Cisco Catalyst 6840-X Series

Contents

Introduction

Cisco's latest high-density switching platform, the Cisco Catalyst® 6840-X Series Switch, is designed for flexible deployment in enterprise campus environments. This 10G/40G-ready platform is ideal for customers who want to introduce premium 10-Gb services in small or midsize campus backbones. This switch is the fixed aspect of a new Cisco® Catalyst 6800 Series of multi-layer switching products with more than 3000 features, including full Multiprotocol Label Switching (MPLS), Virtual Private LAN Services (VPLS), Cisco Locator ID Separation Protocol (LISP), Virtual Switching System (VSS), and security group tag (SGT) support on every port. The switch offers industry-leading table sizes in its form factor with high 10-Gb port density and maturity of software code.

The Catalyst 6840-X offers the following four SKUs:

- 16 and 32 ports of 10-Gb Small Form Factor Pluggable and Small Form Factor Pluggable Plus (SFP/SFP+)
- 24 and 40 ports of 10-Gb SFP/SFP+ with two native uplink ports of 40-Gb Quad SFP (QSFP)

This new system can support a maximum system capacity up to 240 Gbps per chassis, (up to approximately 480 Gbps in a VSS). Figure 1 highlights the number of slots in a Catalyst C6840-X-LE-40G Switch.

Note: Any of these four SKUs can be mixed to form a VSS.

Figure 1. Cisco Catalyst C6840-X-LE-40G

Cisco Catalyst switches are built to make the most of backward compatibility, and provide an architectural foundation for next-generation hardware features and scalability. This platform also runs on the same architecture as the Cisco Catalyst C6880-X, and therefore offers stability with proven operating system software.

If you are already familiar with the Cisco 6500 Catalyst Series with Supervisor Engine 2T, Catalyst 6880-X Switch, and Cisco IOS® Software Release 15.1SY, then you are already familiar with the new Catalyst 6840-X Series.

This white paper provides an architectural overview of the new Catalyst 6840-X chassis, including system design, power, and cooling. It also provides all of the technical details necessary to fully understand this new chassis.

Chassis Overview

Chassis Highlights

This section briefly covers the highlights of the Catalyst 6840-X series chassis.

The Catalyst 6840-X is a **2-slot fixed chassis**, based on the DNA infrastructure and design of Catalyst 6800 Series Switches.

The Catalyst 6840-X chassis includes the following highlights:

- **Four SKUs to choose from:** Select a system that best fits your network scale needs
- **Front-to-back airflow:** The first chassis in Catalyst 6800 Series to offer this feature
- **40-Gb QSFP unlink ports:** 24-and 40-port models come with two native 40-Gb QSFP uplinks for 40-Gb connectivity
- **Smaller-footprint 2-rack unit (RU) form factor:** The Catalyst 6840-X series offer an extra space over the Catalyst 6840-X (5RU)
- **SFP management option:** Use either RJ45 or SFP for port management
- **Up to two "platinum-efficiency" 1100W AC/DC power supplies:** The Catalyst 6840-X series can support 1:1 power redundancy up to 1100W
- **New fan with four high-efficiency 11000-rpm fans:** Provide approximately 200 cfm, capable of cooling the whole system
- **Multi-rate optics:** All 10-Gb ports can also support different speeds of 10M/100M/1G

Other Highlights

- Supports a new multilayer switch feature card (MSFC) route processor complex
  - New Intel-based 2.0-GHz CPU and 4-GB of DDR3 memory
- Supports an improved policy feature card (PFC) forwarding controller
  - Based on PFC4-E (EARL8) with enhancements and fixes
- Supports a new Serial Gigabit Media Independent Interface (SGMII)-based switched Ethernet-out-of-bound channel (EOBC)
  - Dedicated programming channels (rather than bus-based)

Chassis Design

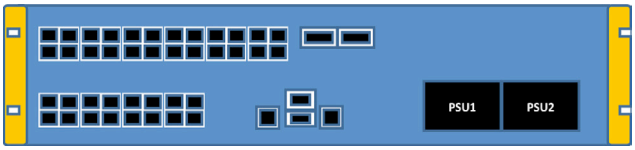
Customers expressed a need for the comprehensive software and hardware features of the Catalyst 6500 and 6800 platforms, but in a smaller (fixed) form factor. This concept was the genesis for the new Catalyst 6840-X chassis design. It is only **2 RUs high**, and chassis dimensions are (H x W x D) 3.5 x 17.35 x 21.52 inches (8.89 x 44.07 x 54.66 cm). Figure 2 shows a mechanical view of the Catalyst C6840-X-LE-40G.

This section briefly covers the high-level system design of the new Cisco Catalyst 6840-X series chassis. Additional details can be found in later sections.

The Catalyst 6840-X chassis comes with **two power supply unit (PSU)** slots, and supports **AC** or **DC** power inputs. There is a **fan tray** with **variable-speed fans in the back of chassis**. The chassis also has built-in RFID for inventory management, Blue Beacon LED for chassis-level identification, and a bi-color LED for system status.

The chassis also includes a front-panel RJ45 console, USB type B connector for a USB console to the route processor (RP), RJ45 and SPF management ports, and a USB 2.0 host port for storage.

Figure 2. Mechanical View of the Catalyst C6840-X-LE-40G



Chassis Power

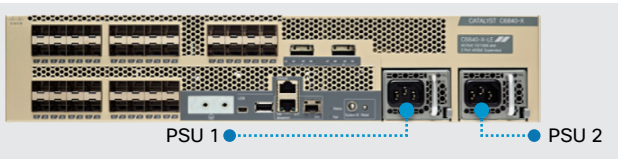
The Catalyst 6840-X chassis supports up to two 1100W AC or DC small form-factor power supply units (PSUs), for a total system capacity of up to 1100W.1100W and 750W PSU are available to use with the system. However, the 40-ports model works only with 1100W PSU.

The system supports either one PSU operating in non-redundant mode or two PSUs operating in redundant mode. Power supplies support for both AC and DC units and supports full online insertion and removal (OIR) capabilities when using redundant mode. This allows the user to easily insert and remove PSUs without loss of service.

Two separate PSUs are located on the bottom right corner (facing) of the chassis.

The PSUs are numbered (1 or 2) from left to right (Figure 3).

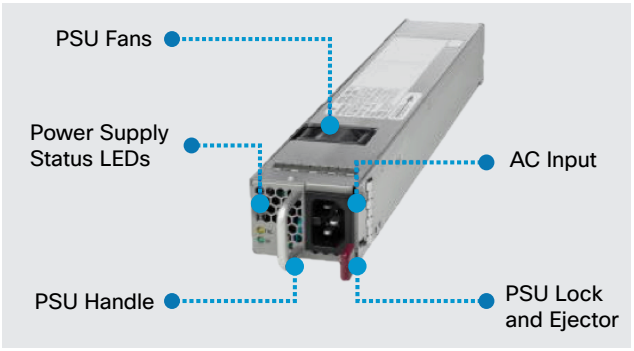
Figure 3. Numbering of Catalyst 6840-X Power Supplies



Power Supply Unit

The maximum power output per 1100W PSU is 1100W at 110V – 220V input and 750W with 750W PSU (Figure 4). Each PSU has been rated as platinum-efficient for greater than 90 percent power efficiency at a 100 percent load. Each PSU has a power holdup time of approximately 20 milliseconds (msec or ms) at a 100 percent load and fully supports OIR. Each PSU comes with front-to-back variable-speed cooling fans.

Figure 4. Catalyst 6840-X 1100W AC PSU



Each PSU has a push-release lock and manual ejector lever for simple, highly secure OIR. Each PSU supports multiple LEDs to determine component and power input and output status (Figure 5).

Figure 5. Catalyst 6840-X AC PSU LEDs

LED	Color	Status	Description
IN		Solid	Input OK
IN		Blinking	Under-Current
Fault		Solid	Malfunction

Note: PSU fault conditions are:

- 5V out of range
- Output stage OT
- Fan fault
- OR-ing fault (output is less than bus voltage)
- OC shutdown
- OT shutdown
- OV shutdown
- Input stage OT
- Fault-induced shutdown occurred
- Thermal sensor fault
- Vout out of range
- Boost Vbulk fault (AC power supply only)

Chassis Cooling

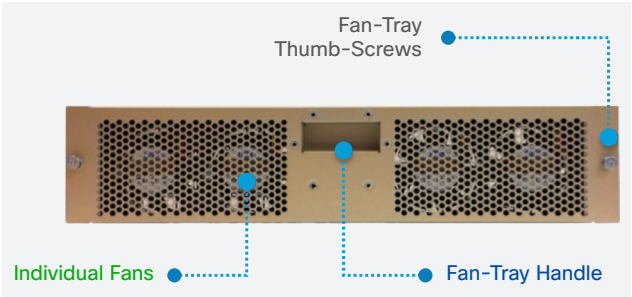
The enhanced fan tray provides enough cooling capacity, of up to 200 cfm, in order to properly cool the whole chassis. This is an enhanced fan-tray assembly with a redundant variable speed of 6000 – 11000 rpm per fan. It supports hardware failure of up to one individual fan; the remaining fans will automatically increase their rpm to compensate and maintain sufficient cooling.

Enhanced Fan Tray

There are a total of four 80-mm fans in the fan tray. Each fan has been rated as high efficiency, for greater than 80 percent power efficiency at a 100 percent load. Each fan has an acoustic noise rating of 67 dB (per ISO-7779) when operating at 14° to 142°F (-10° to 61°C).

The fan tray supports OIR for a minimum of 120 seconds (depending on the ambient temperature). Individual fans cannot be replaced; the whole fan tray must be replaced.

Figure 6. Catalyst 6840-X-LE Fan Tray



The fan tray has a built-in multipoint-control-unit (MCU) controller to receive (for example, fan speed) and send (for example, the temperature status) messages to the baseboard, through the inter-integrated circuit (I2C) bus. The MCU monitors the thermistor mounted on the fan tray and sets fan speeds based on temperature ranges set by the MCU.

The fan tray uses a thumb-screw lock mechanism and fan-tray handle, for simple and highly secure OIR (Figure 6). The fan tray supports front-panel LED to determine fan-tray status.

Figure 7. Catalyst 6880-X Fan-Tray LEDs

LED	Color	Status	Description
FAN		Solid	Fan-Tray OK
FAN		Solid	Fan-Tray Fault

Fan-tray fault conditions are:

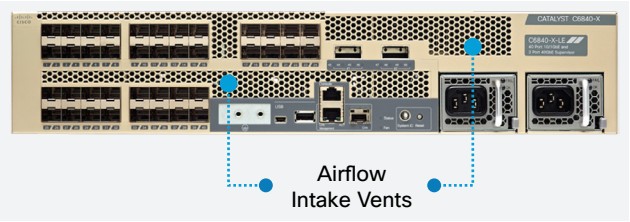
- One or more fans stop spinning
- 42V hot-swappable failures
- 3.3V goes away (from supervisor card)
- +5V supply fails
- PWM signal 0% duty cycle
- Thermistor failure
- System temperature out of range

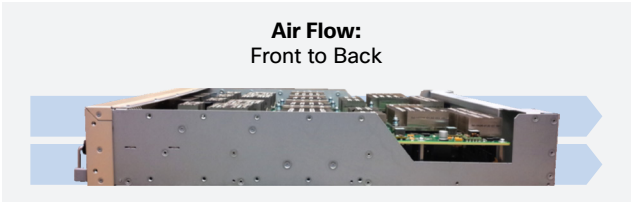
Chassis Air Flow

The Catalyst 6840-X fan tray supports **front-to-back air flow** (Figure 8).

To maintain proper airflow circulation, provide a minimum separation of 6 inches (15 cm) between a wall and the chassis air intake or air exhaust. Also allow for a minimum separation of 12 inches (30.5 cm) between the hot air exhaust on one chassis and the air intake of another chassis.

Figure 8. Catalyst C6840-X-LE-40G Fan-Tray Air Flow

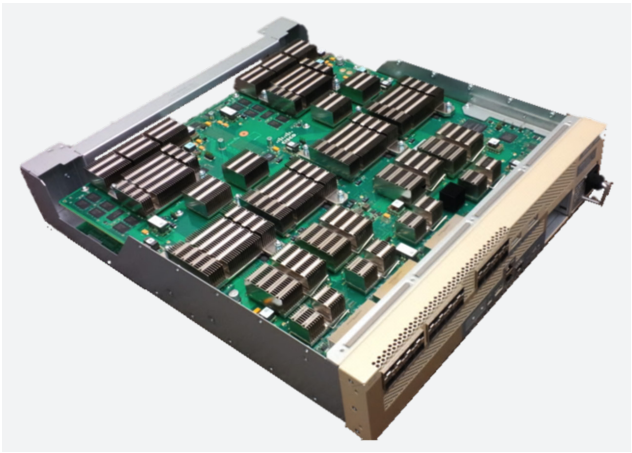




Baseboard Details

The new Catalyst 6840-X chassis is a next-generation small-form-factor fixed platform. The fixed components of the chassis are the system baseboard and daughter card (Figure 9). In many ways, it acts as the Supervisor Engine 2T of the system, but it is also a 16-port multirate SFP+ line card.

Figure 9. Catalyst 6840-X Baseboard and Daughter Card



Baseboard Models

This Catalyst 6840-X has four SKUs (Table 1). The number of interfaces and port type varies among SKUs.

Table 1. Catalyst 6840-X SKUs

SKUs	Details
C6816-X-LE	16x10G (baseboard only)
C6832-X-LE	16x10G+16x10G (baseboard + daughter board)
C6824-X-LE-40G	24x10G+2x40G (daughter board only)
C6840-X-LE-40G	16x10G+24x10G+2x40G (baseboard + daughter board)

Table 2 outlines some high-level differences between these four models.

Table 2. Comparison of the Catalyst 6840-X Baseboards

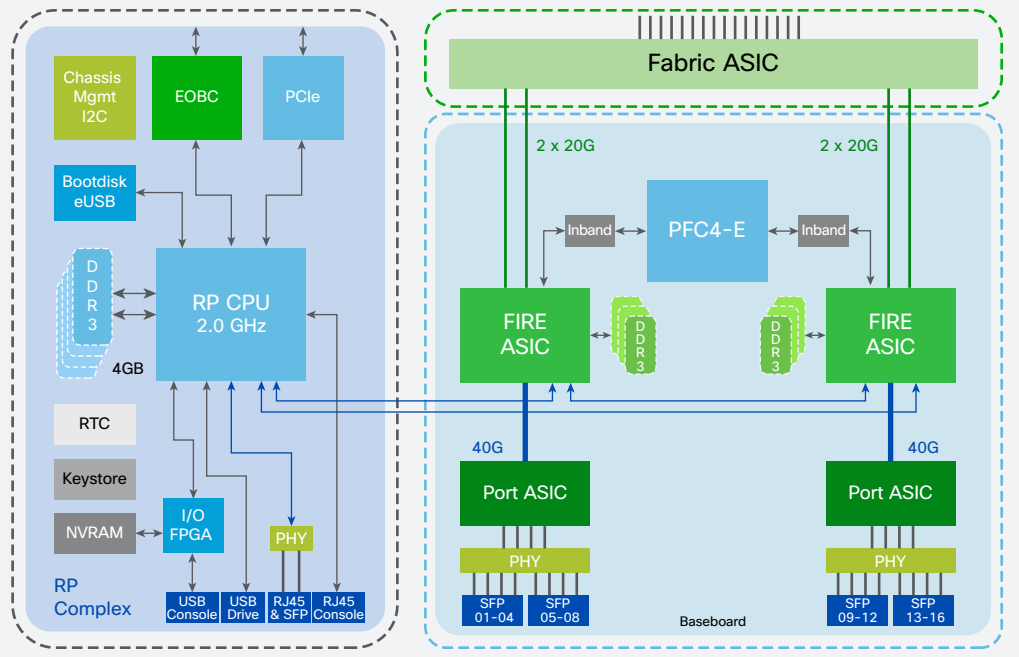
Hardware Option	C6816-X-LE	C6832-X-LE	C6824-X-LE	C6840-X-LE
Number of PFC4-E	1	2	2	3
IPv4/v6 routing capability	256,000/128,000	256,000/128,000	256,000/128,000	256,000/128,000
Multicast routes (IPv4/IPv6)	128,000/64,000	128,000/64,000	128,000/64,000	128,000/64,000
Layer 3 interfaces	128,000	128,000	128,000	128,000
MPLS labels	256,000	256,000	256,000	256,000
Number of adjacencies	256,000	256,000	256,000	256,000
MAC addresses	128,000	128,000	128,000	128,000
Flexible NetFlow	512,000	1,024,000	1,024,000	1,536,000
Security and quality-of-service (QoS) access control lists (ACLs)	64,000	64,000	64,000	64,000

Baseboard Components

The Catalyst 6840-X fixed supervisor slot has three main components (Figure 10):

- Route Processor (RP) complex
- Fabric complex
- Port complex

Figure 10. Catalyst 6840-X Components



Route Processor Complex

The RP complex is the route processor of the system and it is based on the well-known multilayer switch feature card (MSFC-5) on the Supervisor 2T, along with several important new and improved components. As such, it has the same basic design and layout, and serves the same basic functions (Figure 11).

Some highlights include:

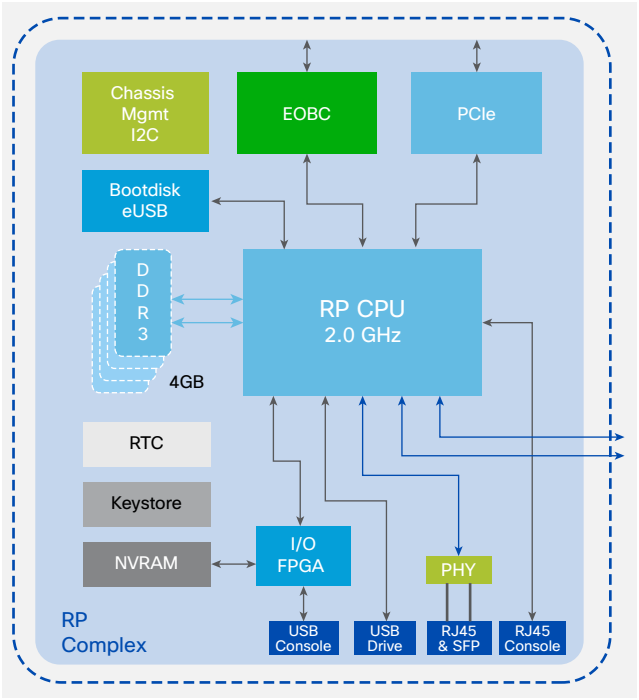
- New 2.0-GHz x86 dual-core CPU
  - Intel® Gladden dual-core CPU running at 2.0 GHz
  - Initial profiling that suggests a nearly two-fold increase in control-plane scale and performance
- 4 GB of 1337-MHz DDR3 ECC SDRAM
  - DDR3 can transfer data at twice the rate of DDR2
- Support for USB Type A file system
  - Supports standard USB thumb drives as a Cisco IOS file system (for example, copy, delete, format, and boot)
- Support for USB Type B serial console
  - Supports standard Type B (mini) serial console cables (in addition to an RJ-45 serial console)
- Compact flash replaced with eUSB
  - Internal Enhanced USB (eUSB) flash will support 8 GB of storage
  - The eUSB design provides a single-level cell (SLC) NAND flash with a sequential read of up to 21 MBps, and write of up to 18 MBps
- New switched EOBC interface
  - Supports the existing EOBC architecture but uses a point-to-point (P2P) switched design for dedicated bandwidth

The Gigabit Ethernet inband MAC driver (on the baseboard and modular port cards) uses three queues per port: high priority, medium priority, and low priority. A weighted-round-robin scheme is used to service these queues.

Each of the queues has an independent buffer ring to reduce packet loss due to data storms. Priorities are defined as follows:

- High priority – Incoming packets with an internal header with the Bridge Protocol Data Unit (BPDU) bit set, or with an internal header class-of-service (CoS) field in the range 5 to 7
- Medium priority – Incoming packets with an internal header CoS field in the range 2 to 4
- Low priority – All other packets (for example, with a CoS field in the range 0 to 1)

Figure 11. Route Processor Complex Block Diagram



Fabric Complex

The fabric complex is the switch fabric of the system. It is based on the well-known 2-Tbps fabric ASIC on the Supervisor 2T. The Catalyst 6840-X local fabric ASIC provides two fabric channels per Fabric Interface and Replication Engine (FIRE) ASIC. Each channel operates at 3.13-GHz clock frequency for 20 Gbps (40 Gbps per FIRE ASIC).

Refer to the Supervisor 2T Architecture White Paper for more details on fabric complex.

Port Complex

The baseboard port complex is based essentially on the C6800-16P10G module. It has the same basic design and layout, and serves the same basic functions (Figure 12).

Some highlights include:

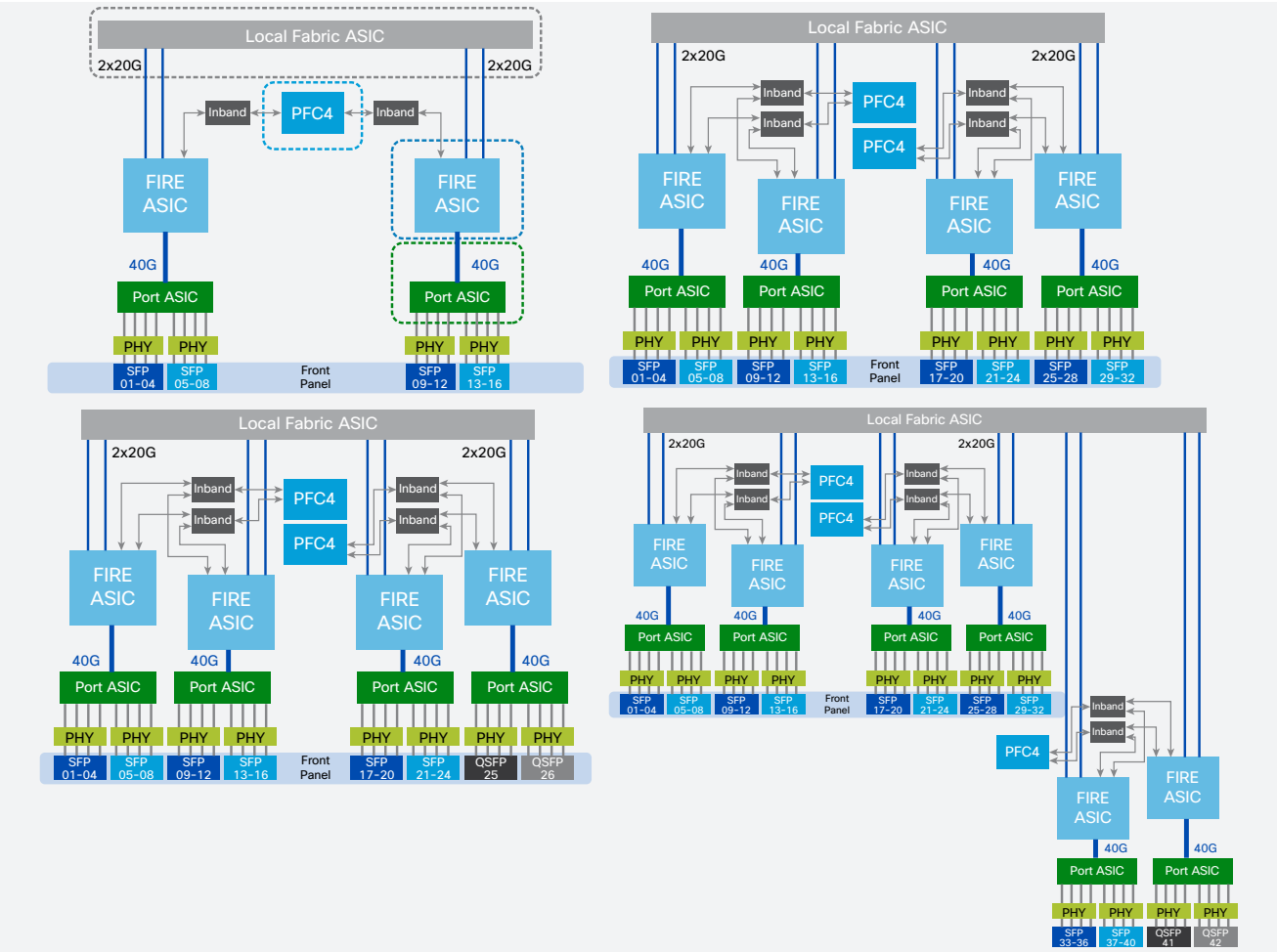
- 16 SFP+ (multirate) Ethernet ports

  The port complex supports 16 SFP+ ports, capable of 10/100/1000, 1 Gigabit Ethernet, and 10 Gigabit Ethernet (they are hardware-capable of 40-Gigabit Ethernet through a CVR-4SFP10G-QSFP reverse adapter).

- 80-Gbps connection to the switch fabric (2:1)
  - This is the same as the C6800-16P10G line card, with 80 Gbps on the backplane and 160 Gbps on the front panel; therefore, it is 2:1 oversubscribed if all ports are used at the line rate

  - The port complex supports a performance-mode configuration that will disable the second half (eight) of the ports, per-port ASIC, and enable 1:1 throughput
- Enhanced DFC4-E forwarding engine
  - The port complex uses an enhanced DFC4-E, which combines the previous Layer 2 and Layer 3 forwarding engine ASICs, improves read/write bandwidth, and includes various other fixes
- New 40-Gbps fabric interface and replication engines
  - The port complex uses an enhanced FIRE ASIC that improves read/write bandwidth, and includes various other fixes.
  - It will offload many of the traditional port ASIC functions (such as packet buffering and links to the CPU) and provides high-speed interfaces: 2-GB DDR3 buffer memory, two 20-Gbps uplinks to the fabric, and one 40-Gbps downlink to the port ASIC
- Gigabit Ethernet links to RP CPU and peer FIRE ASIC
  - These provide multiple 1-Gbps, full-duplex, control-plane paths to the route processor CPU, as well as between separate FIRE ASICs
  - The Gigabit Ethernet links are used to transmit data to and from the CPU, as well as for high-speed external control-plane programming, such as VSS
- 2-GB TX packet buffers per FIRE ASIC
  - 250 MB TX per 10-Gigabit Ethernet port in 2:1 oversubscription mode
  - 500 MB TX per 10-Gigabit Ethernet port in 1:1 performance mode
- Improved 40-Gbps port ASIC
  - The port complex uses one 40-Gb channel between the FIRE ASIC and port ASIC to improve read/write bandwidth
- Improved port ASIC 20-MB RX packet buffers per port ASIC
  - 1.25 MB RX per 10-Gigabit Ethernet port in 2:1 oversubscription mode
  - 2.5 MB RX per 10-Gigabit Ethernet port in 1:1 performance mode

Figure 12. Port Complex Block Diagrams



Packet Walks

This section provides a high-level overview of how packet forwarding is performed on the Catalyst 6840-X baseboard port complex and the 16-port SFP ports baseboard.

Since these are architecturally equivalent (excluding the RP and Link Control Protocol [LCP] complexes, and assuming the same fabric complex), we need to use only a single example.

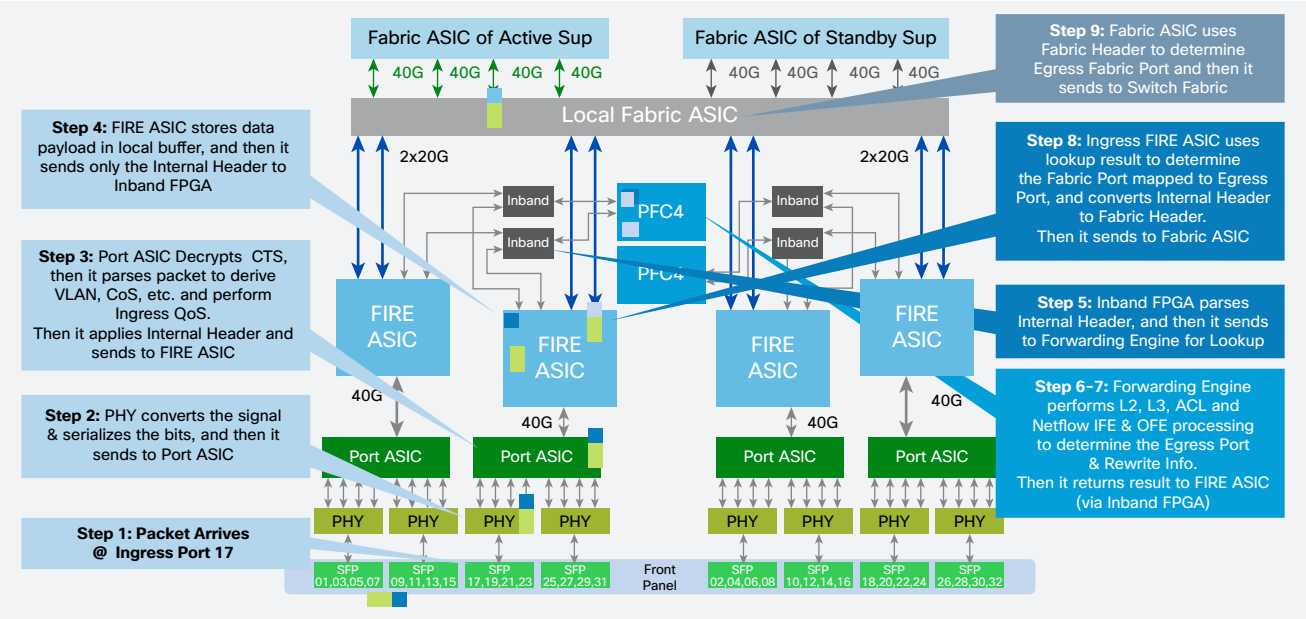
Ingress Forwarding

Following is the basic sequence of events when packets enter Catalyst 6840-X-LE ports.

1. The packet arrives at the ingress port.
2. PHY converts the signal and serializes the bits, and then sends the packet to the port ASIC.
3. The port ASIC parses the packet to derive the VLAN, CoS, etc. and performs ingress QoS. Then it applies an internal header and sends it to the FIRE ASIC.
4. The FIRE ASIC stores the data payload in the local buffer and then sends only the internal header to inband field-programmable gate array (FPGA).
5. The inband FPGA parses the internal header, and then sends it to the forwarding engine (for lookup).
6. The forwarding engine performs Layer 2, Layer 3, ACL, and NetFlow Input Forwarding Engine (IFE) and Output Forwarding Engine(OFE) processing, and determines the egress port and rewrite information.
7. The forwarding engine then returns a new internal header to the FIRE ASIC (through inband FPGA)
8. The ingress FIRE ASIC uses the lookup result to determine the fabric port mapped to the egress port, and converts the internal header to a fabric header. Then it sends to the fabric ASIC.
9. The fabric ASIC uses the fabric header to determine the egress fabric port and then sends the packet to the fabric complex.

Figure 13 shows a visual representation of the ingress packet-forwarding process.

Figure 13. Packet Walk at Ingress



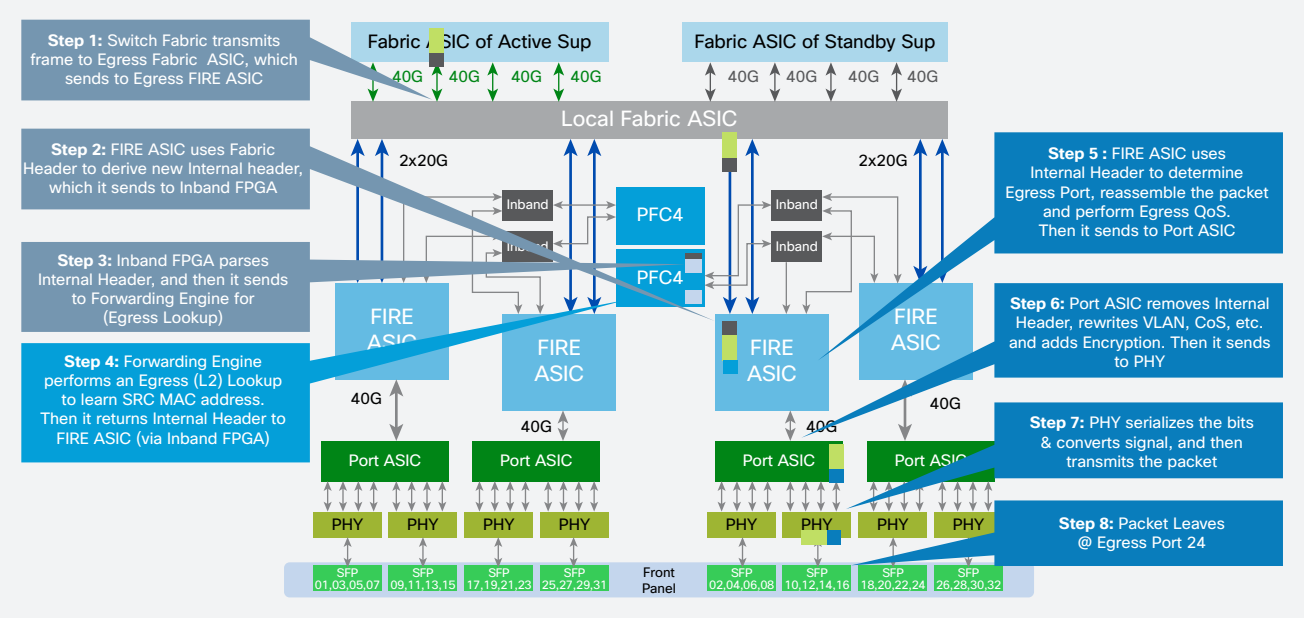
Egress Forwarding

Following is the basic sequence of events when packets exit the Catalyst 6840-X ports.

1. The fabric complex transmits the frame to the egress fabric port, where it is received by the egress FIRE ASIC.
2. The FIRE ASIC uses the fabric header to derive a new internal header, which it sends to the inband FPGA.
3. The inband FPGA parses the internal header, and then sends it to the forwarding engine (egress lookup).
4. The forwarding engine performs an egress (Layer 2) lookup to learn the source MAC address. It then returns the internal header to the FIRE ASIC (via the inband FPGA).
5. The FIRE ASIC uses the new internal header to determine the egress port, reassemble the packet, and perform egress QoS. It then sends to the port ASIC.
6. The port ASIC removes the internal header and rewrites VLAN, CoS, etc. It then sends it to PHY.
7. The PHY serializes the bits and converts the signal, then transmits the packet.
8. The packet leaves at the egress port.

Figure 14 shows a visual representation of the egress packet forwarding process.

Figure 14. Packet Walk at Egress



Port Numbering

Port numbering on the Cisco Catalyst 6840-X series is the following order (Figure 15).

- Native 10-Gb ports
- Native 40-Gb ports
- Converted 10-Gb ports

Figure 15. Port Numbering on the Catalyst 6840-X Series

Converting Native 40G Ports to 8x10G Ports

The Catalyst 6840-X-LE-40G and 6824-X-LE-40G have two native 40-Gb ports. These ports can be converted to 8x10G ports, starting from Cisco IOS Software Release 15.2(2) SY2, by simply changing the operation mode from 'FortyGigabitEthernet mode' to 'TenGigabitEthernet mode'. The following output provides an example.

Note: Changing the operation mode will reset the whole chassis; all 40G ports configuration will be lost.

```
C6840-X-LE#sh hw-module slot 1 operation mode
Module 1 port group 1 is running in TenGigabitEthernet mode
Module 1 port group 2 is running in TenGigabitEthernet mode
Module 1 port group 3 is running in TenGigabitEthernet mode
Module 1 port group 4 is running in TenGigabitEthernet mode
Module 1 port group 5 is running in TenGigabitEthernet mode
Module 1 port group 6 is running in FortyGigabitEthernet mode
C6840-X-LE#
```

```
C6840-X-LE#conf t

Enter configuration commands, one per line.  End with CNTL/Z.

C6840-X-LE(config)#hw-module slot 1 operation-mode port-group 6
TenGigabitEthernet

Operation mode change will reset the whole SUP and all the ports config will be
lost. Do you want to continue with reset? [confirm]

Building configuration...

[OK]

*Mar 11 17:53:10.491: %SYS-5-RELOAD: Reload requested by Exec. Reload Reason: TFR
SKU OPMODE CHANGE.

C6840-X-LE#show ip interface brief

Interface              IP-Address      OK? Method Status                Protocol

TenGigabitEthernet1/40 unassigned      YES NVRAM  administratively down down

TenGigabitEthernet1/43 unassigned      YES unset  up                    up

TenGigabitEthernet1/44 unassigned      YES unset  up                    up

TenGigabitEthernet1/45 unassigned      YES unset  up                    up

TenGigabitEthernet1/46 unassigned      YES unset  up                    up

TenGigabitEthernet1/47 unassigned      YES unset  up                    up

TenGigabitEthernet1/48 unassigned      YES unset  up                    up

TenGigabitEthernet1/49 unassigned      YES unset  up                    up

TenGigabitEthernet1/50 unassigned      YES unset  up                    up
```

Performance Mode

The Cisco Catalyst 6840-X Series has two performance modes: oversubscription mode and performance mode. Note that these modes affect port groups rather than an individual port. By default, the module operates in oversubscription mode since a backplane capacity of 80 GB per DFC4-E feeds 16x10 Gigabit Ethernet ports, resulting in a 2:1 oversubscription ratio. The default port performance mode can be identified using a show command:

```
C6816-X-LE#sh hw-module slot 1 operation mode

Module 1 port group 1 is running in TenGigabitEthernet mode

Module 1 port group 2 is running in TenGigabitEthernet mode

C6816-X-LE#sh hw-module slot 1 oversubscription

port-group      oversubscription-mode

1               enabled

2               enabled
```

Performance mode can be enabled by issuing a configuration command that will result in shutting down certain ports in the selected port group so as to make the port-to-fabric bandwidth come down from 2:1 to 1:1. The second half of the port group will be disabled when configuring performance mode. (for instance, configuring port group 2 on the Catalyst 6816-X-LE will disable ports 13, 14, 15, and 16, and will be available on ports 9, 10, 11, and 12. The ratio will be a 1:1 port-to-fabric bandwidth.

```
C6816-X-LE#conf t

Enter configuration commands, one per line.  End with CNTL/Z.

C6816-X-LE(config)#no hw-module slot 1 oversubscription port-group 2

WARNING: Switch to TRANSPARENT mode on module 1 port-group 2.

*Apr 28 17:13:22.989: %C6K_PLATFORM-6-NON_OVERSUBSCRIPTION_TEN_GIG: Ports 13, 14,
15, 16, of slot 1 disabled to prevent module bandwidth oversubscription.

C6816-X-LE#sh hw-module slot 1 oversubscription

port-group      oversubscription-mode

1               enabled

2               disabled
```

Conclusion

Cisco is extending the scale, performance, and capabilities of the venerable Cisco Catalyst 6500 Series with the introduction of the Catalyst 6800 Series. The new Catalyst 6840-X series chassis provides extremely high levels of scalability and performance, with the size and economics of an innovative fixed architecture. This unique platform offers 10-Gb port density; full IPv4/IPv6 and MPLS/VPLS functionality with large table sizes (up to 256,000 Forwarding Information Base [FIB] entries); and more than 15 years of best-in-class features. With a full suite of Layer 2 and 3, virtualization, security, multicast, IPv6, application visibility, smart operations, and rich media services, the Cisco Catalyst 6840-X delivers superior capabilities from its first day of deployment.

For More Information

- Catalyst 6500 Supervisor 2T Architecture white paper
- Catalyst 6500 Supervisor 2T data sheet
- Catalyst 6880 white paper
- Catalyst 6840 data sheet

C11-737288-00 05/16