

Cisco Catalyst 9000 Switches

A new era of networking

2nd edition



Cisco Catalyst 9000 Switches

A new era of networking

2nd edition

Preface	7
Authors	8
Acknowledgements	10
Organization of this book	11
Intended audience	12
Book writing methodology	13
What is new in this edition of the book?	14
Introduction	15
Executive summary	16
Industry trends	19
Business use cases	21
Catalyst 9000 family of switches	25
Overview	26
Catalyst 9200 Series switches	29
Catalyst 9300 Series switches	33
Catalyst 9400 Series switches	38
Catalyst 9500 Series switches	47
Catalyst 9600 Series switches	54
Packaging, licensing and support	60
ASICs - the power of programmable silicon	65
What is an ASIC?	66
Why programmable ASICs?	69
Cisco UADP - programmable ASIC silicon	73
The Cisco UADP family	79

Cisco IOS XE	85
Cisco IOS evolution	86
Cisco IOS XE architecture	89
Cisco IOS XE benefits	92
High availability	95
Overview	96
Catalyst 9200 Series high availability	98
Catalyst 9300 Series high availability	102
Catalyst 9400 Series high availability	108
Catalyst 9500 Series high availability	115
Catalyst 9600 Series high availability	116
StackWise Virtual	120
Graceful insertion and removal	126
Patching Cisco IOS XE	129
Security and identity	131
Overview	132
Encrypted traffic analytics	133
Trustworthy solutions	136
MACsec encryption	139
QoS and queuing	145
Overview	146
Buffers and queues	148
QoS and queuing in the Cisco UADP ASIC	153
Hierarchical QoS	159

QoS for StackWise Virtual	164
QoS for overlay technologies	165
Application visibility and control	169
Overview	170
Application recognition	171
Application control	174
IoT	175
Overview	176
Power over ethernet innovations	177
Audio video bridging - AVB	180
Cisco DNA Service for Bonjour	182
User-centric platform design	185
Overview	186
RFID	187
Blue beacon	188
Bluetooth console	189
WebUI	190
Flexible templates	191
Programmability and automation	193
Overview	194
Device provisioning	196
Open programmable device APIs	198
Model-driven telemetry	208
Scripting	211

Configuration management tools	213
Cisco DevNet	215
Application hosting	217
Application hosting operation	218
Hardware resources	220
Campus network design	221
Overview	222
Physical infrastructure	225
Multi-layer campus	230
Collapsed core	233
Routed access	235
Campus MPLS	237
Campus wireless	239
Software-Defined Access	243
Appendix	247
References	248
Acronyms	250

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

Another group of authors contributed to the most recent update to this book, which was completed in April 2019.

- Gary M Davis - Customer Experience
- Ginger Liu - Product Management
- Jay Sharma - Technical Marketing

- Jeffrey Meek - Marketing
- Kenny Lei - Technical Marketing
- Minhaj Uddin - 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 Sejung 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)
- Faith Bosworth (Facilitator)
- Henrik van Leeuwen and Lennart Wolfert (Illustrators)
- 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.

Cover photo: Satpal Ranu

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. The book is organized into sections, with each section having multiple chapters.

First we introduce the Cisco® Catalyst® 9000 switching family, review the business drivers for enterprises, and illustrate how Catalyst 9000 switches address the challenges faced by enterprise IT. Next we review the architectural foundations of the Catalyst 9000 switching platform, both from a hardware perspective with the innovative Cisco Unified Access Data Plane (UADP) ASIC, as well as the cutting-edge capabilities provided by Cisco IOS® XE software. These foundational elements enable the Catalyst 9000 switching family to address the many demands placed on enterprise networks today.

How Catalyst 9000 switches 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. These capabilities are reviewed in the chapter *[Programmability and automation](#)*. Then an examination of application hosting on the Catalyst 9000 switching family is provided. Finally, the book examines the present state and future evolution of network design, and how Catalyst 9000 switches lead 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 the innovative Cisco Catalyst 9000 family of switches, and how they help to solve the many challenges that networking professionals face today.

Catalyst 9000 switches provide 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 switching 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 networking that the Catalyst 9000 switching family 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 switching family. The idea of this book is to present readers with the current challenges in enterprise networking and explore how the Catalyst 9000 switching platform solves these challenges. Catalyst 9000 switches provide cutting-edge hardware and software capabilities, easily adapting to future protocols and network architectures without losing sight of simplicity and security. This book explores this powerful next generation networking platform - the basis for the new era of intent-based networking.

A group of 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 switches bring 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.

#Cat9K

#CiscoDNA

#NewEraOfNetworking

What is new in this edition of the book?

This book has been updated to reflect several of the new and improved features that are available with Cisco Catalyst 9000 switches.

Here are the highlights of this revision:

Catalyst 9200 Series switches - the latest addition to the Catalyst 9000 fixed enterprise access-layer switching portfolio. The Catalyst 9200 Series offers full PoE+, power and fan redundancy, stacking bandwidth support up to 160 Gbps, fixed or modular uplinks, and Layer 2 and Layer 3 features. This includes updates to multiple chapters, including UADP 2.0 mini, Cisco IOS XE Lite, high availability (StackWise®-160/80), user design, security, QoS, etc.

Catalyst 9600 Series switches - the next generation purpose-built 40G/100G modular enterprise core/distribution switching platform. The Catalyst 9600 Series provides security, resiliency and performance at scale with a comprehensive set of industry leading, Layer 2 and Layer 3 features. This includes updates to multiple chapters, including UADP 3.0, high availability (SSO and StackWise® Virtual), user design, security, and QoS.

XFSU - Extended FSU builds on the standalone Catalyst 9300 Series FSU capability and provides a mechanism to upgrade and downgrade the software image by segregating the control plane and data plane update, to allow less than 30 seconds of traffic impact during the upgrade.

gRPC - gRPC remote procedure call is an open source remote procedure call (RPC) system initially developed by Google.

The book's revised edition of the Catalyst 9000 switching family addresses all of the above areas and capabilities.

Continue reading to know more!

Introduction

Executive summary

The world is rapidly changing. The demands of ubiquitous mobility, evolving IoT, cloud adoption and rapidly advancing security threats are making IT managers rethink how their networks are designed and implemented. Enterprises of all sizes around the world are replacing their legacy systems with new and evolving digital technologies to create a competitive advantage, enable higher productivity and lower operating costs. As more businesses embrace this change, networks have to adapt. Businesses cannot build networks the same way they have for the past 30 years. Organizations need to create flexible networks that can constantly learn, adapt, protect and evolve.

A new era of intent-based networking

Cisco Digital Network Architecture (Cisco DNA[™]) and Cisco 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. The Cisco Catalyst 9000 family of switches are the next generation in the legendary Catalyst portfolio of enterprise LAN access, distribution and core switches. These are the first purpose-built platforms designed to take advantage of Cisco DNA and SD-Access.

Cisco Catalyst 9000 switches have been designed as the foundation for an entirely new era of intent-based networking. 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 and modular LAN switches can run a single software code base 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.

Catalyst 9000 switches are based on two foundational aspects:

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

DIAGRAM Foundational attributes of Catalyst 9000 switches



x86 Multi-core
CPU

Secure containers
Application Hosting



UADP
ASIC

Programmable pipeline
Flexible tables



Open Cisco
IOS XE

Model-Driven APIs
Streaming telemetry

Catalyst 9000 switches are built on a common ASIC architecture powered by the Cisco Unified Access Data Plane (UADP) ASIC. This serves as an innovative, programmable and flexible silicon foundation for the platform. The Cisco UADP ASIC enables network infrastructures to adapt to new technologies, trends and business needs over time. Catalyst 9000 switches are also built on a standard multi-core 64-bit x86 CPU architecture. Note that Catalyst 9200 Series switches use an ARM CPU integrated into the UADP, for greater cost efficiency and lower power consumption. A common CPU architecture provides predictable software processing and control-plane management, providing the horsepower to tackle next-generation network architectures and providing a platform for application hosting.

Every Catalyst 9000 switch runs on the open and modular Cisco Internet Operating System - Cisco IOS XE. This improves portability across Cisco enterprise platforms (including Catalyst switches, ISR/ASR routers, access points 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 switches are built on common hardware and software, and the foundation for the new era of intent-based networking.

Industry trends

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

Mobility trends and considerations

The need for untethered, uninterrupted access enabled by new wireless and mobility technologies is driving the enterprise network infrastructure market. BYOD and mobile applications make it possible for workers to access corporate data from nearly anywhere from nearly any device. This creates significant challenges for back-end IT infrastructures and those who manage the network. Mobility is now not just a cost of doing business, but a strategic business asset, making it an integral part of the future enterprise network.

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.

The Cisco “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.

Cloud trends and considerations

Enterprises are augmenting internal IT with cloud services, be it on-premises, 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-premises 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 networking leadership with breakthrough innovations in mobility, IoT, cloud and security.

Enabling the mobility use case

Wired and wireless networks have historically been built and operated by different teams. Catalyst 9000 switches with SD-Access deliver 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 SD-Access architecture delivers simplified and seamless roaming for devices across the network. Catalyst 9000 switches with SD-Access support the embedded wireless LAN controller capability for small, branch deployment. With SD-Access, access points connect directly to Catalyst 9000 switches for data plane forwarding directly in hardware.

Furthermore, SD-Access 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 Catalyst 9000 family of switches deliver industry leading multigigabit (mGig) and PoE capacity allowing customers to build the densest wireless environments, leveraging Wi-Fi 6 (802.11ax) and 802.11ac Wave 2 and future wireless innovations. For more information, refer to <https://cs.co/wirelessbook>

└ the bottom line

Catalyst 9000 switches offer the optimal foundation for converging wired and wireless access.

Enabling the IoT use case

There are many new devices being connected to the network such as sensors, alarm systems, HVAC systems, badge readers, etc. which have not traditionally been connected or have been using proprietary protocols. 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.

For example, many devices are starting to advertise their services using the Bonjour (mDNS) protocol. Cisco 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 Cisco DNA Centertm.

Some IoT devices, such as LED lighting, require 'always-on' power over ethernet (PoE). Catalyst 9000 switches support perpetual PoE and fast PoE to keep the lights on, even while the switch reloads.

For professional media and audio applications, Catalyst 9000 access switches support audio video bridging (AVB) and IEEE 1588 timing.

↳ the bottom line

Catalyst 9000 switches are the ideal platform for connecting the Internet of Things.

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.

Catalyst 9000 switches support application hosting with local storage enabling fog computing and network function virtualization. This supports distributed intelligent

agents embedded into the network for analytics, assurance, security, and cloud-connected applications. Customers are able to host third-party applications on Catalyst 9000 switches, making this the most flexible platform in the industry. For more information, refer to <https://cs.co/programmabilitybook>

└ the bottom line

Catalyst 9000 switches are 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 switches 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. Catalyst 9000 switches can detect and mitigate malware hiding in encrypted traffic using encrypted traffic analytics (ETA). Even better, ETA detects anomalies in encrypted traffic without decrypting it.

Catalyst 9000 switches collect metadata in hardware about all 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 family of switches, 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 solutions security technology protects the network switches themselves. A holistic approach provides comprehensive verification of hardware and software integrity by securing the device, network communications, and hosted applications.

└ the bottom line

Catalyst 9000 switches provide the most secure switching environment in the networking industry.

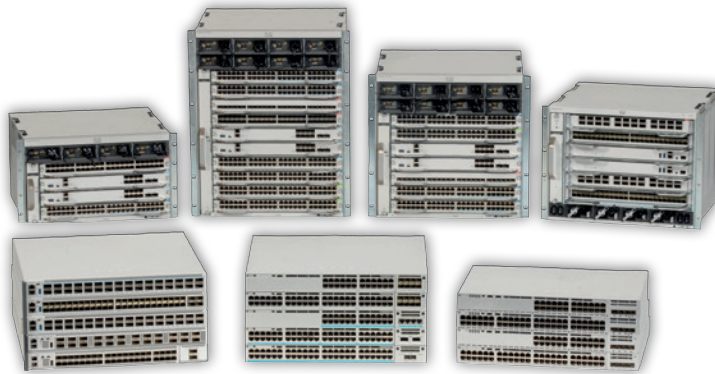
Catalyst 9000 family of
switches

Overview

The Cisco Catalyst 9000 switching family is the next generation in the legendary Cisco Catalyst portfolio of enterprise LAN access, distribution, and core switches. Catalyst 9000 switches extend Cisco networking leadership with breakthrough innovations in mobility, security, IoT, and cloud.

Designed from the ground up for higher performance, greater flexibility, security and resiliency, Catalyst 9000 switches start with feature parity from day one, while adding many new features and functionality. Leveraging the programmability of the Cisco UADP ASIC and Cisco IOS XE as well as the policy-based automation and assurance of Cisco Software-Defined Access, the Catalyst 9000 switching family addresses the challenges of today's always-on network, to help you focus on the needs of the business, not on the network - now and in the future.

DIAGRAM Catalyst 9000 family of switches



Catalyst 9000 switches are built on a common architecture with a strong hardware and software foundation. This commonality and consistency brings simplicity and ease of

operations for network architects and administrators, reducing total operational cost and creating a better user experience.

Common hardware

The Catalyst 9000 switching hardware uses a common design, both internally and externally.

Internally the hardware uses a common ASIC, the Cisco **Unified Access Data Plane (UADP)** ASIC, providing flexibility for packet handling. The hardware also uses a common control plane CPU. For the first time in the history of Catalyst switches, Cisco has also introduced an onboard x86-based CPU to allow the switch to host additional applications (beyond those normally possible on a network switch).

Externally the hardware is designed by one of the best industrial design firms in the world - Pininfarina, designer of the famous Ferrari. This level of design focus brings an enhanced user experience for the Catalyst 9000 switching family. It provides ergonomic design and common attributes that simplify device operations. More details are provided in the chapter [*User-centric platform design*](#).

Common software

The Catalyst 9000 family of switches run a common operating system, the Cisco Internet Operating System IOS XE. Cisco IOS XE is an enhanced, open and programmable OS. With a 30 year history and thousands of features, Cisco IOS XE is arguably the most feature-rich OS in the networking industry. Having a common code base shared across Catalyst 9000 switching platforms enables end-to-end feature support and feature parity throughout the network.

The strong hardware and software foundation of the Catalyst 9000 family of switches enables it to face the challenges of the modern enterprise network, while bringing consistency and simplicity for customers.

The Catalyst 9000 switching family has five members, broadly segregated into two types of network design models:

- Simple branch deployment
- Business-critical deployment

More details are provided in the chapter [Campus network design](#).

- **Catalyst 9200 Series** - fixed, stackable access (simple branch)
- **Catalyst 9300 Series** - fixed, stackable access (business-critical)
- **Catalyst 9400 Series** - modular access and distribution (business-critical)
- **Catalyst 9500 Series** - fixed core and distribution (business-critical)
- **Catalyst 9600 Series** - modular core and distribution (business-critical)

These platforms are discussed in further detail in the following chapters.

Catalyst 9200 Series switches

Catalyst 9200 Series switches focus on offering right-sized switching for simple branch deployments, extending the power of intent-based networking and Catalyst 9000 hardware and software innovations to a broader set of deployments. With its family pedigree, Catalyst 9200 Series switches offer simplicity without compromise – it is secure, always on and IT simplified. Catalyst 9200 Series switches offer full PoE+ capability, power and fan redundancy, stacking bandwidth up to 160 Gbps, modular uplinks, Layer 3 feature support and cold patching. Catalyst 9200 Series switches are purpose-built for cost-effective branch-office access.

DIAGRAM Catalyst 9200 Series switches



Platform overview

All Cisco Catalyst 9200 Series switches are fixed configuration 1RU switches with dual power supplies and redundant fans. The Catalyst 9200 Series offers two model options: Catalyst 9200 switches with fixed uplinks and fans (C9200L SKUs) and Catalyst 9200 switches with modular uplinks and fans (C9200 SKUs), offering various levels of scale and redundancy. Both models have 24 and 48 port copper options with three configurations:

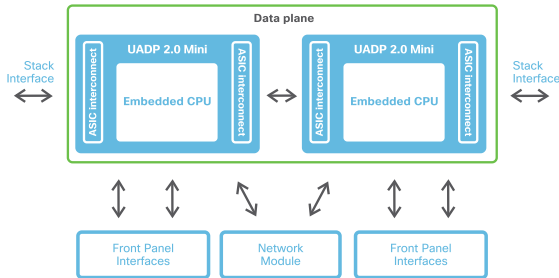
- **Data-only 1G models** - optimized for devices that primarily require data connectivity such as printers and desktop computing. All ports in these models support 10/100/1000Mbps.

- **PoE+ 1G models** - provide all capabilities of the data-only models with added support for PoE (15.4W) and PoE+ (30W) power to devices such as access points, IP phones capable of consuming power over ethernet. All ports can provide PoE+ power simultaneously with dual power supplies.
- **Multigigabit models** - provide multiple speeds up to 10Gbps on mGig ports, supporting Wi-Fi 6 and 802.11ac Wave 2 access points and desktops that demand higher speeds above the 1Gbps. All ports (mGig and 1G) on these models are PoE+ capable. Full PoE+ is supported with dual power supplies.

Architecture

Catalyst 9200 Series switches have been optimized for a simple branch deployment with an architecture that is simple yet powerful. Catalyst 9200 Series switches also have an embedded 4-core ARM CPU on the Cisco UADP 2.0 mini ASIC which is utilized to run the operating system. The Cisco IOS XE operating system has been optimized for the Catalyst 9200 Series as Cisco IOS XE Lite, providing a smaller image size and faster boot time accommodating the optimized hardware without compromising on the benefits of Cisco IOS XE.

Non-mGig models are powered by a single UADP 2.0 mini ASIC, whereas the mGig models have two UADP 2.0 mini ASICs. All ports on Catalyst 9200 Series switches are line rate for all packet sizes.

DIAGRAM Series switch architecture

StackWise-160/80

Catalyst 9200 Series switches provide the ability to stack up to eight switches, providing a centralized control plane while allowing distribution of the data plane. This allows network engineers to manage, configure and troubleshoot a stack of switches as one logical unit. The modular uplink C9200 models support a stacking bandwidth of 160 Gbps, whereas fixed uplink C9200L models have a stacking bandwidth of 80Gbps.

Additional details on benefits and operations of StackWise-160/80 can be found in the chapter [High availability](#).

Network modules

Catalyst 9200 Series modular uplink C9200 models have an optional slot for uplink network modules. The fixed uplink C9200L has fixed uplink configuration for each model. There are three variants of uplink modules which can be used not only to provide connectivity to an uplink switch but can also be used to connect hosts.

- 4x 1G SFP ports
- 4x 10G SFP ports

- 2x 25G SFP ports
- 2x 40G QSFP ports

Uplink modules are field replaceable units (FRU) which enable a swap of network modules without interrupting switch operations, thereby providing investment protection without compromising on availability.

Power supply and fan

All Catalyst 9200 Series switch models support dual redundant power supplies. The data-only model uses 125W AC power supply; 24 port PoE+ and mGig models use 600W AC and 48 port PoE+; and mGig models use 1000W AC supply. All three power supplies are highly efficient, rated by 80Plus as Silver (125W) and Platinum (600W, 1000W) efficiency. Such high efficiency along with innovations such as EEE leads to significantly lower cost of ownership.

Catalyst 9200 Series switch models are equipped with dual variable-speed fans. All models are capable of cooling the switch with a single fan in case of one fan failure. On modular uplink C9200 models, the fans can be replaced on an operational switch. On fixed uplink C9200L models, the fans are fixed.

Catalyst 9300 Series switches

Cisco Catalyst 9300 Series switches are the leading business-critical stackable enterprise fixed access switching platform. At 480 Gbps of stacking bandwidth and up to eight devices in a stack, it is the industry's highest-density stacking bandwidth solution.

DIAGRAM Catalyst 9300 Series switches



Platform overview

All models of Catalyst 9300 Series 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 configurations, each with 24 port and 48 port copper options:

- **Data-only models** - Optimized for devices such as desktops and printers that just need data connectivity from 10 Mbps to 1 Gbps
- **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+
- **Universal PoE (UPOE) 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 per switch is 1800W
- **Multigigabit (mGig) models** - Provide connectivity at multiple speeds up to 10 Gbps on mGig ports. Wireless access points supporting Wi-Fi 6 (802.11ax) and 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 three mGig variants:
 - **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 full mGig speeds
 - **48 port 5G mGig** - All 48 ports support 100 Mbps, 1 Gbps, 2.5 Gbps and 5 Gbps

Architecture

Catalyst 9300 Series 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
- ASIC interconnect / Stack interface

UADP ASIC

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

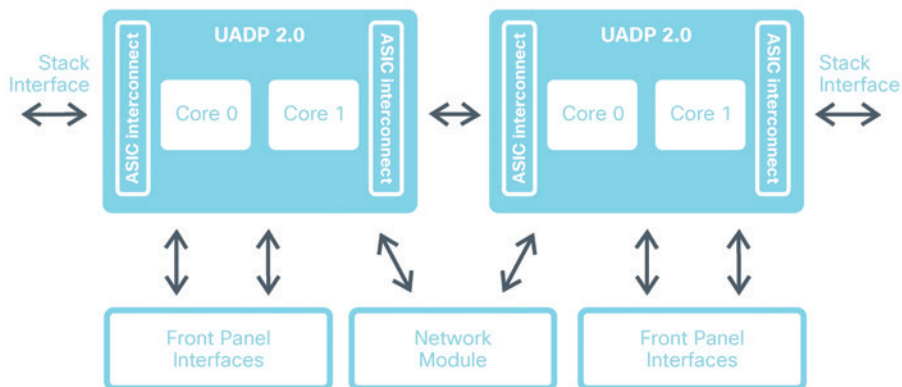
UADP 2.0 ASIC is built using 28 nanometer (nm) technology with two cores, with each core capable of supporting up to 80 Gbps of bandwidth for a total of 160 Gbps. Switches equipped with the UADP 2.0 ASIC support a total of up to 32K IPv4 / 16K IPv6 hardware tables, up to 18K of security ACL TCAM, and 2 x 8MB of buffer.

X86 CPU complex

Catalyst 9300 Series switches are all equipped with the same CPU, system memory and flash storage. Catalyst 9300 Series switches come with a 1.8 Ghz x86 quad core CPU, 8GB DDR4 RAM, and 16 GB of internal flash storage. For application hosting or general purpose storage, these switches support USB 3.0 SSD storage

ASIC interconnect

Catalyst 9300 Series switches consist of an internal stack interface (240 Gbps) acting as ASIC interconnect on the switches with dual ASICs. This interface is internally used for transporting traffic between front-panel ports of two ASICs on the same switch.

DIAGRAM Catalyst 9300 Series switch architecture

StackWise-480

Catalyst 9300 Series switches provide the ability to stack up to eight switches using special cables on the back, 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. the chapter [High availability](#) provides additional details on the operation of StackWise-480.

Network modules

All Catalyst 9300 Series switches have an optional slot for uplink network modules. There are four variants of uplink modules. The ports on these modules can be used for both uplink and downlink connectivity.

- 4x 1G RJ45 ports (10/100/1000Mbps)
- 4x mGig RJ45 ports (no PoE)
- 8x 10G SFP ports

- 2x 25G SFP ports
- 2x 40G QSFP ports

Catalyst 9300 Series switches are compatible with Catalyst 3850 Series uplink modules. However, Catalyst 9300 Series switch uplink modules are not compatible with Catalyst 3850 Series switches.

Power supply and fan

Catalyst 9300 Series 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®

Catalyst 9300 Series switches provide 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. StackPower can be deployed in two modes: power-sharing and redundant mode. Additional details are provided in the chapter [*High availability*](#).

Catalyst 9400 Series switches

Cisco Catalyst 9400 Series switches are the leading business-critical modular enterprise switching access platform. Catalyst 9400 Series switches provide unparalleled investment protection with a flexible chassis architecture capable of supporting up to 9 Tbps of system bandwidth. They also offer unmatched power delivery for high-density Power over Ethernet deployments, delivering 60W UPOE to endpoints. Catalyst 9400 Series switches deliver state-of-the-art high availability with capabilities such as dual supervisors and N+1/N+N power supply redundancy. The platform is campus-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 Series switches



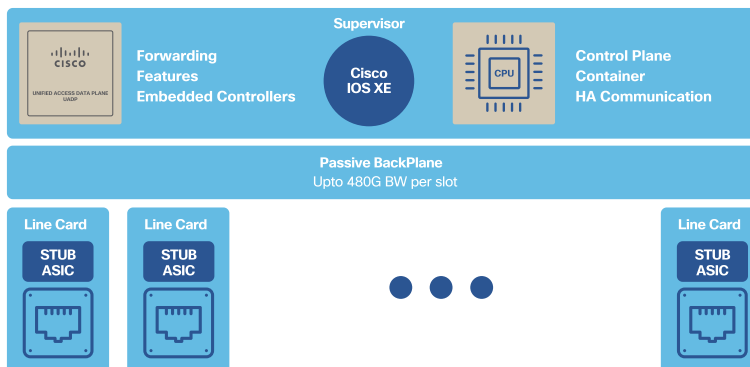
Platform overview

Catalyst 9400 Series 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. Each chassis is designed to support up to 720G of bandwidth between the two supervisor slots, allowing future supervisors to support multiple 100G ports. With the growing need for increased PoE, the chassis has the capability of providing more than 4,800W of PoE power per slot.

Architecture

Catalyst 9400 Series switches are 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 Series switch architecture



Supervisors

There are currently three versions of supervisor available for Catalyst 9400 Series switches:

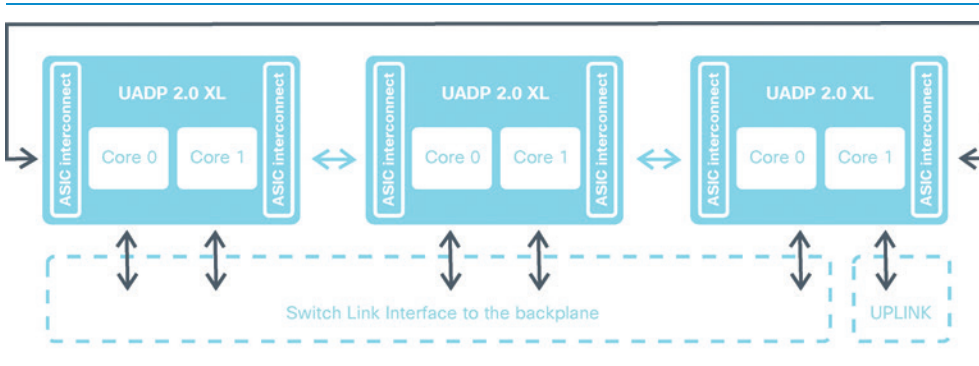
- C9400-SUP-1
- C9400-SUP-1XL
- C9400-SUP-1XL-Y

All Catalyst 9400 supervisors are powered by 3x UADP 2.0 XL ASICs and x86 CPU Processor. The three ASICs are interconnected through a 720G ASIC interconnect for packets passing between the ASICs.

The Catalyst 9400 Series SUP-1 provides 80 Gbps of bandwidth per slot for all chassis models and is optimized for access deployments.

The Catalyst 9400 Series SUP-1XL/1XL-Y provide 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/1XL-Y also adds support for different flexible templates to accommodate various deployment models such as access, distribution, core, SD-Access or NAT.

Cisco UADP use switch link interfaces (SLIs) to connect line card stub devices through the backplane. Each SLI (running at 10G rate with SUP-1/1XL/1XL-Y) 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 Series Supervisor-1XL architecture

Catalyst 9400 Series switches are all equipped with the same CPU, system memory, and flash storage. Catalyst 9400 Series switches come with a 2.4 GHz x86 quad-core CPU, 16GB DDR4 RAM and 16 GB of internal flash storage. For application hosting or general-purpose storage, these switches support USB 3.0 SSD storage and additionally support up to 960GB M2 SATA SSD storage options.

Supervisor uplinks

All supervisors have 8x SFP / SFP+ ports and 2x QSPF+ ports on the front. The architecture of SUP-1 and SUP-1XL provides 80G total uplink bandwidth shared between 1G / 10G / 40G interfaces.

On SUP-1XL-Y, port 1 and 5 of the SFP/SFP+ ports are also capable of 25G (SFP28). On SUP-1XL-Y, the uplink total bandwidth is also 80G shared between 1G / 10G / 25G / 40G interfaces.

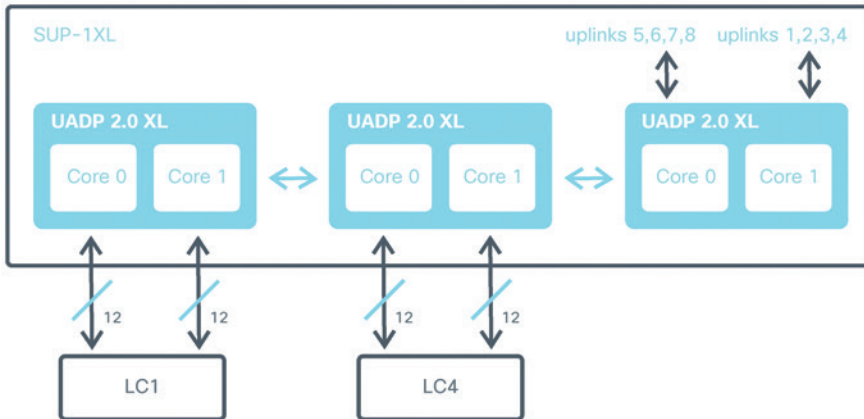
Line cards

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

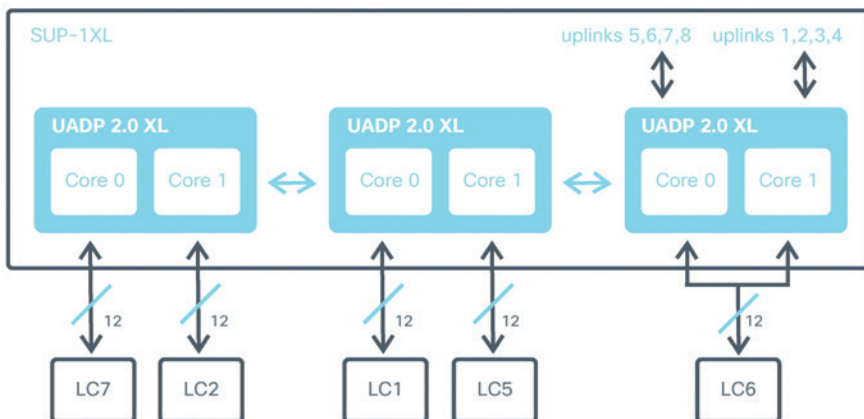
- **Copper RJ45 modules:**
 - **48-port Data-only line card** - All 48 ports on this module support 10 Mbps, 100 Mbps, and 1 Gbps.
 - **48-port PoE+/PoE line card** - All features on the data-only line card with added support for PoE+ (30W) and PoE (15.4W).
 - **48-port UPOE line card** - All features on PoE+/PoE line card with added support of UPOE (60W).
 - All 48 ports within the slot can provide UPOE simultaneously.
 - **48-port mGig line card** - The first 24 ports are the traditional 10/100/1000 Mbps ports and the last 24 ports are mGig ports.
 - 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:**
 - **24-port SFP line card** - Supports 100 Mbps and 1 Gbps speeds.
 - **48-port SFP line card** - All features on 24-port SFP line card with double the density.
 - **24-port SFP+ line card** - Supports 100 Mbps, 1 Gbps and 10 Gbps.

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.

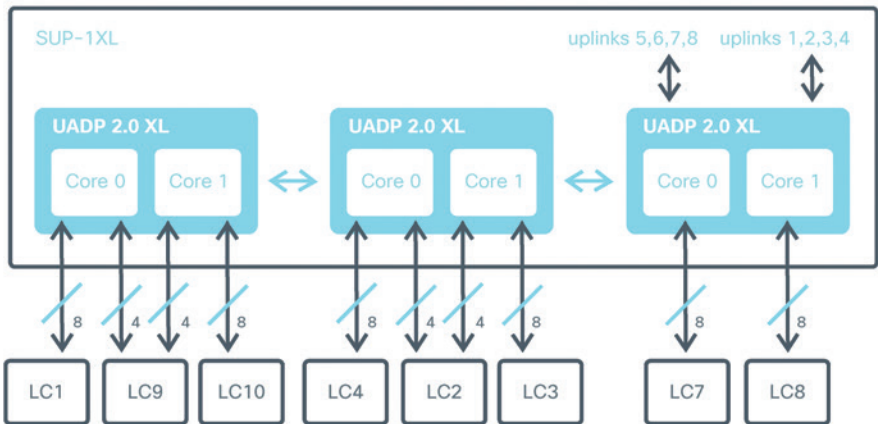
DIAGRAM Catalyst 9400 Series 4 slot chassis

Catalyst 9400 Series 4 slot chassis: 24 SLIs are active for each line card slot. Each UADP services 1 line card.

DIAGRAM Catalyst 9400 Series 7 slot chassis

Catalyst 9400 Series 7 slot chassis: 12 SLIs are active for each line card slot. Each UADP services 2 line cards.

DIAGRAM Catalyst 9400 Series 10 slot chassis



Catalyst 9400 Series 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 Catalyst 9400 Series Supervisor-1, 1XL and 1XL-Y. Also important to note is that line cards are designed to take advantage of higher per-slot bandwidth with future supervisors by running a larger number of SLIs.

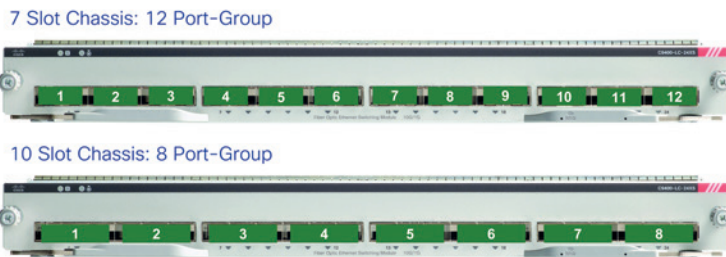
Line card performance mode

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 Catalyst 9400 Series 7 slot and 10 slot SLI port groups



Stateful switchover (SSO)

Catalyst 9400 Series switches support 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 the chapter [High availability](#).

In-service software upgrade (ISSU)

ISSU is an upgrade process available on Catalyst 9400 Series switches 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 the chapter [High availability](#).

Power supply

The power supplies for Catalyst 9400 Series switches 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 Series combines N+1, and N+N redundant options for power supplies. Additional details are provided in the chapter [High availability](#).

Fan tray

The fan tray of Catalyst 9400 Series switches contain multiple individual fans operating in an N+1 redundant mode. Fans operate at variable speeds based on the system temperature and altitude. This makes efficient use of the power and provides lower noise levels. The fan tray on Catalyst 9400 Series switches 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 Series fan tray



Catalyst 9500 Series switches

Cisco Catalyst 9500 Series switches are the industry's first purpose-built business-critical fixed 100G/40G core and distribution layer switches. These switches deliver exceptional table scales and buffering capabilities. Catalyst 9500 Series switches deliver up to 3.2 Tbps 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 Series switches



Platform overview

Catalyst 9500 Series switching platform consists of 1RU fixed configuration switches based on the Cisco Unified Access Data Plane (UADP) ASIC architecture. The platform runs on the Cisco IOS XE operating system. The platform also supports all the foundational high-availability capabilities including dual redundant power supplies and variable-speed highly efficient redundant fans.

Catalyst 9500 Series switches are built on two variants of UADP ASIC: UADP 2.0 XL 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.

100 Gigabit Ethernet switches:

- **C9500-32C** - Catalyst 9500 Series high-performance switch with 32x 100GE ports (2x UADP 3.0)
- **C9500-32QC** - Catalyst 9500 Series high-performance switch with 32x 40GE or 16x 100GE ports (1x UADP 3.0)

40 Gigabit Ethernet switches:

- **C9500-24Q** - Catalyst 9500 Series switch with 24x 40GE ports (4x UADP 2.0 XL)
- **C9500-12Q** - Catalyst 9500 Series switch with 12x 40GE ports (2x UADP 2.0 XL)

25 Gigabit Ethernet switches:

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

10 Gigabit Ethernet switches:

- **C9500-40X** - Catalyst 9500 Series switch with 40x 1/10GE ports (2x UADP 2.0 XL)
- **C9500-16X** - Catalyst 9500 Series switch with 16x 1/10GE ports (1x UADP 2.0 XL)

Architecture

Catalyst 9500 Series 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
- ASIC interconnect

UADP ASIC

Cisco UADP 2.0 XL 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 XL 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.

Cisco UADP 3.0 ASIC is built on 16-nanometer technology using two cores, with each core capable of supporting up to 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

Catalyst 9500 Series switches are all equipped with the same onboard CPU, system memory, and flash storage. Catalyst 9500 Series switches come with a 2.4 Ghz x86 quad core CPU, 16GB DDR4 RAM, and 16 GB 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 960 GB M2 SATA SSD storage options.

ASIC interconnect

Catalyst 9500 Series switches use high-speed ASIC interconnect links for inter-ASIC communication. UADP 2.0 XL has up to 720 Gbps (360 Gbps full duplex) of interconnect bandwidth and UADP 3.0 has up to 1.6 Tbps (800 Gbps full duplex) of interconnect bandwidth between two ASICs. Packets destined to local ports within the ASIC do not use ASIC interconnect links.

DIAGRAM Cisco Catalyst 9500 Series switch block diagram - Cisco UADP 2.0 XL

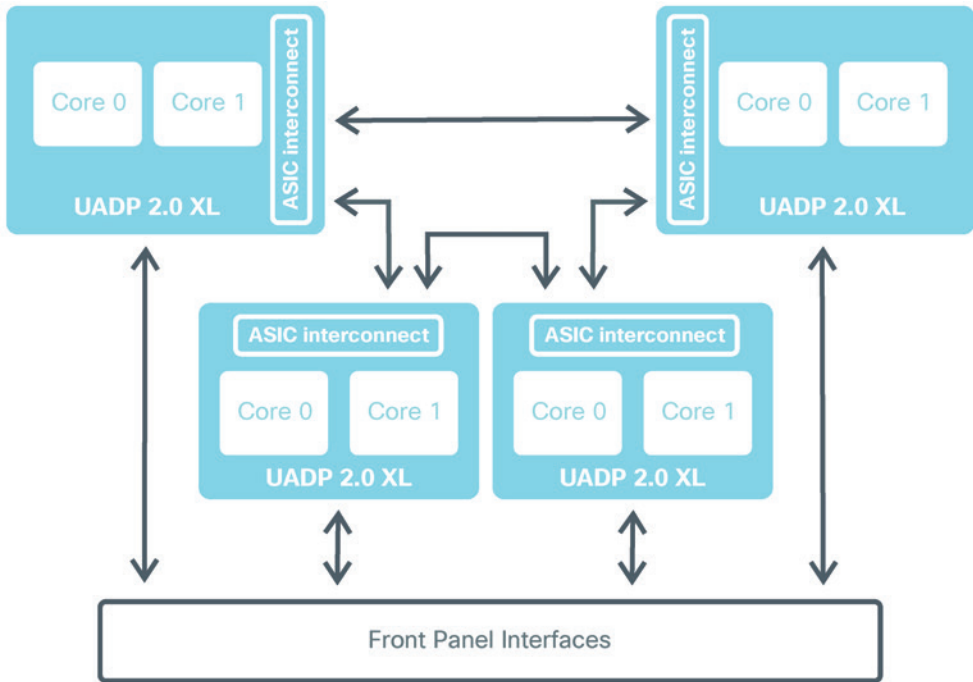
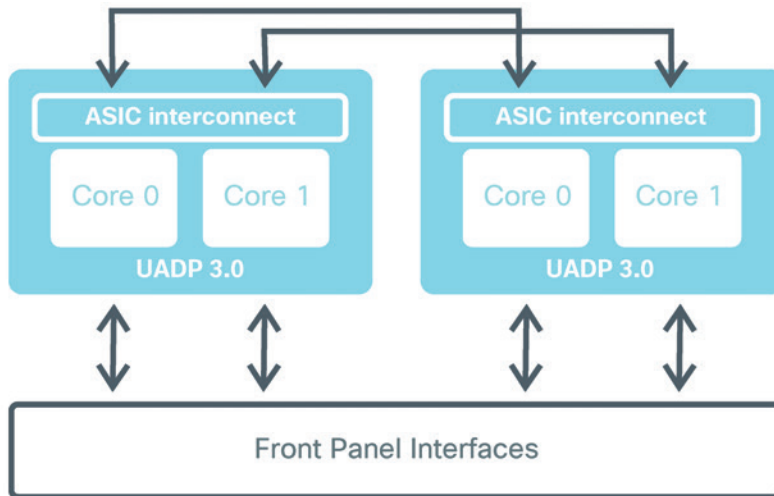


DIAGRAM Cisco Catalyst 9500 Series high-performance switch block diagram - Cisco UADP 3.0



Network modules

Cisco Catalyst 9500 Series switches support 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. There are two network module options available: 8x 1/10G and 2x40G.

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. The network modules support online insertion and removal (OIR).

Power supply

Catalyst 9500 Series switches support 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 backwards compatibility to drive network speeds beyond the current 10G and 40G capabilities while minimizing cost and real estate changes.

Cisco Catalyst 9500 Series 25G & 100G switches

Catalyst 9500 Series 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 <https://cs.co/SFP28>

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 <https://cs.co/QSFP>

Catalyst 9600 Series switches

Cisco Catalyst 9600 Series switches are the next generation of the industry leading business-critical modular enterprise campus core and distribution platform. The Catalyst 9606R chassis is hardware-ready to support a switching capacity of up to 25.6 Tbps. Catalyst 9600 Series switches support granular port densities that fit diverse campus needs, include non-blocking 40 and 100 GE (QSFP28) and 1, 10, and 25GE (SFP28). The platform delivers high availability with field replaceable dual supervisors, redundant power supplies and fans. The platform is campus-optimized with an innovative dual-serviceable fan tray design and side-to-side airflow and is closet-friendly with a ~16-inch depth.

DIAGRAM Cisco Catalyst 9606R chassis



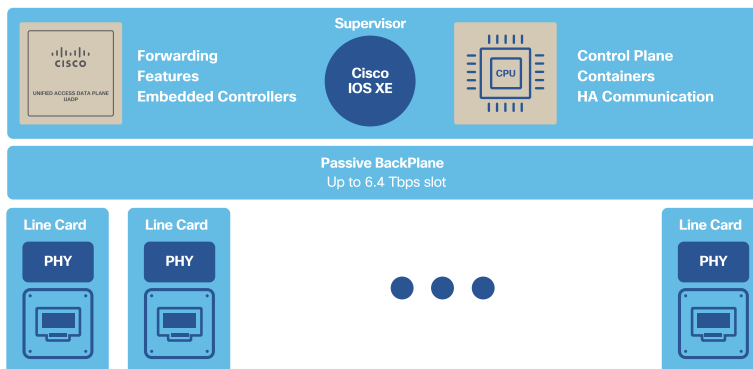
Platform overview

Catalyst 9606R is a 6 slot chassis, with two middle slots dedicated for the supervisors and four slots dedicated for the line cards. Each line card slot has a dedicated total bandwidth of up to 6.4 Tbps. The Catalyst 9606R chassis can provide a maximum of 128 x 40G/100G (QSFP) or 192 x 1G/10G/25G (SFP) ports.

Architecture

Catalyst 9600 Series switches are based on a centralized architecture. All forwarding, security, and queuing are done on the supervisor while the line cards are considered transparent, containing only PHYs and control logics. The simplicity of this centralized design allows easy upgrade of features as well as additional bandwidth by just upgrading the supervisor while keeping the existing line cards. The combination of the centralized architecture and transparent line card also provide uninterrupted supervisor switchover which is the foundation for in-service software upgrade (ISSU).

DIAGRAM Catalyst 9600 Series switching architecture

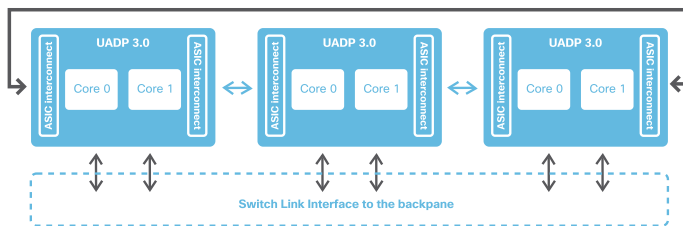


Supervisors

The Catalyst 9600 Series Supervisor-1 is powered by 3x UADP 3.0 ASICs and x86 CPU Processor. The three ASICs are interconnected with 3.2 Tbps ASIC interconnect on each ASIC. The ASIC interconnect uses a broadcast network to ensure traffic reaches all other ASICs. Due to high performance line card requirements, the Supervisor module does not have dedicated uplink ports (any of the line cards ports can be used for uplinks).

The Catalyst 9600 Series Supervisor-1 provides 2.4 Tbps (1.2 Tbps full-duplex) of bandwidth per slot for all the line card slots.

DIAGRAM Catalyst 9600 Series Supervisor-1 architecture



The Catalyst 9600 Series switches are all equipped with the same CPU, system memory, and flash storage. Catalyst 9600 Series switches come with a 2.0 GHz x86 8-Core CPU, 16GB DDR4 RAM, and 16 GB 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.

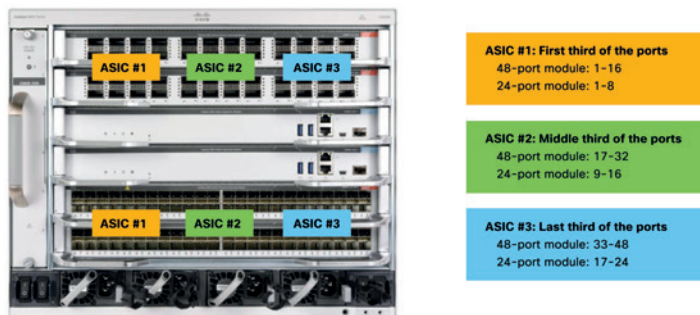
Note Due to the bandwidth requirements of the high-density C9600 line cards, the C9600-SUP-1 does not support uplink ports.

Line cards

Cisco Catalyst 9600 Series switches offer two types of line cards for different connectivity requirements.

- **Fiber QSFP module:**
 - 24-port QSFP+/QSFP28 line card
 - up to 24 ports of 40GE nonblocking
 - up to 12 ports of 100GE nonblocking (upper ports)
- **Fiber SFP module:**
 - 48-port SFP/SFP+/SFP28 line card
 - up to 48 ports 1/10/25GE nonblocking

The diagram below illustrates how the front panel ports are mapped to the 3x UADP 3.0 ASICs.

DIAGRAM Catalyst 9600 Series port mapping

Stateful switchover (SSO)

Catalyst 9600 Series switches support a redundant supervisor 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 the chapter [High availability](#).

In-service software upgrades (ISSU)

ISSU is an upgrade process available on Catalyst 9600 Series switches that allows upgrading supervisor software while traffic forwarding continues. This process is built on top of the stateful switchover. ISSU increases network availability and reduces downtime caused by planned software upgrades. Additional details are provided in the chapter [High availability](#).

Power supply

The power supplies for Catalyst 9600 Series switches come in small form factor while providing high capacity and efficient output. The Catalyst 9606R chassis provides four

power bays. Both AC and DC power supplies are available for Catalyst 9600 Series switches. The platform supports both combined and N+1 redundant mode. Additional details are provided in the chapter [High availability](#).

Fans and fan tray

The fan tray for Catalyst 9600 Series switches contains multiple individual fans operating in an N+1 redundant mode. The fans operate at variable speeds based on the system temperature and altitude. This makes efficient use of the power and provides lower noise level. The fan tray on Catalyst 9600 Series switches 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 9600 Series switch fan tray



Packaging, licensing and support

The Catalyst 9000 switching family uses a simplified licensing model. The previous generation of switches had different license types (LAN Base, IP Base, IP Services, and Enterprise Services) with the added complexity of differences across the multiple Catalyst switch families. The Catalyst 9000 switching 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
- lowers up-front costs by adding more features and support

Software packages and licenses

The Catalyst 9000 switching family offers two mandatory term-based software licensing options:

- **Cisco DNA Essentials** - provides the baseline network functionality used to operate a network and includes the perpetual Network Essentials functionality
- **Cisco DNA Advantage** - includes all the functionality in the Cisco DNA Essentials package (including Network Essentials) plus advanced capabilities such as advanced security, availability, automation, and assurance and the perpetual Network Advantage functionality

Each software feature license option is offered with 3, 5, or 7-year term and includes solution support, the Cisco Enhanced Limited Lifetime hardware Warranty (E-LLW) and Smart Net Total Care™ support for the life of the term. While renewal of the term

license is not mandatory, if the term license is allowed to expire, switch functionality will revert to the perpetual base network functionality and solutions support will end.

Smart Accounts are also mandatory for purchase. If a customer does not have a Smart Account set up prior to the purchase, a new Smart Account must be created at the time of purchase.

Note Cisco DNA Essential license option is not available on the Catalyst 9600 Series

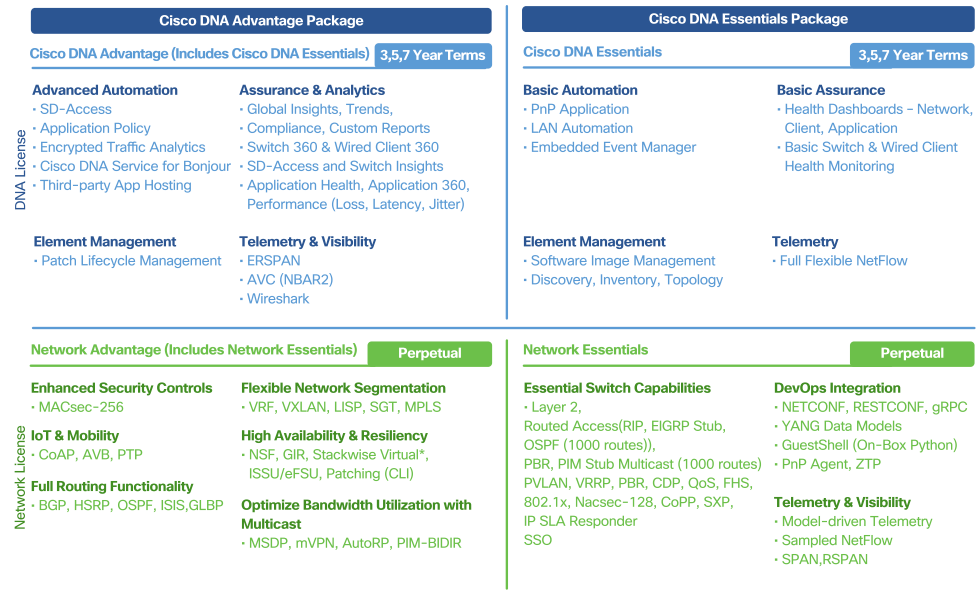
Additional information:

For Smart Account overviews and training sessions, visit <http://cs.co/operationsexchange>

To learn more about end-to-end Smart Account and Smart License management, visit <http://cs.co/smartmanager>

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

DIAGRAM Cisco DNA Advantage and Cisco DNA 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 <https://cisco.com/go/cfn>

Technical support

Technical support for the Catalyst 9000 switching 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 hardware 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](https://cs.co/ELLW). <https://cs.co/ELLW>

- 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 hardware Warranty on [www.cisco.com](https://cs.co/ELLW). <https://cs.co/ELLW>

- Software support for those features enabled through a Cisco DNA Essentials or Cisco DNA Advantage license is included while the subscription is valid. Again, support for these features ends once the subscription expires.

ASICs - the power of
programmable silicon

What is an ASIC?

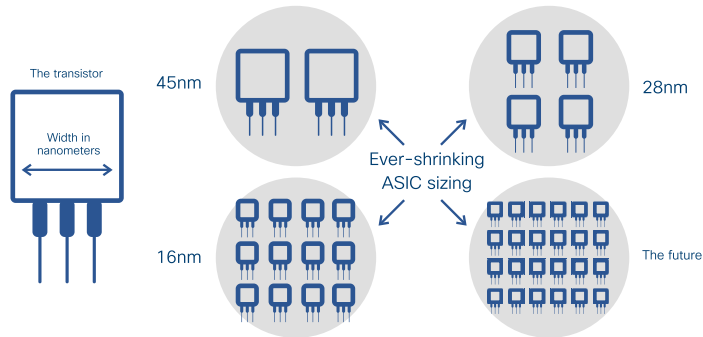
An application specific integrated circuit (ASIC) is a silicon microchip designed for a specific task (e.g. bridging or routing packets), rather than being used for general-purpose processing such as a CPU. ASICs are fundamental to how an Ethernet switch works.

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)

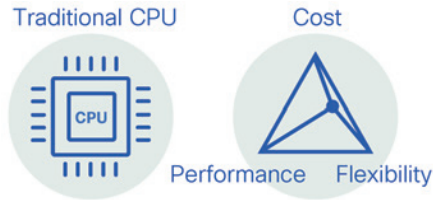
Modern ASICs are generally manufactured at sizes ranging from 45 to 28 nanometers with some newer models as small as 16 nanometers.

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 were 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.

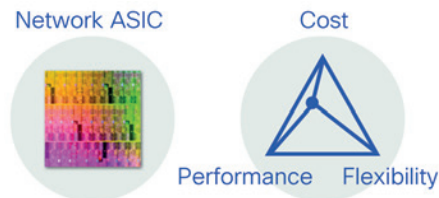
However, this class of networking ASICs are known as fixed ASICs. All aspects of these ASICs (behavior, speed, scale, etc.) are hard-wired (fixed) into them as part of the manufacturing process and cannot be changed without creating a whole new version of the ASIC.

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 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 needed to integrate the ASIC into a network switch, it 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, 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 the network itself, requires hardware replacement to adopt innovative capabilities that the new software-defined world demands.

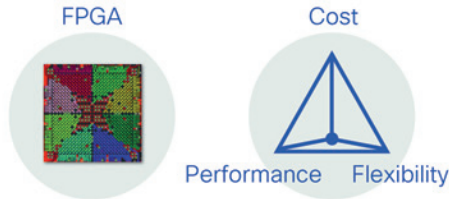
Programmable ASICs

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 Series 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 Cisco **Unified Access Data Plane (UADP)** ASIC.

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

Cisco UADP - programmable ASIC silicon

UADP - Cisco's flexible, programmable switching ASIC

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

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

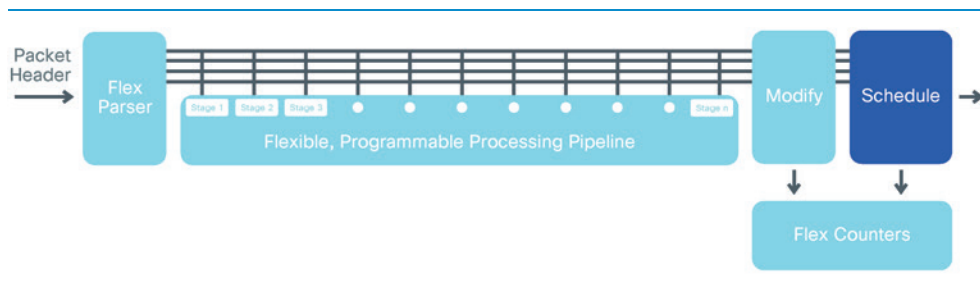
Flexibility has been designed into every aspect of the Cisco UADP from the beginning. In UADP, almost every processing stage that would be present in a fixed-configuration chip has been replaced with a flexible counterpart.

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 Cisco UADP ASIC contains a reprogrammable FlexParser that can parse a packet for different types of headers.

Unlike the traditional fixed-processing pipeline, the Cisco 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.

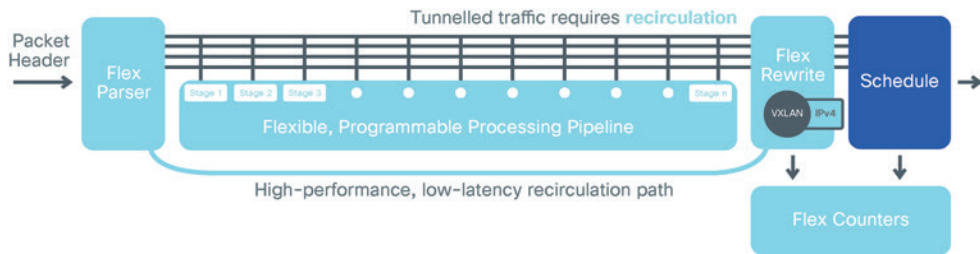
DIAGRAM Programmable Cisco UADP - processing pipeline



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 or remove 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 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. This new header will use 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 again to forward to this new destination, along with any services or policies that may apply to the tunnel.

DIAGRAM Programmable Cisco UADP - packet recirculation

In the Cisco UADP, a packet can be recirculated in approximately 500 nanoseconds (half a microsecond). The bandwidth available for recirculation is flexible, meaning packet recirculation can also use the spare bandwidth not currently being used by the front-panel ports. 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 or 3 passes are normally required.

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 common, it is clear that UADP was purpose-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 does 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, saving the valuable UADP pipeline performance for other functions.

Some examples of micro-engine functions built into the Cisco 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.

The Cisco 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 provides hardware MACsec encryption using the AES algorithm with up to 256-bit keys and 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.

Integrated Flexible NetFlow - Accounting for all traffic flowing through the network is important for multiple use cases. For example, network baselining and capacity planning, or basic application visibility. Using Flexible 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. Catalyst 9000 switches are 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 Flexible NetFlow capability and inspects encrypted transactions, without decryption, 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). By integrating this with Cisco Stealthwatch and cloud-based machine learning capabilities using cognitive threat analytics, a high-accuracy transaction 'fingerprint' analysis can be performed to

determine if the encrypted flow represents 'normal' network 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 Cisco 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. The UADP flexible pipeline can reference up to two packet matches, for multiple parallel actions, without degrading performance. UADP can then apply the appropriate policies configured by the network administrator. Examples include permit/deny, QoS remarking, path selection, packet copy, and other actions.

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 same 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.

Integrated stacking capability

The Cisco 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 a larger, integrated system. 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.

UADP supports the following inter-ASIC connection options:

- stacking multiple switches together via external cables, to build an integrated stack.
- linking multiple ASICs together on the baseboard, within a fixed switch.

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

Programming UADP with microcode

Cisco IOS XE uses multiple layers of software. Some software layers are more closely associated with the hardware than others. For example, features that a user interacts with (e.g. the CLI) are typically at a higher layer and less dependent on the hardware. Meanwhile, 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. Refer to the chapter [Cisco IOS XE](#) for more details on the software architecture.

The microcode for programming the Cisco UADP ASIC is included in the Cisco IOS XE image. Any changes in the microcode come with the image that runs on Catalyst 9000 switches. From a user perspective, the microcode upgrades are seamless and do not require any additional tasks to enable.

The Cisco UADP family

The history of Cisco UADP ASIC began in 2013 when Cisco introduced Catalyst 3850 Series switches. 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 ASIC.

Cisco 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 built on a 65 nanometer (nm) process, while the latest UADP 3.0 was built on 16 nm. Cisco 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.

Cisco UADP 1.0/1.1

Cisco UADP 1.0 is a single-core 65nm 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 Series switches use UADP 1.0.

By 2015, a newer 36nm version of the same ASIC design (version 1.1) was introduced. The main elements and architecture of the ASIC remained essentially unchanged, but with several important improvements were introduced. The key difference between Cisco UADP 1.1 and 1.0 is the use of a dual-core architecture inside the ASIC.

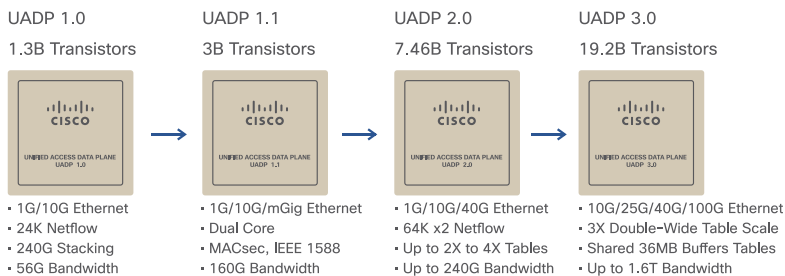
Unlike UADP 1.0, the UADP 1.1 has two ASIC cores and 3 billion transistors. The result is similar to using two UADP 1.0 chips in a single ASIC package. UADP 1.1 also provides higher aggregate bandwidth and performance of up to 160Gbps (80Gbps per core), as well as some new and updated micro-engines. 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 Catalyst 3850 and 3650 MultiGigabit and SFP+ switches use UADP 1.1.

Cisco UADP 2.0

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

The Cisco UADP 2.0 is a dual-core 28nm ASIC 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 (using purpose-built SDM templates), giving the option to deploy the same device in multiple network areas, as discussed in the chapter [Campus network design](#).

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



Cisco UADP 2.0 ASICs have three variants: UADP 2.0, 2.0 XL and 2.0 mini.

Both 2.0 and 2.0 XL have the same architecture, but the UADP 2.0 bandwidth, table scale and overall performance has been optimized for business-critical access layer switches. The Catalyst 9300 Series switches use UADP 2.0.

Cisco 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-generation Catalyst 9500 Series switches and the Catalyst 9400 Series Supervisor-1 and Supervisor-1XL use UADP 2.0 XL.

Cisco UADP 2.0 mini has a modified single-core architecture with an integrated quad core ARM CPU, and the bandwidth, table scale, overall performance and power consumption has been optimized for simple access layer switches. UADP 2.0 mini supports up to 80Gbps of inter-ASiC connectivity, for platforms with lower bandwidth requirements. Catalyst 9200 Series switches use UADP 2.0 mini.

TABLE Cisco UADP 2.0, 2.0 XL and 2.0 mini comparison

	Cisco UADP 2.0 mini	Cisco UADP 2.0	Cisco UADP 2.0 XL
MAC entries	16K-32K	32K	64K
Host routes	8K-10K	24K	48K
LPM routes	3K-4K	8K	64K
Multicast routes	1K	8K	16K
QoS entries	1K	5K	18K
ACL entries	1K	5K	18K
NetFlow entries (per ASiC)	16K	64K	128K
SGT entries	2K	8K	16K
Buffers	6MB	16MB	32MB
Inter-ASiC bandwidth	80G	240G	720G

Cisco 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. New

technologies are appearing every day and driving new requirements for network performance and scale.

The Cisco UADP 3.0 is a dual-core 16nm ASIC with 19.2 billion transistors, to provide a significant increase of aggregate bandwidth up to 1.6Tbps. Cisco 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. In addition to increased bandwidth and performance, UADP 3.0 also incorporates several new improvements that make it the ideal ASIC for campus core and distribution layer switches.

Cisco UADP 3.0 has larger memory tables and greater reprogramming flexibility, with larger shared packet buffers (36MB) to support the interface speed increases. It also has 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 second generation, high-performance, Catalyst 9500 Series switches and the Catalyst 9600 Series Supervisor-1 use UADP 3.0.

TABLE Cisco UADP 2.0 XL and 3.0 comparison

	Cisco UADP 2.0 XL (distribution SDM template)	Cisco UADP 3.0 (core SDM template)
MAC entries	64K	32K
Host routes (IPv4/v6)	48K/24K	212k/212K
LPM routes (IPv4/v6)	64K/32K	
Multicast routes (IPv4/v6)	16K/8K	32K/32K
QoS entries	18K	16K
ACL entries	18K	27K
NetFlow entries (per ASIC)	128K	64K
SGT entries	16K	32K
Buffers	32MB	36MB
Inter-ASIC bandwidth	720G	1600G

Cisco IOS XE

Cisco IOS evolution

The history of the Cisco Internetwork Operating System (Cisco IOS) goes back to Cisco's first product, the AGS multi-protocol routers launched in 1986. At the time Cisco 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 Cisco IOS and over the last 30 years and as the industry has gone through different transitions, Cisco IOS has evolved into a more feature rich OS.

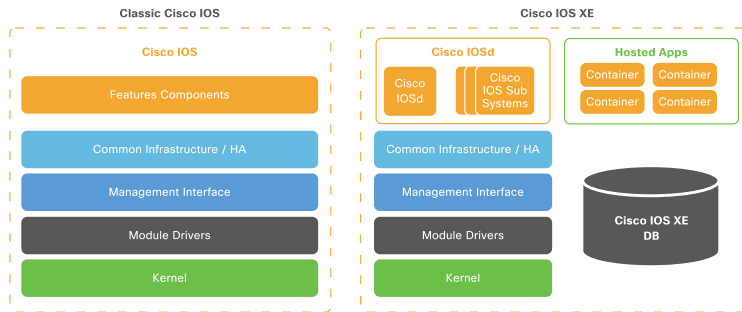
Over time, Cisco IOS software has branched out into many different versions for different products. Meanwhile, the Cisco product portfolio has also expanded into various kinds of switches and routers. Purpose-built network areas have evolved, such as data centers and service providers and new operating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

A few years ago, Cisco introduced Cisco IOS XE, designed to restructure the monolithic code infrastructure of Cisco IOS into a more modular and modern software architecture. With Cisco IOS XE, the OS is 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. Cisco IOS now runs as a modular process on top of the Linux kernel (known as Cisco 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 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 Cisco IOS to Cisco IOS XE comparison



The latest Cisco IOS XE software addresses several key customer needs:

- A common OS for enterprise networks
- Rapid introduction of new features and technologies
- A secure, trustworthy OS to protect the network
- Modularity and high availability
- Programmability and automation

Considering these customer needs, Cisco IOS XE 16.1 was first introduced on Catalyst 3850 / 3650 Series switches, which use the programmable Cisco UADP 1.0 ASIC. Since then, several other enterprise network platforms, including Catalyst 9000 switches, have also adopted Cisco IOS XE for its software flexibility and scalability.

This unified OS software release brings multiple advantages:

- Consistent experience across platforms
- Fewer software images to manage

- Faster certification of software features
- 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 family of switches have taken this one step further. The entire Catalyst 9000 switching family runs on the same code base and same image release. This provides for a faster delivery of innovation, along with consistent feature behavior, bringing added value by simplification of software image selection, deployment, and use.

Note Catalyst 9200 Series has a different binary based on optimised version of Cisco IOS XE called Cisco IOS XE Lite.

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 containers or virtual machines (VMs) for hosting Cisco and third-party applications.

The Cisco IOS XE architecture has three significant enhancements:

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

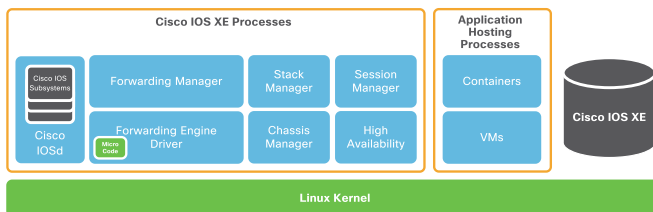
Modular OS

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

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

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

DIAGRAM Modular Cisco IOS XE



Cisco IOS XE Database

Previously, Cisco 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 Cisco 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 Cisco IOS XE.

This standard Cisco IOS XE database makes system data easier to express as data models. Cisco 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 the chapter [Programmability and automation](#).

Containers and VMs

Cisco IOS XE supports containers and VMs hosting capability on Catalyst 9000 switches. Please refer to the chapter [Application hosting](#).

Cisco IOS XE Lite

With Cisco IOS XE being modular, feature sets not applicable on a simple branch (such as MPLS, BGP and ETA) could be removed from operating system libraries. Cisco IOSd and bash were optimized to further reduce the footprint in Cisco IOS XE Lite. Even with a 50% decrease in image size, core benefits of Cisco IOS XE are maintained and all modern features such as streaming telemetry, failure isolation and status recovery are available.

Cisco IOS XE benefits

With the modern ever-changing software-defined environment, it is imperative that the OS software foundation of Catalyst 9000 family of switches 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. The modularity, standard database, object-based models, and containers of Cisco IOS XE 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 solutions.

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 Cisco IOS faster compared to bundle mode. Install mode is the recommended mode, and advanced high-availability features such as ISSU, patching, and xFSU 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 solution, Catalyst 9000 switches verify the

authenticity of the platform, prevent malicious code execution, establish run-time defenses, and secures communication. For more information, please refer to the chapter [*Trustworthy solutions*](#).

High availability

Overview

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

- **High availability on the Catalyst 9200 Series** - stacking, power
- **High availability on the Catalyst 9300 Series** - stacking, power, fast software upgrade, extended fast software upgrade
- **High availability on the Catalyst 9400 Series** - chassis, supervisors, power, ISSU
- **High availability on the Catalyst 9500 Series** - power, StackWise Virtual, ISSU
- **High availability on the Catalyst 9600 Series** - chassis, supervisors, power, StackWise Virtual, ISSU
- Graceful insertion and removal
- Cisco IOS XE patching

The Catalyst 9000 family of 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 Catalyst 9000 switch series utilizes SSO and NSF in slightly different ways for high availability.

Catalyst 9200 Series high availability

Catalyst 9200 Series switches deliver access layer redundancy with features such as StackWise-160/80 and power supply redundancy.

Catalyst 9200 Series switch stacks

Fixed configuration Cisco Catalyst 9200 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 Lite on Catalyst 9200 Series switches support stacking individually among C9200 and C9200L models for up to eight members but mixed stacking between them is not supported. They are physically connected in a ring with special stacking cables connected to the back of each switch using their stacking ports.

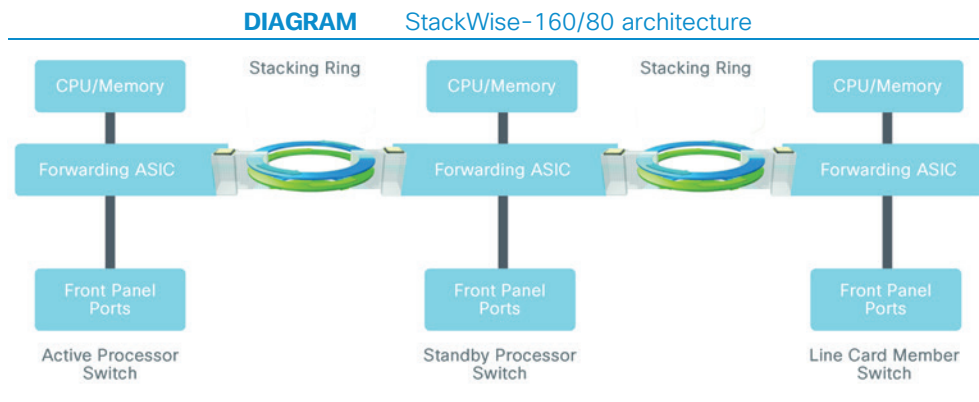
Catalyst 9200 Series switch stacks deliver deterministic and non-blocking switching performance for up to 416 ports. The switching performance delivers hardware-accelerated, integrated network services such as:

- PoE up to 15.4W
- PoE+ up to 30W
- Quality of service (QoS)
- Access control lists
- Flexible NetFlow

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

StackWise-160/80 architecture

Catalyst 9200 Series switches enable stacking using a stack-ring fabric known as either StackWise-160 or StackWise-80. StackWise-160 is supported on C9200 switch models with the support of 160 Gbps stack bandwidth. StackWise-80 is supported on C9200L switch models with the support of 80 Gbps stack bandwidth. The fabric consists of two counter-rotating rings (40/20 Gbps/ring), and the system's throughput is a function of the aggregated throughput of these rings (80/40 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.



StackWise-160/80 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-160/80 SSO support

Catalyst 9200 Series 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-160/80 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
Services	QoS, ACL, PBR, NetFlow, port security

Power redundancy

Cisco Catalyst 9200 Series switches have two power supply bays with the support of hot-swappable power supplies. The power supplies operate in two modes based on the PoE and non-PoE models. PoE models support 1+1 combined mode, where system and PoE power is shared by both power supplies. If one of the 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:

- Low priority ports
- High priority ports
- Switches

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

Non-PoE models support 1:1 Redundant mode where the second power supply is completely redundant. In the event of active power supply failure, the redundant power supply becomes active immediately.

Catalyst 9300 Series high availability

Catalyst 9300 Series switches deliver access layer redundancy with features such as StackWise-480, StackPower, fast software upgrade (FSU) and extended FSU.

Catalyst 9300 Series 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 9300 Series switches support mixed stacking between any Catalyst 9300 Series 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 Series switch stacks deliver deterministic and non-blocking switching performance for up to 448 ports. The switching performance delivers hardware-accelerated, integrated network services such as:

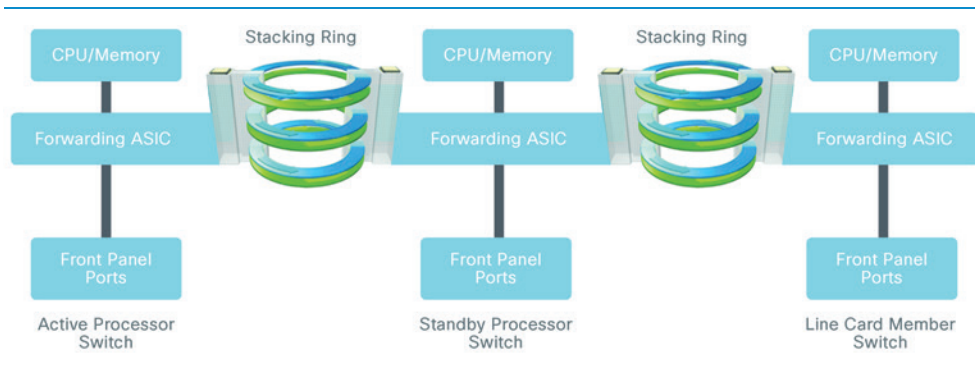
- PoE up to 15.4W
- PoE+ up to 30W
- Universal PoE (UPOE) and 802.3bt up to 60W
- Quality of service (QoS)
- Access control lists
- Flexible NetFlow

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

StackWise-480 architecture

Catalyst 9300 Series 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 Series 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, Flexible 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

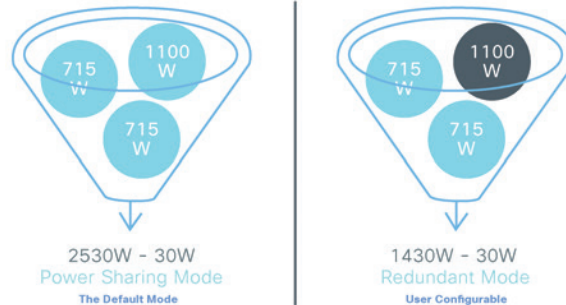
Cisco StackPower has two modes of operation: shared and redundant.

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:

- Low priority ports
- High priority ports
- 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 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 Series 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 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.

Extended fast software upgrade

Extended FSU provides a mechanism to upgrade and downgrade the software image by segregating the control plane and data plane update. It updates the control plane by leveraging the NSF architecture with a flush and re-learn mechanism to reduce the impact on the data plane, allowing less than 30 seconds of traffic impact during the upgrade cycle.

Further, it can also be used to manage switch reloads without performing a software upgrade with less than 30 seconds of traffic impact. Extended FSU is not supported on a StackWise-480 configuration.

Catalyst 9400 Series high availability

Cisco Catalyst 9400 Series switches provide several features to minimize outages:

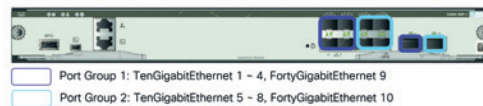
- Dual supervisor stateful switchover with non-stop forwarding
- Supervisor engine uplink redundancy
- StackWise Virtual
- In-service software upgrade
- Power supply redundancy
- Power priority

Dual supervisor SSO with NSF

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.

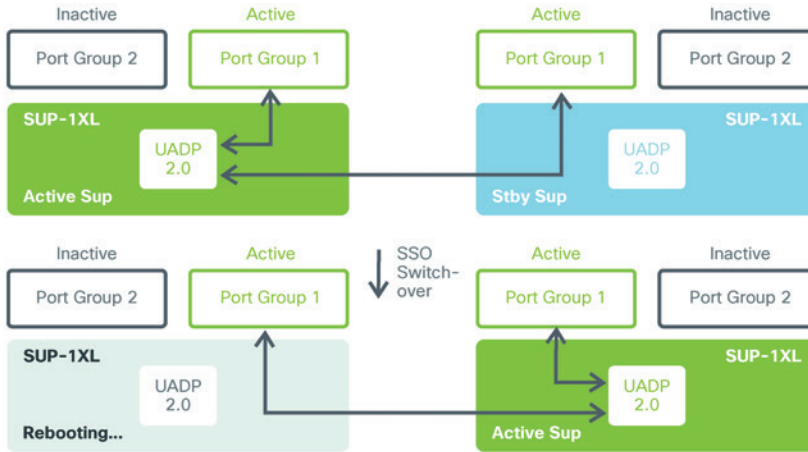
Each port group supports a maximum of 40 Gbps of traffic. The supervisor provides flexibility of using those 40 Gbps with either (4) 10 GE interfaces or (1) 40 GE interface.

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 Series switch 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

StackWise Virtual

StackWise Virtual feature allows the merging of two physical switches together into a single, logical switch. StackWise Virtual also supports in-service software upgrade (as described in the later section).

ISSU with dual supervisors

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 Series 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

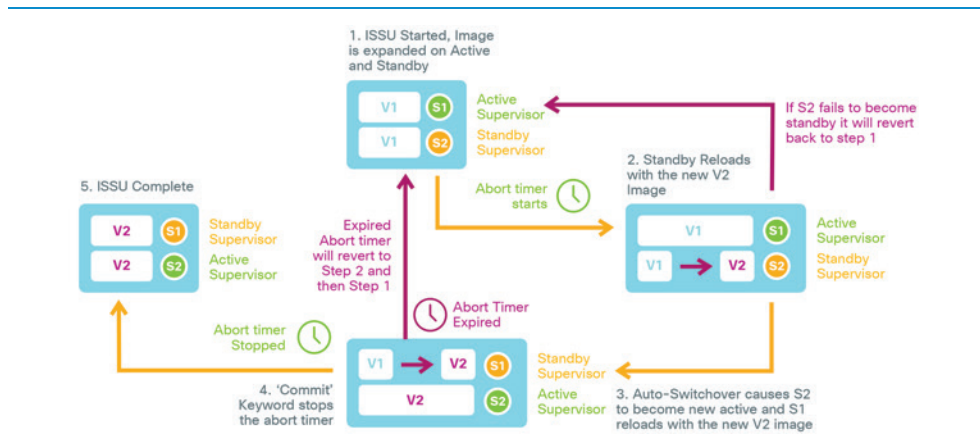
ISSU process

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

Catalyst 9400 Series switches support 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.

Catalyst 9500 Series high availability

Cisco Catalyst 9500 Series switches provide several features to minimize outages:

- StackWise Virtual
- ISSU with StackWise Virtual
- Power supply redundancy

StackWise Virtual

StackWise Virtual feature allows the merging of two physical switches together into a single, logical switch (as described in the later section).

ISSU with StackWise Virtual

StackWise Virtual supports in-service software upgrades to provide sub-second convergence during software upgrades and downgrades (as described in the later section).

Power redundancy mode

Cisco Catalyst 9500 Series switches have two power supply bays. The power supplies can operate in a 1+1 redundant mode, providing protection against a single power supply failure. If one of the active power supplies fails, the backup power supply is activated.

Catalyst 9600 Series high availability

Cisco Catalyst 9600 Series switches provide several features to minimize outages:

- Dual supervisor stateful switchover with non-stop forwarding
- StackWise Virtual
- In-service software upgrade
- Power supply redundancy
- Power priority

Dual supervisor SSO with NSF

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.

StackWise Virtual

StackWise Virtual feature allows the merging of two physical switches together into a single, logical switch. StackWise Virtual also supports In-service software upgrade (as described in the later section).

ISSU with dual supervisors

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 9600 Series 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

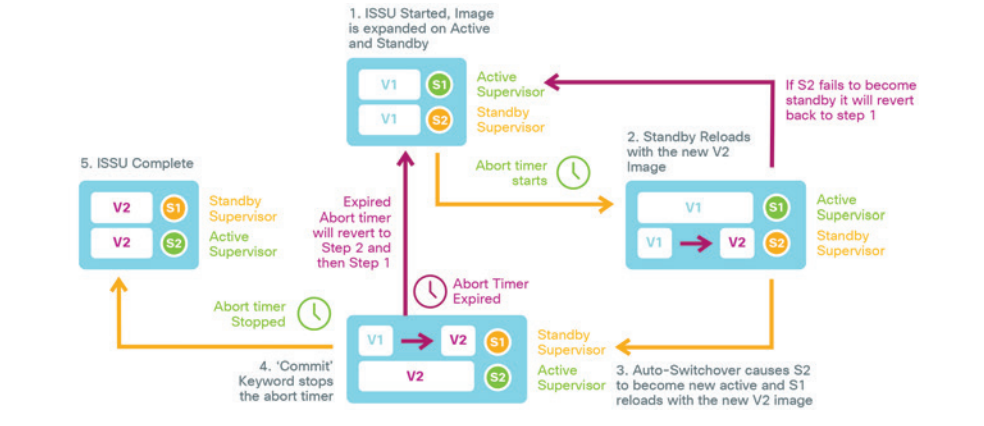
ISSU process

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 9600 Series 6-slot chassis has four power supply 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 N+1 configuration. N+1 provides protection against a single power supply 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.

Power priority

Catalyst 9600 Series switches support power priority for line card slots. If the system requires more power than the available system power, due to 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 powers off line cards starting from the bottom slots and then works its way up to the top.

StackWise Virtual

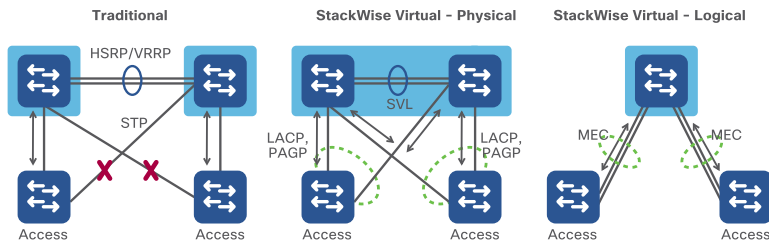
The StackWise Virtual feature on Cisco Catalyst 9400, 9500 and 9600 Series switches allow the merging of two physical switches together into a single, logical switch. The two switches operate as one; they share the same configuration and forwarding state. This is analogous to the virtual switch system (VSS) feature in the previous generation of the Catalyst switching line.

Note Catalyst 9400, 9500, and 9600 Series 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 (NSF) and in-service software upgrade (ISSU), 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

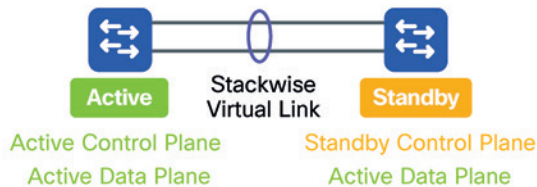


StackWise Virtual architecture

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:

- StackWise Virtual link
- Dual-active detection link
- 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 supported 10/25/40/100G 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 automatically to resolve the dual-active state upon the recovery of SVL Links. Dual-active detection can be deployed using either a dedicated link on any 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.

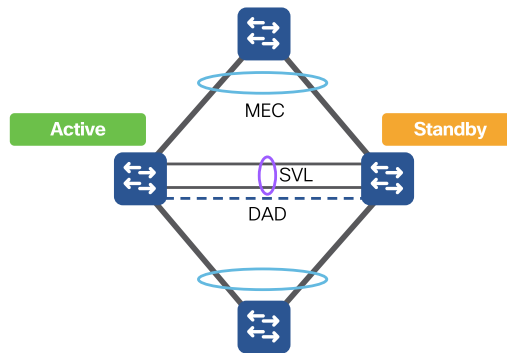
Note

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

ISSU with StackWise Virtual

StackWise Virtual supports in-service software upgrades. ISSU helps network administrators avoid a network outage when performing an Cisco IOS XE software upgrade on a StackWise Virtual pair. ISSU supports upgrades, downgrades, and rollbacks.

ISSU prerequisites

Before an ISSU can be performed on a Catalyst 9400, 9500 or 9600 Series 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

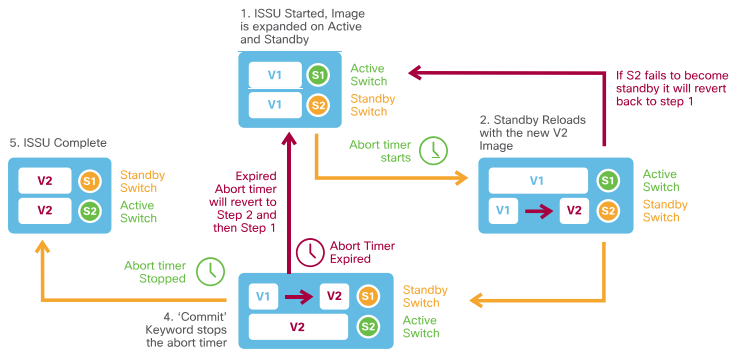
ISSU process

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

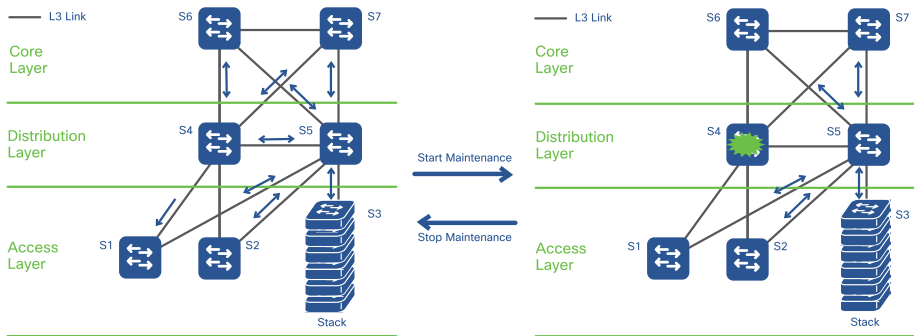
During the switchover, there will be a sub-second traffic convergence

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. GIR is not available on Cisco Catalyst 9200 Series switches.

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 the 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 switch 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
BGP	Withdraw BGP Advertised routes	Re-insert routes into BGP
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.

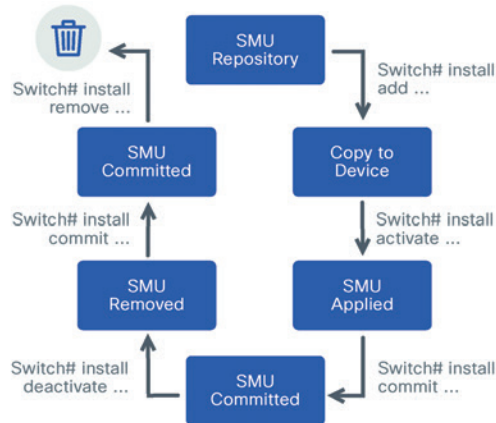
Catalyst 9000 switches running Cisco IOS XE also provide system-generated snapshots to record the state of a switch before and after maintenance. Snapshots are useful for verifying that a switch is operating correctly when it returns to service.

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 one. 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 Software maintenance upgrade workflow

Two types of patching are available:

- **Cold patching** - Application of a cold patch requires reloading the switch in order to fix the issue. Cisco IOS XE Lite switches support cold patching only.
- **Hot patching** - Hot patches do not require a system reload. Cisco IOS XE switches only need 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 family of switches. Security features from prior Catalyst switching families have been carried forward into the Catalyst 9000 family of switches.

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 family of switches, 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 packets

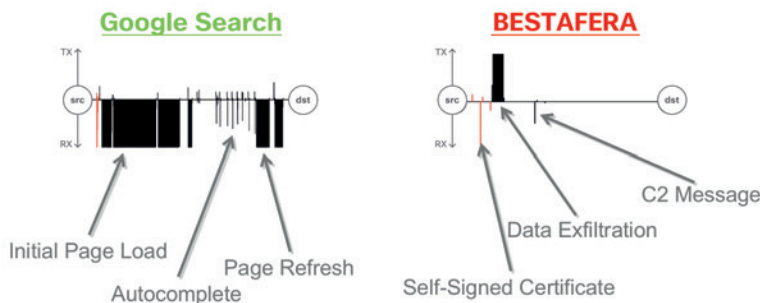
Initial data packets (IDP) are the first packets between two hosts. In the case of ETA, they occur during the handshake used to set up a secure session. Even for an encrypted

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 inter-arrival time between packets in a flow. Length is measured in bytes and inter-arrival 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 cognitive threat 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 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 Catalyst 9000 switches

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 Cisco 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 solutions

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 family 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 switch 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 core components, Catalyst 9000 switches provide 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. The process 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 encryption

Why was MACsec encryption 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 through 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 link encryption 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).

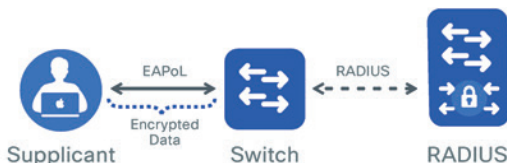
Today, MACsec encryption is enabled between switch-to-switch, user or server host-to-switches, or router-to-switches, to ensure data is protected.

Topologies

Host-to-switch - usually these are front panel ports

A host is required to run Cisco Anyconnect client to carry out software encryption / decryption on the host. Today Anyconnect supports up to AES 128 Bit encryption.

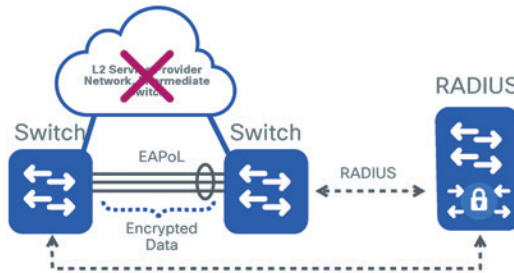
DIAGRAM Host-to-switch MACsec encryption



Host-to-switch topology:

- does not support manual key exchange
- supports 802.1X 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 on a Catalyst 9400 Series chassis

DIAGRAM Switch-to-switch MACsec encryption

Switch-to-switch topology:

- supports manual key exchange configuration
- does not support 802.1X 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

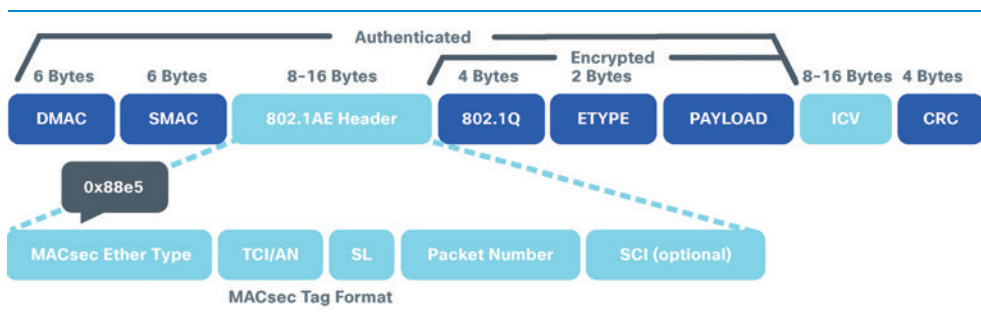
MACsec encryption requires hardware support on the switches to process encryption / decryption at line rate. The Catalyst 9000 switching family has an integrated block in the UADP ASIC which does line rate MACsec encryption and decryption at any speed.

TABLE MACsec encryption support by platform

	MACsec	Catalyst 9200 Series	Catalyst 9300 Series	Catalyst 9400 Series	Catalyst 9500 Series	Catalyst 9600 Series
Switch-to-switch	128 Bits SAP	Line rate	Line rate	Line rate	Line rate	Line rate
	128 Bits MKA	Line rate	Line rate	Line rate	Line rate	Line rate
	256 Bits MKA	n/a	Line rate	Line rate	Line rate	Line rate
Host-to-switch	128 Bits MKA	Line rate	Line rate	Line rate	Line rate	Line rate
	256 Bits MKA	n/a	Line rate	Line rate	Line rate	Line rate

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 Catalyst 9000 switches

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 <http://cs.co/9009E5dtV> on Catalyst 9000 switches, 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 <http://cs.co/9006E5dtA>, please refer to cisco.com.

Buffers and queues

The Cisco UADP ASIC provides a shared packet buffer complex (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 Mini	UADP 2.0	UADP 2.0 XL	UADP 3.0
Buffer MB per core	N/A	8	16	36
Buffer MB per ASIC	6	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, Catalyst 9000 switches use dynamic threshold scaling (DTS).

In DTS the hardware buffer is split into multiple segments:

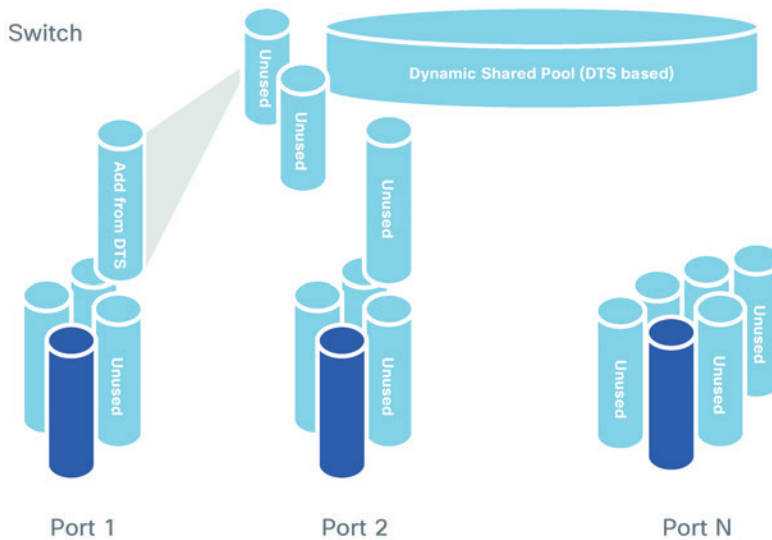
- Ingress buffers (11%) are used for packets scheduled towards the stack and ASIC interfaces
- Egress stack buffers (25%) are used to receive traffic from the stack ports
 - The buffer is sized to accommodate up to eight stack members

- 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:

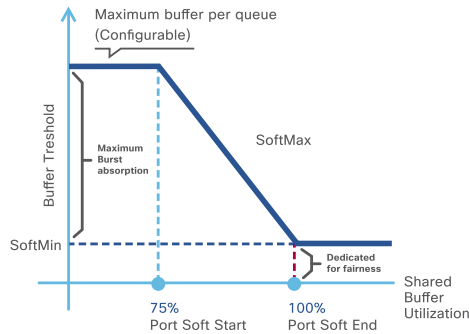
- Per-port buffers are split into dedicated (hard) and shared (soft) categories:
 - Shared buffers are good for absorbing packet bursts
 - Dedicated buffers are good for predictable performance
- Dedicated buffers are used first, followed by the shared buffers:
 - Dynamic threshold algorithm (DTA) is 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 <http://cs.co/DTSBuffer> configuration, please refer to cisco.com

Hardware queues

By default, Catalyst 9000 switches implement 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 switch 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 UADP 2.0, UADP 2.0XL and UADP 3.0-based platforms add the weighted random early discard (WRED) algorithm to the 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 Cisco UADP ASIC

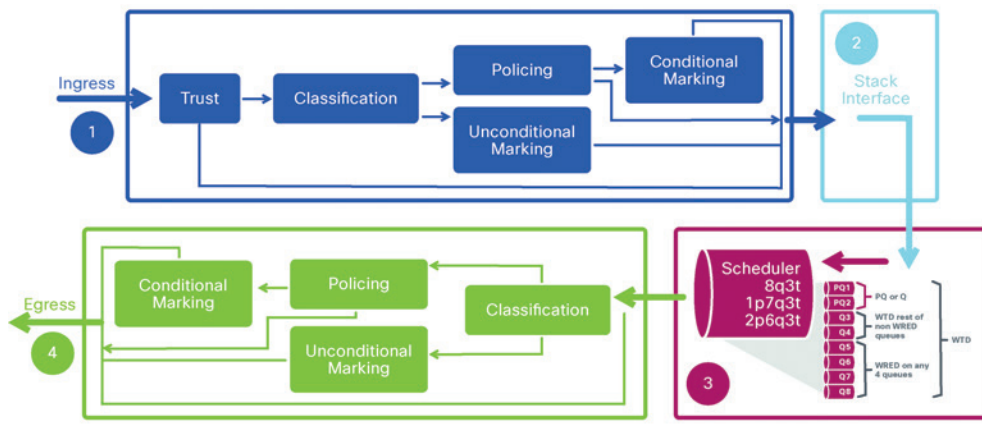
This section describes how the various QoS and queuing processes are applied to a packet as it traverses the UADP.

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 Cisco UADP ASIC QoS and queuing packet 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 queuing 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

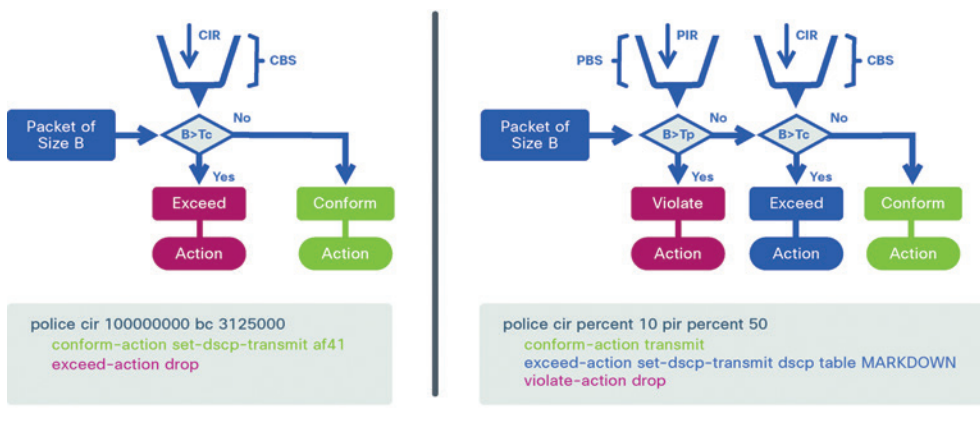
- COS
- EXP
- TCP/UDP Ports
- NBAR protocols
- Per VLAN

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.

Catalyst 9000 switches supports both 1R2C and 2R3C policers.

DIAGRAM Catalyst 9000 switch policer types



Cisco UADP ASIC interconnect queuing

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

The Cisco 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 the ASIC interconnect.

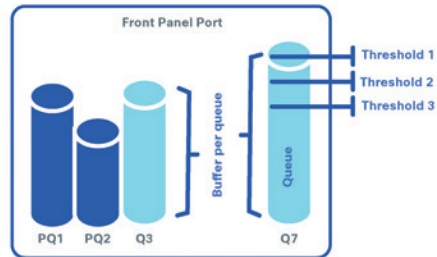
Egress queuing and scheduling

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. Catalyst 9000 switches 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 Catalyst 9000 switching ports, each queue has three configurable threshold values.

Shapers

Shapers typically delay excess traffic using a buffer or queuing 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 on Catalyst 9000 switches.

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.

Catalyst 9000 switches support four HQoS combinations:

- Port shaper
- Aggregate policer
- Per-port, per-VLAN policy
- 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 family of 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 <http://cs.co/QoSSV>, 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 Cisco 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.

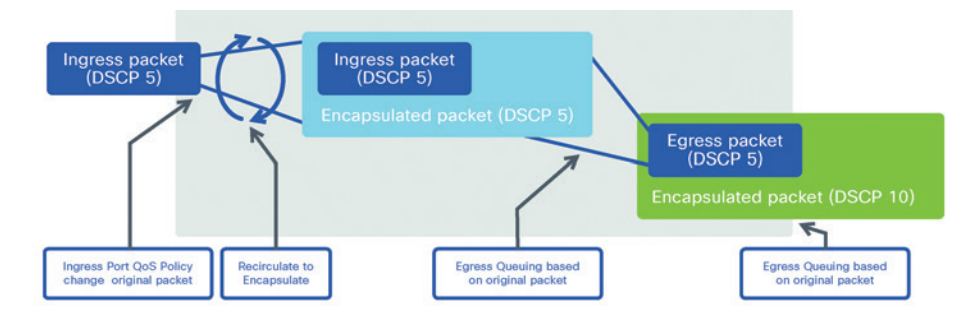
GRE and VXLAN encapsulation in the Cisco UADP ASIC

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



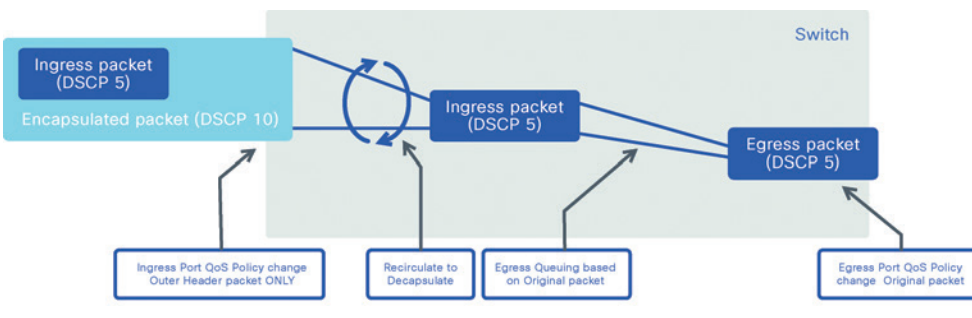
Note The queuing actions are applied before egress policing / marking actions. Refer to the earlier sections for more detail on egress QoS behavior in UADP.

GRE and VXLAN decapsulation in the Cisco 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 <http://cs.co/QoS GRE> (RFC 2784 <http://tools.ietf.org/html/rfc2784>), please refer to cisco.com.

For more information on QoS on VXLAN overlays <http://cs.co/QoS VXLAN> (RFC 7348 <http://tools.ietf.org/html/rfc7348>), 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 two modes used to map DSCP or ToS to EXP within Catalyst 9000 switching family:

- **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 copied to IP ToS/DSCP
Pipe	MPLS EXP set by SP QoS policy	MPLS EXP may be changed by SP QoS policy	Original ToS/DSCP preserved (egress queuing based on MPLS EXP)
Short-pipe (Note: not supported on Catalyst 9000 switches)			Original ToS/DSCP preserved (egress queuing based in ToS/DSCP)

For more information on MPLS QoS <http://cs.co/QoS MPLS> (RFC 5462 <http://tools.ietf.org/html/rfc5462>), please refer to cisco.com.

Application visibility and 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.

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
- 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 (such as identifying traffic as either business relevant or irrelevant)

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 Series switch stack, NBAR2 detection runs on each stack member, scaling the solution as more members are added.

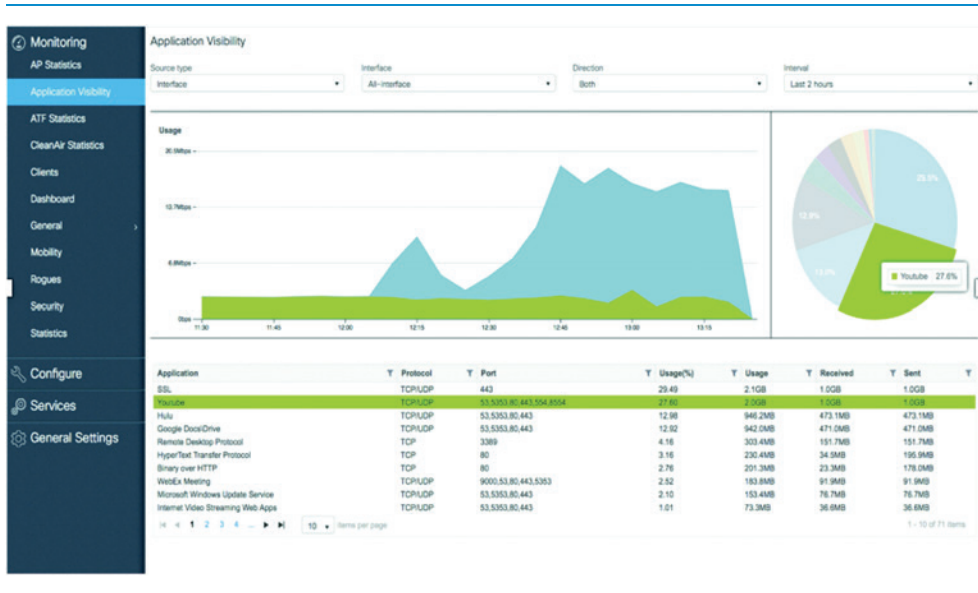
NBAR2 can be activated for application visibility explicitly on an interface by enabling protocol discovery, and for application policy by attaching a QoS policy that contains a match protocol classifier.

Application monitoring

Monitoring of applications is available through the CLI or WebUI.

- The CLI provides accumulated statistics over time.
 - The CLI command is: `show ip nbar protocol-discovery top-n`
- The switch WebUI provides statistics for up to the last 48 hours in 5-minute granularity

DIAGRAM Catalyst 9000 AVC monitoring via WebUI



- 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 Series switching 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 control 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 (such as business relevant marked traffic)

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 Cisco UADP ASIC.

IoT

Overview

The Internet of things (IoT) 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), Cisco DNA Service for Bonjour, and application hosting. Please see the chapter [*Application hosting*](#).

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, connecting 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. PoE offers convenience for device attachment and use, along with resiliency 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 9200, 9300 and 9400 Series 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 Series switching 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 Cisco UPOE (Universal PoE). UPOE provides for up to 60W of PoE power. As of Sept 2018, IEEE 802.3bt was ratified and approved. The standard introduces two additional types of power: 60W (Type3) and 90W (Type4).

Both Catalyst 9300 Series switches and 9400 Series switching line cards support 802.3bt (Type 3) power options. Cisco provides PoE, PoE+, and 802.3bt 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

New innovations - fast PoE, perpetual PoE and 2-Event Classification

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.

With 2-Event classification, Catalyst 9000 switches can negotiate power faster without using protocols CDP / LLDP by using Layer 1 negotiation. This allows for end-points to negotiate power in under a second, leading to faster onboarding of devices on the network.

Audio video bridging - AVB

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

The AVB standard is supported on Catalyst 9300 and Catalyst 9500 (UADP2.0) switches, as well as the Catalyst 3850 and 3650 Series 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 switches offers support for this important set of standards.

Cisco 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 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

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, Cisco DNA Service for Bonjour scales this capability up to network-wide application that runs on Cisco 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 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 service discovery gateway and Cisco DNA Service for Bonjour (running on Cisco 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 of switches, 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 switches add usability improvements on the hardware side such as RFID, blue beacon, 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 Catalyst 9000 switches. Cisco has front-facing ultra-high frequency (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 of switches is integrated with passive serialized global trade item number (SGTIN)-198 bit encoded RFID tags. On Catalyst 9400 and 9600 Series, every component such as supervisor engines, line cards, power supplies and fan tray, has RFID tags. Since Catalyst 9000 switching 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 <http://cs.co/C9KRFID> on cisco.com for more details.

Blue beacon

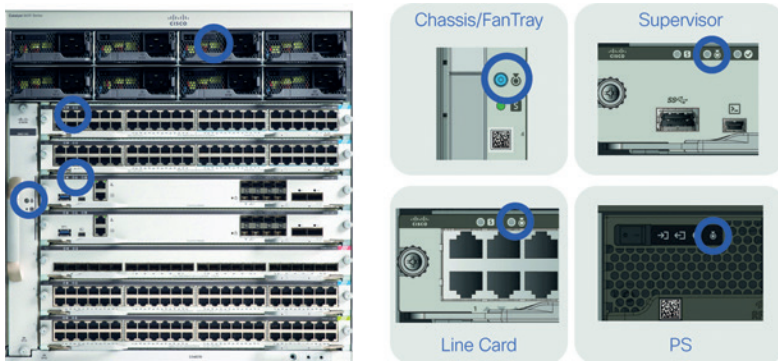
When troubleshooting, configuring, or moving equipment in a large enterprise, it might be difficult to locate a particular device. To help identify the device, Cisco has placed a blue beacon LED on the Catalyst 9000 family of switches. This blue beacon can be turned on and off, either via command-line 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 9200 and 9300 Series switches, blue beacon LEDs are located on both front and back. When Catalyst 9200 and 9300 Series switches are stacked, the blue beacon for every member can be managed individually.

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

On Catalyst 9400 and 9600 Series switches, the fan tray, supervisors, line-card modules, and power supplies (Catalyst 9400 only) each have their own addressable blue beacon.

DIAGRAM Blue beacon on Catalyst 9400 Series switches



Bluetooth console

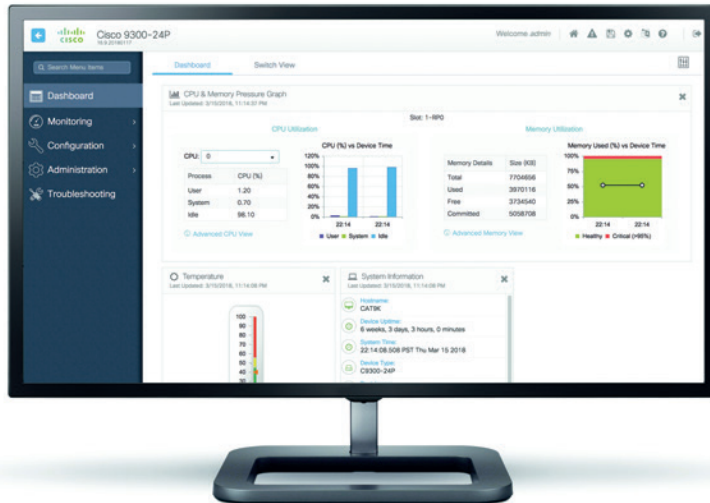
Network administrators often use a console cable for onsite configuration. However, a console cable has distance limitations and is not convenient to use.

Catalyst 9000 switches have 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 (GUI) 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 switching WebUI



Flexible templates

Flexible templates give Catalyst 9000 switches an option to be positioned in different roles in the network design. The Cisco UADP ASIC has capabilities to optimize its hardware table for specific network roles within 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.

Catalyst 9200 and 9300 Series switches support two templates optimized for the access layer deployments.

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

TABLE SDM templates available per platform

Template Name	Access / Distribution	Core	SD-Access	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

Shown below are the SDM templates available on Catalyst 9600 Series and Catalyst 9500 High Performance Series switching models using UADP 3.0.

TABLE SDM templates available per platform

Template Name	Distribution	Core	SD-Access	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

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 Cisco 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 Cisco 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. For more information, refer to <https://cs.co/programmabilitybook>

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 DHCP option 150)
- Python script name (using DHCP 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 Cisco 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 of 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 Cisco DNA Center. The entire process is fully automated.

Open programmable device APIs

A brief history of device automation

Over the years, customers have tried 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. This is because 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.

Application programming interface (API)

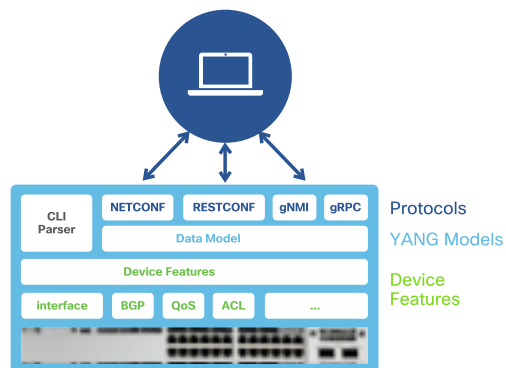
In computer programming, an API is a set of subroutine definitions, protocols, and tools for building application software. A good API makes it easier to develop a computer program by providing all the building blocks, which are then put together by the programmer. Software vendors and customers are 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 such as 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, gNMI, and gRPC can be used to interface with external automation software toolchains.

DIAGRAM Open programmable device APIs



Data models

In programmability, a data model is one of the most important components of open programmable APIs. It 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 <https://tools.ietf.org/html/rfc6020>.

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

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

- **Configuration data models** - the set of writable data required to transform a system from its initial default state into its current state. A configuration model instructs the device to do something and can be mapped to the running configuration of a Cisco IOS XE device
- **Operational data models** - the set of read-only data status information and statistics on a device. An 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 data models** - are defined by standards bodies such as IEEE and IETF or by working groups such as OpenConfig. An advantage of using open models is that they are common across 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 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.

Cisco publishes YANG data models in a common GitHub Repository (<https://github.com/YangModels/yang/tree/master/vendor/cisco>) which is updated every time a new release becomes available.

In addition to downloading data models from the GitHub repository, IOS XE on Catalyst 9000 switches 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 the chapter *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:

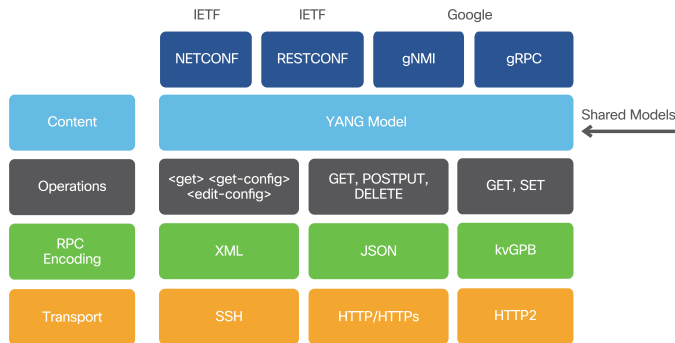
- **PYANG** (<https://github.com/mbj4668/pyang>) - Python library to validate, navigate and automatically build documentation
- **YangExplorer** (<https://github.com/CiscoDevNet/yang-explorer>) - open source tool to easily start exploring the model and automate small tasks
- **YDK** (<https://github.com/CiscoDevNet/ydk-py>) (YANG Development Kit) - software development kit that provides an abstraction layer of the API modeled in YANG
- **Yang Catalog** (<https://yangcatalog.org/>) - 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 switches 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

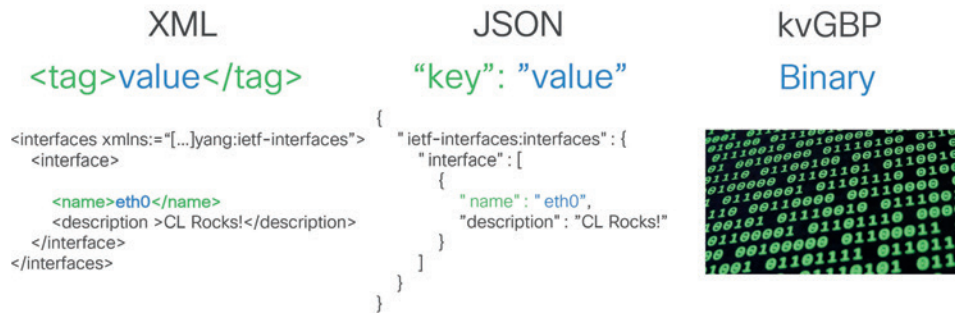
All of the API protocols use Remote Procedure Calls (RPCs) for communication. A client running on an external server sends an RPC message to the network device and the network device replies with an RPC-reply message. The data payload of both RPC and RPC-reply is encoded using a format defined by the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

- **Key-value Google Protocol Buffers** (kvGPB) is the open source version of the encoding format defined by Google. It has 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. The gRPC protocol is available for the transport of configured subscriptions, and can be used with the *yang-push* stream.

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 (<https://tools.ietf.org/html/rfc6241>) defines the basics of NETCONF, and introduced the concept of optional extensions. Examples of NETCONF extensions are the YANG data model language RFC-6020 (<https://tools.ietf.org/html/rfc6020>) and Notifications RFC-5277 (<https://tools.ietf.org/html/rfc5277>).

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 IOS XE config terminal mode

RESTCONF

RESTCONF is a network configuration protocol (like NETCONF) defined by the IETF in the RFC-8040 <https://tools.ietf.org/html/rfc8040>.

Very often RESTCONF is confused with REST. Representational State Transfer (REST) is a framework for client-server communications. While RESTCONF is a network configuration protocol for accessing YANG models, built using the REST framework.

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)

gRPC

gRPC (gRPC Remote Procedure Calls) is an open source remote procedure call(RPC) system initially developed by Google. It uses HTTP/2 for transport, Protocol Buffers as the interface description language, and provides features such as authentication, bidirectional streaming and flow control, blocking or nonblocking bindings, and cancellation and timeouts. It generates cross-platform client and server bindings for many languages.

gRPC supports several built in features inherited from http2, such as compressing headers, persistent single TCP connections, cancellation and timeout contracts between client and server making it efficient, accurate and optimized for low-latency communication.

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 a scalable and efficient manner with a low impact on the network device. Read more in the chapter [Open programmable device APIs](#).

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 switches is based on the standard IETF PUB/SUB (publication/subscription). There are two flavors of MDT: dial-in and dial out.

To receive streaming data from the device, Catalyst 9000 switches support dial-in as well as dial-out. For dial-in, 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. Dial-in is supported using NETCONF and gRPC.

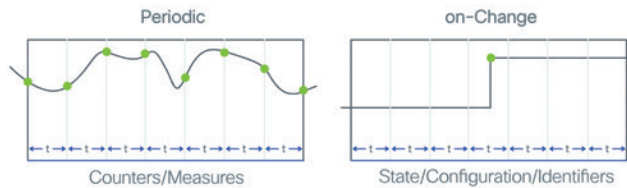
Dial-out telemetry is pushed from the device to the collector based off a configured model. The subscriptions need to be statically configured on the network device using any of the available interfaces (CLI, APIs, etc.) and the device will open a session with the application. If the session goes down, the device will try to open a new session. The subscriptions automatically reconnect to the receiver after a stateful switchover. Dial-out is supported through gRPC. The table below highlights the key differences between a dial-in and dial-out mechanisms.

Dial-in (dynamic)	Dial-out (static or configured)
Telemetry updates are sent to the initiator/subscriber.	Telemetry updates are sent to the specified receiver/collector.
Life of the subscription is tied to the connection (session) that created it, and over which telemetry updates are sent. No change in the running configuration is observed.	Subscription is created as part of the running configuration; it remains the device configuration till the configuration is removed.
Dial-in subscriptions need to be re-initiated after a reload, because established connections or sessions are killed during stateful switchover.	Dial-out subscriptions are created as part of the device configuration, and they automatically reconnect to the receiver after a stateful switchover.
Subscription ID is dynamically generated upon successful establishment of a subscription.	Subscription ID is fixed and configured on the device as part of the configuration.

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 known as the ELK stack) as well as messaging brokers such as Apache Kafka, ActiveMQ and RabbitMQ.

Scripting

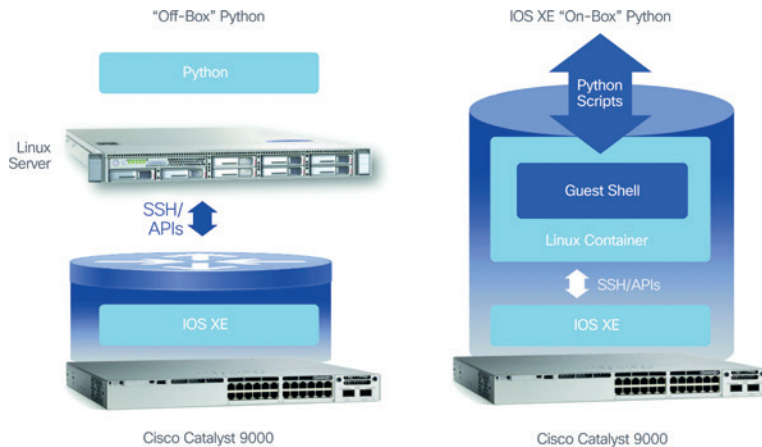
Scripts have been used for ages to quickly and easily automate small tasks. Over the years, some scripting languages have become more popular than others, such as Perl, TCL, and JavaScript. In recent years, Python became the most popular scripting language. One of the main reasons for its ever-growing 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 switch 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 switch 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, Catalyst 9000 switches are 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.

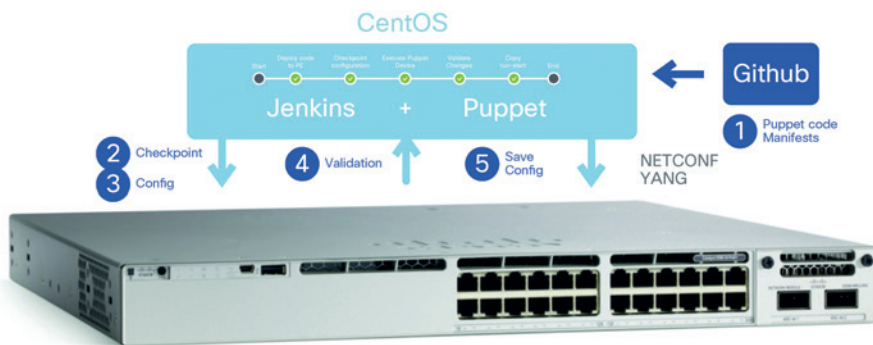
Ansible is based on an agent-less architecture which has been a key factor in its success. Ansible already supports Cisco devices, such as Catalyst 9000 switches running Cisco IOS XE, and provides intent-based modules to configure interfaces, VLANs, VRFs, users, etc as well as CLI and NETCONF-based configurations.

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 switches CI/CD with Puppet

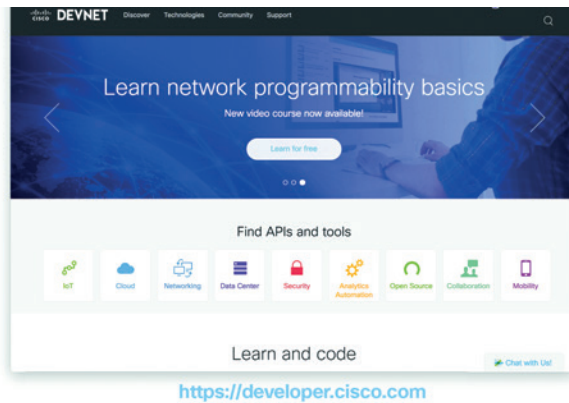


CI/CD is a new configuration management practice created to reduce the time needed to deploy new changes in production and to increase the confidence that changes will be successful. It enables this through an automated testing process and integration with source control systems for an easy rollback 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 <https://developer.cisco.com/> 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.

Application hosting on Catalyst 9000 switches opens up new innovation opportunities by converging network connectivity with a distributed application runtime environment, including hosting applications developed by partners and developers.

For maximum flexibility, and in isolation from the main operating system, the Cisco application framework on Catalyst 9000 switches manages applications containerized in Docker containers.

Cisco IOS XE 16.12.1 introduces native Docker container support on Catalyst 9000 switches. This will enable customers to build their own applications or bring their own applications. Developers can easily create, deploy and move the applications across Docker hosts. Developers don't have to reinvent the wheel by rewriting the app every time there is an infrastructure change. Once written for Docker, it will work well with any infrastructure with docker environment.

Type of applications

Two types of networking applications are supported on Catalyst 9000 switches:

- **Control plane** - applications that access the management interface
- **Data plane** - applications that access the front-panel data ports

The new front panel port "AppGigabitEthernet" interface is introduced on Catalyst 9300 Series switches with Cisco IOS XE 16.11.1 release. It is an internal hardware data port which is hardware-switched to the front-panel data ports. With this

"AppGigabitEthernet" interface, Catalyst 9300 Series switches can support both control plane and data plane applications.

Catalyst 9400 Series switches currently support control plane applications only, and will support data plane applications in the future release. Cisco Catalyst 9600 and 9500 Series switches only support control plane applications and do not have "AppGigabitEthernet" interface capability.

	Catalyst 9300 Series	Catalyst 9400 Series	Catalyst 9500 Series	Catalyst 9600 Series
Mgmt port	Supported	Supported	Supported	Supported
Front panel (Trunk/VLAN)	Supported	Future release	NA	NA

Application hosting management

Hosted applications can be managed through command line interface (CLI) and Cisco IOS XE WebUI until Cisco IOS XE 16.11.1 release.

With Cisco IOS XE 16.12.1 and later software releases, the Cisco DNA Center will provide a centralized user interface to deploy and manage the entire lifecycle of the applications and will be the Cisco recommended solution.

Hardware resources

Cisco IOS XE running on the Catalyst 9000 switches 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 Cisco IOS XE run-time processes ensuring both its integrity and performance.

Applications must reside in one of the external SSD storage (USB 3.0 or M2 SATA) options provided by the Catalyst 9000. Applications have no access to the internal device flash storage, which is reserved for Cisco 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 3.0 Storage (GB)	M2 SATA Storage (GB)
Catalyst 9300 Series	2	1 x 1.8GHz	120	N/A
Catalyst 9400 Series	8	1 x 2.4GHz	N/A	240/480/960
Catalyst 9500 Series (with UADP 2.0 XL)	8	1 x 2.4GHz	120	N/A
Catalyst 9500 Series (with UADP 3.0)	8	1 x 2.4GHz	N/A	240/480/960
Catalyst 9600 Series	8	1 x 2.0GHz	NA	240/480/960

Note Catalyst 9200 does not support application hosting.

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 switch 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 switched 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 controller

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 plan
- **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 switches are designed to address one or more design model, all share a common set of capabilities:

- **Catalyst 9600 Series switches** - modular form factor, with 1/10/25G SFP modules and 40/100G QSFP modules
 - Built with sufficient ASIC bandwidth to support medium to large-sized campus core and distribution design

- **Catalyst 9500 Series switches** - 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 switches** - modular form factor, with 10/100/1000/mGig RJ45, 1/10/25G 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 switches** - fixed, stackable form factor, with 10/100/1000/mGig RJ45, 1/10/25G SFP modules and 40G QSFP uplink ports
 - Built for high-density user access with sufficient ASIC bandwidth and table scalability to support business critical branch and medium to large campus access designs
- **Catalyst 9200 Series switches** - fixed, stackable form factor, with 10/100/1000/mGig RJ45 and 1/10/25G SFP module ports
 - Built for simplified branch deployment with optimized ASIC bandwidth and table scalability to support small to medium campus and branch access designs

↳ the bottom line

Catalyst 9000 switches are the most flexible, highest scaling and resilient platforms built for a new era of intent-based networking.

While most of this book focuses on the specific details of individual Catalyst 9000 switches, the following chapters are dedicated to how an end-to-end Catalyst 9000 switching 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.

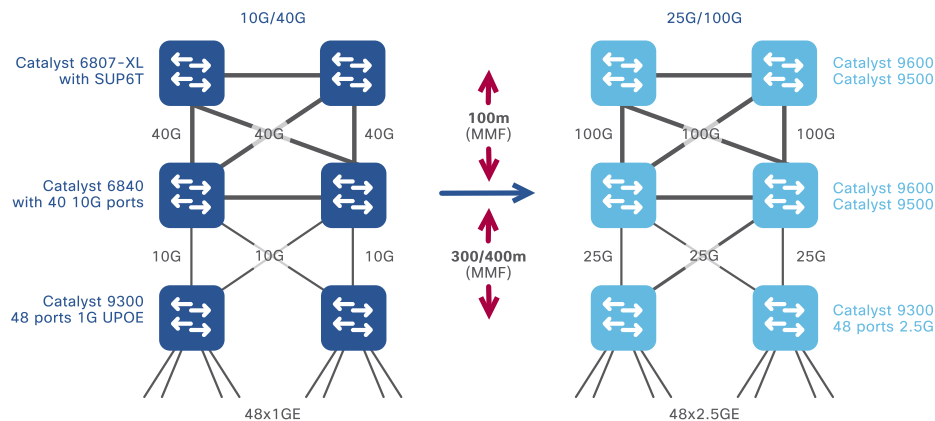
Use Case 1: Speed transition with similar cable distances

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

Innovative Cisco SFP-10/25G-CSR-S optics with traditional cable lengths of up to 300/400m over OM3/4 MMF achieve speeds of 25G.

DIAGRAM Speed transition with similar cable distance



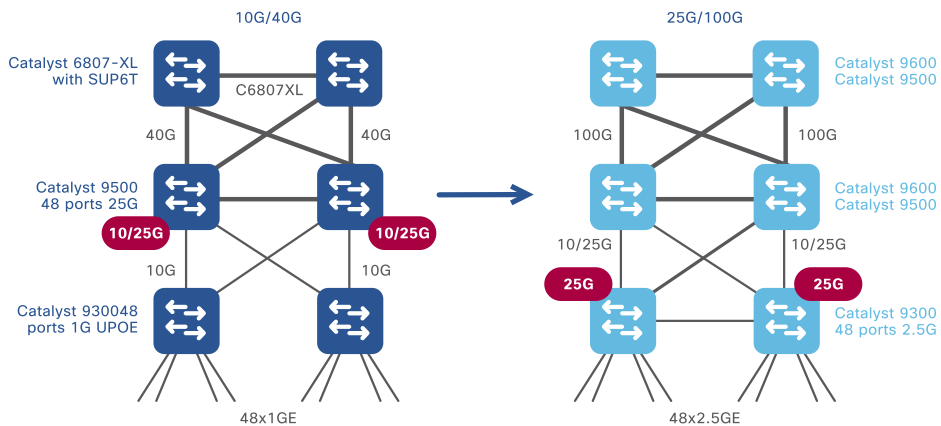
Use Case 2: Speed migration with dual-rate optics

The Cisco 25G portfolio provides backwards 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

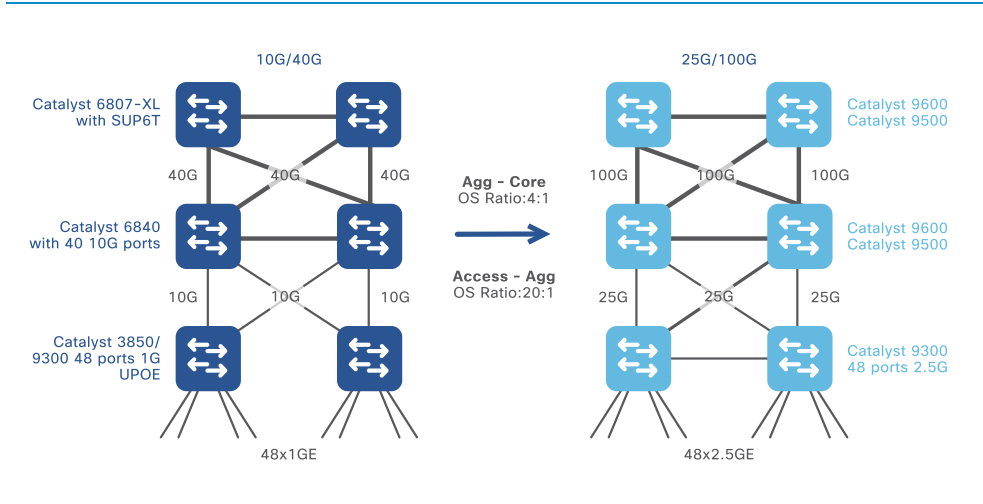


Use Case 3: Speed transition with similar oversubscription ratios

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:

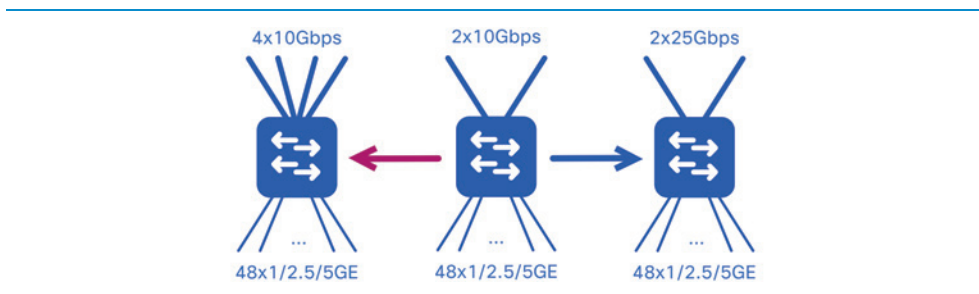
- More uplinks can be added (with link-aggregation or ECMP)
- Replace existing cabling and transceivers to support higher speeds

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.

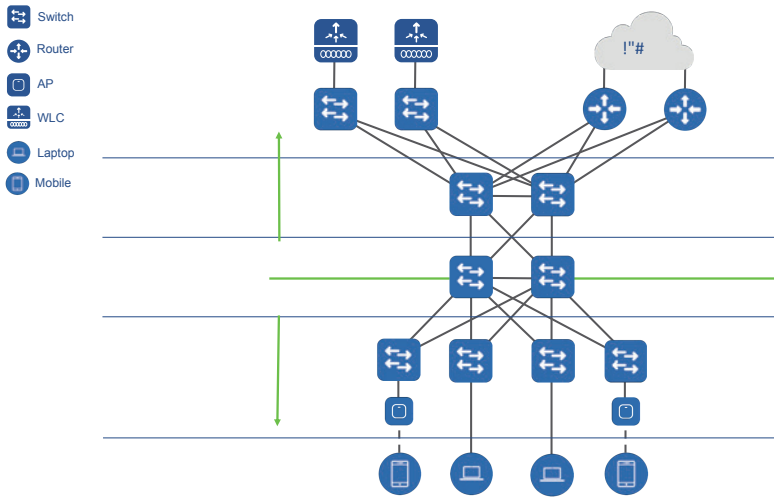
DIAGRAM Comparing higher speed vs. load sharing



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

Catalyst 9500 Series 10 / 40G models are based on UADP 2.0 XL 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 and all Catalyst 9600 Series 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

Catalyst 9500 and 9600 Series switches are optimized for the core and distribution layers.

Catalyst 9200, 9300 and 9400 Series switches employ flavors of UADP 2.0 ASIC forwarding engines, and implement the access SDM template to set the optimal ASIC table allocation for switched access role. Catalyst 9200 and 9300 Series switches, having a stackable fixed-form factor is ideal for providing switch-level redundancy. Catalyst 9400 Series switches, by contrast, is a modular switch providing the highest levels of network availability by delivering supervisor, line card and power redundancy in a chassis. All three switching series provide capabilities to support small, medium and large wiring closets.

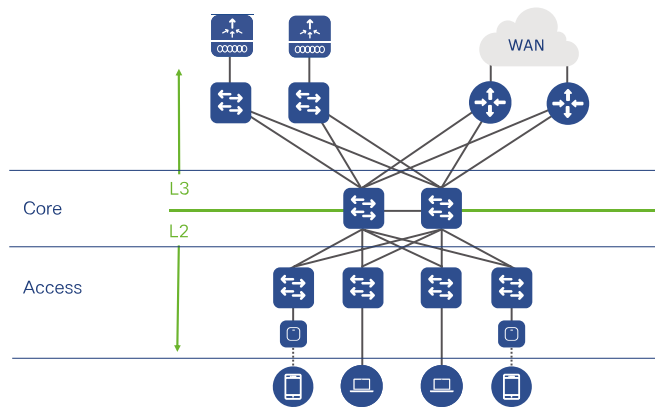
└ the bottom line

Catalyst 9200, 9300 and 9400 Series switches 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.

Catalyst 9500 Series switches use the Cisco UADP programmable ASIC in a fixed-form factor design, and Catalyst 9600 Series switching models use the Cisco UADP programmable ASIC in a modular design. The difference between them is a matter of

port speed, density and in-chassis high availability, 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**

Catalyst 9500 and 9600 Series switches are optimized for the collapsed core layer.

There are no changes to the access layer in this design, and the recommendations for using Catalyst 9200, 9300 and 9400 Series switches in these roles still apply.

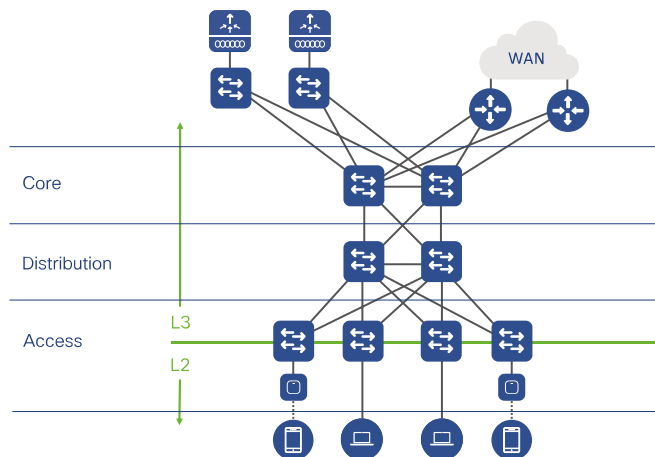
↳ **the bottom line**

Catalyst 9200, 9300 and 9400 Series switches 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 switching platforms remains the same. Catalyst 9600 Series switches serve best as a modular core or distribution and Catalyst 9500 Series switches serve best as fixed-form distribution or core, and Catalyst 9200, 9300 and 9400 Series switches are best suited as access layer switches. In addition to an SDM template that optimizes Catalyst 9400 Series switches as an L2 switch, there is also a template that reprograms them to be optimized for L3 access.

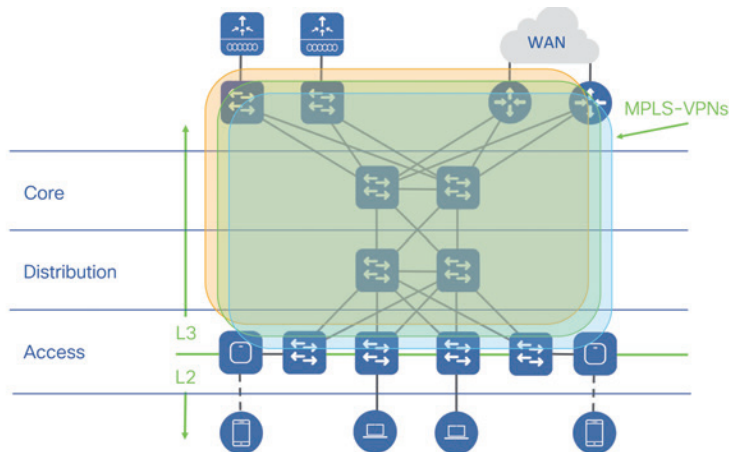
└ the bottom line

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 the chapter [*Software-Defined Access*](#).

Catalyst 9000 switches running Cisco IOS XE supports campus MPLS. The same recommendations apply as with a routed access network, with the addition of MPLS feature support. Catalyst 9600 and 9500 switches are designed for core and distribution layer services, but Catalyst 9300 and 9400 Series switches are best suited as access devices. Catalyst 9200 Series switches do not support MPLS.

↳ the bottom line

All business-critical Catalyst 9000 switches are optimized for Campus MPLS networks.

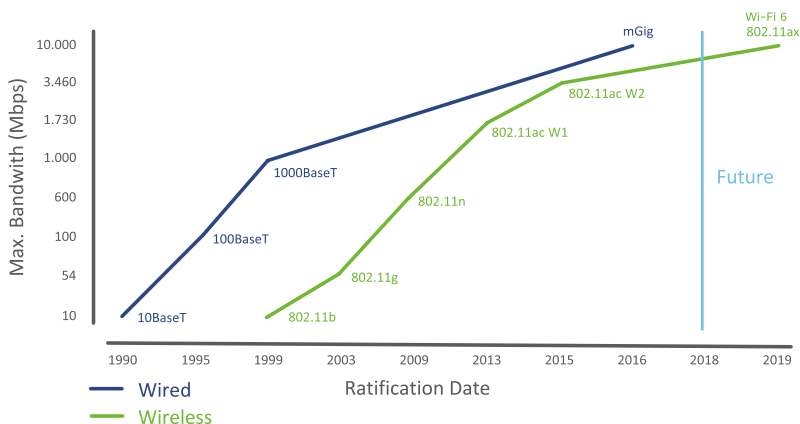
Campus wireless

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 Wi-Fi 6 (802.11ax) and 802.11ac Wave 2 now provide multigigabit 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.

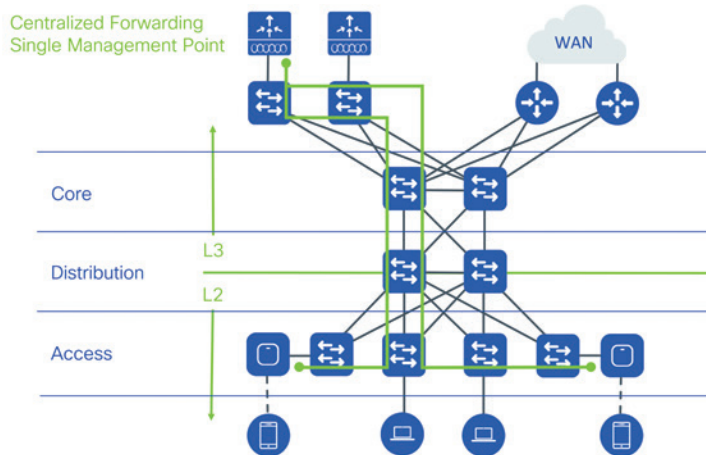
↳ the bottom line

As Wi-Fi 6 and 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 10 Gbps links, and APs are connecting a small number of clients at 10 or 100 Mbps rates. With the adoption of Wi-Fi 6 and 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**

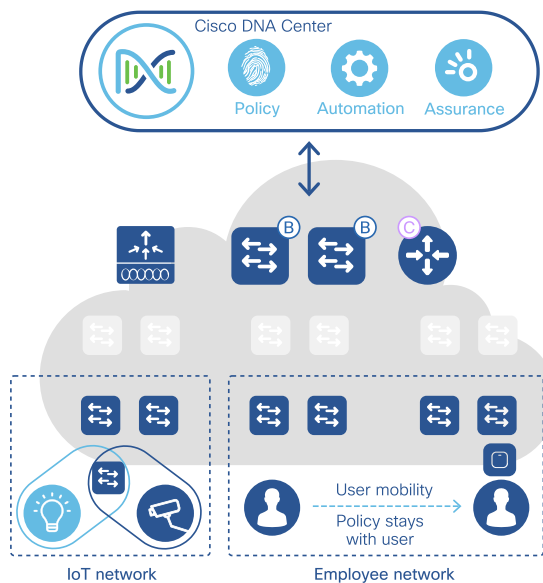
All Catalyst 9000 switches are optimized for campus wireless networks.

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. For more information, refer to <https://cs.co/wirelessbook>.

Software-Defined Access

Cisco's Software-Defined Access (SD-Access) solution is a programmable network architecture that provides software-based policy and segmentation from the edge of the network to the applications. Cisco SD-Access is implemented via Cisco Digital Network Architecture Center (Cisco 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 Cisco 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.

Catalyst 9000 switches 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

Cisco SD-Access wireless

Cisco 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** - Cisco 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 the Cisco Wireless e-book <http://cs.co/sdabook>.

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

Catalyst 9500 and 9600 Series switches can be configured to use the SD-Access 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

Catalyst 9500 and 9600 Series switching platforms are optimized for the SD-Access fabric border role.

Catalyst 9400 Series switches with Sup-1XL and Sup-1XL-Y can be configured to use the Cisco SD-Access SDM template which provides the optimal ASIC table allocation for an SD-Access fabric edge role. All Catalyst 9200, 9300, and 9400 Series switching platforms support 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 9200, 9300 and 9400 Series switching platforms are optimized for the Cisco SD-Access fabric edge role.

Appendix

References

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

Overview of the Cisco Catalyst 9000 switching family:

<https://www.cisco.com/c/en/us/products/switches/catalyst-9000.html>

Overview of Cisco Catalyst 9000 switches:

<https://www.cisco.com/c/en/us/products/switches/catalyst-9200-series-switches/index.html>

<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>

<https://www.cisco.com/c/en/us/products/switches/catalyst-9600-series-switches/index.html>

Cisco Catalyst 9000 switching white paper:

<https://www.cisco.com/c/en/us/products/switches/catalyst-9200-series-switches/white-paper-listing.html>

<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>

<https://www.cisco.com/c/en/us/products/switches/catalyst-9600-series-switches/white-paper-listing.html>

Cisco Live On-Demand Library:

<https://www.ciscolive.com/global/on-demand-library.html?#/>

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 Series switching architecture](#)

[BRKARC-3873: Catalyst 9400 Series switching architecture](#)

[BRKARC-2007: Catalyst 9500 Series switching architecture](#)

[BRKCRS-3300: Cisco IOS XE: enabling the digital network architecture](#)

[BRKCRS-2810: Cisco SD-Access – A Look Under the Hood](#)

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

DevNet, the Cisco Developers Network:

<https://developer.cisco.com/>

<https://developer.cisco.com/site/ios-xe>

Catalyst 9000 Switches automation:

<https://github.com/CiscoDevNet/catalyst9k-network-automation>

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 Cisco IOS XE](#)

Acronyms

AAA - Authentication, Authorization and Accounting	CAGR - Compound Annual Growth Rate
ACK - acknowledgment	CAPWAP - Control And Provisioning of Wireless Access Points
ACL - Access Control List	Cisco DNA - Cisco Digital Network Architecture
AES - Advanced Encryption Standard	CDP - Cisco Discovery Protocol
AOC - Active Optical Cables	CEF - Cisco Express Forwarding
API - Application Programming Interface	CI/CD - Continuous Integration, Continuous Delivery
AQM - Active Queue Management	CIR - Committed Information Rate
AR - Augmented Reality	CLI - Command Line Interface
ARP - Address Resolution Protocol	CoS - Class of Service
ASIC - application-specific integrated Circuit	CPU - Central Processing Unit
AVB - Audio Video Bridging	CSMA/CD - Carrier Sense Multiple Access with Collision Detection
AVC - Application Visibility and Control	CTA - Cisco Trust Anchor
BGP - Border Gateway Protocol	CTA - Cognitive Threat Analytics
BOOTP - Bootstrap Protocol	CoA - Change of Authorization
BPDU - Bridge Protocol Data Units	DAC - Direct Attach Copper
BUM - Broadcast Unknown unicast Multicast	DAD - Dual-Active Detection Link
BYOD - Bring Your Own Device	DHCP - Dynamic Host Configuration Protocol
CAF - Cisco Application Framework	

DIY - Do-It-Yourself	ERSPAN - Encapsulated Remote Switched Port Analyzer
DMZ - Demilitarized Zone	ETA - Encrypted Traffic Analytics
DNS - Domain Name System	FIB - Forwarding Information Base
DNS-AS - DNS as Authoritative Source	FIFO - First In First Out
DPI - Deep Packet Inspection	FNF - Flexible NetFlow
DSCP - Differentiated Services Code Point	FPGA - Field Programmable Gate Array
DTLS - Datagram Transport Layer Security	FSU - Fast Software Upgrade
DTP - Dynamic Trunk Protocol	FTP - File Transfer Protocol
DTS - Dynamic Threshold Scheduler	GIR - Graceful Insertion and Removal
EAP - ECC - Error-Correcting Code	gNMI- google Network Management Interface
EAPoL - Extensible Authentication Protocol over LAN	GPE - Generic Protocol Extension
EARL - Encoded Address Recognition Logic	GPO - Group Policy Object
ECMP - Equal-Cost Multipathing	GRE - Generic Routing Encapsulation
ECN - Explicit Congestion Notification	gRPC - google Remote Procedure Call
EFC - Egress Forwarding Controller	Gbps - Gigabits per second
EGR - Egress Global Resolution	HA - High Availability
EIGRP - Enhanced Interior Gateway Routing Protocol	HQoS - Hierarchical QoS
ELLW - Enhanced Limited-Lifetime Warranty	HSRP - Hot Standby Router Protocol
EQS - Egress Queuing Scheduler	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

kvGPB - Key0Value Google Protocol Buffers

LACP - Link Aggregation Control Protocol

LAN - Local Area Network

LDP - Label Distribution Protocol

LED - Light-Emitting Diode

LISP - Locator/ID Separation Protocol

LSC - Locally Significant Certificates

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

NIST - National Institute of Standards and Technology	PSU - Power Supply Unit
NSF - Non-Stop Forwarding	PXE - Preboot Execution Environment
NSH - Network Services Headers	PnP - Network Plug and Play
NSO - Network Service Orchestrator	QSA - QSFP to SFP Adapter
OIR - Online Insertion and Removal	QSFP - Quad Small Form-factor Pluggable
OM - Optical Multimode	QoS - Quality of Service
onePK - One Platform Kit	RADIUS - Remote Authentication Dial-In User Service
OS - Operating System	RED - Random Early Discard
OSI - Open Systems Interconnection model	REST - REpresentational State Transfer
OSPF - Open Shortest Path First	RFC - Request for Comments
P2P - peer-to-peer	RFID - Radio-Frequency IDentification
PAgP - Port Aggregation Protocol	RNG - Random number generators
PBC - Packet Buffer Complex	RPC - Remote Procedure Call
PBR - Policy-Based Routing	RPM - RPM Package Manager
PDU - Protocol Data Unit	rpm - Revolutions per minute
PE - Provider Edge router	RSPAN - Remote Switched Port Analyzer
PHY - PHYsical layer	SAP - Security Association Protocol
PIR - Peak Information Rate	SATA - Serial AT Attachment
PMK - Pair-wise Master Key	SD-Access - Software-Defined Access
POE - Power over Ethernet	SDK - Software Development Kit
pps - packets per second	SDN - Software Defined Networking
	SDM - Switch Database Manager

SFP - Small Form-factor Pluggable	SSO - Stateful Switchover
SGACL - Scalable Group Access Control List	ST - Service Template
SGFW - Security Group Firewall	STP - Spanning Tree Protocol
SGT - Scalable Group Tag	SUDI - Secure Unique Device Identifier
SGTIN - Serialized Global Trade Item Number	SVL - Stackwise Virtual Link
SIP - Session Initiation Protocol	SW - Software
SKU - Stock Keeping Unit	SXP - SGT Exchange Protocol
SLI - Switch Link Interface	TAC - Technical Assistance Center
SMC - Stealthwatch Management Console	TCAM - Ternary Content Addressable Memory
SMF - Single Mode Fiber	TCO - Total Cost of Ownership
SMU - Software Maintenance Update	TCP - Transmission Control Protocol
SNMP - Simple Network Management Protocol	TFTP - Trivial File Transfer Protocol
SPAN - Switched Port Analyzer	TLS - Transport Layer Security
SPLT - Sequence of Packet Lengths and Times	TTL - Time to Live
SQS - Stack Queue Scheduler	ToS - Type of Service
SRAM - Static Random-Access Memory	UADP - Unified Access Data Plane
SRP - Spatial Reuse Protocol	UDLD - Unidirectional Link Detection
SSD - Solid-State Drive	UDP - User Datagram Protocol
SSH - Secure SHell	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

XPS - Expandable Power System

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

Jay Sharma
Jeff Meek
Gary M Davis
Ginger Liu

