be split into four main parts:

1. Ingress classification, policing and marking.
2. Queueing to the stack interface.
3. Egress queueing and scheduling.
4. Egress classification, policing and marking.

The diagram below depicts these four parts, and the remainder of this section explores each step in detail.

DIAGRAM UADP ASIC QoS and QueuingPacket Walk



Ingress Classification, Policing and Marking

This subsection refers to the first processing block in the QoS and queuing packet walk diagram above.

QoS processing begins when a port receives a packet and checks its markings. Ingress packet markers are trusted by default. If the default trust behavior is undesired, the UADP can classify the packet and apply policing or remarking at line rate.

The table below highlights the default trust and queueing behavior.

TABLE Default Trust and Queuing Behavior

Incoming Packet	Outgoing Packet	Trust Behavior	Queuing Behavior
Layer 3	Layer 3	Preserve DSCP/Precedence	Based on DSCP
Layer 2	Layer 2	Not applicable	Based on CoS
Tagged	Tagged	Preserve DSCP and CoS	Based on DSCP (trust DSCP takes precedence)
Layer 3	Tagged	Preserve DSCP, CoS is set to 0	Based on DSCP

The Catalyst 9000 family of switches classifies traffic based on the types below and can use logical constructs, i.e. AND or OR, between multiple classification parameters:

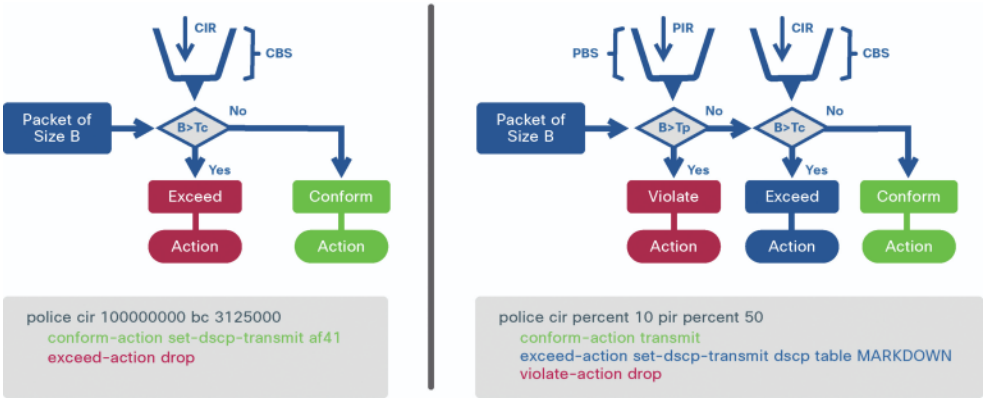
- ACLs
- DSCP
- IP Precedence
- COS
- EXP
- TCP/UDP Ports
- NBAR protocols
- Per VLAN

Policers and Burst Rates

Rate and Burst are the two key parameters which are used implicitly in the policing configuration. With a single-rate two-color policer, the rate, also referred to as the Committed Information Rate (CIR), is defined as the maximum amount of data that can be forwarded in a given interval. The burst is an indication of how much of a CIR can be exceeded. A dual-rate three-color policer may also specify Peak Information Rate (PIR), which is the peak rate allowed above CIR. The max burst is an indication of how much PIR can exceed.

The Catalyst 9000 series supports both 1R2C and 2R3C policers.

DIAGRAM Catalyst 9000 Policer Types



UADP ASIC Interconnect Queuing

This subsection refers to the second processing block in the QoS and queuing packet walk diagram above.

The UADP ASIC interconnect is a point-to-point connection between multiple UADP ASICs. These connections can be on the same switch or to a stack cable leading to a separate switch. An ingress queuing scheduler (IQS) performs congestion management

and provides scheduling and queuing for packets destined to other UADP ASICs. Packets with priority labels are enqueued first on to ASIC Interconnect.

Egress Queueing and Scheduling

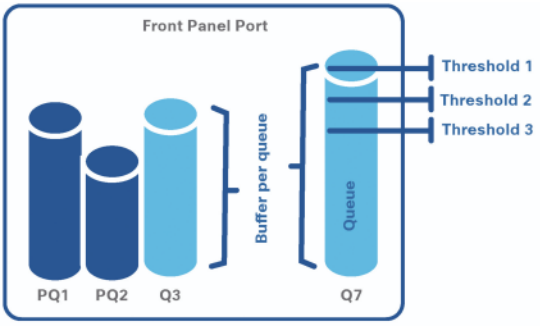
This subsection refers to the third processing block in the QoS and queuing packet walk diagram above.

As discussed earlier, Scheduler is a function of the EQS block which provides multiple port queues, buffers and thresholds allocated to the queues.

Port Queues

With Catalyst 9000 switches, each port supports up to eight egress queues, two of which can be configured as priority queues. Weighted Round Robin (WRR) techniques are employed to empty the transmit queue in proportion to the assigned weights. The Catalyst 9000 manages buffers and congestion using DTS and WRED, respectively, as described in the previous section.

DIAGRAM Port Queuing Parameters



Queue Buffers

Each queue needs to consume a certain amount of buffer (memory) space to store the transit data. The deeper the queue, the more traffic it can hold. Usage of buffers induces latency since they are queuing the packets to be transmitted.

Queue Thresholds

Thresholds are arbitrary internal levels assigned by the switch port that define utilization points in the queue at which the congestion management algorithm can start dropping data. The priority of the packet determines which data is eligible to be dropped when a threshold is reached. DSCP or COS are used to assign traffic to each threshold. In this way, when the threshold is exceeded, the congestion management algorithm immediately knows which packets with which priority values are eligible to be dropped. On the Catalyst 9000 ports, each queue has three configurable threshold values.

Shapers

Shapers typically delay excess traffic using a buffer or queueing mechanism to hold packets and shape the flow when the data rate of the source is higher than expected. In contrast, policers will drop the traffic right away while shapers will try to buffer it first. Shapers are applied on the hardware queues in the Catalyst 9000.

Egress Classification, Policing and Marking

This subsection refers to the fourth processing block in the QoS and queuing packet walk diagram above.

Egress classification, policing and marking are exactly the same as ingress processing with one key difference: the original header can be used only if the ingress path has not modified the header's QoS marker; otherwise, the ingress marking is used.

Hierarchical QoS

Hierarchical QoS (HQoS) allows two policy levels to be configured for QoS thus allowing for greater policy granularity. Hierarchical policies can be viewed as a parent policy at the top level and a child at the bottom level. Administrators can use HQoS to:

- Allow a parent class to shape multiple queues in a child policy.
- Apply specific policy map actions to the aggregate traffic.
- Apply class-specific policy map actions.

One of the key advantages of Catalyst 9000 switches is the support for HQoS in hardware.

The Catalyst 9000 supports four HQoS combinations:

1. Port shaper.
2. Aggregate policer.
3. Per-port, per-VLAN policy.
4. Parent using shaper.

Port Shaper

An HQoS port shaper applies a shaper to all egress traffic using class-default. Within this shaped bandwidth, additional child policies can be specified.

The following CLI example demonstrates a HQoS port shaper configuration:

```
policy-map PARENT
    class class-default
     shape average percent 10
     service-policy CHILD
```

```
policy-map CHILD
    class VOICE
     priority level 1
     police rate percent 20
    class C1
     bandwidth remaining percent 10
    class C2
     bandwidth remaining percent 20
    class C3
     bandwidth remaining percent 70
```

Notes on HQoS port shapers:

- Only the ”class-default” class can be used in the parent policy.
- Only one or two priority queues are allowed in the child policy.
- Different bandwidth per class in the child policy is permitted.

Aggregate Policer

An HQoS aggregate policer applies to all egress traffic using class-default. Within this policed bandwidth, additional child policies can be specified.

The following CLI example demonstrates a HQoS aggregate policer configuration:

```
policy-map PARENT
   class class-default
    police cir percent 30
    service-policy CHILD

policy-map CHILD
   class C1
    set dscp 10
   class C2
    set dscp 20
   class C3
    set dscp 30
```

Notes on the HQoS aggregate policer:

A table-map can be used as a set action in the child policy.

Per-port, per-VLAN Policy

Multiple HQoS parent policers are applied with each policer matching a VLAN as its classifier. Within each VLAN's individual policed bandwidth, additional child policies may be applied.

The following CLI example demonstrates a HQoS per-port, per-VLAN configuration:

```
policy-map PARENT
   class vlan10
    police rate percent 10
    service-policy CHILD
   class vlan20
    police rate percent 20
    service-policy CHILD
   class vlan30
    police rate percent 30
    service-policy CHILD

policy-map CHILD
   class C1
    set dscp 10
```

Notes on the HQoS per-port, per-VLAN policy:

- A table-map can be used as a set action in the child policy.
- Multiple classes under a parent policy are permitted.
- Shaping can be used instead of per-VLAN classification.

Parent Using Shaper

Multiple HQoS shapers are applied under the parent policy, with each shaper matching a traffic class. Within each individually-shaped bandwidth, additional child policies may be applied.

The following CLI example demonstrates a HQoS Parent using Shaper configuration:

```
policy-map PARENT
   class C1
    shape average percent 10
    service-policy CHILD
   class C3
    shape average percent 20
    service-policy CHILD
   class class-default
    shape average percent 30
    service-policy CHILD

policy-map CHILD
   class C1
    police rate percent 10
    set dscp 10
```

Notes on the HQoS parent using shaper:

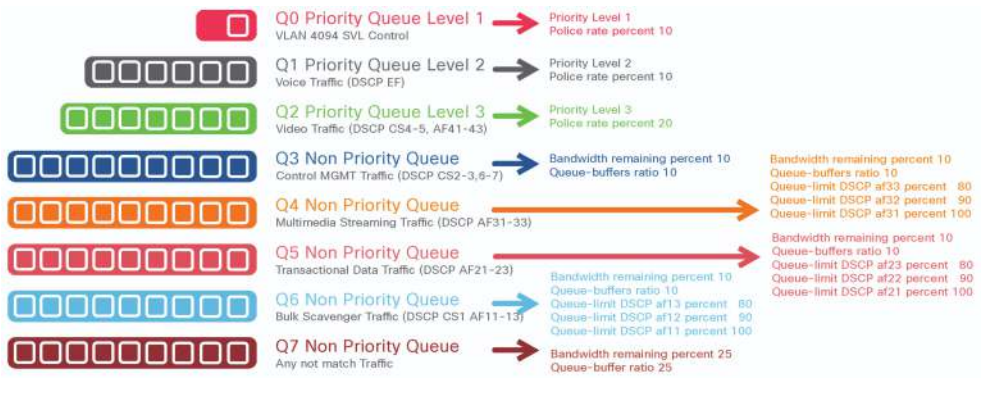
Table-map can be used as a set action in the child policy.

QoS for Stackwise Virtual

Catalyst 9000 series switches running Stackwise Virtual follow the same rules as a standalone switch, except for the special ports used to form the Stackwise Virtual Link (SVL). The SVL is treated as an internal system link. As a result, its configuration, mode of operation, resiliency, Quality of Service (QoS) and load sharing, all follow a special set of rules. The SVL port QoS and queuing mechanism are hard-coded. The SVL port will not apply any custom QoS or queuing policy.

The following diagram illustrates the default policy applied on the SVL.

DIAGRAM Stackwise Virtual Link QoS



For more information on *QoS for Stackwise Virtual*, refer to cisco.com.

QoS for Overlay Technologies

Modern switching networks use virtual network overlays to support mobility, segmentation, and programmability at scale. Overlays are a key enabler of SD-Access.

GRE and VXLAN QoS

GRE and VXLAN are overlay technologies which encapsulate the original IP packet / frame with an outer IP packet header and without modifying the original payload.

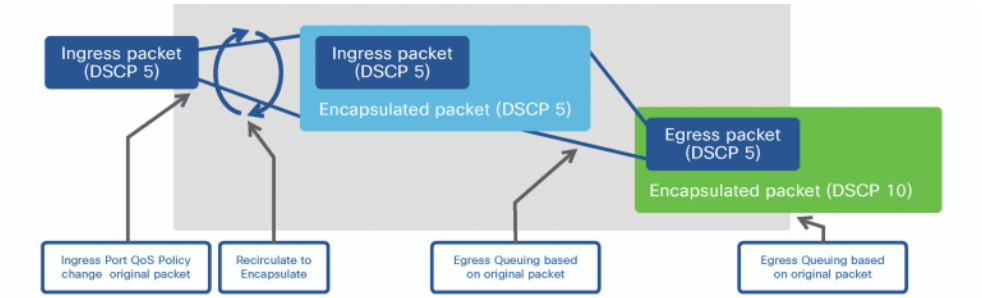
GRE and VXLAN Encapsulation in UADP ASIC

In GRE encapsulation, the original IP packet is encapsulated with the new IP and GRE header and the ToS byte value from the inner IP header is copied to the outer IP header. GRE interfaces do not support QoS policies on ingress.

In VXLAN encapsulation, the original L2 frame is encapsulated with new IP and VXLAN Header and the ToS byte value from the inner IP header, which is part of the original L2 frame, is copied to the outer IP header.

The egress queuing is based on the original (copied) header values.

DIAGRAM QoS Marking for GRE / VXLAN Overlay Encapsulation



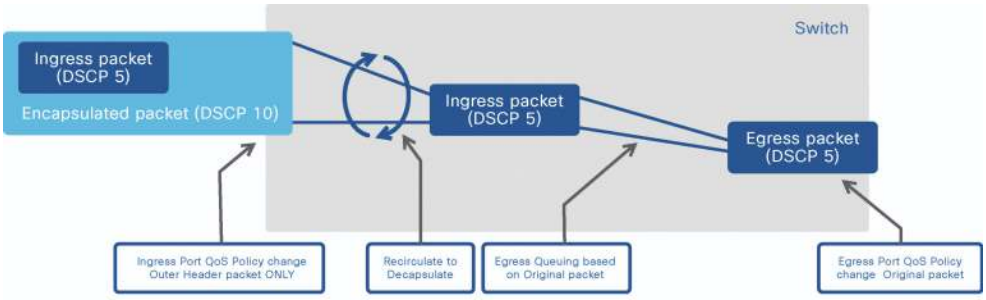
Note The queueing actions are applied before egress policing / marking actions. Refer to the earlier sections for more detail on egress QoS behavior in UADP.

GRE and VXLAN Decapsulation in UADP ASIC

In both cases, when a QoS policy is applied on ingress interface (where packets arrive encapsulated), only the outer header is used for classification and QoS only affects the outer header. The inner packet will not be changed and retains the original marking. The actual GRE/VXLAN tunnel interfaces do not support QoS Policies on egress.

When a QoS policy is applied on egress interface (where the original [decapsulated] packet leaves), the policy will affect the original packet.

DIAGRAM QoS marking for GRE / VXLAN Overlay Decapsulation



For more information on *QoS on GRE overlays* (*RFC 2784*), please refer to cisco.com.

For more information on *QoS on VXLAN overlays* (*RFC 7348*), please refer to cisco.com.

MPLS QoS

MPLS overlays use the MPLS experimental bits (EXP) field in the MPLS header for QoS treatment. In an IP network, the DSCP (6-bit field) defines a class and drop precedence. The EXP bits can be used to carry some of the information encoded in the IP DSCP and can also be used to encode the dropping precedence.

By default, Cisco IOS XE copies the three most significant bits of the DSCP or the ToS field of the IP packet to the EXP field in the MPLS header. However, you can also set the EXP field by defining a mapping between the DSCP or ToS and the EXP bits.

There are three modes used to map DSCP or ToS to EXP:

- Uniform mode (default) – has only one layer of QoS, end-to-end.

  The ingress PE router copies the DSCP from the incoming IP packet into the MPLS EXP bits of the imposed labels. As the EXP bits travel through the core, the bits may or may not be modified by intermediate P routers. In case of modification, the new EXP value is copied back to DSCP bits of the IP Packet.

- Full Pipe mode – uses two layers of QoS - (1) an underlying QoS for the data, which remains unchanged when traversing the ASIC core; and (2) per-hop QoS, which is applied to the outer header, separate from the underlying IP packets.

When an IP packet reaches the edge of the MPLS network, the egress PE router classifies the newly exposed IP packets for outbound queuing based on the EXP bits from the removed MPLS label. The inner IP packet DSCP bits are not modified.

Below is a brief description of the various MPLS QoS modes:

TABLE MPLS QoS Modes

Tunneling Mode	IP to Label	Label to Label	Label to IP
Uniform	Copy ToS/DSCP into MPLS EXP (may be changed by SP also)	MPLS EXP may be changed by SP QoS policy	MPLS EXP copied to IP ToS/DSCP
Pipe	MPLS EXP set by SP QoS policy		Original ToS/DSCP preserved (egress queuing based on MPLS EXP)
Short-Pipe (Note: not supported on Catalyst 9000)			Original ToS/DSCP preserved (egress queuing based in ToS/DSCP)

For more information on MPLS QoS (RFC 5462), please refer to cisco.com.

Application Visibility & Control

Overview

Network engineers are asked to enforce end-to-end business-aligned policies to achieve target performance as well as quickly isolate and resolve application performance problems. Network engineers need detailed oversight of the different types of applications running on the network to optimize business-relevant traffic performance.

The Cisco Catalyst 9000 Switches support the Application Visibility and Control (AVC) solution which truly offers innovative and powerful capabilities for application awareness for business-critical applications in enterprise networks. The Cisco AVC solution leverages multiple technologies to recognize, analyze, and control more than 1400 applications, including voice and video, email, file sharing, gaming, peer-to-peer (P2P), and cloud-based applications.

Cisco AVC has three main functions:

- recognition for granular detection of applications in the network beyond Layer 4,
- ability to prioritize business-relevant versus business irrelevant traffic, and
- control by prioritizing application bandwidth, especially for business-relevant traffic.

Application Recognition

The technology used in AVC to identify applications is next-generation Network-Based Application Recognition (NBAR2).

NBAR2

NBAR2 provides native stateful deep packet inspection (DPI) capabilities, enhancing the application recognition engine to support over 1400 applications (including 140 encrypted applications).

NBAR2 provides powerful capabilities, including:

- Categorizing applications, such as category, sub-category, and application group.
- Field extraction of data such as HTTP URL, SIP domain, and mail server.
- Customized definition of applications, based on ports, payload values, or HTTP URL/Host.
- Customizable attributes for each protocol.

AVC can be configured on wired access ports for both standalone and stacked switches. Catalyst 9000 switches use performance-optimized NBAR2 in which only a few packets are needed to identify the application, reducing the impact on the switch CPU. In a Catalyst 9300 stack, NBAR2 detection runs on each stack member, scaling the solution as more members are added.

NBAR2 can be activated for application visibility explicitly on an interface by enabling protocol discovery, and for application policy by attaching a QoS policy that contains a match protocol classifier.

Application Monitoring

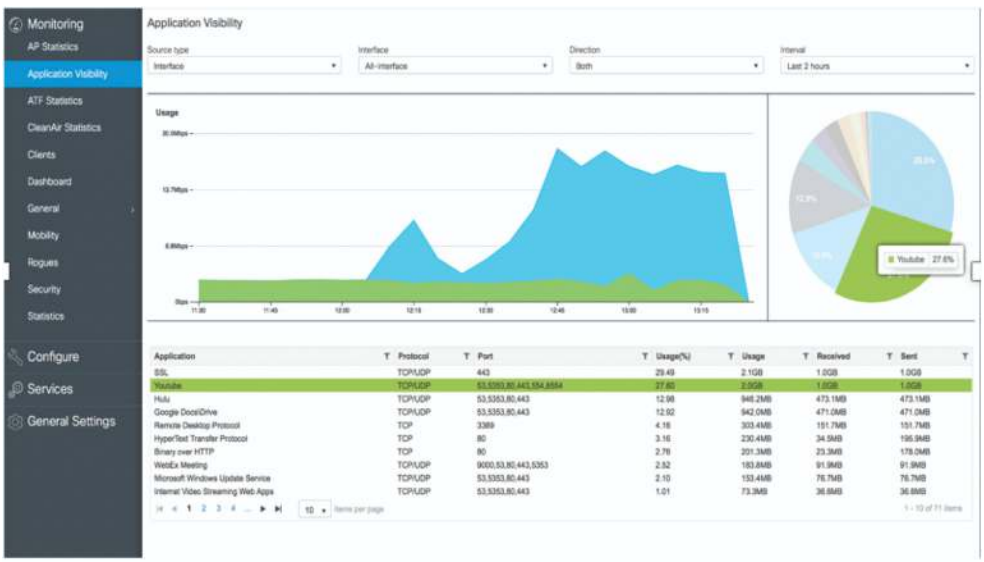
Monitoring of applications is available through the CLI or WebUI.

1 The CLI provides accumulated statistics over time.

- The CLI command is: *show ip nbar protocol-discovery top-n*

2 The Switch WebUI provides statistics for up to the last 48 hours in 5-minute granularity.

DIAGRAM Catalyst 9000 WebUI



3 AVC monitoring with Flexible Netflow.

Flexible NetFlow (FNF) can be configured on an interface to provide application statistics for the interface. The AVC solution is compatible with NetFlow v9 and IPFIX.

FNF enables customizing traffic analysis parameters according to specific requirements.

In Catalyst 9000 switches, FNF collection is performed in hardware with the UADP ASIC. In a Catalyst 9300 stack of switches, FNF collection runs on each member of the stack, scaling the solution as more members are added.

Application Control

Cisco Quality of Service (QoS) provides prioritization, shaping, and rate-limiting of traffic which is used by the contol portion of AVC.

QoS can place designated applications into specific QoS classes/queues. This enables:

- Placing high priority, latency-sensitive traffic into a priority queue.
- Guaranteeing a minimum bandwidth for an individual application or for a group of applications within a QoS traffic class.

QoS can use application information provided by NBAR2 in managing network traffic. The QoS class-map statements enable matching to NBAR2-supported applications and L7 application fields (such as HTTP URL or Host), as well as to NBAR2 attributes. Class-map statements can coexist with all other traditional QoS match attributes, such as IP, subnet, and DSCP.

In Catalyst 9000 switches, QoS enforcement is performed in hardware with the UADP ASIC.

IoT

Overview

The Internet of Things is one of the fastest growing industry trends and it is driving important innovations in existing technologies such as Power over Ethernet (PoE) as well as new technologies and solutions such as Audio Video Bridging (AVB), DNA Service for Bonjour, and Application Hosting. Application Hosting is covered in a dedicated section.

For more information on IoT, please visit *www.cisco.com/go/iot* for more details.

Power over Ethernet Innovations

Power over Ethernet (PoE) is used ubiquitously in network deployments today.PoE serves as a foundational technology in many modern network deployments, allowing devices such as IP Phones, Access Points, IP-based cameras, and other devices. PoE not only provides data connectivity over their Ethernet connecting cable, but also with the power allowing the device to operate. PoE removes the need for wall sockets to power each PoE-enabled device and eliminates the cost of additional electrical cabling and circuits.

From the original Cisco proprietary Inline Power (ILP) implementation, which was limited to 7 watts (7W) of maximum PoE power, PoE has now been standardized as IEEE 802.3af (typically known simply as "PoE", supporting up to 15.4W) and IEEE 802.3at (known as "PoE+", supporting up to 30W). The increase in total available PoE power, as well as the standardization of the PoE approach, has allowed the proliferation of a large and thriving ecosystem of PoE-powered devices.

Today, PoE power is typically used throughout all enterprise network environments, big and small. The convenience that PoE provides for device attachment and use, along with the resiliency it creates for the powered device infrastructure. For example, one UPS backup power supply in a central wiring closet now protects all downstream powered devices from interruption.

Catalyst 9300 and 9400 switches support both PoE and PoE+. Depending on the capacity and density of power supplies in use, these levels of power may be available on all or a subset of the ports. Cisco's innovation with StackPower in the Catalyst 9300 platform also allows for sharing PoE power across multiple stackable switches. This provides a significant benefit in the number of powered devices and high availability on the network.

Current developments with PoE - UPoE

Cisco has pioneered PoE since its inception, driving new advances to the standard and setting the stage for the next phase of PoE innovation. As powered device requirements

push beyond the 30W maximum power defined by 802.3at, Cisco has led the way with the definition and deployment of UPoE (Universal PoE). UPoE provides for up to 60W of PoE power. While UPoE is a Cisco-proprietary solution, it nevertheless is a very useful innovation for powered devices that require more PoE power, such as virtual desktop terminals, IP turrets, compact switches, building management gateways, LED lights, wireless access points, and IP phones.

Both Catalyst 9300 switches and Catalyst 9400 line card models that support UPoE power options are available. Cisco provides for PoE, PoE+, and UPoE power options, not just on 1G copper switch ports but also on mGig ports, allowing devices to be accommodated that need both higher throughput and a higher level of PoE power.

For more details on PoE, please visit *www.cisco.com/go/poe* for more details.

Note The PoE hardware on the Catalyst 9300 switches and Catalyst 9400 are 802.3bt capable.

New Innovations - Fast PoE and Perpetual PoE

As the use of PoE has proliferated, so have the uses cases in which it is used. One of the newer and more innovative uses of PoE is actually to power building lighting fixtures. For example, using PoE+ or UPoE ports as the single power source for commercial and industrial lighting.

Why would one do this versus the traditional method of providing power for lights?

- Lower-powered lighting fixtures can operate more efficiently using 60W UPoE power than traditional lighting systems
- PoE-delivered power uses low-cost Cat5e/6/6a cabling as compared to more expensive traditional electrical wiring
- PoE-delivered DC power is more efficient for use with many new electrical fixtures compared with AC-delivered power

Additional uses to which PoE is being used include powering IoT devices in building systems, such as building controls, thermostats, HVAC control systems, door locks, badge readers, and many similar items of critical building infrastructure.

To support these business critical deployments, Cisco has created two new deployment modes for PoE - Fast PoE and Perpetual PoE.

The goal of Fast PoE is to provide PoE power rapidly during the switch boot up process. Rather than waiting for the entire Cisco IOS XE control plane to load, Fast PoE aims to provide PoE power to attached devices within less than 30 seconds after power is applied to the switch. This is especially important to help bring IoT and other similar devices online as quickly as possible after a power outage, rather than waiting several minutes for a full reload of the switch to complete.

Perpetual PoE has a similar but slightly different goal. The aim of Perpetual PoE is to keep PoE power available to downstream devices even during an Cisco IOS XE control plane reload, ensuring continuity of power for attached devices. For example, if the switch was to be reloaded (as during a software upgrade), it is undesirable to power down critical attached devices such as lighting fixtures during the reload cycle.

AVB - Audio Video Bridging

In the past, audio and video (AV) deployments have traditionally relied on analog, point-to-point infrastructures for implementation and deployment. With the migration of AV to digital transmission, these infrastructures have largely retained their point-to-point nature. This deployment model has traditionally resulted in very cumbersome and expensive deployments that created significant operational challenges. Proposed solutions to these digital implementation issues were often non-standard, expensive, and came with a significant operational burden.

Ethernet was widely viewed as a new way forward for AV implementation - one that could offer a common medium for digital AV data interchange, and do so flexibly and inexpensively. Ethernet was not, however, designed for the low-latency, predictable, lossless requirements of digital AV.

This is the genesis of the Audio Video Bridging (AVB) set of standards. These consist of the following major areas:

- **IEEE 802.1Qat**: Stream Reservation Protocol (SRP) / Multiple Stream Reservation Protocol (MSRP). These provide an end-to-end admission control system required for the carriage of AV streams, ensuring the availability of resources such as bandwidth and latency.
- **IEEE 802.1Qav**: Forwarding and Queuing for Time-Sensitive Streams (FQTSS). Provides an AV traffic scheduling mechanism.
- **IEEE 802.1AS**: Generalized Precision Time Protocol (gPTP). Provides synchronization and timing for time-sensitive applications on L2 devices. gPTP is based on the IEEE 1588 Precision Time Protocol (PTP) standard.
- **IEEE 802.1BA**: Defines profiles for features, options, configurations, defaults, protocols, and procedures for AVB devices.

For more information on AVB, please visit *www.cisco.com/go/avb* for more details.

The AVB standard is supported on Catalyst 9000 series switches, as well as the Catalyst 3850 and 3650 platforms. With an AVB-capable endpoint and switch, analog AV signals are aggregated at AVB endpoints and transmitted on the Catalyst-based infrastructure. As a system, AVB is a cost-effective and flexible solution to collapse AV infrastructures onto reliable, simple Ethernet media.

AVB is an important part of the future of digital media production, and the Catalyst 9000 series offers support for this important set of standards.

DNA Service for Bonjour

The Apple Bonjour protocol is widely used in campus environments such as education and retail for device discovery and simplified connectivity to network services. Based on the Multicast DNS (mDNS) standard, Bonjour is widely used with many devices and service types, including, for example, many Apple devices, to provide easy discovery of devices and simplified device attachment.

The Bonjour protocol is optimized for "plug and play" use in home and small office deployments, and almost too simple for enterprise use. The mDNS protocol uses link-layer multicast for device and services discovery, and is inherently not routable, being limited to the local L2 broadcast domain (i.e. one hop only). This, of course, limits the deployability and use of the Bonjour protocol in larger enterprise networks, which are inherently based on routed infrastructures.

For more information on Bonjour and mDNS, please visit *www.cisco.com/go/mdns* for more details.

To address this challenge, Cisco introduced Service Discovery Gateway (also known as Bonjour Gateway) services into various switching and routing platforms with Cisco IOS XE. The goal of the Service Discovery Gateway capability on these platforms is to allow reachability to Bonjour-announced services, even when the Bonjour client and the offered service are located in different L3 IP subnets on the same network device.

To extend this, DNA Service for Bonjour scales this capability up to network-wide application that runs on DNA Center, and critically, provides policy-based access to Bonjour services across the entire enterprise network. One of the significant challenges that must be addressed is limiting advertisement of the many various services and devices, based on policy. For example, in a public school, it may be desirable for the teacher to have the ability to connect to an Apple TV device to display classroom content, while also ensuring that students attached to the same network do not have this capability.

Note Policy-based filtering of Bonjour mDNS advertisement is not supported by the native Bonjour protocol on its own.

However, by combining Cisco's Service Discovery Gateway and DNA Service for Bonjour (running on DNA Center), an administrator can enable scalable Bonjour services across their entire enterprise environment, with a powerful set of policy-based access controls to Bonjour-announced resources.

User Centric Platform Design

Overview

In the Catalyst 9000 family, the usability of the platforms has been a key design consideration. This is very important since the cost of operating the network is much higher than the costs of the initial purchase of the products. Catalyst 9000 adds usability improvements on the hardware side such as RFID, Blue Beacon, and Bluetooth Console, and software improvements such as WebUI and Flexible Templates.

RFID

Inventory management for a large number of switches is time-consuming. Labeling/tagging each switch and manually entering all the information into an inventory database can delay the provisioning of the device.

Keeping track of network devices and their components has now been made much easier with the Catalyst 9000. Cisco has front-facing UHF Radio Frequency Identification (RFID) technology to provide the latest auto-ID capabilities for asset management, location, and tracking. The RFID tag does not need to be visible to be read with a scanner as long as the tag is within 6 ft from the scanner.

The Catalyst 9000 family is integrated with passive Serialized Global Trade Item Number (SGTIN)-198 bit encoded RFID tags. On Catalyst 9400, every component, such as supervisor engines, line cards, power supplies and fan tray, has RFID tags. Since Catalyst 9000 RFID tags are passive tags, an additional power source is not required. The process is powered by the signal from the RFID reader and multiple tags can be read simultaneously.

The RFID tags contain the required device information for inventory management, including serial number, Product ID and manufacturing date, etc. The Cisco RFIDs provide a user partition, where the inventory managers can store their own custom data with password-protection. Cisco IOS XE cannot read or change any information in the RFID tag, which is only manageable by an RFID system.

Please refer to the *RFID white paper* on Cisco.com for more details.

Blue Beacon

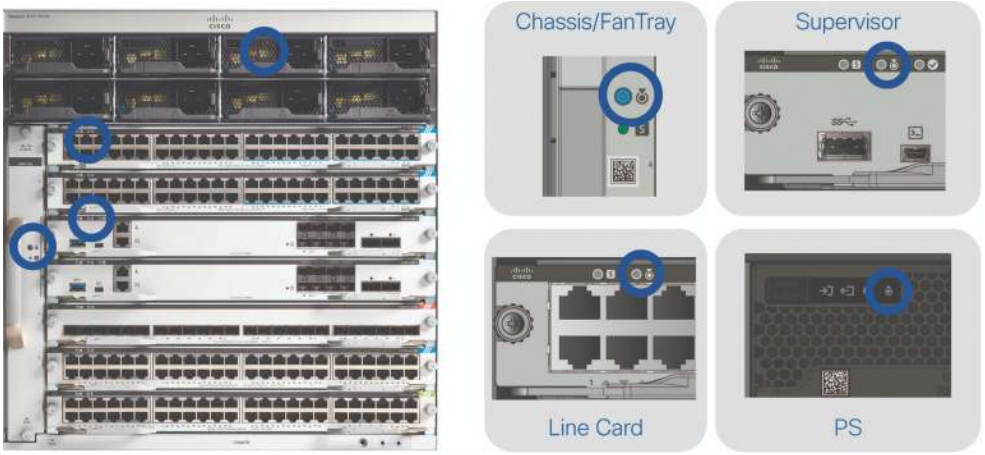
When troubleshooting, configuring, or moving equipment in a large enterprise, it might be difficult to locate a particular device. To help identify the device, Cisco has placed a blue beacon LED on the Catalyst 9000 family. This blue beacon can be turned on and off, either via Command-Line Interface (CLI) or manually via a button on the device. Every time the blue beacon is turned on/off, the device will generate a Syslog message.

On Catalyst 9300 series, blue beacon LEDs are located on both front and back. When Catalyst 9300 switches are stacked, the blue beacon for every member can be managed individually.

On Catalyst 9500 series, there is only one blue beacon LED, located on the front of the chassis.

On Catalyst 9400 series, the fan tray, supervisors, line-card modules, and power supplies each have their own addressable blue beacon.

DIAGRAM Blue Beacon on Catalyst 9400



Bluetooth Console

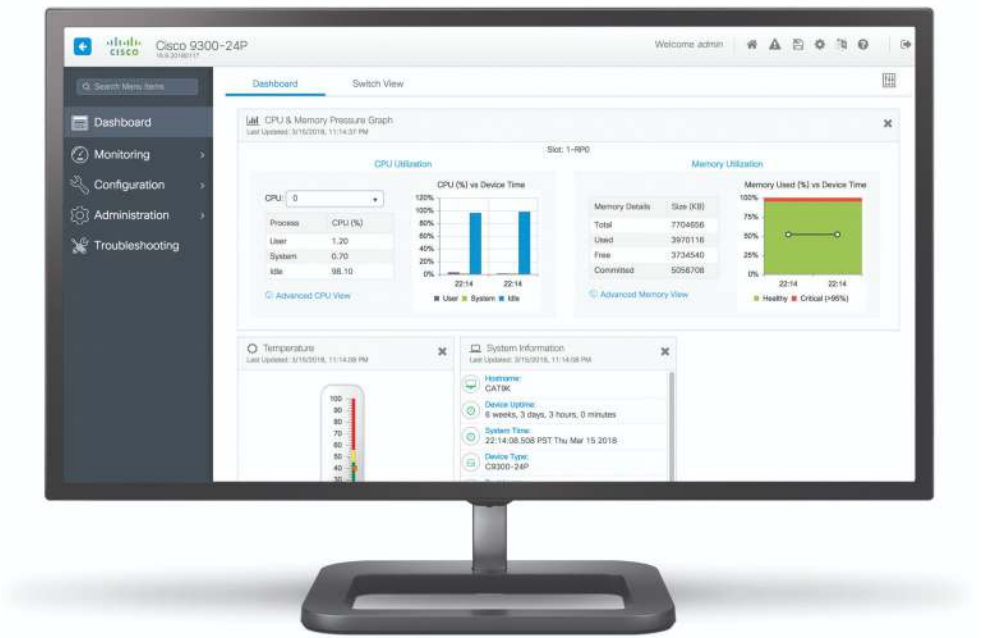
Network administrators often use a console cable for onsite configuration. However, a console cable has distance limitations and is not convenient to use.

The Catalyst 9000 series has introduced optional Bluetooth console functionality to provide wireless console access. A Bluetooth dongle needs to be connected via the front panel USB port to enable a wireless interface which has the same capabilities as the wired Ethernet management interface. Bluetooth can be used for CLI access via SSH or Telnet, configuration via the WebUI, or to transfer images or config files.

WebUI

WebUI is a Graphical User Interface device-management tool that provides the ability to configure and monitor a device. The WebUI tool is embedded in the system image at any license level. To enable WebUI on a device, configure the HTTP/S server and local or external server authentication.

DIAGRAM Catalyst 9000 WebUI



Flexible Templates

Flexible templates give Catalyst 9000 switches an option to be positioned in different roles in the network design. The UADP ASIC has capabilities to optimize its hardware table for specific network roles in the network. For example, it is possible to reduce the entries for Security ACLs and use it for QoS ACLs instead. Flexible templates are represented in Cisco IOS XE as Switching Database Manager (SDM) templates.

Note Catalyst 9300 supports only one default template optimized for the access layer.

Shown below are the SDM templates available on Catalyst 9400 using Sup-1XL and Catalyst 9500 using UADP 2.0 XL.

TABLE SDM Templates available per platform

	Access / Distribution	Core	SDA	NAT
Purpose	Maximizes system resources for MAC and security	Maximizes system resources for unicast and multicast routing	Maximizes system resources to support fabric deployment	Maximizes system resources for Layer 3 and NAT to support collapsed core WAN deployments
Longest Prefix Match (v4 / v6)	64,000 / 32,000	64,000 / 32,000	64,000 / 32,000	64,000 / 32,000
Host routes (v4 / v6)	Up to 112,000 / Up to 56,000	Up to 96,000 / Up to 48,000	Up to 144,000 / Up to 72,000	Up to 112,000 / Up to 56,000
Multicast (v4 / v6)	16,000 / 8,000	32,000 / 16,000	16,000 / 8,000	32,000 /16,000
MAC address	64,000	16,000	16,000	16,000
QoS ACL	18,000	18,000	18,000	3,000
PBR / NAT	2,000	2,000	2,000	16,000

Shown below are the SDM templates available on Catalyst 9500 using UADP 3.0.

TABLE SDM Templates available per platform

	Distribution	Core	SDA	NAT
Purpose	Maximizes system resources for MAC and security	Maximizes system resources for unicast and multicast routing	Maximizes system resources to support fabric deployment	Maximizes system resources for Layer 3 and NAT to support collapsed core WAN deployments
Longest Prefix Match / Host routes (IPv4 and IPv6)	114,000	212,000	212,000	212,000
Multicast (IPv4 and IPv6)	16,000	32,000	32,000	32,000
MAC address	80,000	32,000	32,000	32,000
QoS ACL	16,000	16,000	16,000	8,000
Security ACL	27,000	27,000	27,000	20,000
PBR / NAT	3,000	3,000	2,000	15,500

Programmability and Automation

Overview

The following factors influence decision making from a configuration and operational point of view:

- Network infrastructures are growing rapidly in terms of the number of devices and applications.
- There is a need for more rapid innovation.
- There is a requirement to reduce OPEX and to increase productivity.
- There can be a lack of confidence that changes will be successful, usually due to insufficient testing.
- There are too many manual processes.

All these factors lead to a growing need for automation at every level, from device provisioning to fully automated configuration, management, monitoring, and troubleshooting of network devices and network infrastructures.

Programmability is a very loosely-defined term that arrived when Software Defined Networking (SDN) was introduced several years ago. In this book, the term programmability, and specifically network device programmability, is defined as the set of features provided by the network device Operating System to enable automation.

Cisco Solution vs Do-It-Yourself (DIY)

The programmability features of Cisco IOS XE are very flexible and can be used in three main areas:

- **Cisco Solutions**: as outlined elsewhere, Cisco DNA Center and SD-Access provide a turnkey solution to automate and assure an entire campus network of wired and wireless devices. Cisco itself makes extensive use of programmability in DNA Center.

- **3rd-party integrations**: 3rd-party software vendors can build their own network management tools and systems using the available open data models, APIs, and tools without direct interaction with Cisco.
- **Do-It-Yourself (DIY)**: customers or partners can directly access the network device to build their own custom solution to automate every phase of device lifecycle.

Device Provisioning

Cisco IOS XE provides several options for automatic, accurate, consistent, repeatable provisioning process at a lower operating cost, and in a shorter deployment time than a traditional manual process:

- Preboot Execution Environment (PXE).
- Zero Touch Provisioning (ZTP).
- Cisco Network Plug and Play (PnP).

Preboot Execution Environment

Preboot Execution Environment (PXE) is a very common provisioning process used by system administrators to provision servers based on standard protocols such as BOOTP, DHCP, and TFTP. When the device boots up, instead of using the pre-loaded image, it sends a DHCP request to look for a PXE server. The PXE server will send an image to the device and the device boots up using the image just downloaded.

Cisco IOS XE provides PXE based on iPXE. iPXE is an Open Source version of the PXE created to support additional protocols such as HTTP. It can run on both wired and wireless connections. The PXE process is frequently described as Network Boot.

Zero Touch Provisioning

When a device that supports Zero Touch Provisioning (ZTP) boots up and fails to find the startup configuration (during day zero install or after erasing the configuration and rebooting), the device enters the ZTP mode. The device locates a DHCP server, which provides the following details:

- IP address and Gateway.
- DNS server.

- IP address or URL of a TFTP or HTTP server (using DCHP option 150).
- Python script name (using DCHP option 67).

The device bootstraps itself using the IP address provided and enables Guestshell. The device then downloads the Python script from the TFTP or HTTP server, to configure the device. Guestshell provides the environment for the Python script to run.

The provisioning logic implemented in the downloaded Python script is flexible and allows partial or full configuration of the devices in one or several phases, as well as image upgrade, patch management etc. Using this option, you can roll out hundreds of switches which are fully customizable, without any manual configuration.

Cisco Network Plug and Play

Cisco Network Plug and Play (PnP) is the device provisioning turn-key solution integrated with many Cisco products and solutions such as APIC-EM and DNA Center.

The Cisco PnP provides a simple, secure, unified, and integrated solution for enterprise network customers to ease new branch or campus device rollouts, or for provisioning updates to an existing network Cisco routers, switches, and wireless devices with a zero-touch deployment experience.

The Cisco PnP architecture and boot sequence are similar to ZTP, but instead of downloading a customized Python script, the configuration, image upgrades, and patches are managed using the provided GUI in APIC-EM or DNA Center. The entire process is fully automated.

Open Programmable Device APIs

Over the years, customers have been trying to build levels of automation based on the Command Line Interface (CLI) using different scripting languages such as Perl and TCL. CLI scripting has several limitations, including lack of transaction management, no structured error management, and an ever-changing structure and syntax of commands that make scripts fragile and costly to maintain. These are all side-effects of the fact that CLIs are designed to be used by humans and are not a programmatic interface.

SNMP was designed to overcome the drawbacks of automation based on CLI and was meant to be used for both configuration and operations, but in practice, it is used mostly for device monitoring only. This is because the lack of a defined discovery process makes it hard to find the correct MIB modules. There are also a lack of MIBs with write permissions, limitations inherent in the use of the UDP protocol, and no useful standard security and commit mechanisms.

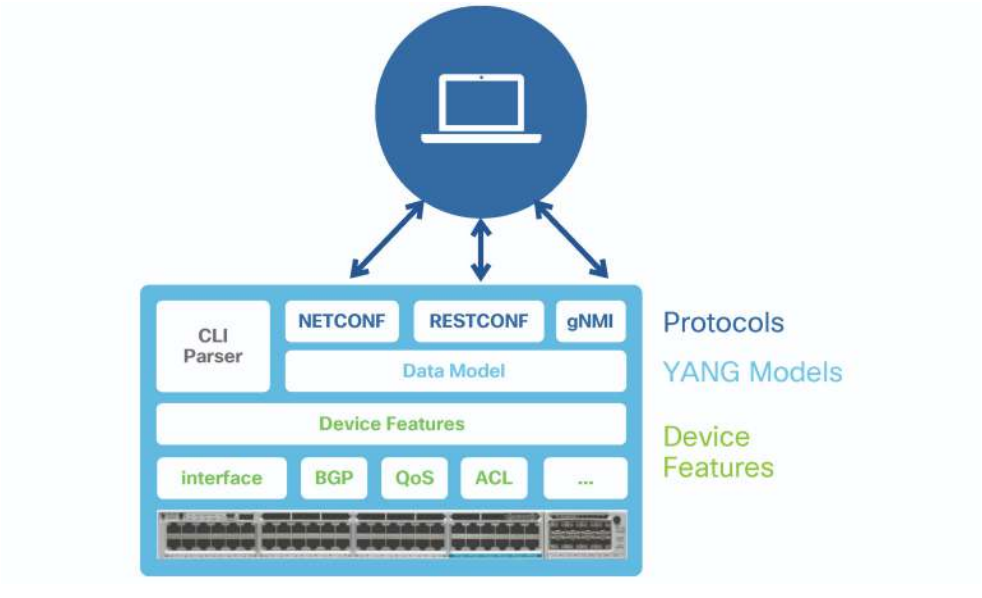
In computer programming, an Application Programming Interface (API) is a set of subroutine definitions, protocols, and tools for building application software. A good API makes it easier to develop a computer program by providing all the building blocks, which are then put together by the programmer. Software vendors and customers are looking for APIs which can provide key features such as structured data, error handling, and a variety of API management tools.

Network vendors such as Cisco have tried to introduce many different APIs over the years, from the very first NETCONF implementation in 2006, to WSMA, onePK, NX-API, and others. These APIs have not been widely adopted by customers for many different reasons, principally that they are vendor-proprietary.

These are the key reasons why Cisco has decided to build new APIs based on open standards like NETCONF and YANG data models for all of the main Operating Systems - Cisco IOS XE, Cisco NX-OS and Cisco IOS XR.

The diagram below illustrates an open standard API, with a common YANG data model infrastructure, built on top of the device-level features, to define both the device configuration and operational state. Different protocols such as NETCONF, RESTCONF, and gNMI can be used to interface with external automation software toolchains.

DIAGRAM Open Programmable Device APIs



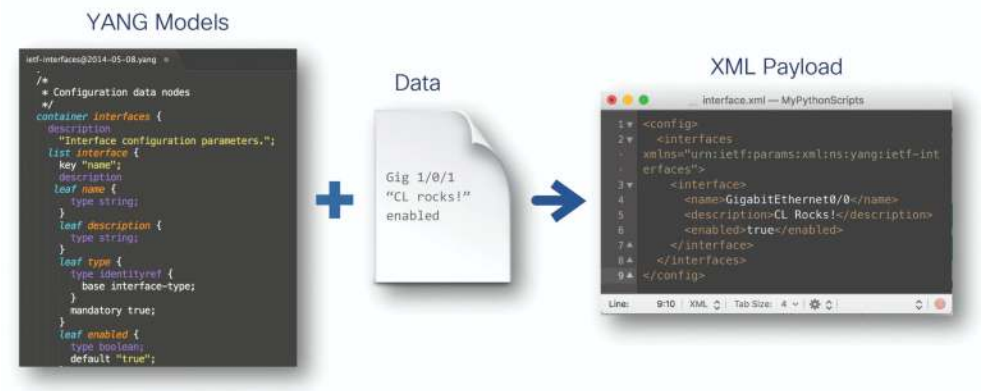
Data Models

A data model is one of the most important components of open programmable APIs. It precisely defines the data structure, syntax and semantics of a given feature and is meant to solve the issue of unstructured data provided by CLIs.

YANG

Yet Another Next-Generation (YANG) is a data modeling language developed by the IETF to enable the reuse of data models across equipment from different network vendors. It is widely used by network operators to automate the configuration and monitoring of network devices. YANG is defined in RFC 6020.

DIAGRAM YANG Models Example



As shown above, YANG data models can be considered as templates. YANG models need actual data in order to build operations that can be exchanged with a network device, in order to retrieve or change the device configuration or state.

Data Model Types

Note The IETF standards make a distinction between configuration and operational data models.

A configuration data model is the set of writable data required to transform a system from its initial default state into its current state. Essentially a configuration data model instructs the device to do certain things and can be easily mapped to the running configuration of a Cisco IOS XE device.

An operational data model is the set of read-only data status information and statistics on a device. The operational data model is comprised of what the device is actually doing and is mapped to the information traditionally provided by show commands.

Both configuration and operational data models can be further classified as either native or open data models.

Native data models are defined by networking vendors such as Cisco and are specific to an Operating System. For example, there are native models for Cisco IOS XE and native models for Cisco NX-OS.

Open models are defined by standards bodies such as IEEE and IETF or by working groups such as OpenConfig. A big advantage of using open models is that they are common across different Operating Systems and vendors, providing a consistent way to manage all devices.

Open models are a subset of native models. That is because the process of defining an open model is typically slower than the same process for native models. The main reason for this is that in order to define an open model, many parties need to agree upon the content and model structure, whereas vendors have flexibility to define native models as they wish.

The approach followed by Cisco IOS XE is to provide a comprehensive list of IOS XE native data models and also provide open models as they become available. Open models in Cisco IOS XE are mapped to corresponding IOS XE native models. Therefore Cisco

customers are free to decide whether to use the open or the native model to manage a given feature.

Among the various open data models in existence, OpenConfig data models have recently become popular. OpenConfig is an informal working group of big network operators, led by Google, sharing the same goals of automating their networks.

Cisco publishes YANG data models in a common *GitHub Repository* which is updated every time a new release becomes available.

In addition to downloading the data models from the GitHub repository, Cisco IOS XE on Catalyst 9000 allows you to download them directly from the device as well. This feature is very useful, especially if the data models have been updated on the device using the patching process described in Chapter 6 High Availability.

Tools, SDKs, Resources

Building API operations by hand is possible, but not very practical. This is why many tools have been created to navigate through the data models, validate them, and to build API operations.

These are some of the most common tools available today:

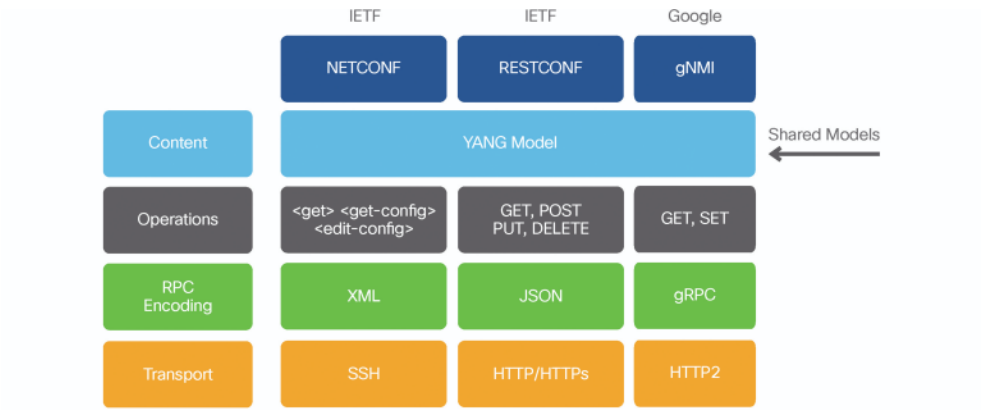
- **PYANG**: Python library to validate, navigate and automatically build documentation.
- ***YangExplorer***: open source tool to easily start exploring the model and automate small tasks.
- **YDK** (YANG Development Kit): software development kit that provides an abstraction layer of the API modeled in YANG.
- ***Yang Catalog***: a reference for all YANG modules available in the industry.

Device API Protocols

The various interface protocols supported by Cisco IOS XE on Catalyst 9000 devices share a common data model infrastructure. From a purely technical point of view, all of the protocols can be used to manage the same YANG data model, but of course, that is not an easy-to-maintain approach.

The diagram below shows the main differences between the various API protocols at each layer of the stack, starting with the Transport (SSH vs HTTP vs HTTP2), the different encoding formats used for the Remote Procedure Calls (RPCs), as well as different sets of operations.

DIAGRAM Comparison of Device API Protocols



Note The YANG data model infrastructure is common across all of the Cisco IOS XE API Protocols

Encoding Formats

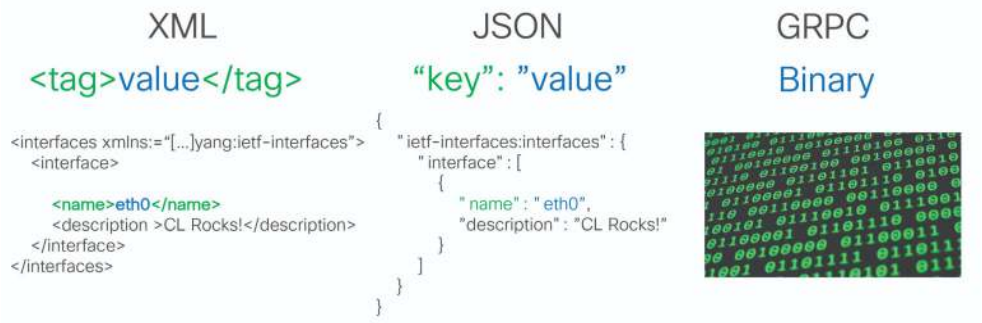
All of the API protocols use Remote Procedure Calls (RPCs) for communication. A client running on an external server sends an RPC message to the network device and the network device replies with an RPC-reply message. The data payload of both RPC and RPC-reply is encoded using a format defined by the protocol. The most common encoding formats are:

- **Extensible Markup Language (**XML) is a text-based, human-readable format where the information exchanged is rendered using tags. It is very similar to HTML but while HTML is used to build web pages, XML is used to describe data. NETCONF uses XML.

- **JavaScript Object Notation** (JSON) is an alternative to XML. It is also text-based and human-readable, but instead of using tags, it uses <key:value> pairs which make the data easier to parse. That is the reason JSON is usually preferred over XML. RESTCONF supports both JSON and XML.

- **Google Remote Procedure Call** (gRPC) is the open source version of the encoding format defined by Google for their own applications.It is proven to be very efficient and scalable. It is a binary format, making it non-human readable. It relies on protocol buffers to encode on the client side and decode on the server side. While gNMI supports a variety of encoding formats, gRPC is the suggested and most used format.

DIAGRAM RPC Encoding Formats



NETCONF

NETCONF is a popular network configuration protocol defined by the IETF to help network operators to manage their networks. While most people think it is a new protocol, it has actually been around for a long time and Cisco started supporting NETCONF in IOS more than a decade ago.

Since then, NETCONF has evolved considerably with the definition of a new set of IETF Requests For Comments (RFC). RFC-6241 defines the basics of NETCONF, and introduced the concept of optional extensions. Examples of NETCONF extensions are the YANG data model language (RFC-6020) and Notifications (RFC-5277).

The NETCONF stack includes SSH transport, messages in the form of generic RPCs and encoded using the XML format.

The main NETCONF operations are:

- **`<get>`**: to retrieve running configuration and device state information. Can be easily mapped to an IOS XE "show" command.
- **`<get-config>`**: to retrieve all or part of a specified configuration. Similar to an IOS XE "show run" command.

- **<edit-config>**: to change all or part of a configuration. Has the same behavior as using Cisco IOS XE config terminal mode.

RESTCONF

RESTCONF is a network configuration protocol (like NETCONF) defined by the IETF in the RFC-8040.

The RESTCONF stack includes HTTP/HTTPS transport, messages in the form of generic RPCs (like NETCONF) and encoded using either XML or the JSON format.

The RESTCONF stack includes HTTP/HTTPS transport, messages in the form of generic RPCs (like NETCONF) and encoded using either XML or the JSON format.

The RESTCONF operations are defined by the REST framework and can be easily mapped to the corresponding operations in NETCONF:

- **GET**: to retrieve a resource. Similar to a NETCONF <get-config>, <get>.
- **POST**: to create a new resource. Similar to an <edit-config> with operation="create" option.
- **PUT**: to create or modify a resource. Similar to an <edit-config> with operation="create/replace".
- **DELETE**: to delete a resource. Similar to an <edit-config> with operation="delete".

The target resource in all operations is defined using a standard URL, commonly used to reference web pages on the Internet.

gNMI

The **Google Network Management Interface (**gNMI) is an alternative protocol to NETCONF and RESTCONF for accessing YANG models.

While both NETCONF and RESTCONF are standards-based protocols defined by the IETF, gNMI is an open source network management protocol developed by Google.

gNMI operations are:

- **CAP**: sent to the network device on first connect to discover device capabilities.
- **GET**: to retrieve the device configuration or state. Includes attributes such as Prefix, Paths and Data Type (Oper, Config, All).
- **SET**: to change the device configuration. Has the same attributes as GET except for Type (Update, Replace, Delete).

Model-Driven Telemetry

Network monitoring challenges

Automation solutions based on CLI and SNMP have, over time, proven to be incomplete, inefficient, and hard to scale and maintain. New requirements in terms of speed, scale, fault isolation, forensic analysis, and near real-time data availability, are making legacy monitoring solutions insufficient for most organizations.

The new paradigm

Model-Driven Telemetry (MDT) has been designed to overcome the drawbacks and shortcomings of legacy monitoring solutions described above. MDT provides structured data in the form of the same YANG data models defined in Chapter 12.3 Open Programmable Device APIs, in a scalable and efficient manner with a low impact on the network device.

MDT provides a new approach for network monitoring in which operational data is streamed continuously from network devices to an external collector and provides real-time access to statistics for monitoring the network. The data can be streamed on-change, or at a periodic time interval.

Compared to an SNMP get request, which is a **pull** mechanism, MDT uses a **push** mechanism.

The Cisco IOS XE MDT implementation on Catalyst 9000 is based on the standard IETF PUB/SUB (publication/subscription).

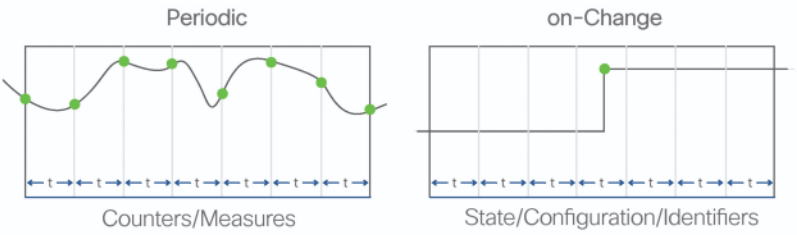
To receive streaming data from the device, an external collector must set up a subscription to a data set in a YANG model. This can be created via an RPC subscription message. Once enabled on the device, it continuously sends data for the lifetime of that

subscription. Publication is by streaming data to a destination using periodic or on-change notifications.

For periodic notifications, the data is streamed out to the destination at the configured interval. The period is the time expressed in 1/100 of seconds between updates. This is a similar concept to SNMP "get" requests, but with all the benefits of the MDT push-based model at a much higher rate. Data that is expected to change at a high frequency is best consumed via periodic publications, such as interface counters and CPU utilization.

For on-change notifications, the data is published only when a change occurs. On-change publication is best utilized for data that is not expected to change at a high frequency and where it is important to know immediately when a change has occurred.

DIAGRAM Periodic vs. On-Change Subscriptions



A similar approach is used by Cisco DNA Center Assurance to collect streaming data from MDT-capable devices and correlate with other data sources.

Developing a DIY data collector is a complex task, but customers willing to build their own collector can make use of Open Source tools such as TCollector, collectd and the Elastic Stack (commonly know as the ELK stack) as well as messaging brokers such as Apache Kafka, ActiveMQ and RabbitMQ.

Scripting

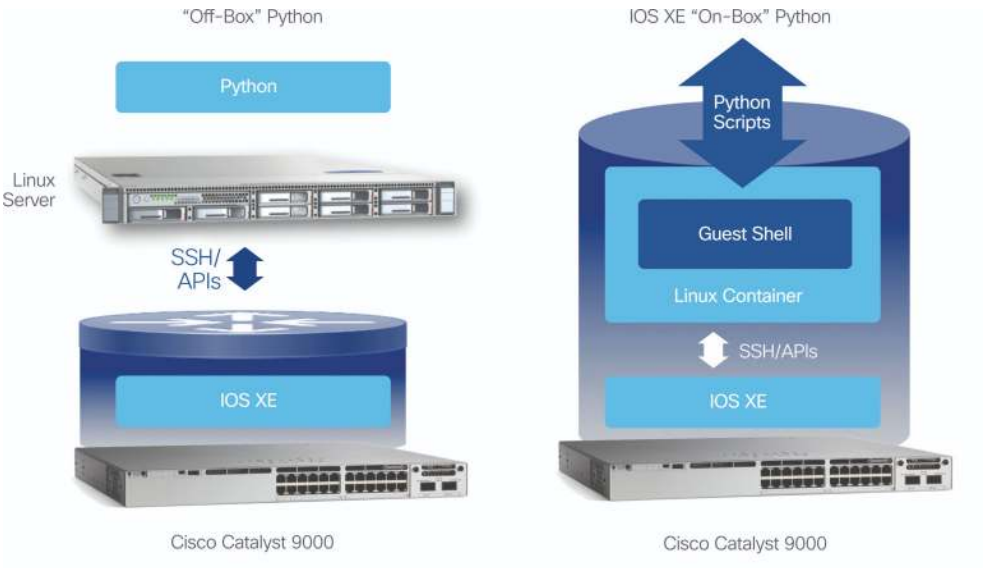
Scripts have been used for ages to quickly and easily automate small tasks. Over the years, some scripting languages have become more popular than others, such as Perl, TCL, and JavaScript. In recent years, Python became the most popular scripting language. One of the main reasons for its ever-growing popularity is that Python is easy to get started with. It provides an interactive shell, allowing a quick way to execute scripts line by line and is more human-readable than most of the other scripting languages.

Maybe the most important reason is the very extensive list of publicly available Python libraries for most of the Operating Systems and ease of installing, updating and uninstalling the libraries using the Python package manager (pip). There is also a Python virtual environment, that helps with the management of application library dependencies.

Off-box vs On-box Python

Python scripts can be used to automate a Catalyst 9000 running Cisco IOS XE in two different ways, commonly named off-box and on-box scripting.

DIAGRAM Cisco IOS XE Off-Box and On-Box Python



- **With Off-box**, the Python script is executed in an external server and it connects to the Cisco IOS XE device using an SSH connection for CLI based automation or via the open APIs (NETCONF/RESTCONF/gNMI).
- **With On-box**, the Python script is executed inside the Catalyst 9000 in a built-in Linux container named Guestshell. From the Guestshell environment, Python scripts can access the underlying Cisco IOS XE using the same mechanism used by off-box Python scripting.

By providing off-box and on-box scripting options, the Catalyst 9000 is a perfect fit for centralized and distributed Python-based application architectures.

Configuration Management Tools

Configuration management tools automate systems and applications in a consistent fashion at scale. Such tools have been used by system administrators for more than a decade. Lately, an increasing number of customers are using (or considering) configuration management tools to automate their networks as well. Configuration management tools enable a variety of advantages:

- A consistent approach across different vendors and Operating Systems.
- Easy integration with version control systems.
- A simple way to collect hardware and software device facts
- Provides an intent-based configuration approach.
- No changes are made if the system, application, or device is already in the desired state.

In the networking space, configuration management tools initially were used to automate data center networks only, but now enterprise networks are starting using them as well. Among others, Ansible and Puppet are the most popular tools. Configuration management tools can be classified into two groups, based on their architecture: agent-based or agent-less. Agent-based architectures require a software agent to be installed on the managed device, whereas agent-less architectures do not need one. This is a key difference, particularly in the enterprise space where customers tend to be more conservative about installing 3rd-party software on devices.

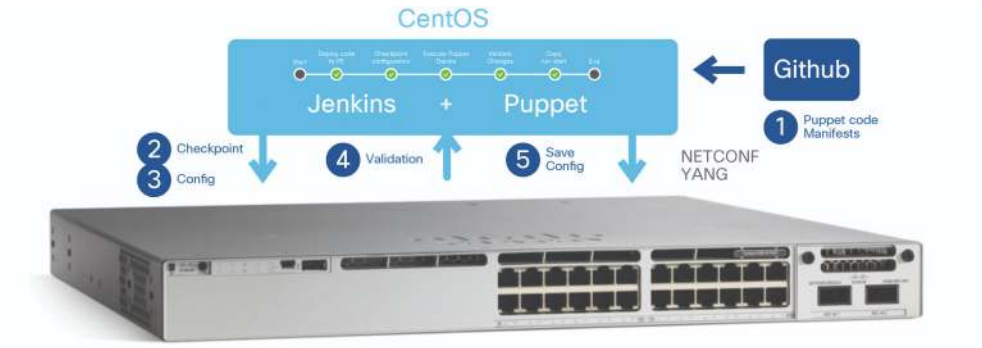
Ansible is based on an agent-less architecture which has been a key factor in its success. Ansible already supports Cisco devices, such as Catalyst 9000 switches running Cisco IOS XE and provides intent-based modules to configure interfaces, VLANs, VRFs, users, etc as well as CLI and NETCONF-based configurations.

Puppet traditionally requires an agent on the managed device (agent-based architecture) but Puppet is currently developing an agent-less architecture specifically

for network devices. It will also support Cisco devices running Cisco IOS XE, and it will be based on the NETCONF APIs.

Configuration Management Tools are a key component of another emerging trend among many customers embracing the DevOps and NetOps culture: Continuous Integration and Continuous Deployment or CI/CD.

DIAGRAM Cisco Catalyst 9000 CI/CD with Puppet

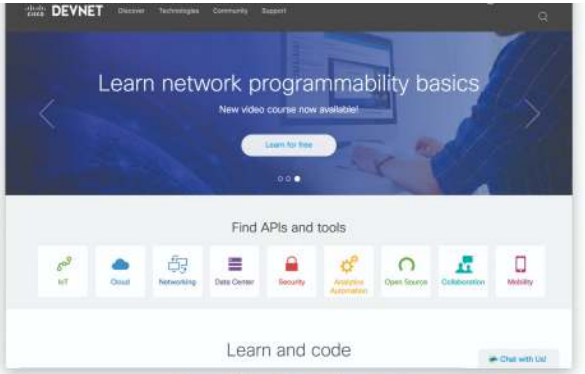


CI/CD is a new configuration management practice created to reduce the time needed to deploy new changes in production and to increase the confidence that changes will be successful. It enables this through an automated testing process and integration with source control systems for an easy rollback in case failures do occur. Customers implementing CI/CD frameworks have experienced lower development cycles, faster pace of innovation, and a lower total IT cost.

Cisco DevNet

Cisco DevNet is an initiative created to allow customers and partners to easily start learning the latest programmability technologies provided by the main Cisco Operating Systems. Cisco DevNet provides self-explanatory learning labs, video courses, device sandbox to test the programmability technologies on simulated devices and/or real hardware, along with API documentation, community support and more.

DIAGRAM Cisco DevNet



Cisco DevNet is completely free - customers and partners just need to sign up and start learning.

Application Hosting

Application Hosting Operation

Applications are used in enterprise networks for a variety of business relevant use cases. Examples of enterprise applications include administrative tools such as performance monitors and protocol analyzers and security toolsets such as intrusion detection services. Traditionally, such applications would operate on an external physical or virtual server.

Cisco built the Cisco Application Framework (CAF) to manage containerized applications running on any network device. CAF is also known as Cisco IOx. Originally created to host Internet of Things (IoT) applications, the Catalyst 9000 leverages IOx for enterprise applications within a campus environment.

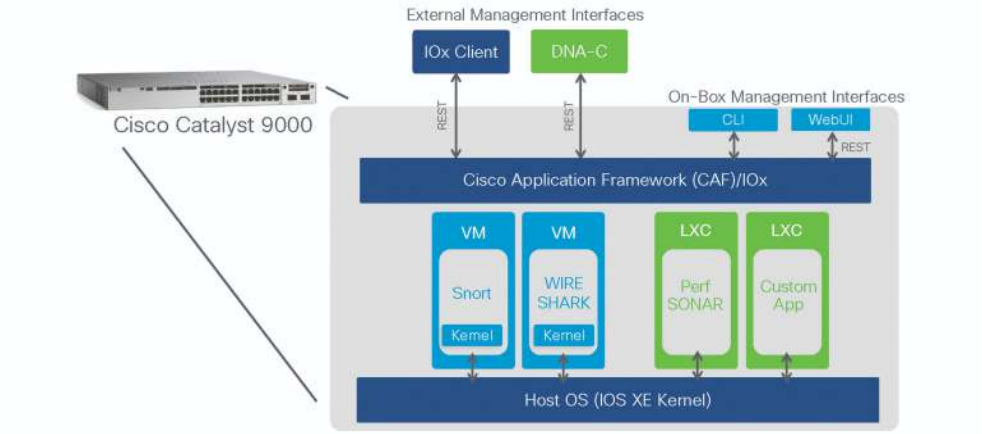
Cisco IOx on a Catalyst 9000 supports applications containerized in KVM-based virtual machines and LXC Linux containers. While native Docker containers are not yet supported, Docker tools can be used to easily build IOx applications in LXC format. Cisco IOx empowers operators to deploy any containerized application on their network devices.

Though Cisco will periodically publish certain services for application hosting, Cisco encourages and supports the deployment of any KVM-based VM or Linux LXC that fits within the IOx framework.

The Catalyst 9000 provides several options to manage hosted applications:

- **Cisco DNA Center**: Cisco's SDN controller used to run all LAN, WAN and WLAN enterprise devices.
- **IOx Client**: a Python-based tool capable of building and managing IOx applications.
- **WebUI:** the Cisco IOS XE GUI to configure and monitor the device, support IOx applications as well.
- **Command Line Interface**: a set of Cisco IOS XE console commands for managing IOx and IOx applications.

DIAGRAM Catalyst 9000 Application Hosting Framework



Anatomy of an IOx Application

An IOx application is packaged in a standard Linux tar archive format containing several files, including the application descriptor, one or more disk images, and, optionally, certificate files and other auxiliary files. The application descriptor is a file, written in YAML format, that includes:

- Application information: name, description, version, and author.
- Application type: LXC or VM.
- Hardware resources required: CPU, memory and storage.
- A list of virtual network interfaces used by the application.
- The disk image files used to load the application itself.
- Startup tasks or pre-execution scripts.

Step-by-step instructions on how to build *IOx applications* are available on the Cisco DevNet website.

Hardware Resources

Cisco IOS XE running on the Catalyst 9000 reserves dedicated memory and CPU resources for application hosting. By reserving memory and CPU resources, the switch provides a separate execution space for user applications. It protects the switch's IOS XE run-time processes ensuring both its integrity and performance.

Applications must reside in one of the external SSD storage (USB or M2 SATA) options provided by the Catalyst 9000. Applications have no access to the internal device flash storage, which is reserved for IOS XE for integrity reasons.

Note The external SSD storage is shared (not reserved) between Cisco IOS XE and hosted applications.

TABLE Catalyst 9000 Application Hosting Resources

Platform	Memory (GB)	CPU (cores)	USB Storage (GB)		M2 SATA Storage (GB)
			USB 2.0 Front	USB 3.0 Back	
Catalyst 9300	2	1 x 1.8GHz	16	120	N/A
Catalyst 9400	8	1 x 2.4GHz	16	N/A	960
Catalyst 9500 (with UADP 2.0)	8	1 x 2.4GHz	16	120	N/A
Catalyst 9500 (with UADP 3.0)	8	1 x 2.4GHz	N/A	N/A	960

Campus Network Design

Overview

Network design is important because individual devices must work together cohesively to optimize a network. While each platform has unique capabilities, and many capabilities are similar, the way various platforms are combined together can result in optimal or sub-optimal network behavior. It is therefore important to choose the best platform for a specific purpose. Cisco provides several different Catalyst 9000 series models to address a range of needs.

A campus network focuses mainly on how humans and their devices communicate with each other, and with services in the data center, private or public cloud services, and the Internet. Campus networks also tend to be geographically diverse, with a variety of unique requirements. The number and types of users and devices, as well as their geographic diversity, influence optimal network design.

└the bottom line
Select the “Right Platform” for the “Right Job”.

Cisco has 30 years of experience designing campus environments. Over that period, Cisco has developed and rigorously tested various designs from which five distinct campus design models have emerged:

- **Multi-Layer** - a multi-tier design that uses routing between core and distribution, with a switching access.
- **Collapsed Core** - a two-tier design for small sites, with a routed core and a switched access.
- **Routed Access** - a multi-tier design, which utilizes routing end-to-end through all layers.
- **Campus MPLS** - a multi-tier design, using MPLS on top of a routed domain to provide segmented services.
- **Software-Defined Access** - a multi-tier network fabric based on the designs above, automated and assured by a contoller.

While each design model has evolved to address a specific set of requirements, all share a common set of characteristics. The common characteristics of campus designs include:

- **Hierarchy** - structured design which allocates defined roles to each layer, and follows a structured cabling plant.
- **Redundancy** - including physical redundancy (links, chassis, power), data plane redundancy, control plane redundancy.
- **Bandwidth** - sufficient capacity at each network tier to support the aggregate system load.
- **Port Density** - sufficient interfaces at each network tier to support all its connected neighbors.
- **Scalability** - sufficient hardware and software resources to support all interconnections and network services.
- **Wireless LAN** - sufficient wired infrastructure and features to support mobility across the campus.

While each of the Catalyst 9000 platforms is designed to address one or more design model, all share a common set of capabilities:

- **Catalyst 9500 series** - Fixed form factor, with 1/10/25G SFP and 40/100G QSFP onboard and module ports.
  - Built with sufficient ASIC bandwidth and table scale to support medium to large-sized campus cores and distribution designs.
- **Catalyst 9400 series** - Modular form factor, with 10/100/1000/mGig RJ45, 1/10G SFP modules and 40G QSFP Supervisor ports.
  - Built for high-density user access with sufficient ASIC bandwidth and table scale to support large campus access designs or medium-sized distribution designs.

- **Catalyst 9300 series** - Fixed, stackable form factor, with 10/100/1000/mGig RJ45 and 1/10/25G SFP or 40G QSFP module ports.
  - Built for high-density user access with sufficient ASIC bandwidth and table scalability to support large campus access designs.

└the bottom line
Catalyst 9000 switches are the most flexible, highest scaling, and resilient platforms for the new era of networking

While most of this book focuses on the specific details of individual Catalyst 9000 platforms, the following chapters are dedicated to how an end-to-end Catalyst 9000 series solution can simplify and optimize overall network design.

Physical Infrastructure

This section discusses critical design considerations for the physical infrastructure, in order to provide sufficient bandwidth and port density for an optimal campus network.

The Need for Speed

It is expected that IP traffic will grow at a Compound Annual Growth Rate (CAGR) of up to 24% from 2016 to 2021. As discussed earlier, increasing deployments of high-speed wireless technologies, the proliferation of IoT devices and sensors, and powerful new endpoints that handle high volume data, are forcing network engineers to redesign their physical infrastructure. Client access speeds are increasing from 1 Gigabit per second (Gbps) to 2.5 Gbps, 5 Gbps and potentially 10 Gbps. These new client access speed requirements demand network infrastructure speeds greater than 10G.

However, moving to higher speeds can bring additional challenges. For instance, changing from existing 10G Small Form factor Pluggable (SFP) ports to 40G Quad Small Form factor Pluggable (QSFP) carries significant costs (transceivers and cables). While an alternative for link-aggregation is one approach, achieving effective load sharing with multiple 10G links depends on the hashing algorithm and traffic patterns.

Backwards compatibility for easier migration is another key consideration for new speed adoption in campus networks. It is beneficial to use a common transceiver form factor, such as SFP and QSFP. Modern 10G SFP+ and the new 25G SFP28 interfaces support backward compatibility with 100M and 1G SFP. Similarly, the new 100G QSFP28 interfaces support backward compatibility with 40G QSFP.

Note Cisco 25G SFP28 transceivers support dual-rate optics to operate at both 10G or 25G speed.

For example, upgrading access to distribution layer links from 10G to 25G is a straightforward migration. Simply remove the 10G SFP+ transceiver and install a new

25G SFP28 transceiver (using the same cable). If the remote side is still using a 10G SFP+ transceiver, the link will operate at 10G. Once a corresponding 25G SFP28 transceiver is installed on the remote side, the link will operate at 25G.

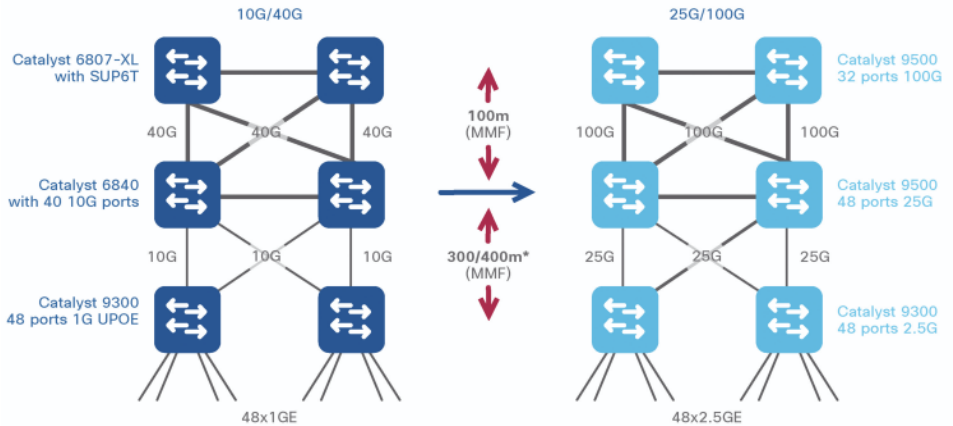
Use Case 1: Speed transition with similar Cable Distances

As access layer bandwidth increases from 10/100/1000M to 2.5G, campus backbones should transition from 10/40G speeds to 25/100G speeds. Customers need newer optics to support cabling distances similar to their existing environment.

the bottom line

The use of innovative Cisco SFP-10/25G-CSR-S transceivers supports traditional cable lengths of up to 300/400m over OM3/4 MMF (depending upon fiber quality) to achieve speeds of 25G.

DIAGRAM Speed transition with similar Cable Distance



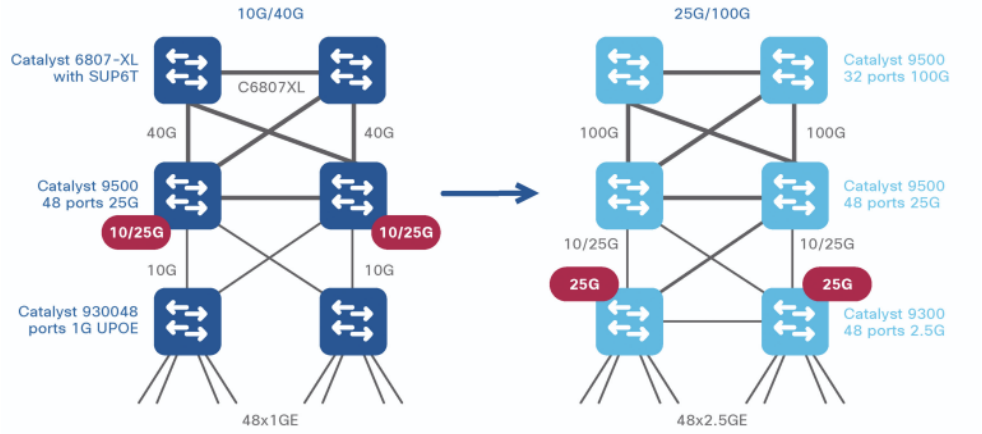
Use Case 2: Speed Migration with Dual-Rate Optics

The Cisco 25G portfolio provides backward compatibility to 10G with transceivers built with dual-rate optics. Cisco dual-rate optics will auto-negotiate to the highest speed supported. For example, if the remote device is only capable of 10Gbps, the two devices will operate at 10Gbps speed.

└ the bottom line

There is no requirement to upgrade the infrastructure to 25G. Transceivers can operate at 10G and be upgraded as part of regular refresh cycles.

DIAGRAM Speed transition with Dual-Rate Optics



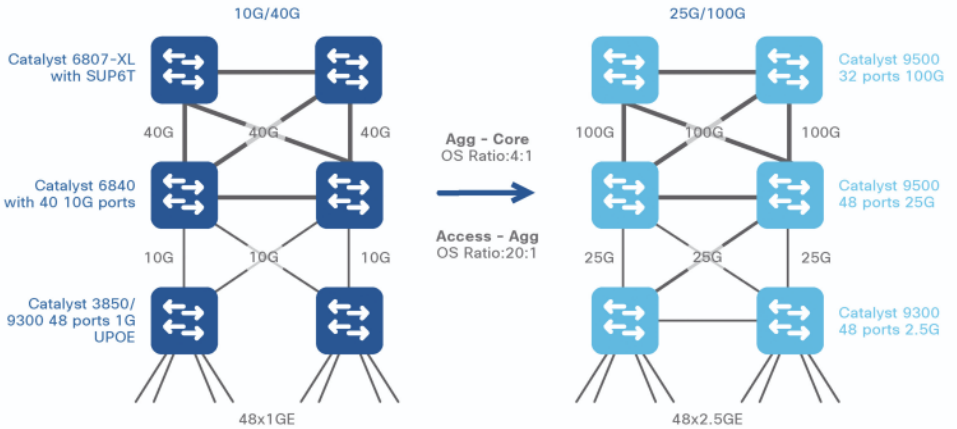
Use Case 3: Speed transition with similar Oversubscription Ratios

In the past, the rule-of-thumb design recommendation for oversubscription was ~20:1 for the access to distribution, ~4:1 for distribution to core, and ~2:1 or 1:1 for the core layer. As access layer bandwidth increases (e.g. 2.5G), there is a corresponding need to upgrade the inter-switch connections to preserve the recommended oversubscription ratios.

└ the bottom line

25G and 100G are natural successors to 10G and 40G, with at least 2.5X bandwidth increase.

DIAGRAM Speed transition with similar Oversubscription Ratios



Use Case 4: Comparing Higher speed vs. Load sharing

Speed upgrades have outpaced the refresh cycles for cabling. To satisfy growing bandwidth requirements, there are two basic approaches:

- more uplinks can be added (with link-aggregation or ECMP)
- replace existing cabling and transceivers to support higher speeds

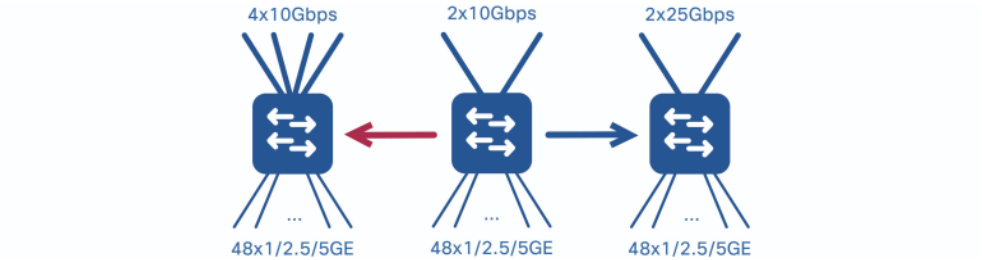
There are significant challenges and cost implications to both. Adding additional links incurs the complexity of achieving effective load-sharing, as well as some QoS

implication. While upgrading to higher link speeds does not require complex load-sharing or QoS challenges, it may require replacing transceivers and/or cabling.

└the bottom line

Catalyst 9000 platforms support a common set of port types and transceivers, with flexible port speeds. This greatly simplifies upgrading existing infrastructure to higher speeds.

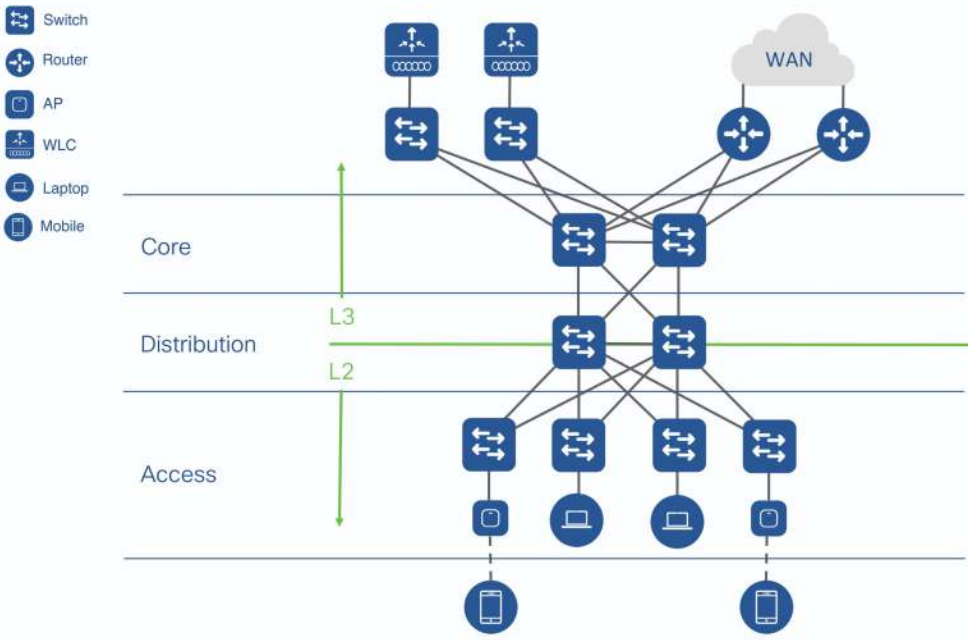
DIAGRAM Comparing Higher speed vs. Load sharing



Multi-Layer Campus

A Multi-Layer LAN deployment is the most commonly-deployed design model. It consists of three layers: core, distribution and access.

DIAGRAM Multi-Layer Campus Design



The core in a Multi-Layer campus design is based upon Layer 3 IP routing and functions as a high-speed interconnection point to other network layers (e.g. DC, WAN, Branch, etc). The distribution layer consists of IP routing upstream to the core and Layer 2 switching downstream to the access layer. Distribution blocks function as an aggregation point for client traffic traveling to the rest of the network; they serve as transition points between switched wiring closets into the routed core. The primary

role of the access layer is simply to connect end-points and switch their traffic into the rest of the network.

There are several main advantages to this design:

- It is a tried-and-true design implemented widely during its twenty-year history.
- The hierarchy of layers assigns specific roles and responsibilities to each connection block.
- It is scalable and modular. Blocks can be added or removed in any layer without major impact on the design of other layers.
- It defines a separation of duties to LAN switches that architects use to evaluate products and build network policies.
- It allows for spanning VLANs (Layer 2 domains) across multiple access wiring closets if needed, providing design flexibility.

There are some disadvantages, however, to this design. The distribution layer, in particular, introduces complexity because of its role as an L2/L3 interchange. L2 networks are considered flood domains for BUM traffic (broadcast, unknown unicast, multicast) and subject to loops during link failures and reconvergence. If not tuned correctly, L2 networks simply fail when BUM traffic consumes processing resources or a network loop blocks flows as participating switches reoptimize a path.

Tools such as Rapid Per-VLAN Spanning Tree (RPVST), Spanning Tree Protocol (STP) guards, IGMP snooping and storm-control are used to handle these situations, but administrators are required to enable and monitor these features. A similar situation exists when passing traffic across the access layer switched border to the routed network. Sub-optimal paths and asymmetric flows with consequent traffic flooding can occur in the multi-layer design if Spanning Tree switching mismatches IP routing (i.e. if the L2 and L3 topologies are incongruent). Administrators need to configure switches so that STP root switches are also the IP default gateway, using a First-Hop Routing Protocol (FHRP), even in the event of a switch failure.

Note Catalyst 9000 series switches continue to support multi-layer networks, and their unique features are optimized for roles of each layer, using resource templates.

The Catalyst 9500 10/40G models are based on UADP 2.0 and use the distribution SDM template, by default, to set the optimal ASIC table allocation for a distribution layer role. In addition, the Catalyst 9500 10/25/40/100G models are based on UADP 3.0, and employ the core SDM template by default, providing the optimal ASIC table allocation for a core switch.

└the bottom line
The Catalyst 9500 Series platforms are optimized for the core and distribution layers.

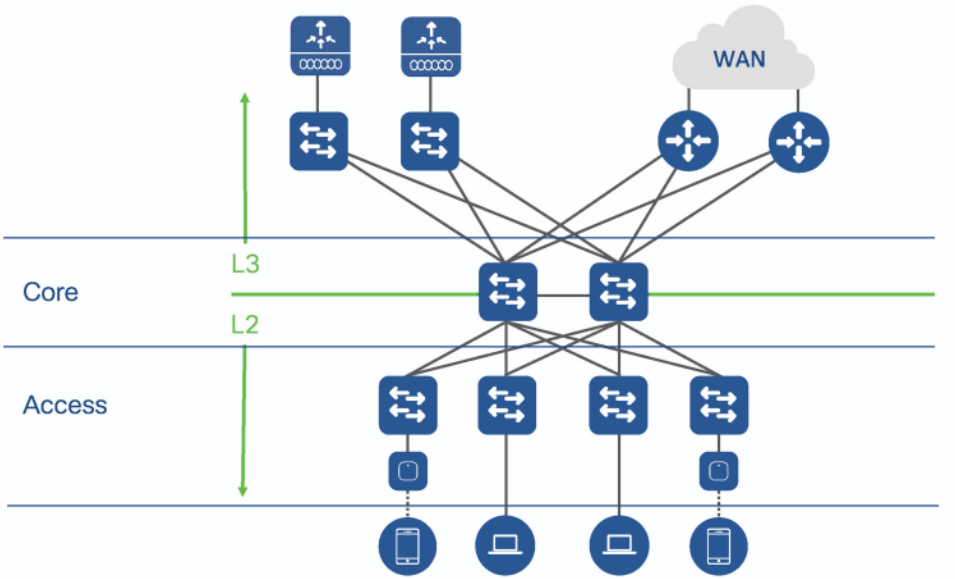
The Catalyst 9300 and 9400 series switches both employ UADP 2.0 forwarding engines, and implement the access SDM template to set the optimal ASIC table allocation for switched access role. The Catalyst 9300, having a stackable fixed-form factor is ideal for providing switch-level redundancy. The Catalyst 9400, by contrast, is a modular switch providing the highest levels of network availability by delivering supervisor, line-card and power redundancy in a chassis. Both switching lines provide capabilities to support small, medium and large wiring closets.

└the bottom line
The Catalyst 9300 and 9400 Series platforms are optimized for the access layer.

Collapsed Core

A collapsed core design is based on the same principles of a multi-layer LAN for in a small campus network, with the core and distribution layers collapsed into each other.

DIAGRAM 2-Tier Collapsed Topology



All of the advantages of a multi-layer design also apply to a collapsed core, which requires less network infrastructure and is, therefore, a right-sized and cost-effective solution for small sites.

The same drawbacks of the multi-layer also apply to this design. In this case, the distribution layer complexity described previously has added into the core.

The Catalyst 9500 series models use the UADP programmable ASIC in a fixed-form factor design. The difference between models is a matter of port speed and density, but in all cases they can use the distribution SDM template to optimize forwarding tables for use as a collapsed core switch.

└the bottom line
The Catalyst 9500 Series platforms are optimized for the collapsed core layer.

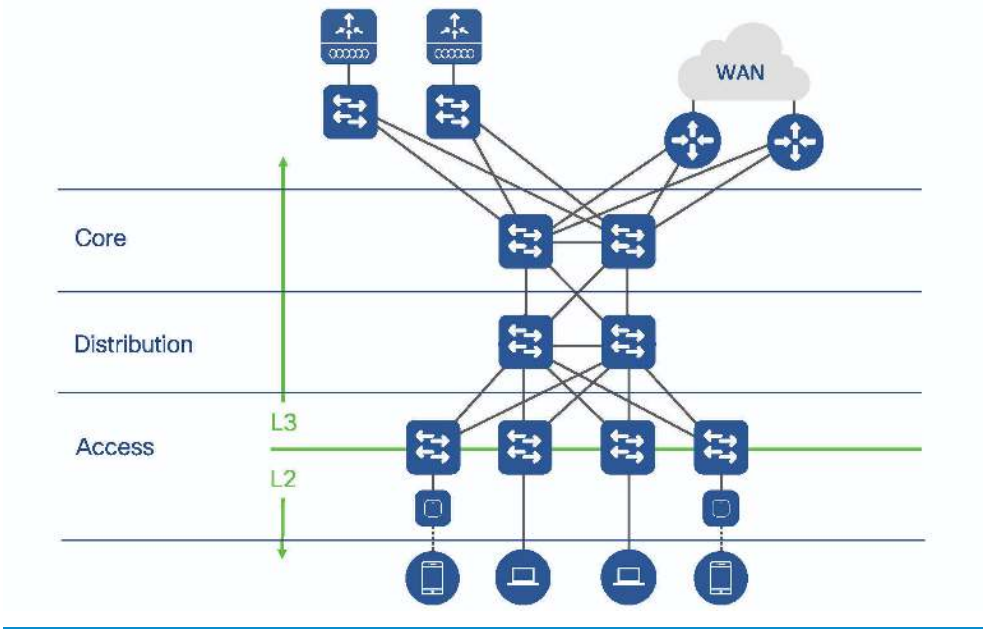
There are no changes to the access layer in this design, and the recommendations for using the Catalyst 9300 and 9400 in these roles still apply.

└the bottom line
The Catalyst 9300 and 9400 Series platforms are optimized for the access layer.

Routed Access

The routed access design uses the same physical network topology as a traditional multi-layer architecture. This is true regardless of whether a full three-tier architecture or the collapsed core method is employed. The difference is the placement of the Layer 2 and Layer 3 boundaries. As the name implies, in a routed access design the L3 boundary moves down to the access layer, and VLANs are locally contained within each access layer switch. These switches then connect upstream to the campus network using routed uplinks.

DIAGRAM Routed Access Design



The use of a routed access design affords a number of benefits:

- It reduces deployment and management complexity since all network links are routed connections with consistent configurations.
- It eliminates the need to configure 802.1Q trunks or tune Spanning Tree and first-hop routing protocols between layers.
- It simplifies network operation and troubleshooting since a single control protocol manages the network's behavior.
- It significantly reduces failure domains, by moving L2 domains to the access layer and isolating STP domains to individual switches.
- This design allows better utilization of all available network paths. Routed access networks do not impose STP blocking and instead use equal-cost multipathing (ECMP) to automatically distribute flows across all connections from the access layer upstream.

Routed access networks do have some drawbacks. It is not possible to span VLANs across a campus network in a routed access design model. Although best practice campus design seeks to eliminate large L2 domains because of the risk they pose to network stability, there are times when it is necessary to interconnect systems that operate in the MAC-layer only. Another drawback of routed access design is that if ACLs are implemented, they must be distributed to the access layer, rather than centrally at the distribution layer in the hierarchical design.

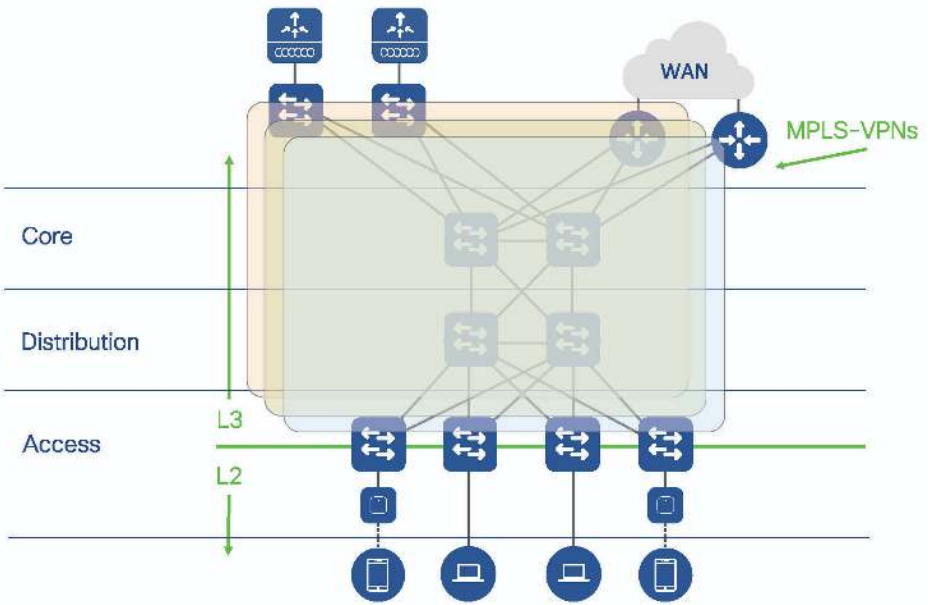
A routed access network utilizes the same topology as a hierarchical one. Consequently, the positioning of Catalyst 9000 series platforms remains the same. Catalyst 9500 switches serve best as fixed-form distribution and core, and Catalyst 9300s and 9400s are best suited as access layer switches. In addition to an SDM template that optimizes a 9300 or 9400 as an L2 switch, there is also a template that reprograms them to be optimized for L3 access.

└the bottom line
All Catalyst 9000 switches are optimized for routed access networks.

Campus MPLS

The campus MPLS design builds upon the previous network designs and delivers virtual separation of routing domains. This requires overlaying MPLS-VPN technology on top of all routed points within the network. For a multi-layer campus, MPLS extends from the distribution blocks through the network core; in a Routed Access design, MPLS spans the entire LAN.

DIAGRAM Campus MPLS



The main advantage of MPLS-VPN into campus designs is to provide macro-level network segmentation. MPLS-VPNs use Virtual Routing and Forwarding (VRF) instances in order to separate routing domains, for example, to keep guest traffic from being able to reach any private corporate segments. Campus MPLS networks also allow network administrators to manipulate MPLS protocols, for example, to improve the efficiency of

label switching over basic routing, and the ability to steer traffic through a network (called MPLS-TE or traffic engineering) to optimize available paths.

However, MPLS adds complexity on top of an IP network as it requires an additional control plane (based on Multi-Protocol BGP) in order to exchange the necessary VRF information between Provider Edge (PE) devices in the MPLS architecture. Creating and managing multiple, virtual routing domains requires appropriate design by architects, as well as management expertise of operators. During the design phase, equipment selected for the LAN must be evaluated for its ability to handle L2, L3, and also MPLS protocols. During operations, network administrators must understand MPLS protocols and their rules, and be able to troubleshoot multiple, concurrent routing planes.

A simplified version of a segmented network approach that uses VRFs, without MPLS tagging, can be achieved by using a VRF-Lite end-to-end deployment. VRF-Lite leverages 802.1q trunks and VLAN IDs between switches to transport and differentiate traffic residing in different VRFs. While possible on a small scale, VRF-Lite rapidly becomes too difficult to manage at any scale above 8-10 VRFs. Organizations wishing to provide segmentation within their campus or WAN, and needing to scale beyond this number of VRFs, would traditionally opt for the more complex (but scalable) MPLS-VPN solution.

The enterprise requirements for segmentation are becoming more stringent. Although MPLS-VPN furnishes macro-segmentation, it does not address micro-segmentation using access control policies. Security teams now demand policy enforcement between hosts within the same routing domains. The ability to combine macro-segmentation (using VRFs) and micro-segmentation (using Scalable Group Tags [SGTs]) is discussed further in Chapter 14.8 Software-Defined Access.

The entire Catalyst 9000 series supports campus MPLS. The same recommendations apply as with a routed access network, with the addition of MPLS feature support. The Catalyst 9500 platforms are designed for core and distribution layer services, but the Catalyst 9300 and 9400 switching lines are best suited as access devices.

└the bottom line
All Catalyst 9000 switches are optimized for Campus MPLS networks.

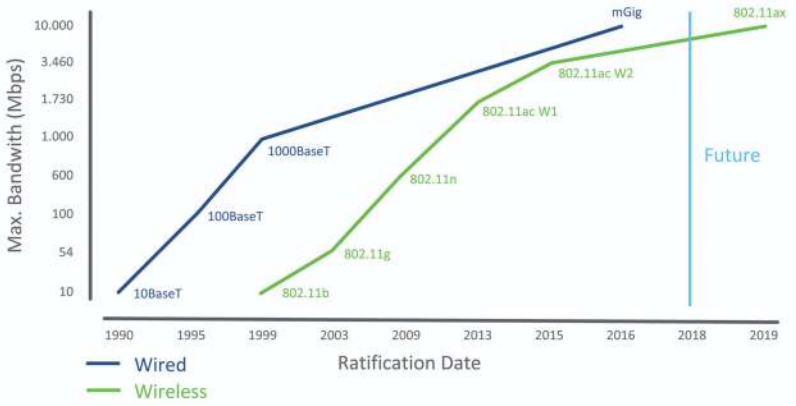
Campus Wireless

Catalyst 9000 switches provide a variety of unique capabilities and innovations to deliver both wired LAN access and optimal wireless access.

802.11 wireless LANs (also known as Wi-Fi) are an access layer technology. Wi-Fi is fast becoming the default choice for users to connect their client machines. People want to move about and take their computers and phones with them, so they can get work done faster. For users, mobility is a powerful tool for productivity and efficiency.

Modern wireless deployments can now offer link speeds comparable to, or even in excess of, what may be available on the wired infrastructure. Wireless deployments using 802.11ac Wave 2 (and in future 802.11ax) now provide multi-Gigabit wireless link speeds.

DIAGRAM Wired and Wireless Evolution



Businesses recognize these mobility trends and are transitioning to wireless-only offices, not only to meet the needs of their mobile workforce but also to optimize their

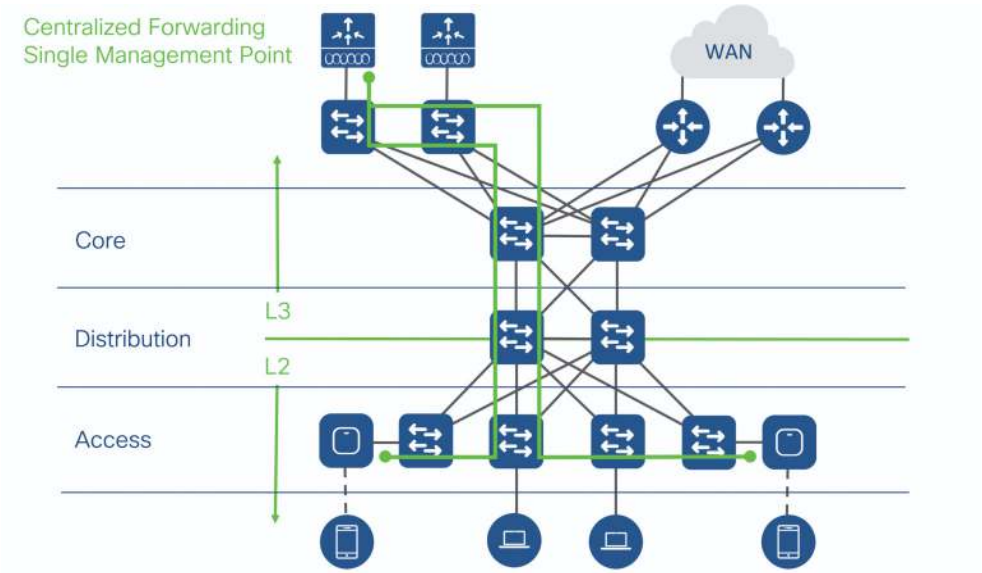
budgets and achieve the right balance between mobile and fixed endpoints. Businesses must also realize that this trend to wireless access also requires a more flexible wired infrastructure.

└the bottom line
As 802.11ac Wave 2 adoption grows, switches must support higher connection speeds such as mGig, 25G or 100G.

In a multi-layered campus design, a centralized wireless network uses a Wireless LAN Controller (WLC). WLCs usually connect at either the campus distribution or core layer, in a service block connected to the core, or even a remote data center. Wireless Access Points (APs) connect to the access layer switches, normally using PoE and high-speed copper Ethernet.

In this design, the WLC becomes the central point-of-management. All configuration and monitoring of wireless APs takes place on the WLC. APs then tunnel all data traffic they receive to the WLC, requiring it to make all forwarding and policy decisions. In other words, the WLC becomes the L2 wireless access boundary, which is connected to a local L3 routing border. This technique allows wireless clients to L2 roam between APs but appear to the network as if they are connected within the same L3 subnet.

DIAGRAM Centralized WLAN Design



A centralized wireless network design enables network administrators to configure hundreds of APs from a single management point. The centralized design also solves the roaming challenges of 802.11 networks. Instead of extending wireless networks across routed boundaries, a WLC consolidates them at a single point in the network.

One challenge with this approach is that wireless LANs are still managed separately from the rest of the wired network. Ideally, network administrators should be able to define a single policy for all endpoints and for the policy to be enforced the same way, regardless of the access medium. This separation causes unnecessary duplication of effort and potential configuration errors.

Another shortcoming deals with network scalability. When APs forward all data to central WLCs, the controllers must be able to handle the entire traffic load. This has not been a problem when WLCs are connected using multiple 1 or 10Gbps links, and APs are connecting a small number of clients at 10 or 100Mbps rates. With the adoption of 802.11ac Wave 2, WLCs will need to support multiple 10 Gbps links in order to keep up

with hundreds of clients that can connect at more than 1Gbps rates. This requires higher bandwidth and greater port density of the switches connected to the WLC.

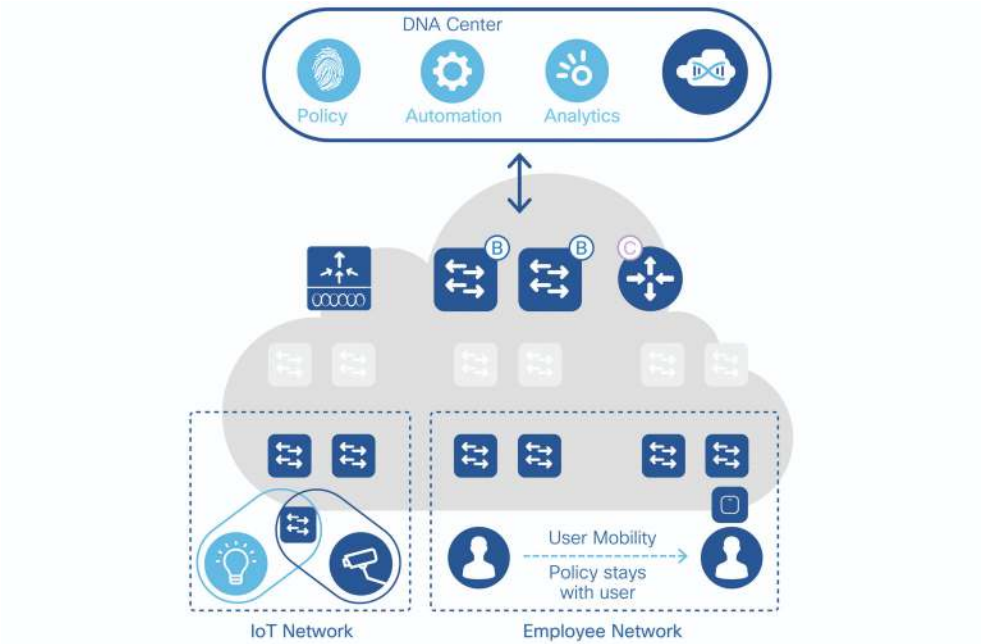
└the bottom line
All Catalyst 9000 switches are optimized for Campus Wireless networks.

Software-Defined Access (SD-Access) offers an innovative new approach for both wired and wireless LAN deployments. SD-Access retains the benefits of centralized management while adding common policy and addressing scale.

Software-Defined Access

Cisco's Software-Defined Access (or SD-Access) solution is a programmable network architecture that provides software-based policy and segmentation from the edge of the network to the applications. SD-Access is implemented via Cisco Digital Network Architecture Center (DNA Center) which provides design settings, policy definition and automated provisioning of the network elements, as well as assurance analytics for an intelligent wired and wireless network.

DIAGRAM SD-Access Solution Overview



SD-Access creates a logical overlay fabric network to provide the benefits of host mobility, segmentation, and group-based policy regardless of their location on campus, fully automated and assured by Cisco DNA Center.

The Catalyst 9000 series platforms participate in the physical and network layers of Cisco SD-Access. The SD-Access network layer (or fabric) is comprised of two main components:

- **Network Underlay** - is analogous to an existing Layer 3 routed hierarchical network, but with a simplified focus on transporting data packets between network devices for the fabric overlay. From a design perspective, this is the equivalent of a Routed Access design.
- **Fabric Overlay** - is a primarily a logical (tunneled) network, that virtually interconnects all of the network devices (to form a "fabric"). The fabric overlay creates a virtual environment to enable macro and micro-level segmentation, group-based security and application policy, as well as dynamic host mobility services for wired and wireless clients.

SD-Access Wireless

SD-Access treats wireless data exactly the same as wired data. This approach enables a common policy across both mediums. The control plane of the wireless network remains centralized. But for the wireless data plane, SD-Access moves forwarding to a local, distributed forwarding model via the switch infrastructure. Instead of tunneling all client traffic to the controller, each access point constructs a VXLAN tunnel directly to the fabric edge switch to which it attaches. The switch terminates this traffic and then provides full treatment for the wireless traffic just as it would for a wired host, including group-based policy structured around VNs and SGTs.

Some of the key SD-Access wireless benefits are:

- **Centralized control plane** - Wireless infrastructure operations occur within a centralized wireless network such as AP management, RRM, client onboarding and roaming.

- **Distributed data plane** - Wireless data traffic is distributed to the fabric edge switches for optimal performance and scalability.
- **Seamless L2 roaming** - Clients can roam seamlessly within VNs stretched across a campus while retaining the same IP address and group policy.
- **Policy simplification** - SD-Access breaks the dependency between policy and network constructs and abstracts it for application across wired and wireless end-points.

For more information on Cisco SD-Access, please visit *www.cisco.com/go/sdaccess* for more details.

Since SD-Access creates an overlay, the network devices need to support new encapsulations and protocols, such as LISP, VXLAN and SGT.

The Catalyst 9500 series can be configured to use the SDA SDM template, which provides the optimal ASIC table allocation for an SD-Access Fabric Border role. The flexible UADP ASIC provides VXLAN-GPO frame encapsulation and integrated SGT-based ACL classification and enforcement, as well as many evolving SD-Access capabilities.

└the bottom line
The Catalyst 9500 Series platforms are optimized for the SD-Access Fabric Border role.

The Catalyst 9300 and 9400 series can be configured to use the SDA SDM template which provides the optimal ASIC table allocation for an SD-Access Fabric Edge role. The flexible UADP ASIC provides VXLAN-GPO frame encapsulation and integrated SGT-based ACL classification and enforcement, including direct AP VXLAN tunnels for SD-Access Wireless.

└the bottom line
The Catalyst 9300 and 9400 Series platforms are optimized for the SD-Access Fabric Edge role.

Appendix

References

Additional websites which offer more details information about the Catalyst 9000 family and its capabilities:

Overview of the Cisco Catalyst 9000 family:
https://www.cisco.com/c/en/us/products/switches/catalyst-9000.html

Overview of Cisco Catalyst 9000 Series switches:
https://www.cisco.com/c/en/us/products/switches/catalyst-9300-series-switches/index.html
https://www.cisco.com/c/en/us/products/switches/catalyst-9400-series-switches/index.html
https://www.cisco.com/c/en/us/products/switches/catalyst-9500-series-switches/index.html

Cisco Catalyst 9000 Series switches white paper:
https://www.cisco.com/c/en/us/products/switches/catalyst-9300-series-switches/white-paper-listing.html
https://www.cisco.com/c/en/us/products/switches/catalyst-9400-series-switches/white-paper-listing.html
https://www.cisco.com/c/en/us/products/switches/catalyst-9500-series-switches/white-paper-listing.html

Cisco Live On-Demand Library:
https://www.ciscolive.com/global/on-demand-library/?#/

Search for the session IDs shown below:
BRKARC-2035: The Catalyst 9000 Switch Family – An Architectural View
BRKARC-3467: Cisco Enterprise Silicon - Delivering Innovation for Advanced Routing and Switching
BRKARC-3863: Catalyst 9300 Switching Architecture
BRKARC-3873: Catalyst 9400 Switching Architecture
BRKCRS-3300: IOS XE: Enabling the Digital Network Architecture

Additional websites which offer more details information about programmability and automation on Catalyst 9000 family:

DevNet, the Cisco Developers Network:
https://developer.cisco.com/
https://developer.cisco.com/site/ios-xe

Cisco Live On-Demand Library:
https://www.ciscolive.com/global/on-demand-library/?#/

Search for the session IDs shown below:
BRKCRS-1450: Introduction to Catalyst Programmability
BRKCRS-2451: Scripting Catalyst switches—tools and techniques beyond the basics
BRKCRS-2004: Application Hosting and Model-Driven Telemetry on Open IOS XE

Acronyms

AAA - Authentication, Authorization and Accounting

ACK - acknowledgment

ACL - Access Control List

AES - Advanced Encryption Standard

AOC - Active Optical Cables

API - Application Programming Interface

AQM - Active Queue Management

AR - Augmented Reality

ARP - Address Resolution Protocol

ASIC - Application-Specific Integrated Circuit

AVB - Audio Video Bridging

AVC - Application Visibility and Control

BGP - Border Gateway Protocol

BOOTP - Bootstrap Protocol

BPDU - Bridge Protocol Data Units

BUM - Broadcast Unknown unicast Multicast

BYOD - Bring Your Own Device

CAF - Cisco Application Framework

CAPWAP - Control And Provisioning of Wireless Access Points

CDP - Cisco Discovery Protocol

CEF - Cisco Express Forwarding

CI/CD - Continuous Integration, Continuous Delivery

CLI - Command Line Interface

CoS - Class of Service

CPU - Central Processing Unit

CSMA/CD - Carrier Sense Multiple Access with Collision Detection

CTA - Cisco Trust Anchor

CTA - Cognitive Threat Analytics

CoA - Change of Authorization

DAC - Direct Attach Copper

DAD - Dual-Active Detection Link

DHCP - Dynamic Host Configuration Protocol

DIY - Do-It-Yourself

DMZ - Demilitarized Zone

DNA – Digital Network Architecture

DNS - Domain Name System

DNS-AS - DNS as Authoritative Source

DPI - Deep Packet Inspection

DSCP - Differentiated Services Code Point

DTLS - Datagram Transport Layer Security

DTP - Dynamic Trunk Protocol

DTS - Dynamic Threshold Scheduler

EAP - ECC - Error-Correcting Code

EAPoL - Extensible Authentication Protocol over LAN

EARL - Encoded Address Recognition Logic

ECMP - Equal-Cost Multipathing

ECN - Explicit Congestion Notification

EFC - Egress Forwarding Controller

EGR - Egress Global Resolution

EIGRP - Enhanced Interior Gateway Routing Protocol

ELLW - Enhanced Limited-Lifetime Warranty

EQS - Egress Queuing Scheduler

ERSPAN - Encapsulated Remote Switched Port Analyzer

ETA - Encrypted Traffic Analytics

FIB - Forwarding Information Base

FIFO - First In First Out

FNF - Flexible NetFlow

FPGA - Field Programmable Gate Array

FSU - Fast Software Upgrade

FTP - File Transfer Protocol

GIR - Graceful Insertion and Removal

gNMI- google Network Management Interface

GPE - Generic Protocol Extension

GPO - Group Policy Object

GRE - Generic Routing Encapsulation

gRPC – google Remote Procedure Call

Gbps - Gigabits per second

HA - High Availability

HQoS - Hierarchical QoS

HSRP - Hot Standby Router Protocol

HTTP - Hypertext Transfer Protocol

HVAC - Heating Ventilation and Air Conditioning

HW - Hardware

IBNS - Identity-Based Networking Services

IDP - Initial Data Packet

IEEE - Institute of Electrical and Electronics Engineers

IETF - Internet Engineering Task Force

IFC - Ingress Forwarding Controller

IGMP - Internet Group Management Protocol

IGR - Ingress Global Resolution IPsec - Internet Protocol security

ILP - Inline Power

IPFIX - IP Flow Information Export

IPTV - Internet Protocol Television

IQS - Ingress Queuing Scheduler

IS-IS - Intermediate System to Intermediate System

ISE - Identity Services Engine

ISSU - In-Service Software Upgrade

IoT - Internet of Things

JSON - JavaScript Object Notation

KVM - Kernel-based Virtual Machine

LACP - Link Aggregation Control Protocol

LAN - Local Area Network

LDP - Label Distribution Protocol

LED - Light-Emitting Diode

LISP - Locator/ID Separation Protocol

LXC - LinuX Container

MAB - MAC Authentication Bypass

MAC - Media Access Control

MACsec - Media Access Control security

MDT - Model Driven Telemetry

MEC - Multi-Chassis EtherChannel

MFIB - Multicast Forwarding Information Base

mGIG - multigigabit

MKA - MACsec Key Agreement

MLD - Multicast Listener Discovery

MMF - Multi-Mode Fiber

MPLS - Multiprotocol Label Switching

MSB - Most Significant Bits

MTU - Maximum Transmission Unit

Mbps - Megabits per second

NAT - Network Address Translation

NBAR - Network-Based Application Recognition

NSF - Non-Stop Forwarding

NSH - Network Services Headers

NSO – Network Service Orchestrator

OIR - Online Insertion and Removal

OM - Optical Multimode

onePK - One Platform Kit

OS - Operating System

OSI - Open Systems Interconnection model

OSPF - Open Shortest Path First

P2P - peer-to-peer

PAgP - Port Aggregation Protocol

PBC - Packet Buffer Complex

PBR - Policy-Based Routing

PDU - Protocol Data Unit

PE - Provider Edge router

PHY - PHYsical layer

PMK - Pair-wise Master Key

POE - Power over Ethernet

pps - packets per second

PSU - Power Supply Unit

PXE - Preboot Execution Environment

PnP - Network Plug and Play

QSA - QSFP to SFP Adapter

QSFP - Quad Small Form-factor Pluggable

QoS - Quality of Service

RADIUS - Remote Authentication Dial-In User Service

RED - Random Early Discard

REST - REpresentational State Transfer

RFC - Request for Comments

RFID - Radio-Frequency IDentification

RNG - Random number generators

RPC – Remote Procedure Call

RPM - RPM Package Manager

rpm - Revolutions per minute

RSPAN - Remote Switched Port ANalyzer

SAP - Security Association Protocol

SATA - Serial AT Attachment

SD-Access - Software Defined Access

SDK - Software Development Kit

SDM - Switch Database Manager

SFP - Small Form-factor Pluggable

SGACL - Scalable Group Access Control List

SGFW - Security Group Firewall

SGT - Scalable Group Tag

SGTIN - Serialized Global Trade Item Number

SIP - Session Initiation Protocol

SKU - Stock Keeping Unit

SLI - Switch Link Interface

SMC - Stealthwatch Management Console

SMF - Single Mode Fiber

SMU - Software Maintenance Update

SNMP - Simple Network Management Protocol

SPAN - Switched Port Analyzer

SPLT - Sequence of Packet Lengths and Times

SQS - Stack Queue Scheduler

SRAM - Static Random-Access Memory

SRP - Spatial Reuse Protocol

SSD - Solid-State Drive

SSH - Secure SHell

SSO - Stateful Switchover

ST - Service Template

STP - Spanning Tree Protocol

SUDI - Secure Unique Device Identifier

SVL - Stackwise Virtual Link

SW - Software

SXP - SGT Exchange Protocol

TAC - Technical Assistance Center

TCAM - Ternary Content Addressable Memory

TCO - Total Cost of Ownership

TCP - Transmission Control Protocol

TFTP - Trivial File Transfer Protocol

TLS - Transport Layer Security

TTL - Time to Live

ToS - Type of Service

UADP - Unified Access Data Plane

UDLD - Unidirectional Link Detection

UDP - User Datagram Protocol

UHF - Ultra High Frequency

UI - User Interface

UPoE - Universal Power Over Ethernet

URL - Uniform Resource Locator

USB - Universal Serial Bus

VLAN - Virtual LAN

VM - Virtual Machine

VN - Virtual Network

VR - Virtual Reality

VRF - Virtual Routing and Forwarding

VRRP - Virtual Router Redundancy Protocol

VTP - VLAN Trunk Protocol

VXLAN - Virtual eXtensible LAN

WAN - Wide Area Network

WLC - Wireless LAN Controller

WRED - Weighted Random Early Discard

WRR - Weighted Round Robin

WSMA - Web Services Management Agent

WTD - Weighted Tail Drop

XFP - 10 Gigabit Small Form Factor Pluggable

XML - Extensible Markup Language

YANG - Yet Another Next Generation

YDK -YANG Development Kit

ZTP - Zero Touch Provisioning

Bob Sayle
Dave Zacks
Dimitar Hristov
Fabrizio Maccioni
Ivor Diedricks
Jay Yoo
Kenny Lei
Mahesh Nagireddy
Minhaj Uddin
Muhammad Imam
Sai Zeya
Shawn Wargo