Cisco Catalyst 6500 Supervisor 2T Architecture

White Paper

Author: Carl Solder – CCIE #2416
Distinguished Technical Marketing Engineer
Cloud Switching Services Technology Group

Reviews: Patrick Warichet – CCIE #14218
Shawn Wargo
Technical Marketing Engineers
Cloud Switching Services Technology Group

Introduction

The Cisco Catalyst 6500 Supervisor Engine 2T is the latest addition to the Catalyst 6500 family of Multi-Layer Switching Supervisor Engines. It offers much higher levels of forwarding performance, increases the scalability of many previously supported features, and introduces a host of new hardware-enabled functions beyond all previous Catalyst 6500 Supervisor models.

This white paper will provide an architectural overview of the new Supervisor 2T. It will explore the physical layout of the Supervisor 2T, provide details about its updated hardware components, and give an overview of its newly introduced features.

High-Level Description of Supervisor 2T

The Supervisor 2T is made up of four main physical components:

- The baseboard
- The 5th generation Multi-Layer Switching Feature Card (MSFC5)
- The 4th generation Policy Feature Card (PFC4)
- The 2 Tbps Switch Fabric

The Supervisor baseboard forms the foundation upon which many of the purpose-built daughter cards and other components are placed. It houses a multitude of application-specific integrated circuits (ASICs), including the ASIC complex that makes up the primary two Terabit (2080 Gbps) crossbar switch fabric, as well as the port ASICs that control the front-panel 10 GE and GE ports.

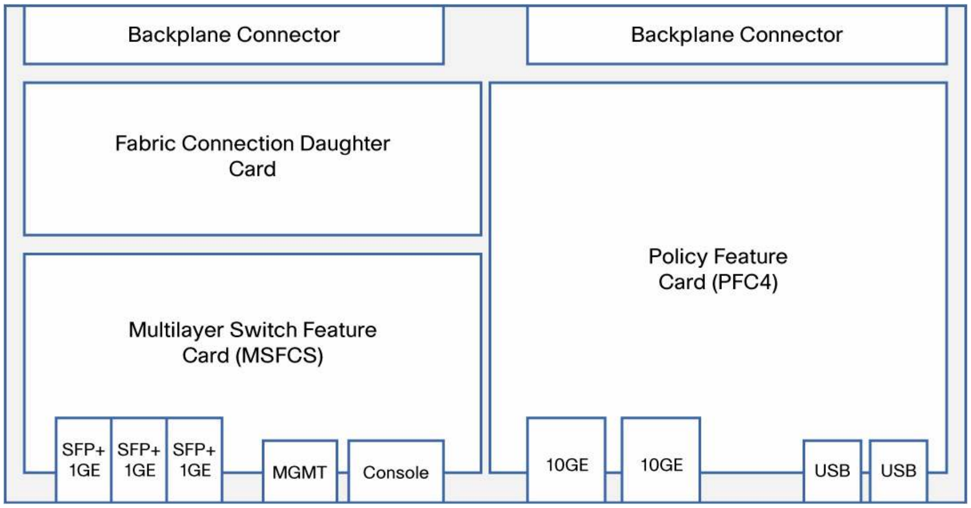
The MSFC5 is a daughter card that holds the CPU complex, which serves as the control plane for the switch. The control plane handles the processing of all software-related features. One major difference from earlier versions of the MSFC is that this version combines what were previously two separate CPU complexes into one. More details on this new CPU complex will be explored later in this paper.

The PFC4 is another daughter card that incorporates a special set of ASICs and memory blocks, which provide hardware-accelerated data-plane services for packets traversing the switch. It introduces a number of scalability enhancements, by increasing the size of many of the memory tables used by many of the hardware-accelerated features. The PFC4 also introduces a number of new hardware-accelerated features, such as Cisco TrustSec (CTS) and Virtual Private LAN Service (VPLS).

The 2 Tbps Switch Fabric provides 26 dedicated 20 Gbps or 40 Gbps channels to support the new 6513-E chassis (in addition to all existing E series chassis models). On the Supervisor 720, the switch fabric supported 18 fabric channels, which were used to provide two fabric channels per slot on all slots (with the exception of the 6513 chassis). With the new 6513-E chassis, the 2T Switch Fabric is capable of supporting dual fabric channels for all linecard slots (Slots 7 and 8 are reserved for the Active and Standby Supervisors).

A high-level overview of the Supervisor 2T board layout is shown in the diagram below.

Figure 1. Supervisor 2T Board Layout


Backplane Connector
Backplane Connector
Fabric Connection Daughter Card
Policy Feature Card (PFC4)
Multilayer Switch Feature Card (MSFCS)
SFP+ 1GE
SFP+ 1GE
SFP+ 1GE
MGMT
Console
10GE
10GE
USB
USB

A summary of the Supervisor 2T critical features is listed in the table below:

Table 1. Important Baseboard Features of Supervisor 2T

Feature	Description
Switch fabric type	2080 Gbps (2 Tbps) crossbar switch fabric
Forwarding engine daughter card	PFC4 or PFC4XL
CPU daughter card	MSFC5
Uplink ports	2 x 10 GE (X2 optic support) 3 x GE (SFP support)
USB ports	2 x USB (1 x Type-A and 1 x Type-B)
Management Ports	Serial console port (RJ-45) Connectivity management processor Ethernet port (RJ-45)
Management LED	Blue beacon LED
Media slot	Compact flash slot (Type II)
Forwarding performance	Up to 60 Mpps for L2, IPv4, and MPLS traffic Up to 30 Mpps for IPv6 traffic

The following sections provide more details on each of the major components of the Supervisor 2T.

System Level Requirements

The Supervisor 2T is designed to operate in any E-Series 6500 chassis. The Supervisor 2T will not be supported in any of the earlier non E-Series chassis. The table below provides an overview of the supported and non-supported chassis for Supervisor 2T.

Table 2. Chassis Options for Supervisor 2T

Supported Chassis	Non Supported Chassis
6503-E, 6504-E, 6506-E, 6509-E, 6509-V-E, 6513-E	6503, 6506, 6509, 6509-NEB, 6509-NEB-A, 6513, 7603, 7603-S, 7604, 7606, 7606-S, 7609, OSR-7609, 7609-S, 7613

For the E-Series chassis, a corresponding E-Series fan (or high-speed HS fan) for that chassis is required to support Supervisor 2T. While the 2500 W power supply is the minimum-sized power supply that must be used for a 6, 9, and 13-slot chassis supporting Supervisor 2T, the current supported minimum shipping power supply is 3000 W.

The 6503-E requires a 1400 W power supply and the 6504-E requires a 2700 W power supply, when a Supervisor 2T is used in each chassis. Either AC or DC power supply options can be used. The chassis types, as well as corresponding fan and power supply options that can be used with a Supervisor 2T, are detailed in the following table.

Table 3. Supported Chassis, Fan and Power Supply for Supervisor 2T

Supported Chassis	Supported Fan	Supported Power Supply
6503-E	WS-C6503-E-FAN	PWR-1400-AC
6504-E	FAN-MOD4-HS	PWR-2700-AC/4 PWR-2700-DC/4
6506-E	WS-C6506-E-FAN	WS-CAC-2500W (now End of Sale) WS-CDC-2500W WS-CAC-3000W WS-CAC-4000W-US WS-CAC-4000-INT PWR-4000-DC WS-CAC-6000W PWR-6000-DC WS-CAC-8700W
6509-E	WS-C6509-E-FAN	WS-CAC-2500W (now End of Sale) WS-CDC-2500W WS-CAC-3000W WS-CAC-4000W-US WS-CAC-4000-INT PWR-4000-DC WS-CAC-6000W PWR-6000-DC WS-CAC-8700W
6509-V-E	WS-C6509-V-E-FAN	WS-CAC-2500W (now End of Sale) WS-CDC-2500W WS-CAC-3000W WS-CAC-4000W-US WS-CAC-4000-INT PWR-4000-DC WS-CAC-6000W PWR-6000-DC WS-CAC-8700W
6513-E	WS-C6513-E-FAN	WS-CDC-2500W WS-CAC-3000W WS-CAC-4000W-US WS-CAC-4000-INT PWR-4000-DC WS-CAC-6000W PWR-6000-DC WS-CAC-8700W

The installation of the Supervisor 2T into a given chassis is always performed in specific slots. Keyed guide pins are used between the chassis and Supervisor connectors to manage this. Each chassis has two slots reserved for Supervisor use. These are called out in the following table.

Table 4. Chassis Supervisor Slots

	6503-E	6504-E	6506-E	6509-E	6509-V-E	6513-E
Slot 1	**Sup or L/C**	**Sup or L/C**	Linecard	Linecard	Linecard	Linecard
Slot 2	**Sup or L/C**	**Sup or L/C**	Linecard	Linecard	Linecard	Linecard
Slot 3	Linecard	Linecard	Linecard	Linecard	Linecard	Linecard
Slot 4		Linecard	Linecard	Linecard	Linecard	Linecard
Slot 5			**Sup or L/C**	**Sup or L/C**	**Sup or L/C**	Linecard
Slot 6			**Sup or L/C**	**Sup or L/C**	**Sup or L/C**	Linecard
Slot 7				Linecard	Linecard	**Sup Only**
Slot 8				Linecard	Linecard	**Sup Only**
Slot 9				Linecard	Linecard	Linecard
Slot 10						Linecard
Slot 11						Linecard
Slot 12						Linecard
Slot 13						Linecard

The Supervisor 2T provides backward compatibility with the existing WS-X6700 Series Linecards (with the exception of the WS-X6708-10G, which will be replaced by the new WS-X6908-10G, discussed later), as well as select WS-X6100 Series Linecards only. There is no support for the WS-X62xx, WS-X63xx, WS-X64xx, or WS-X65xx Linecards.

Note: All WS-X67xx Linecards equipped with the Central Forwarding Card (CFC) are supported in a Supervisor 2T system, and will function in centralized CEF720 mode.

There is no compatibility between the Supervisor 2T and earlier generations of Distributed Forwarding Cards (DFCs), such as DFC, DFC2, or DFC3x. The DFC is used to accelerate forwarding performance for the system as a whole, and uses the same forwarding architecture that is found on the PFC. The PFC4 architecture introduces a number of changes that differentiate it significantly in operation from earlier PFC/DFC models.

These changes require that only the DFC4 can interoperate with the PFC4. Any existing WS-X67xx Linecards that have a DFC3x will have to be upgraded with the new DFC4. WS-X67xx Linecards with a DFC4 installed will function in distributed dCEF720 mode.

Note: Any newly purchased WS-X6700 Series Linecards that are shipped with a DFC4 or DFC4XL pre-installed have been renamed as WS-X6800 Series Linecards, to clearly separate the performance differences.

Note: Due to compatibility issues, the WS-X6708-10G-3C/3CXL cannot be inserted in a Supervisor 2T system, and must be upgraded to the new WS-X6908-10G-2T/2TXL.

The Supervisor 2T Linecard support also introduces the new WS-X6900 Series Linecards. These support dual 40 Gbps fabric channel connections, and operate in distributed dCEF2T mode.

To summarize, the following linecards are supported by Supervisor 2T:

Table 5. Supervisor 2T-Compatible Linecards

Linecard Family	Linecard	Linecard Description
6900 Series Linecards (dCEF2T)	WS-X6908-10G-2T WS-X6908-10G-2TXL	8-port 10 GE (DFC4/DFC4XL)
	WS-X6904-40G-2T WS-X6904-40G-2TXL	4-port 40 GE or 16-port 10 GE (DFC4/DFC4XL)
6800 Series Linecards (dCEF720)	WS-X6816-10T-2T WS-X6816-10T-2TXL	16-port 10 G Base-T (DFC4/DFC4XL)
	WS-X6848-TX-2T WS-X6848-TX-2TXL	48-port 10/100/1000 RJ-45 (DFC4/DFC4XL)
	WS-X6848-SFP-2T WS-X6848-SFP-2TXL	48-port GE SFP (DFC4/DFC4XL)
	WS-X6824-SFP-2T WS-X6824-SFP-2TXL	24-port GE SFP (DFC4/DFC4XL)
6700 Series Linecards (CEF720)	WS-X6704-10GE	4-port 10 GE (CFC or DFC4/DFC4XL)
	WS-X6748-GE-TX	48-port 10/100/1000 (CFC or DFC4/DFC4XL)
	WS-X6748-SFP	48-port GE SFP (CFC or DFC4/DFC4XL)
	WS-X6724-SFP	24-port GE SFP (CFC or DFC4/DFC4XL)
6100 Series Linecards (Classic)	WS-X6148A-RJ-45	48-port 10/100/1000 RJ-45 (Upgradable to PoE 802.3af)
	WS-X6148A-45AF	48-port 10/100 RJ-45 with PoE 802.3af
	WS-X6148-FE-SFP	48-port 100 BASE-X (SFP)
	WS-X6148A-GE-TX	48-port 10/100/1000 Ethernet RJ-45
	WS-X6148A-GE-45AF	48-port 10/100/1000 Ethernet RJ-45 with PoE 802.3af
	WS-X6148E-GE-AT	48-port 10/100/1000 RJ-45 with PoE 802.3af and PoE+ 802.3at

Note: Any existing WS-X67xx Linecards can be upgraded by removing their existing CFC or DFC3x and replacing it with a new DFC4 or DFC4XL. They will then be operationally equivalent to the WS-X68xx linecards but will maintain their WS-X67xx identification.

Table 6. DFC4 Field Upgradable Linecard

Linecard	Linecard Description
WS-X6718-10G-3C	16-port 10 GE (DFC3/DFC3XL)
WS-X6716-10T	16-port 10 G Base-T (DFC3/DFC3XL)
WS-X6724-SFP	24-port GE SFP (DFC3/DFC3XL or CFC)
WS-X6748-SFP	48-port GE SFP (DFC3/DFC3XL or CFC)
WS-X6748-GE-TX	48-port 10/100/1000 (DFC3/DFC3XL or CFC)
WS-X6704-10GE	4-port 10 GE (DFC3/DFC3XL or CFC)

Applicable linecards (from the table above) utilizing a DFC3x will need to upgrade to a DFC4 in order to operate in a Supervisor 2T chassis. This level of interoperability is summarized in the table below.

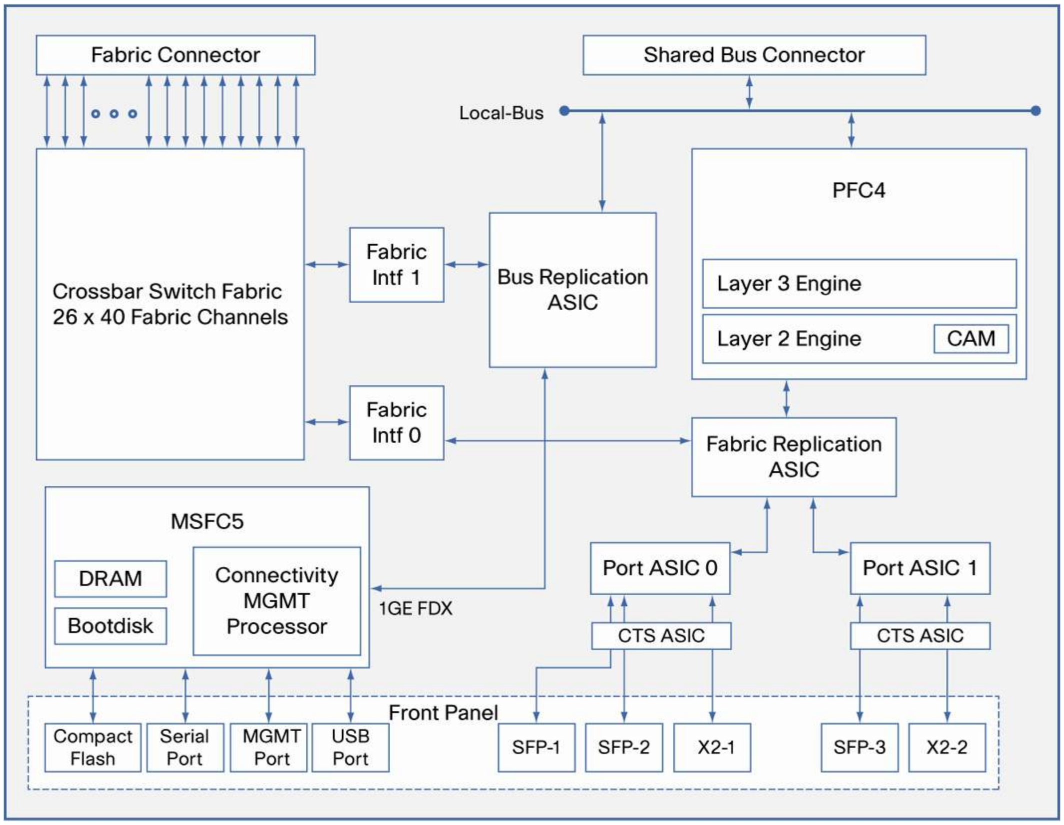
Table 7. PFC/DFC Interoperability Matrix

	PFC3A	PFC3B	PFC3BXL	PFC3C	PFC3CXL	PFC4	PFC4XL
DFC3A	√	PFC3B operates as PFC3A	PFC3BXL operates as PFC3A	PFC3C operates as PFC3A	PFC3CXL operates as PFC3A	Not compatible	Not compatible
DFC3B	DFC3B operates as DFC3A	√	PFC3BXL operates as PFC3B	PFC3C operates as PFC3B	PFC3CXL operates as PFC3B	Not compatible	Not compatible
DFC3BXL	DFC3BXL operates as DFC3A	DFC3BXL operates as DFC3B	√	DFC3BXL operates as DFC3B and PFC3C operates as PFC3B	PFC3CXL operates as PFC3BXL	Not compatible	Not compatible
DFC3C	DFC3C operates as DFC3A	DFC3C operates as DFC3B	PFC3BXL operates as PFC3B and DFC3C operates as DFC3B	√	PFC3CXL operates as PFC3C	Not compatible	Not compatible
DFC3CXL	DFC3CXL operates as DFC3A	DFC3CXL operates as DFC3B	DFC3CXL operates as DFC3BXL	DFC3CXL operates as DFC3C	√	Not compatible	Not compatible
DFC4	Not compatible	Not compatible	Not compatible	Not compatible	Not compatible	√	PFC4XL operates as PFC4
DFC4XL	Not compatible	Not compatible	Not compatible	Not compatible	Not compatible	DFC4XL operates as DFC4	√

Supervisor Architecture

The major processing blocks of the Supervisor 2T include the crossbar switch fabric, the MSFC5, the PFC4, the fabric interface ASICs, and the bus and fabric replication ASICs. Each of these block elements and their interconnection can be seen in the following diagram.

Figure 2. Supervisor 2T High-Level Block Diagram



Theory of Operation

All packet processing is performed in a specific sequence through the different ASIC blocks. A high-level packet walk is provided for packets that ingress and egress the local ports on the Supervisor below.

Ingress Packet Processing

1. Packets arriving on either the 1 G or 10 G ports have preliminary checks performed on them (by the PHY), such as cyclic redundancy checks (CRCs), before being forwarded to the Cisco TS ASIC.
2. The CTS ASIC performs ingress 802.1ae decryption (if enabled) and extracts the Security Group Tag (SGT) for Roles-Based ACL (RBACL) processing (if enabled). If CTS is disabled, this ASIC is passive. The packet is then forwarded to the port ASIC.
3. The port ASIC will store the packet in its local packet buffer and then applies an internal packet header, which contains information about the source port, VLAN, and more.
4. The packet is then forwarded to the fabric interface and replication ASIC, where information from the packet header is forwarded to the PFC4 or DFC4 for forwarding lookup processing.
5. The Layer 2 and Layer 3 processing is performed by the PFC4 (if a CFC is used, or local uplink ports) and the forwarding result of is sent back to fabric interface and replication ASIC.

6. Additional lookups may be performed by the fabric interface and replication ASIC (if this packet needs replication services, such as SPAN, multicast, and others).
7. The packet is then forwarded to the egress port ASIC (if the destination is local), or to the switch fabric (if the destination is a port on a remote fabric-capable linecard), or to the bus replication ASIC (if the packet destination is a linecard with connectivity only to the bus).

Egress Packet Processing (Packets Received from Crossbar)

1. Packets are received from one of the switch fabric channels.
2. The switch fabric sends the packet to the fabric replication ASIC.
3. The fabric replication ASIC stores the packet and sends the packet header information to the PFC4 for forwarding lookup processing.
4. The PFC4 performs the lookup and sends the results back to the fabric replication ASIC.
5. The fabric replication ASIC performs an additional lookup, if replication services are required.
6. The fabric replication ASIC then forwards the packet (and replicated packets, if applicable) to the port ASIC, which contains destination information from PFC4.
7. The port ASIC stores the packet in the packet buffer and performs egress checks on the packet.
8. The port ASIC performs the packet rewrite and then forwards the packet to the CTS ASIC.
9. If CTS is enabled, the CTS ASIC performs 802.1ae encryption, and the packet is forwarded to the Physical Layer Protocol (PHY), to be transmitted onto the cable.

Egress Packet Processing (Packets Received from Shared Bus)

1. Packets are received from the bus and are placed on the local bus.
2. The Bus Replication ASIC stores the packet and sends packet header information to the PFC4 for forwarding lookup processing.
3. The PFC4 performs the lookup and sends the results back to the bus replication ASIC.
4. The bus replication ASIC performs an additional lookup, if replication services are required.
5. The bus replication ASIC forwards the packet (and replicated packets, if applicable) to the port ASIC, which contains destination information from PFC4.
6. The port ASIC stores the packet in the packet buffer, and performs egress checks on the packet.
7. The port ASIC performs the packet rewrite and forwards the packet to CTS ASIC.
8. If CTS is enabled, the CTS ASIC performs 802.1ae encryption, and the packet is forwarded to the PHY, to be transmitted onto the cable.

The following sections provide more detail for each processing block.

The Fabric and Bus Connector

The Catalyst 6500 supports two different switch backplanes, the crossbar switch fabric (top left in the above diagram), and bus backplanes (top right in the above diagram). The crossbar switch fabric is the high-capacity backplane that is used by the CEF720 and CEF2T generation of linecards to optimize switching performance. This backplane refers to the 2 Terabit backplane that is contained within the Supervisor's name. A second backplane (referred to as the "bus" backplane) is also present to support WS-X61xx linecards, supported service modules, and linecards that do not utilize a local DFC4 for forwarding.

The crossbar switch fabric provides a set of fabric channels (or data paths) that are assigned to the slots in the chassis where linecards are inserted. This is referred to collectively as the crossbar switch backplane. This array of fabric channels provides an any-to-any (full-mesh) connection option for the attached linecard to forward data over a dedicated path to any other linecard installed in the chassis.

The bus backplane is a 16 Gbps (full duplex) shared data bus that is used to provide a connection between attached “classic” linecards. The data bus operates at 62.5 Mhz and is 256 bits wide. The bridge ASIC provides the interface through which those classic linecards can communicate with the PFC4 and MSFC5 for data processing services.

Crossbar Switch Fabric

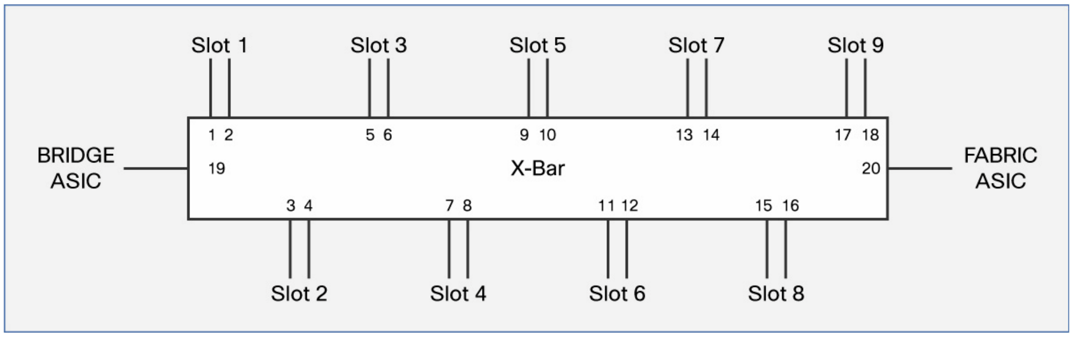
The crossbar switch fabric on the Supervisor 2T provides 2080 Gbps of switching capacity. This capacity is based on the use of 26 fabric channels that are used to provision data paths to each slot in the chassis. Each fabric channel can operate at either 40 Gbps or 20 Gbps, depending on the inserted linecard. The capacity of the switch fabric is calculated as follows:

26 x 40 Gbps = 1040 Gbps
1040 Gbps x 2 (full duplex) = 2080 Gbps

The 2080 Gbps number is a marketing number (common among all switch vendors) and is used in all literature to denote that full duplex transmission allows data traffic to be transmitted and received simultaneously. While the switch fabric capacity is documented as a full duplex number, note that the per-slot capacity of the E-Series chassis is NOT a full duplex number.

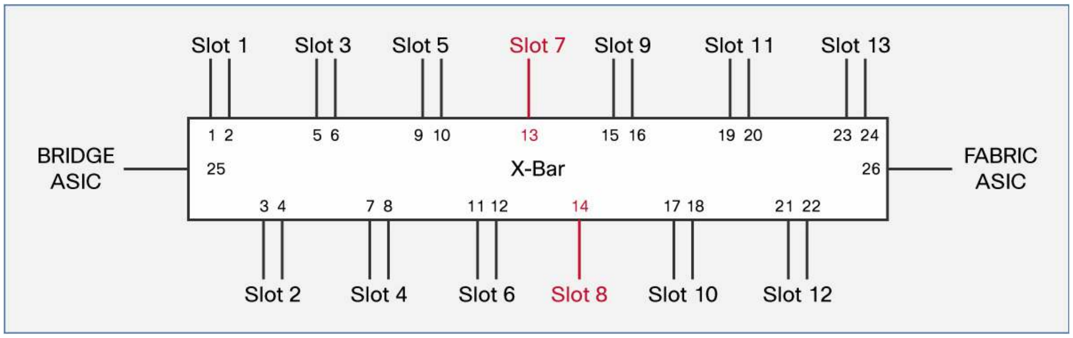
The 80 Gbps per slot nomenclature represents 2 x 40 Gbps fabric channels that are assigned to each slot providing for 80 Gbps per slot in total. If marketing math were used for this per slot capacity, one could argue that the E-Series chassis provides 160 Gbps per slot.

Figure 3. Fabric Channel Layout in 6509-E



For every chassis (except the 6513-E), there are enough fabric channels to provide dual fabric channels to each linecard slot, including the two Supervisor slots. The exception is the 6513-E. For the 6513-E chassis, there are dual fabric channels provided for Slots 1 through 6 and for Slots 9 through 13. Slots 7 and 8 are designated Supervisor-only slots. If a linecard is inserted in either of these Supervisor-only slots, it will not be powered up.

Figure 4. Fabric Channel Layout in 6513-E



Fabric Replication ASIC

This ASIC is used to provide a number of important functions. First and foremost, it receives packets from the front panel GE and 10GE ports, extrapolates valuable information from packet headers, and forwards this information to the PFC4 for packet lookup and associated forwarding services processing (Security, Quality of Service, NetFlow, and others). When packets return from packet lookup processing, this ASIC will perform packet rewrites according to the lookup result.

Another important processing role performed by this ASIC is multicast replication. This includes IGMP snooping for Layer 2 packets, as well as multicast expansion for Layer 3 multicast packets. Additionally, other replication services are supported to provision switched port analyser (SPAN, ER-SPAN, and more) capabilities.

New capabilities also include support for Cisco TrustSec (CTS) and Virtual Switch Link (VSL). Given that, the front panel 10 GE ports can be used to become part of a VSL that facilitates the creation of a Virtual Switching System (VSS) domain.

Port ASIC

There are two port ASICs on the Supervisor used to provision the front panel 2 x 10GE and 3 x 1 GE ports. One port ASIC supports a single 10 GE port and a single GE port. The other port ASIC supports a single 10 GE port and two GE ports. The following list defines this port ASIC’s capabilities:

- Per-port VLAN translation
- VSL support (10 GE ports only)
- Cisco TrustSec support (802.1ae link layer encryption)
- Jumbo frames (up to 9216 bytes)
- Flow control
- 1P3Q4T (one strict priority queue, three normal round robin queues, and four Weighted Random Early Detection [WRED] thresholds per normal queue) queue structure for GE ports (this is the TX queue structure)
- 1P7Q4T (one strict priority queue, seven normal round robin queues, and four WRED thresholds per normal queue) queue structure for 10 GE ports (this is the TX queue structure)
- 1Q8T (one normal round robin queues and eight WRED thresholds for that queue) queue structure for 1 GE ports (this is the RX queue structure)
- 2Q4T (two normal round robin queues and four WRED thresholds per normal queue) queue structure for 10 GE ports (this is the RX queue structure)
- 256 MB total queue buffer (split among the front panel 10 G and 1 G ports)
- DWRR, WRR, and SRR scheduling schemes
- WRED and Tail Drop congestion management

- 802.1Q VLAN encapsulation
- ECC protection

Bridge ASIC

The bridge ASIC primarily serves as the gateway for linecards using the bus to connect to the control plane and data plane (MSFC5 and PFC4 respectively). It provides a connection into the bus backplane and receives packets from linecards, which it will forward to the MSFC5 or PFC4 for processing. It provides a packet buffer, as well as flow control to manage data flows from the linecards. Once packet processing is complete for the packet, the bridge ASIC will send the results of the forwarding operation back over the bus to the classic linecards.

MSFC5/PFC4

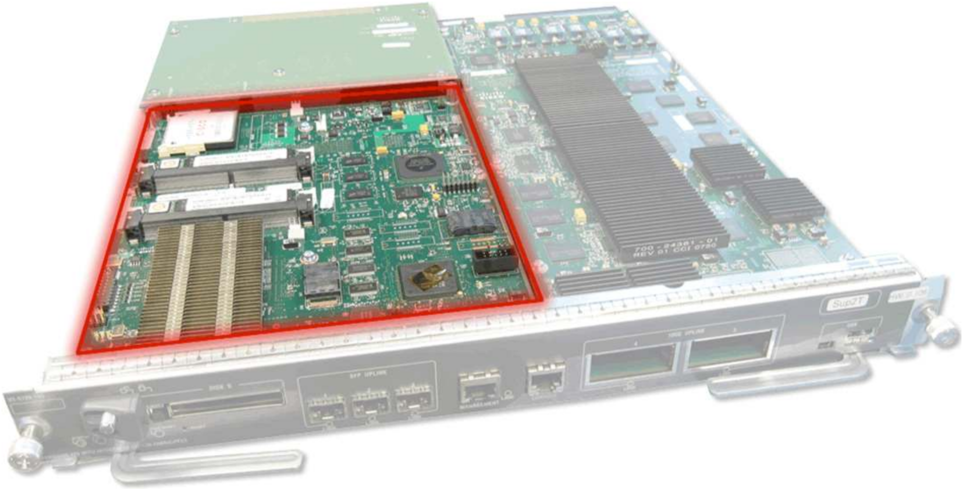
Both of these will be discussed individually in more detail later in this paper.

MSFC5

The MSFC5 is a next-generation CPU daughter card for the Supervisor 2T. It is not an optional daughter card and will be present on every Supervisor 2T. The MSFC5 cannot be installed on any other Supervisor 32 or Supervisor 720, and is designed for the exclusive use on the Supervisor 2T.

The MSFC5 performs control plane services for the switch. Control plane functions typically process all those features and other processes that are not handled directly in hardware by purpose built ASICs. The MSFC5 CPU handles Layer 2 and Layer 3 control plane processes, such as the routing protocols, management protocols like SNMP and SYSLOG, and Layer 2 protocols (such as Spanning Tree, Cisco Discovery Protocol, and others), the switch console, and more.

Figure 5. MSFC5 on Supervisor 2T



On previous generations of the MSFC, there were two main CPU complexes that resided on the MSFC. These CPU complexes were known as the Route Processor (RP) and Switch Processor (SP) complex. The RP complex was responsible for performing Layer 3 control plane services, IOS configuration and associated management of the configuration, Address Resolution Protocol (ARP) processing, Internet Control Message Protocol (ICMP) processing and more.

Its other main function was to create the CEF forwarding tables that are programmed into the PFC hardware memory tables (through the SP). The SP complex was responsible for performing Layer 2 control plane services, managing system power as well as programming various hardware elements in the switch. The IOS image that ran on these previous MSFCs, while downloaded from http://www.cisco.com as one binary image file, was in fact two discrete images: one that ran on the RP CPU complex and one that ran on the SP CPU complex.

The most important enhancement of the MSFC5 is the move from the dual CPU complex (RP/SP) to a single CPU complex that combines the RP and SP complexes into one. As such, this also introduces a new IOS image for the Supervisor 2T that combines the previous two running images into one.

Another valuable enhancement of this new MSFC5 is the introduction of a Connectivity Management Processor (CMP). The CMP is a stand-alone CPU that the administrator can use to perform a variety of remote management services. Examples of how the CMP can be used include:

- System recovery of the control plane
- System resets and reboots
- The copying of IOS image files should the primary IOS image be corrupted or deleted

The CMP and the RP share the same console through a programmable multiplexor. By default, the firmware programs the multiplexor so the RP console is active on the front panel. The multiplexor intercepts specific escape sequences that instruct it to switch from one console to the other.

If the sequence (Ctrl-C, Shift-M) is used three consecutive times, the multiplexor will switch the console to the CMP. If the sequence (Ctrl-R, Shift-M) is used three consecutive times, the multiplexor will switch back to the RP console.

External IP connectivity to the CMP is provided by a new 10/100/1000 RJ-45 management interface on the front panel. This port can then be configured with an IP address, gateway, and connection method (for example, Telnet and SSH). The user can then access and control the system remotely, even if the RP is currently down.

The specifications for the MSFC5 CPU complex as compared to the MSFC3 (Supervisor 720) are shown in the following table:

Table 8. MSFC5 vs. MSFC3

Feature	MSFC3 (Supervisor 720-10G)	MSFC5 (Supervisor 2T)
CP CPU speed	SP CPU - 600 Mhz RP CPU - 600 Mhz	Dual core with each core @ 1.5 Ghz
Number CPU cores	1	2
Connectivity Management Processor (CMP) CPU	No CMP	Single core @ 266 Mhz 32 MB boot flash 256 MB system memory
NVRAM	2 MB	4 MB
OBFL flash	N/A	4 MB
Boot disk	SP CPU - 1 GB RP CPU - 64 MB	CF based - 1 GB
CF card on Supervisor front panel	Yes - 1 CF Slot	Yes - 1 CF Slot
DRAM	SP - Up to 1 GB RP - Up to 1 GB	2 GB (non-XL) 4 GB (XL)

The MSFC5 uses DDR-II memory for improved speed of access. Two memory sockets are present on the board and are currently loaded with a single 2 GB DIMM for the non-XL version and two 2 GB DIMM for the XL version of the Supervisor 2T. An OBFL (On Board Failure Logging) flash is also present, providing 4 MB of memory for logging temperature readings and other diagnostic information. A battery-powered NVRAM memory block of 4 MB is also present on the board, which stores the startup-config, VLAN database, and other information necessary to boot the system.

The MSFC5 has two temperature sensors, each located on one side of the board. The right hand side sensor is designated the “inlet” sensor, while the sensor on the left of the board is designated the “outlet” sensor. These sensors provide information that is fed into the environmental subsystem, which uses this information to assess temperature thresholds.

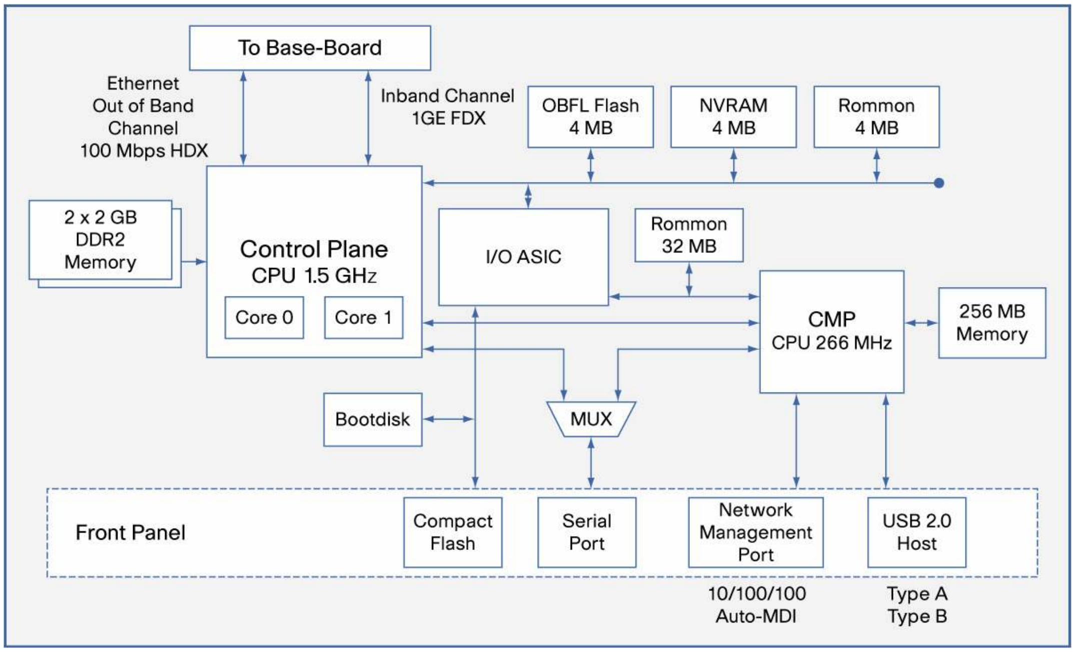
The MSFC5 complex also supports an on-board Compact Flash (CF) Boot Disk, which supports a standard CF Type-II disk, with disk densities of up to 8 GB. The default Boot Disk shipped with the MSFC5 is 1 GB. The CF Boot Disk can be used as a storage space for IOS images and configuration files as well as other user files that the administrator may wish to have stored local to the switch.

Along with the Boot Disk, the MSFC5 also manages the Supervisor 2T front panel CF (Compact Flash) slot. Like the Boot Disk, this CF slot supports CF Type-II disks with densities of up to 8 GB being supported. The external CF Slot has an LED visible from the front panel that indicates if the CF slot is in use.

The CMP, as introduced above, is a new enhancement that makes its debut on the Supervisor 2T. In many respects, it is a CPU complex within the larger MSFC5 CPU complex. As such, it has its own CPU and memory separate from that found on the MSFC5. The CMP supports a front panel 10/100/1000 management port. This port supports auto negotiation, detection and correction of pair swaps (MDI crossover), and support for jumbo frames up to 9216 bytes in size. It also uses a bi-color LED on the front panel port to show activity and link status. In addition to the front panel Ethernet port, the CMP also provides a direct interface to the front USB ports, allowing USB console connection to the CMP.

A detailed view of the MSFC5 board is shown below:

Figure 6. MSFC5 Block Diagram



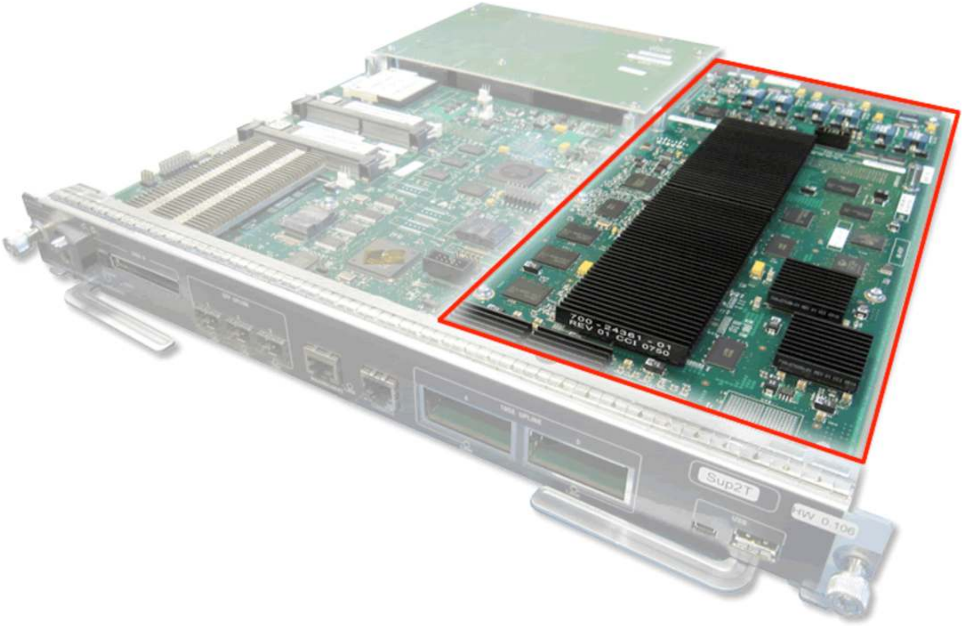
PFC4 and DFC4

The PFC4 provides hardware accelerated forwarding for packets traversing the switch. This includes forwarding for IPv4 unicast/multicast, IPv6 unicast/multicast, Multi-Protocol Label Switching (MPLS), and Layer 2 packets. Along with forwarding, the PFC4 is also responsible for processing a number of services that are handled during the forwarding lookup process. This includes, but is not limited to, the processing of security Access Control Lists (ACLs), applying rate limiting policies, quality of service classification and marking, NetFlow flow collection and flow statistics creation, EtherChannel load balancing, packet rewrite lookup, and packet rewrite statistics collection.

The DFC4 is a daughter card that is located on selected linecards. The DFC4 contains the same ASIC functional blocks that are found on the PFC4. The main purpose of the DFC4 is to provide local forwarding services for the linecard, offloading this function from the PFC4. Using DFC4s helps scale performance of the overall chassis. Without any DFC4s present in the chassis, the maximum performance of a chassis with Supervisor 2T is 60 Mpps (for IPv4 forwarding). Each DFC4 adds an additional 60 Mpps of forwarding performance to the aggregate forwarding capacity of the chassis. All of the information written about the PFC4 below equally applies (function, scalability, and performance) to the DFC4.

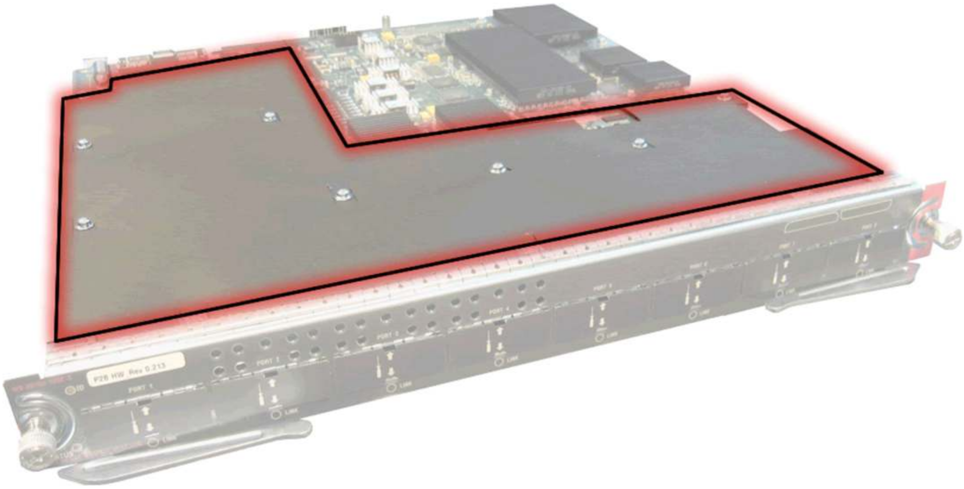
The PFC4 is located on the right-hand side of the Supervisor 2T baseboard, as shown in the diagram below.

Figure 7. Policy Feature Card 4 on the Supervisor 2T



The location of the DFC4 on a linecard is shown in the following diagram. In the photo below, the DFC4 is located underneath a protective cover that protects the daughter card from getting damaged when the linecard is inserted or removed from a chassis.

Figure 8. DFC on the WS-X6908-10G-2T Linecard



PFC4 and PFC4XL Feature Review

The PFC4 offers a number of enhancements and new capabilities over earlier generations of the PFC. These hardware feature capabilities are summarized in the following table.

Table 9. New PFC4 Enhancements

Functional Area	Feature
Forwarding	Increased forwarding performance of up to 60 Mpps for L2, IPv4 and MPLS forwarding and up to 30 Mpps for IPv6 forwarding
	Improved EtherChannel load balancing utilizing an 8-bit hash
	Increased multicast routes to 256 K
	Support for 16 K bridge domains
	Support for 128 K logical interfaces
	Increased MAC address table to 128 K
	IGMPv3 snooping in hardware
	PIM registers in hardware
	IPv6 MLDv2 snooping in hardware
	IPv6 in IPv6 tunnelling
	IPv6 in IPv4 tunnelling
	Full Pipe tunnel mode
	Tunnel source address sharing
Network Security	Cisco TrustSec (CTS) with support for RBACL (note that 802.1ae link layer encryption is a function of the port ASIC and not the PFC4)
	Hardware control plane policing (including multicast CoPP) and an increase in the number of hardware rate limiters
	Shared ACL table (with QoS) now up to 256 K entries
	L2+L3+L4 ACLs
	1:1 mask ratio with ACE values
	ACL dry run and hitless commit
	Classify using more fields in IP header
	Source MAC + IP Binding
	Drop on source MAC miss
	Per protocol per drop
	Ipv4 and IPv6 uRPF

Functional Area	Feature
	Up to 16 path lookup for uRPF
NetFlow Services	Egress NetFlow
	Increased NetFlow table size to 512 K (non XL) or 1 M (XL)
	Improved NetFlow hash
	Flexible NetFlow and NetFlow v9 support in hardware
	Hardware-sampled NetFlow
Quality of Service (QoS)	Per-port -per-VLAN policies
	Distributed policing (up to 4 K policers)
	Increased scalability for aggregate policing (up to 16 K policers) and microflow policing (up to 128 policers)
	Increased number of flow masks to reduce feature incompatibilities
	Egress microflow policing
	Increase in DSCP mutation maps
	Packet or byte-based policing
Virtualization	Layer 2 over GRE
	MPLS aggregate labels increased up to 16 K
	Native H-VPLS
	MPLS over GRE
	16 K EoMPLS tunnels

The following sections provide a brief overview of each of the enhancements noted in the above table, divided into Layer 2 and Layer 3 functional groups.

Layer 2 - Increased MAC Address Support

A 128 K MAC address table is standard on both models of the PFC4. The MAC address table has been enhanced to support additional fields in a MAC address table entry, such as the bridge domain (BD) that the MAC address is a part of.

Layer 2 - Bridge Domains

The bridge domain is a new concept that has been introduced with PFC4. A bridge domain is used to help scale traditional VLANs, as well as to scale internal Layer 2 forwarding within the switch. Bridge domains are mapped to each VLAN that a user configures, as well as other resources such as an EoMPLS/VPLS tunnel, L3 sub-interfaces, and multicast Egress replication-mode. In essence, a bridge domain is equal to a VLAN (which is a 12-bit ID, allowing 4096 unique values). Bridge domains use a 14 bit ID (12 bits are VLAN ID), for a total of 16 K bridge domains supported in hardware by the PFC4.

Layer 2 - Increased Logical interfaces

The PFC4 introduces the concept of a Logical Interface (LIF), which is a hardware-independent interface (or port) reference index associated with all frames entering the forwarding engine. A LIF is a 72-bit internal address comprised of the bridge domain (BD), the source port index and other associated information (for example, protocol type) used by the PFC4 to facilitate forwarding decisions. This allows Layer 2 forwarding characteristics to be logically separated from Layer 3 characteristics. The PFC4 adds hardware support for up to 128 K LIFs.

Layer 2 - Improved EtherChannel Hash

Skew tables have been introduced to overcome issues with distributing traffic over an odd number of links (3, 5, 6, and 7) that form part of an EtherChannel link bundle. Previous PFC3x engines offered balanced distribution across even numbered link (2, 4, and 8) bundles. However, link bundles with 3, 5, 6, or 7 member ports could end up with uneven distribution. The use of skew tables, as part of the processing logic for selecting a link in a bundle, helps alleviate those earlier problems.

Inputs (fields) used for the hash itself have also been extended. Options for including the input and output interfaces into the hash, along with the VLAN ID, provide for more granular link selection. These options also help overcome earlier issues with link selection for multicast traffic over VLAN trunks, where one link in the bundle could be favored over others.

Layer 2 - VSL Support

VSL is a technology that allows two physical Catalyst 6500 switches to operate together as a single logical unit. In this mode, the Supervisor Engine 2T in one chassis is the SSO Active and the Supervisor Engine 2T in the other chassis is the SSO Standby. As with stand-alone (non-VSS) HA, the Active Supervisor is responsible for the control plane. While only one of the two control planes is active in a virtual switch setup, both data planes are active. Having both data planes active allows the virtual switch system to optimize hardware forwarding and use both PFC4 engines to perform hardware-enabled tasks for each chassis. Support for VSL is included with the PFC4 and mirrors that found in the previous PFC3C/PFC3C-XL Supervisor Engines 720.

Layer 2 - Per Port-Per-VLAN

This feature is designed for Metro Ethernet deployments where policies based on both per-port and per- VLAN need to be deployed. Typically these scenarios define a network deployment model where different VLANs carrying different customer traffic is carried on the same physical interface. Earlier PFC engines only allowed port-based or VLAN-based polices to be applied to a port at any one time. A port-based policy would disregard VLAN information, and likewise a VLAN-based policy would disregard port information. Support for this new interface type in the PFC4 allows per-port-per-VLAN policies to be applied where both the VLAN and port can be considered in the assigned policy.

Layer 3 - Increased Layer 3 Forwarding Performance

The PFC4 forwarding engine now supports forwarding performances of up to 60 Mpps for both Layer 2 and IPv4 Layer 3 forwarding. Technically, it is actually a 120 Mpps forwarding engine, but since each packet passes through the PFC4 twice (once for ingress processing and once for egress processing, which is discussed later), the effective forwarding performance equates to 60 Mpps.

Due to the length of the address, IPv6 requires an additional internal cycle, which makes the effective IPv6 forwarding performance 30 Mpps. Note that the PFC3B/XL forwarding engines supported forwarding performance up to 30 Mpps for IPv4 and 15 Mpps for IPv6, while the PFC3C/XL forwarding engines supported forwarding performance up to 48 Mpps for IPv4 and 24 Mpps for IPv6.

Layer 3 - uRPF for IPv6

Unicast Reverse Path Forwarding (uRPF) is a tool that can be used to protect against address spoofing. The uRPF check performs a lookup into the FIB, using the source address as the lookup index (as opposed to the destination address). The lookup is performed to determine if the packet arrived on an interface where the source address of the packet matches the network found through that interface. If the packet had a source IP address of “A”, but it arrived on an interface from where network “A” does not exist, then the packet is deemed to have been spoofed and will be dropped.

The PFC4 extends support for uRPF to include IPv6, which was not available in the PFC3x. More importantly, PFC4 now supports 16 prefixes during both an IPv4 and an IPv6 uRPF lookup compared to only supporting the lookup of 2 prefixes with the PFC3x.

Layer 3 - Tunnel Source Address Sharing

With the PFC3x family, each tunnel was required to use a unique source IP address as the tunnel source. If two or more tunnels were configured using the same source address, packets for the second and subsequent tunnels configured would be switched in software. With the PFC4, this scenario has been enhanced and now multiple tunnels can be configured to share the same source address, while the tunnelled packets are switched in hardware.

Layer 3 - IPv6 Tunnelling

The PFC3x did not support IPv6 tunnelling options where the outer header was an IPv6 header. This behavior has changed with PFC4. The PFC4 now supports the following new IPv6 tunnelling options:

- **ISATAP (Intra-Site Automatic Tunnel Address Protocol):** ISATAP tunnelling protocol allows IPv6 packets to be globally routed through the IPv6 network, while also being automatically tunneled through IPv4 clouds locally within a site. ISATAP tunnelling uses an address format with specially constructed 64-bit EUI-64 interfaces ID. PFC4 fully supports ISATAP tunnels, whereas PFC3x offered only partial support.
- **IPv6 tunnel for GRE packets:** PFC4 supports IPv6 GRE tunnelling mode, where both IPv4 and IPv6 packets can be encapsulated in an IPv6 GRE header and tunnelled across an IPv6 network. The PFC3x had limited support for this tunnelling mode. The PFC4 can support up to 256 IPv6 GRE tunnels.
- **IPv6 generic tunnelling:** IPv6 packets can be encapsulated in an IPv6 or IPv4 header and tunnelled across the IPv6 network. The PFC4 offers support for this mode and 256 tunnel source addresses are supported. The PFC3x family did not support this tunnelling option.

Layer 3 - VPLS

Virtual Private LAN Service (VPLS) allows for a Layer 3 MPLS network to appear to operate like a Layer 2 Ethernet-based network. It effectively allows native Ethernet frames to be transported across an MPLS network providing any to any connectivity, just as if it were on a normal Ethernet LAN. Previous PFC engines required an external WAN card, such as the OSM or SIP-400, to provide support for VPLS. With PFC4, VPLS is supported natively, and no external WAN cards are required for this support.

Layer 3 - MPLS over GRE

The PFC4 now supports the MPLS over GRE tunnelling option in hardware. This was not supported in previous PFC3x engines. This capability is important for those customers looking to join MPLS backbones over a common IP network.

Layer 3 - MPLS Tunnel Modes

MPLS provides a number of options for controlling how the EXP (experimental bit) can be set. The EXP value provides a way for a priority value to be assigned to an MPLS packet. The EXP setting is three bits providing 23 values (eight values). The mechanisms used to control the setting of the EXP bit are referred to as tunnel modes, and are defined in RFC3270. The three tunnel modes are called uniform, short pipe, and pipe tunnel modes.

In uniform mode, any changes made to the EXP value of the topmost label on a label stack are propagated both upward, as new labels are added, and downward, as labels are removed.

In short pipe mode, the IP precedence bits in an IP packet are propagated upward into the label stack as labels are added. When labels are swapped, the existing EXP value is kept. If the topmost EXP value is changed, this change is propagated downward only within the label stack, not to the IP packet.

Full pipe mode is just like short pipe mode, except the PHB (per-hop behavior) on the mpls2ip link is selected, based on the removed EXP value rather than the recently exposed type of service value. The underlying type of service in the IP header in the packet is not modified.

Regarding current levels of support for MPLS tunnel modes, previous PFC3x forwarding engines supported short pipe and uniform tunnel modes. The PFC4 now adds support for full pipe tunnel mode.

Layer 3 - Increased Support for Ethernet over MPLS Tunnels

Ethernet over MPLS (EoMPLS) provides a mechanism for two disparate LAN networks to be joined over an MPLS network, thus allowing Ethernet frames to traverse an MPLS cloud. The PFC3x supported a maximum of 4 K

EoMPLS tunnels. The number of tunnels now supported with the PFC4 has been increased to 16 K tunnels in hardware.

Layer 3 - MPLS Aggregate Label Support

The number of MPLS aggregate labels supported on the PFC4 has been increased significantly over what was supported on the PFC3x. Aggregate label support has been increased to 16 K, up from 512 on the previous PFC3x family.

Layer 3 - Layer 2 Over GRE

The PFC4 family adds support for transporting Layer 2 packets over a generic route encapsulation (GRE) tunnel. This is a new capability added into the PFC4 hardware.

Layer 3 - Increased Multicast Routes

The PFC3x supported a maximum of 32 K multicast routes in hardware. The number of multicast routes that is supported in the PFC4 has been increased to 256 K. Note that the initial release will limit the maximum multicast routes to 128 K for IPv4 and 128 K for IPv6.

Layer 3 - PIM Register Encapsulation/De-Encapsulation for IPv4 and IPv6

PIM (Protocol Independent Multicast) is used to build a forwarding path through a network for multicast packets. As part of the process that PIM uses to build this forwarding path, multicast routers and switches will send PIM register packets to a device called the rendezvous point (RP). When these PIM register packets are sent, they are encapsulated in an IPv4 unicast header. If a Catalyst 6500 is configured as a PIM rendezvous point (RP), then the processing of these PIM register packets requires the packets to be de-encapsulated in software.

The PFC4 now supports the ability to both encapsulate and deencapsulate these PIM register packets in hardware, thus avoiding the performance penalty previously incurred on the Supervisor in processing these packets.

Layer 3 - IGMPv3/MLDv2 Snooping

IGMP is a multicast protocol that allows a host to indicate to the network that they wish to receive a multicast stream. The network uses this information to build a multicast topology over which multicast packets can be efficiently forwarded to hosts. IGMPv3 is the latest version of IGMP that allows hosts to specify a list of sources they would receive traffic from enabling the network to drop packets from unwanted sources.

Earlier PFC3x engines support hardware-based IGMP Layer 2 snooping for v1 and v2. IGMPv3 source-specific snooping used a combined software and hardware model, where software is used to track specific sources, which is then tied to a non source-specific hardware entry. The PFC4 now adds hardware-based snooping support for IGMP v3.

MLDv2 snooping is Layer 2 snooping for IPv6 multicast packets. The PFC4 supports hardware-based MLDv2 snooping for IPv6 hosts as well as IGMPv3 snooping for IPv4 hosts.

Layer 3 - Increased Support for NetFlow Entries

Up to 512 K NetFlow entries can now be stored in the PFC4, and up to 1 M NetFlow entries (512 K for ingress NetFlow and 512 K for egress NetFlow) can now be stored in PFC4XL. This is quadruple the number of entries offered on the comparable PFC3x forwarding engines.

Layer 3 - Improved NetFlow Hash

The NetFlow implementation on all PFC forwarding engines uses a hash to both store and retrieve entries in the NetFlow table. With each generation of PFC engine, the efficiency of the hash has improved, allowing a greater percentage of the NetFlow table to be utilized. With PFC2 and PFC3a, the hash efficiency was 50 percent. With subsequent PFC3x, it improved to 90 percent. With PFC4, the hash has been improved yet again to provide a hash efficiency of close to 99 percent.

Layer 3 - Egress NetFlow

Previously, NetFlow was only supported for ingress data traffic. Egress NetFlow provides support for collecting flow statistics for packets after they have had ingress processing applied to them, and prior to transmission out the egress interface or interfaces. This can be of value, especially for users wanting to collect flow statistics for data moving into and out of a VPN, for example.

Layer 3 - Sampled NetFlow

Sampled NetFlow is a new feature in the PFC4 that allows customers to opt for NetFlow records to be created based on a sample of the traffic matching the flow. Sample Netflow uses a 1/N based sampling which inspects one packet every N packets. The PFC3x was capable of performing sampling but the operation was performed after the inspection process. The PFC4 performs the sampling during the inspection process, effectively reducing the amount of NetFlow entries. There are 1 K global NetFlow samplers supported in PFC4.

Layer 3 - MPLS NetFlow

The PFC4 provides support for aggregate label at the Provider Edge (PE). This feature allows the inspection of IP traffic belonging to a particular VPN before it is added with a MPLS label (ip2mpls) and after the last label is removed (mpls2ip). The MPLS NetFlow also allows the inspection of the IP header for non aggregate label at a P device (mpls2mpls).

Layer 3 - Layer 2 Netflow

The layer 2 Netflow feature in the PFC4 allows Netflow lookups for IPv4, IPv6 and MPLS based packets to be performed using the Layer 2 header.

Layer 3 - Flexible NetFlow

The PFC4 now supports Flexible NetFlow (FnF), based on the NetFlow v9 record format. FnF allows users more flexibility in defining which record types they want to use for a v9 record. More importantly, FnF now also includes a number of new field options to allow for collection of MPLS, IPv6, and multicast information in a NetFlow record.

Layer 3 - Distributed Policing

In a PFC3x-based system using DFC3s, an aggregate policer applied on a VLAN that included ports on different DFC3-enabled linecards could not synchronize their token buckets. As such, each DFC3-enabled linecard would maintain its own aggregate policed count, resulting in the true aggregate rate being multiplied by the number of DFC3s which apply the policer.

PFC4 solves this problem with the introduction of the distributed policer. This allows the policing state to be synchronized across multiple DFC4-enabled linecards providing for multi-port multi-module policing. A total of 1 K distributed policers are supported with PFC4.

Layer 3 - DSCP Mutation

Quality of Service (QoS) support in PFC4 is enhanced with its support for multiple ingress and egress Differentiated Services Code Point (DSCP) mutation maps. A DSCP mutation map is a table that defines to what an existing DSCP value in a packet can be modified. DSCP mutation maps facilitate the marking (or reprioritizing of packets) process. Up to 14 ingress DSCP mutation maps and up to 16 egress DSCP mutation maps can be defined in the PFC4.

Layer 3 - Aggregate Policers

An aggregate policer is a rate limiting policy that can be applied to a port, group of ports, VLAN or group of VLANs that limits total traffic through those ports/VLANs to a predetermined bandwidth amount. Traffic in excess of the limit can either be marked down and forwarded, or dropped. The previous PFC3x forwarding engine supported a maximum of 1023 aggregate policers per chassis. PFC4 increases the limit on aggregate policers supported to 6 K.

Layer 3 - Microflow Policers

Like an aggregate policer, a microflow policer is a rate-limiting policy that can be applied to a port, group of ports, VLAN, or group of VLANs that limits total traffic for each flow through those ports or VLANs to a stated bandwidth amount. Traffic in excess of the limit can either be marked down and forwarded, or dropped. The previous PFC3x forwarding engine supported a maximum of 63 microflow policers per chassis. PFC4 increases the limit on microflow policers supported to 127.

Another enhancement to microflow policing is that with PFC4, a microflow policer can now be configured for both ingress and egress. Previously, a microflow policer could only be configured for the ingress direction.

Layer 3 - Policing by Packets or Bytes

The PFC4 now introduces support for a policer to enforce its rate, using either a packet or byte count. Previously, only byte count was supported.

Layer 3 - Cisco TrustSec (CTS)

CTS is an access control architecture where Security policies are enforced based on group membership as opposed to IP or MAC address. Group membership policies are inherently built into each packet by attaching a Security Group Tag (SGT) to each packet that enters a CTS domain. The SGT is assigned by the ingress switch and is used at the egress switch to determine access rights. The SGT is used in conjunction with Roles Based ACL (RBACL).

Layer 3 - Role-Based ACL

RBACL is an integral part of the CTS model that provides for a scalable way to apply access control within a CTS domain. Policies applied using RBACLs are assigned user group policies that encompass multiple end hosts. Data sent by hosts is tagged with an SGT that is carried inside a CTS header. This SGT is assigned by the hardware, and is carried with the packet as it traverses through the network. At intermediate nodes, the SGT can be modified or reassigned using classification ACLs. When the packet arrives at the network edge, the RBACL can be used to enforce Security policies that have been defined by the assigned SGT.

Layer 3 - Layer 2 ACL

PFC4 introduces enhanced support for a Layer 2 ACL that allows inspection of all of the Layer 2 fields. The ACL can inspect source and destination MAC address, Ethertype, VLAN ID, 802.1p user priority (or class of service) bits, and outer and inner tags (for 802.1Q tunnel packets).

Layer 3 - ACL Dry Run

PFC4 introduces the capability to perform a dry run (or pre-commit test) of a Layer 3 ACL. Earlier Ternary Content Addressable Memory (TCAM) implementations attempted to program the ACL without verifying whether enough space was available. The new ACL Dry Run feature allows the user to temporarily use a portion of the ACL TCAM to first test whether the configured ACL will fit. This guarantees Layer 3 ACL hardware programming will complete, prior to commitment.

Layer 3 - ACL Hitless Commit

PFC3x and earlier TCAM programming implementations are executed immediately upon configuration commit. If an ACL is applied to an interface, and then the configuration is changed, TCAM programming must first remove the previous ACL configuration before applying the new. There is a small period of time, while the TCAM programming is still being completed, that the only ACL entry present (default) is implicit deny. During this brief period, packets hitting the interface will be dropped until the updated TCAM programming process is complete.

PFC4 introduces the ability to program the ACL TCAM entries using a special pointer. Using this new ability, the ACL Hitless Commit feature allows the modified ACL to be programmed using new TCAM entries, while the original ACL entry remains intact. Once the new TCAM programming is complete, the old ACL TCAM pointer is removed, and replaced by the new pointer. This eliminates the temporary implicit deny scenario during transition.

Layer 3 - Layer 2 + Layer 3 + Layer 4 ACL

PFC3x provided support for Layer 2 or Layer 3/4 ACLs, but not both at the same time. With this new ACL type, PFC4 allows both Layer 2, 3 and 4 information to be inspected at the same time. Especially useful in wireless networks where mobility can often change the user’s source IP address, the ACL could be built to inspect the source MAC address along with other higher layer information (such as Layer 4 port information) to apply a Security policy.

Layer 3 - Classification Enhancements

The PFC4 engine offers several extensions to classification ACLs. As well as being able to match on the traditional classification options such as IP address, TCP/UDP ports, and others, the PFC4 also offers match on packet length, Time to Live (TTL), IP options, and IPv6 Extended Header. Some worms and other forms of attack sometimes require matching on these fields to make a positive ID.

Layer 3 - Per Protocol Drop (IPv4, IPv6, MPLS)

PFC4 adds support for the ability to only forward protocol traffic if enabled on the interface. The protocols that can be defined at an interface level are IPv4, IPv6, and MPLS. Traffic not matching the defined protocol will be dropped.

Layer 3 - Increase in ACL Label Support

An ACL label is used to group access control entries (ACEs) that are associated with the same access control list. An access control list entry that starts with “access-list 101....” uses the label “101” to indicate which ACL group this entry belongs to. Previously, the PFC3B/XL and PFC3C/XL supported a maximum of 4096 ACL labels. PFC4 increases support for the number of ACL labels to 16 K.

Layer 3 - Increase in ACL TCAM Capacity

The PFC4 forwarding engine implements two banks of TCAMs for classification purposes, providing a total of 256 K access control entries for DFC4XL and 64 K for DFC4. These ACEs can be shared between the Security and QoS for both ingress and egress lookups. There is also a corresponding increase in the mask to entry ratio. This capability is discussed in more detail later in the paper. In summary, this allows for more efficient use of the TCAM space when defining Security policies.

Layer 3 - Source MAC + IP Binding

A binding of IP address, VLAN, and MAC address can be made to facilitate the decision-making process for forwarding packets. This enforcement is performed by the PFC4 in hardware, and is especially useful in protecting against address spoofing. The IP source guard is an example of one feature that makes use of this capability.

Layer 3 - Drop on Source MAC Miss

This feature is another hardware enhancement that is used to further enhance the port Security feature. Port Security can be used to bind MAC addresses to a given port, ensuring that only packets with the defined MAC address are forwarded.

Layer 3 - RPF Check Interfaces

The Reverse Path Forwarding (RPF) check is used to help determine if a packet has been spoofed. It uses a reverse lookup, whereby the source address is used to initiate the lookup. If the packet arrives on an interface where its source address is not seen as existing out that interface, it is deemed a spoofed packet and will be dropped. With an RPF check, multiple paths can be incorporated into the lookup. Previous PFC3x engines supported 2 paths in the RPF lookup. PFC4 increases the number of paths included in the lookup to 16.

Layer 3 - RPF Checks for IP Multicast Packets

RPF checks for IP Multicast packets were previously performed in software, and the correct RPF interface was then programmed into the hardware forwarding entries. PFC4 allows full hardware-based RPF checks for IP Multicast. This capability also allows for dual RPF check, to support PIM sparse-mode shortest path tree (SPT) switchover to occur in hardware.

PFC4 Architecture Enhancement Summary

The following section provides some details for the PFC4 architecture, its performance metrics, and an explanation of some of the feature enhancements that it introduces.

PFC4 Forwarding Architecture and Performance Metrics

The forwarding architecture of the new Supervisor 2T is based on the Cisco Express Forwarding (CEF) architecture, the same forwarding architecture used on the Supervisor 720. One of the primary enhancements of the Supervisor 2T (over the Supervisor 720) is the doubling of the centralized forwarding performance to 60 Mpps.

The following sections provide an overview of the major changes in functionality, scalability, and performance for Layer 2 and Layer 3 forwarding features supported by the PFC4.

Table 10. PFC4 Layer2 and Layer 3 Feature Comparison with Previous Generations of PFCs

Feature	PFC3B Sup 720-3B	PFC3BXL Sup 720-3BXL	PFC3C Sup 720-10G-3C	PFC3CXL Sup 720-10G-3CXL	PFC4 Sup 2T	PFC4XL Sup 2T-XL
IPv4 forwarding	30 Mpps	30 Mpps	48 Mpps	48 Mpps	60 Mpps	60 Mpps
IPv6 forwarding	15 Mpps	15 Mpps	24 Mpps	24 Mpps	30 Mpps	30 Mpps
MPLS forwarding	30 Mpps	30 Mpps	48 Mpps	48 Mpps	60 Mpps	60 Mpps
Layer 2 forwarding	30 Mpps	30 Mpps	48 Mpps	48 Mpps	60 Mpps	60 Mpps
EoMPLS imposition	30 Mpps	30 Mpps	48 Mpps	48 Mpps	60 Mpps	60 Mpps
EoMPLS disposition	15 Mpps	15 Mpps	24 Mpps[**]	24 Mpps[**]	30 Mpps	30 Mpps
FIB TCAM	256 K	1 M	256 K	1 M	256 K	1 M
Adjacency Table	1 M	1 M	1 M	1 M	1 M	1 M
MAC (CAM)	64 K (32K)[*]	64 K (32K)[*]	96 K (80K)[*]	96 K (80K)[*]	128 K	128 K
EtherChannel hash	3 bits	3 bits	3 bits	3 bits	8 bits	8 bits

[*] The number outside of the brackets represents the maximum hardware capacity. The number inside the brackets represents the average utilization expected by users based on hash-based table programming. Note that average utilization can get up to the maximum hardware limit but that will depend on result of hashing.

[**] Those numbers are for underlying IPv4 traffic only.

MAC address learning continues to be performed in hardware, as was done with the Supervisor 720. The MAC (CAM) table in the Supervisor 2T has been increased in size to 128 K and by using a new architecture the efficiency has been improved to 99 percent.

The aggregate forwarding performance of the chassis will multiply by 60 Mpps for each DFC4-based linecard that is installed in the chassis. This facilitates an aggregate system performance up to 720 Mpps, assuming 6513-E.

PFC4 Security and QoS Architecture

Security and QoS functionality have been enhanced on the Supervisor 2T. One of the main enhancements has been the move to a consolidated TCAM (memory) bank for holding Security and QoS ACLs that have been defined in Security and QoS Policies in the switch configuration. In previous PFC3x forwarding engines, there were two separate TCAM banks used for this purpose, each 32 K in size. One TCAM bank was used for Security ACLs and the other bank was used for QoS ACLs.

With Supervisor 2T, up to 256 K entries are available for use (this is in the PFC4XL). By default in the PFC4XL, 64 K entries are reserved for QoS, while 192 K entries are available for Security ACLs and related feature ACLs (such as NAT, WCCP, and others). Based on user-specified software configuration, up to 128 K entries can be made available for QoS ACLs.

Another major improvement is the ACE to mask ratio. An ACE is one ACL permit/deny statement that exists within the framework of an ACL. In the ACE statement, the mask is what is used to identify what portion of the address space should be used to match on the incoming or outgoing address. Let’s look at an example to better understand how the ACE breaks down into the different functional components.

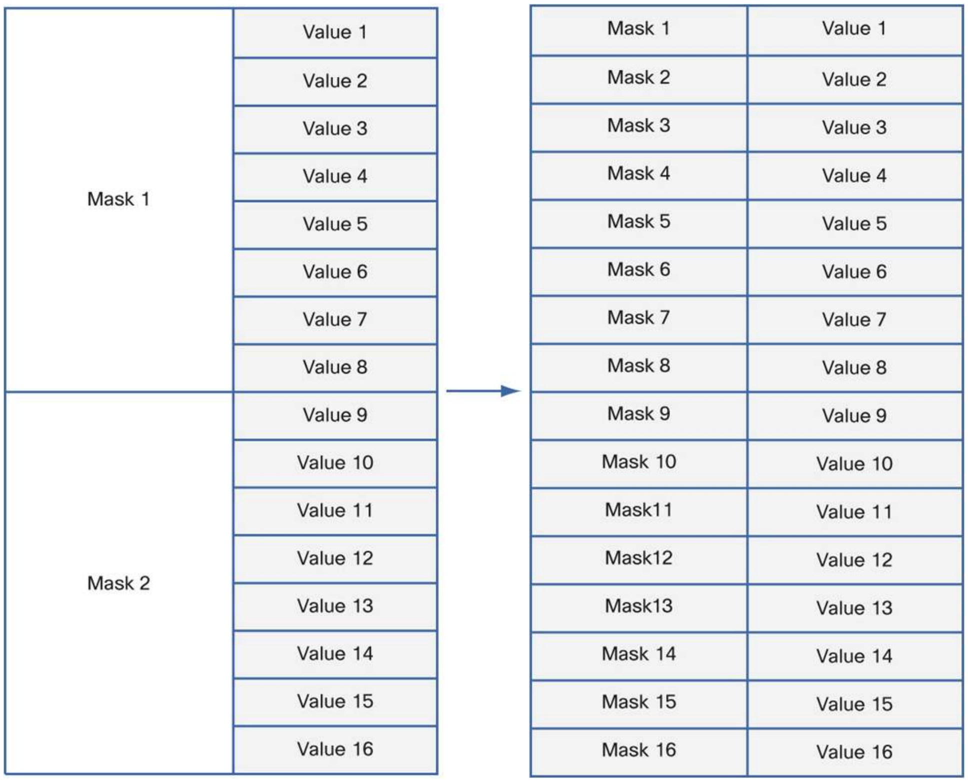
access-list 101 permit ip 10.1.1.0 0.0.0.255 any

In the above example, the entire line is one ACE. This ACE typically forms part of a larger ACL, consisting of multiple configuration lines. The mask is the “0.0.0.255” part of the ACE, while the value is the “10.1.1.0” part of the ACE example above. In this example, the mask signifies that only the first 24 bits of the IP address should be used to match on classified packets.

With PFC3x forwarding engines, the hardware tables support 32K ACEs and 4 K masks, which yields an 8:1 ratio of ACEs to masks. Depending on how a customer might build their ACL, there is potential for masks to be consumed before ACEs were consumed leading to a potential inefficient use of the Security TCAM. With Supervisor Engine 2T, the mask to value ratio has changed and now supports a 1:1 ratio providing 1 mask for each ACE (value). This should increase the flexibility for how customers can deploy ACLs and minimize any potential inefficient use of table entries.

The diagram below shows how the masks and values are represented in the hardware TCAMs with PFC3x on the left and Supervisor 2T on the right.

Figure 9. ACL TCAM Mask Layout Before (on PFC3x) and After (on PFC4)



QoS on the Supervisor 2T now offers support for distributed policing. For a Supervisor 720-based system, a rate-limiting policy applied to a VLAN that had member ports spread across DFC3-enabled linecards would result in each DFC3 maintaining its own token bucket. In other words, a rate-limiting policy of 2 Gbps would result in each DFC3 maintaining its own 2 Gbps rate limiting count. With the Supervisor 2T, synchronization of the rate limiting policy occurs between participating DFC4-enabled linecards, ensuring the aggregate traffic load for traffic in that VLAN is truly limited to the configured rate.

The following table provides a summary of the enhancements incorporated into the Supervisor 2T for Security and QoS.

Table 11. PFC4 Security and QoS Comparison with Older PFCs

Feature	PFC3B	PFC3BXL	PFC3C	PFC3CXL	PFC4	PFC4XL
Security ACL entries	32 K	32 K	32 K	32 K	Up to 48 K*	Up to 192 K*
Security ACL labels	4 K	4 K	4 K	4 K	16 K	16 K
Security ACL masks	4 K	4 K	4 K	4 K	Up to 48 K	Up to 192 K
ACE/mask ratio	8:1	8:1	8:1	8:1	1:1	1:1
ACL LOUs	64	64	64	64	104	104
ACL L4OPs	10 per ACL	10 per ACL	10 per ACL	10 per ACL	10 per ACL	10 per ACL
Cisco TrustSec	No	No	No	No	Yes	Yes
Role-based ACL	No	No	No	No	Yes - up to 32 K	Yes - up to 64 K
IPv6 uRPF	No	No	No	No	Yes	Yes
IPv4 uRPF load sharing paths	Up to 6	Up to 6	Up to 6	Up to 6	16	16
QoS ACL entries	32 K	32 K	32 K	32 K	Up to 32 K*	Up to 128 K*
QoS ACL labels	4 K	4 K	4 K	4 K	Up to 16 K*	Up to 16 K*
QoS ACL masks	4 K	4 K	4 K	4 K	Up to 32 K*	Up to 128 K*
Distributed policers	No	No	No	No	Up to 4 K	Up to 4 K
Egress microflow policing	No	No	No	No	Yes	Yes
Number of aggregate policers	1023	1023	1023	1023	16 K	16 K
Number of microflow policers	63	63	63	63	127	127
Packet or byte-based policing	No	No	No	No	Yes	Yes

* ACL TCAM is combined for Security and QoS in PFC4 and PFC4XL

The front panel ports on the Supervisor 2T have different QoS configurations, based on the configured mode of the ports. When the 10 G switch ports are setup as a VSL, the QoS configuration is different than when those ports are not configured for VSL. The ports can also be configured to operate in 10 G mode only (effectively disabling the 3 x 1 GE ports). This also has an effect on the QoS configuration depending on in which mode the associated switch port is operating. The QoS setup for the front panel ports is shown in the following table.

Table 12. Supervisor 2T Front Panel Port QoS Configuration

	10 GE Ports	3 x 1 GE Ports
No VSL (10 G and 1 G ports active)	4 Queues	4 queues
No VSL (10 GE mode only)	8 Queues	Shutdown
VSL (10 G and 1 G ports active)	4 Queues	4 queues
VSL (10 G mode only)	8 Queues	Shutdown

PFC4 NetFlow

The support for NetFlow in the PFC4 has been enhanced on both the scalability and functionality side. Of perhaps the most significance is support for true egress NetFlow. With the change in how a packet is processed and the new pipeline processing method that the PFC4 uses to process a packet, there are two separate processing paths that a packet will take: one for ingress services and one for egress services.

While this is discussed in more detail later in this paper, the PFC4 now allows for both ingress and egress NetFlow services to be performed on all packets. One of the biggest benefits of egress NetFlow is the ability to account for packets that are encapsulated or de-encapsulated from tunnels and those packets entering or leaving an MPLS cloud. Another example is to account for egress Multicast packets which are replicated (number of outgoing interfaces [OIFs]) from a single ingress packet.

Support for Flexible NetFlow (FnF) is now built into hardware. FnF offers a more flexible method to create flow monitors that allow for the collection of data that fits user specified templates. In this manner, an administrator can create a flow monitor to collect IPv6 specific information on one interface, while on another interface create a separate flow monitor to collect IPv4 multicast specific information.

Cisco TrustSec (CTS)

This architecture uses access control, authentication, and encryption to build a scalable, highly secure network. There are three important elements of the Cisco TS architecture that is part of the hardware capabilities in the Supervisor 2T:

- Support for SGT and Destination Group Tag (DGT) tagging
- Role Based ACL (RBACL) ink layer encryption (IEEE 802.1ae)

The support for IEEE 802.1ae link layer encryption is specific to the port ASIC that is located on the Supervisor 2T baseboard, and is not part of the PFC4 or PFC4XL capability.

The Security Group Tag (SGT) and Destination Group Tag (DGT) are tags that are inserted into a packet and are used to define the Security policies that should be applied to this packet as it traverses the CTS cloud. Cisco TrustSec uses an eight-byte header and contains sixteen bits that are used to indicate the SGT or DGT for that packet. RBACL provides a means to provide classification of packets using the SGT/DGT to apply Security policies.

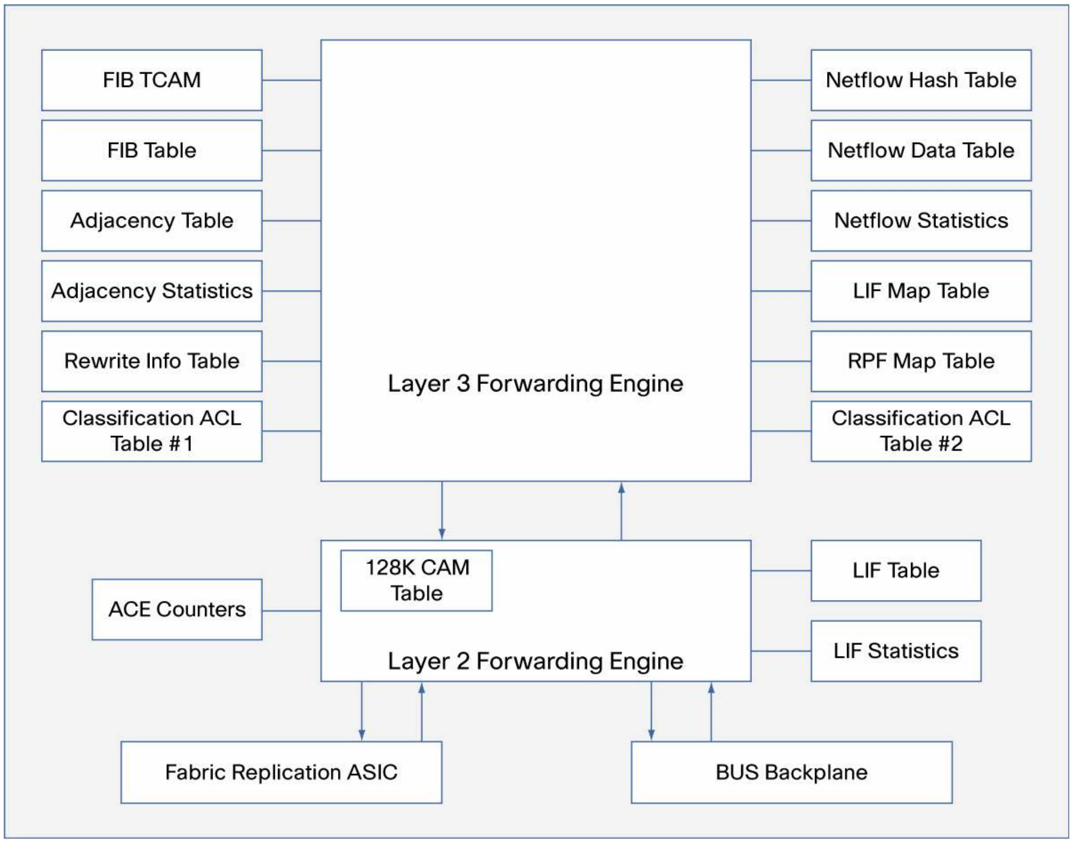
The PFC4 provides support for SGT, DGT, and RBACL in the following manner:

- Both SGT and DGT assignment can be performed by the PFC4
- The SGT can be derived during packet processing on ingress from the input packet or from an ingress ACL
- The DGT can be derived from the destination IP lookup (in the FIB), from the NetFlow process, or the ingress ACL
- RBACL is supported on the egress interface
- CTS Tunnel encapsulation

PFC4 Architecture

The PFC4 is made up of two main ASIC processing blocks, along with a number of high-speed memory blocks that serve to provide the hardware acceleration of selected features. One ASIC block performs Layer 3 services, while the other ASIC block performs Layer 2 services. A high-level view of the PFC4 is shown in the following diagram.

Figure 10. PFC4 Functional Blocks



At the center of the PFC4 complex are the two forwarding engines. These two ASIC complexes are responsible for the forwarding of all Layer 2 and Layer 3 packets in hardware. Attached to each of these ASIC blocks are a series of tables that are used to store information that facilitates the forwarding of packets in hardware.

The following sections provide more details about each of these two forwarding engines and the associated tables with which they interface.

Layer 2 Forwarding Engine

This engine is responsible for Layer 2 packet processing and supports a number of enhancements beyond those found on the Layer 2 forwarding engines in previous PFC3x complexes. Integrated into the forwarding engine ASIC is a MAC address table containing 128 K entries. The MAC address table consists of two banks of 4 K lines with 16 entries per line (2 x 4 K x 16 = 128 K entries). Each entry in the MAC address table is 115 bits wide, and contains forwarding and aging information related to a destination MAC entry and an associated bridge domain pair.

Rather than running a hash operation to find a pointer into the first bank, and then run the hash again to derive a pointer into the second bank, two simultaneous hash functions are performed at the same time to provide a pointer into each of the memory banks. In this manner, the Layer 2 lookup performance is maximized.

Prior to PFC4, every interface in the system was identified by a VLAN ID, including internal uses such as L3 sub-interfaces, VPNs, tunnels, and egress multicast replication-mode. This restricted the total number of unique interfaces to 4096. The new Bridge Domain (BD) concept is one of the more significant enhancements introduced with the PFC4, and is designed to increase scaling of VLANs internally to the switch. When a user creates a VLAN, it maps internally to a unique bridge domain. Support for 16 K bridge domains is built into PFC4 hardware. However, at FCS, only 4 K VLANs will be supported by the software running on the Supervisor 2T.

All frames entering the Layer 2 forwarding engine are associated with a Logical Interface (LIF), which is, in essence, a map to a port index and VLAN pair on which the frame entered the switch. A LIF database of 512 K entries (each comprised of BD, LIF, and control bits) resides in the Layer 2 forwarding engine. Each LIF entry is ultimately used to facilitate Layer 3 processing whenever the packet is passed to the Layer 3 forwarding engine. Along with the LIF database is a LIF statistics table. This table maintains diagnostic VLAN counters, along with byte and frame count statistics per ingress and egress LIF, and consists of one million entries.

The Layer 2 forwarding engine maintains a set of Access Control Entry (ACE) counters. When the Layer 3 forwarding engine performs classification processing, it will communicate with the Layer 2 forwarding engine to update ACL counters when a hit is registered against an ACE (such as a line in an ACL list).

The following table provides a summary of the main differences between Layer 2 support in the PFC4 and previous PFC3x versions.

Table 13. PFC4 Layer 2 Forwarding Engine Features

Layer 2 Feature	PFC3B/PFC3BXL	PFC3C/PFC3CXL	PFC4/PFC4XL
MAC address table	64 K	96 K	128 K
Number of VLANs	4 K	4 K	16 K (bridge domains)
VPLS forwarding and learning	No	No	Yes
Source MAC miss redirection	No	No	Yes
EtherChannel hash	3 bits	3 bits	8 bits
ACE counters	32 K (on L3 ASIC)	32 K (on L3 ASIC)	256 K
LIFs*	4 K	4 K	128 K
Physical interfaces	4 K	4 K	16 K
LIF/VLAN statistics	VLAN stats: 4 K x 6 counters	VLAN stats: 4 K x 6 counters	LIF stats: 1 M counters
Layer 2 rate limiters	4	4	20 ingress/6 egress
VSL support	No	Yes	Yes

* For PFC3x, the logical interfaces and VLANs shared the same 4 K pool

Even for Layer 3 routed flows, the Layer 2 forwarding engine performs a number of critical services prior to handing the packet over to the Layer 3 forwarding engine for Layer 3 processing. Included in its processing list are the following functions:

- Performs CRC error checking
- Performs LIF and BD (bridge domain) lookup
- Maintains LIF statistics
- Performs Layer 2 MAC table lookup
- Calculates Result Bundle Hash (RBH) or EtherChannel load-balancing hash
- Determines 16 static MAC match conditions for system frames (such as CDP, BPDU, and more)
- Performs IGMP/MLD/PIM snooping
- Performs hardware rate limiting
- Provides ACE counters

Layer 3 Forwarding Engine

The Layer 3 forwarding engine performs Layer 3+ services, including IPv4, IPv6, and MPLS forwarding lookups, as well as Security, QoS, and NetFlow policies on packets traversing the switch. There have been a number of enhancements incorporated into this PFC4 layer3 forwarding.

From a capacity standpoint, fundamentally faster packet processing, support for more NetFlow entries, and more ACLs form only part of the wide range of features whose operational limits have been increased. Some new features have also been introduced, such as support for egress NetFlow, egress microflow policing, and distributed policing.

A summary of the major changes for this Layer 3 ASIC, compared to earlier PFC3x found on previous Supervisor engines, is detailed in the table below.

Table 14. PFC4 Layer 3 Forwarding Engine Features

Layer 3 Feature	PFC3B/PFC3BXL	PFC3C/PFC3CXL	PFC4/PFC4XL
FIB table	Up to 1 M (XL)	Up to 1 M (XL)	Up to 1 M (XL)
Adjacency table	1 M	1 M	1 M
Adjacency statistics	512 K	512 K	512 K
CEF load sharing paths	16	16	16
Total SVI	4 K	4 K	128 K (64 K non XL)
Number of VPNs	4 K	4 K	16 K
MPLS aggregate VPN labels	512	512	16 K
Location of aggregate VPN label	L2 forwarding engine	L2 forwarding engine	L3 forwarding engine
NetFlow entries	Up to 256 K (XL)	Up to 256 K (XL)	1 M (XL) (Ingress: 512 K) (Egress: 512 K)
Egress NetFlow	No	No	Yes
NetFlow flow masks	4	4	80 - 32 for IPv4, 32 for IPv6, 8 for L2, 8 for MPLS
Flexible NetFlow	No	No	Yes
Copy-based NetFlow	No	No	Yes
Sampling in hardware	No	No	Yes
Number of NetFlow samplers	N/A	N/A	1 K
MPLS over GRE	No	No	Yes
Label operation in one pass	Push 3 Pop 2	Push 3 Pop 2	Push 5 Pop 1
Number of EoMPLS VC	4 K	4 K	128 K
MPLS QoS modes	Uniform, half pipe	Uniform, half pipe	Uniform, half pipe, pipe
ACL labels	4 K	4 K	16 K
Security ACLs	32 K	32 K	48 K (non-XL default) 192 K (XL default)
ACL counters	32 K	32 K	256 K
ACL LOU	64	64	104
QoS ACLs	32 K	32 K	16 K (non-XL default) 64 K (XL default)

Layer 3 Feature	PFC3B/PFC3BXL	PFC3C/PFC3CXL	PFC4/PFC4XL
Port ACLs	2 K	2 K	8 K
ACL accounting statistics	None	None	4 K
RPF interface check	2	2	16
Hardware rate limiters	8 (L3)	8 (L3)	32 (L3)
Aggregate policers	1023	1023	16 K
Aggregate policer profile	N/A	N/A	1 K
Microflow policer buckets	Up to 256 K	Up to 256 K	512 K IFE and 512 K OFE*
Shared microflow policers	63	63	512
Egress microflow policing	No	No	Yes
Distributed policers	No	No	4 K
Packet or byte-based policing	No	No	Yes
QoS policy groups	0	0	128
DSCP mutation maps	1	1	14 Ingress 16 Egress

* IFE and OFE are defined in the next section

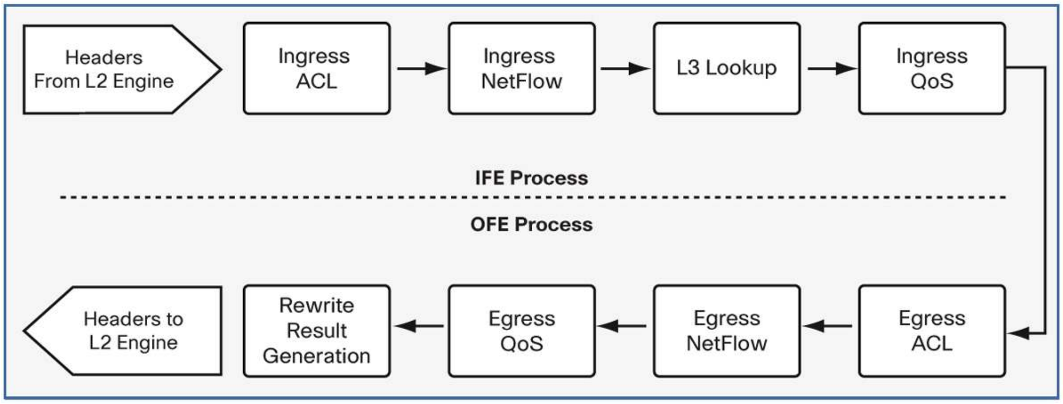
Layer 3 Forwarding Engine Processing Paths

The Layer 3 forwarding engine has two basic processing paths (also referred to as pipelines): one for ingress (or input) forwarding (IFE) and one for egress (or output) forwarding (OFE). These two pipelines perform the following functions:

- The IFE pipeline performs ingress functions, including input classification, input QOS, ACLs, RPF checks, ingress NetFlow, and L3 FIB-based forwarding.
- The OFE pipeline performs egress functions, including adjacency lookup, egress classification, and rewrite instruction generation.

When a packet header enters the L3 ASIC, the IFE pipeline is the first pipeline to process the packet. After completion of IFE processing, the header is then passed onwards to the OFE pipeline, along with the results of the IFE processing. This can be seen in the following diagram.

Figure 11. Layer 3 Forwarding Engine Processing Pipelines



The processing cycles for IFE and OFE processing are always performed in an IFE/OFE order. At the completion of OFE processing, the Layer 3 forwarding engine will collate the results and hand the packet back to the Layer 2 forwarding engine for onward processing. The processing performed by this ASIC is displayed in the following table.

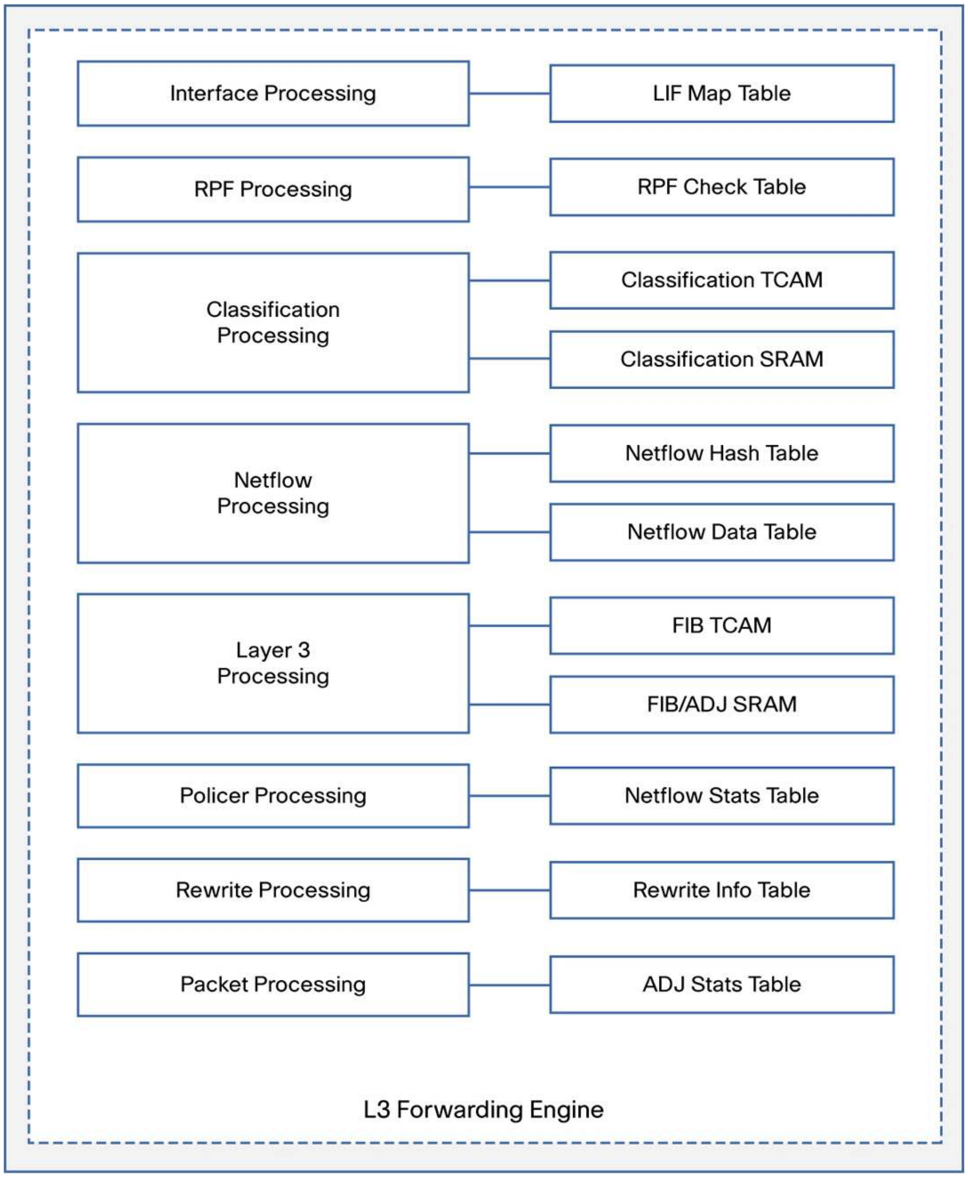
Table 15. PFC4 Layer 3 Forwarding Engine IFE and OFE Functions

Function	IFE (Ingress Processing)	OFE (Egress Processing)
CRC check on incoming frames from the L2 engine	Yes	N/A
Checks IFE processing result before performing OFE processing	N/A	Yes
Ingress LIF map table lookup	Yes	N/A
Egress LIF map table lookup	N/A	Yes
RPF check	Yes	No
Security ACL classification	Yes	Yes
Security ACL classification based on SGT	Yes	Yes
Security ACL classification based on DGT	No	Yes
RBACL - generation of SGT/DGT	Yes	No
QoS ACL classification	Yes	Yes
ACL redirects	Yes	Yes
Aggregate policing	Yes	Yes
Microflow policing	Yes	Yes
Distributed policing	Yes	Yes
Ingress DSCP mutation	Yes	N/A
Egress DSCP mutation	N/A	Yes
ACL-based accounting	Yes	Yes
NetFlow flow creation	Yes	Yes
NetFlow redirects (WCCP, NAT, TCP Intercept, etc)	Yes	Yes
MTU check	No	Yes
TTL check	No	Yes
Generates rewrite information	Performed independently of IFE and OFE	
Update adjacency statistics	Performed independently of IFE and OFE	
Update accounting statistics	Performed independently of IFE and OFE	
Execute CPU Rate limiters	Performed independently of IFE and OFE	

PFC4 (and PFC4XL) Functional Elements

The Layer 3 forwarding engine contains a number of functional elements, all of which work together to provide Layer 3 processing for packets. The main functional elements are executed in the order shown and can be seen in the following diagram.

Figure 12. Layer 3 Forwarding Engine Functional Elements



Each of these functional elements is detailed in the following sections.

Interface Processing and LIF Map Table

A LIF is a new concept introduced with the PFC4. The LIF helps enable per-port-per-VLAN interface processing, which separates Layer 2 characteristics from Layer 3 characteristics. The LIF map table (which is separate from, but is mapped to, the LIF table located on the Layer 2 forwarding engine) contains 128 K entries, and helps scale the number of Layer 3 interfaces and sub-interfaces that can be supported by the PFC4. There are two LIF map tables used by the Layer 3 forwarding engine: one for ingress LIFs and one for egress LIFs.

Built within the LIF map table entry are a number of fields that define the operating characteristics associated with each LIF. Some of the information contained within a LIF entry is used as an input for subsequent table lookups by other Layer 3 forwarding engine processes.

Examples of the information contained within a LIF map table entry include the following:

- Security Group Tag (SGT) - applicable for CTS processing
- Tunnel interface information (GRE, EoMPLS, IPv6, and others)
- RPF lookup requirements
- MPLS and EoMPLS interface information
- MPLS VPN ID
- IPv4/IPv6 VPN ID
- Trusted or un-trusted state (from a QoS perspective)
- ACL label

For ingress processing, the LIF table facilitates the following functions:

- Helps derive per logical interface configured parameters such as ACL labels, VPN, and more (some of this information forms part of the lookup key in other tables)
- Holds information that is used for PACL (Port ACL) lookup
- Supports IP multicast filtering
- Filters at Layer 3 for MPLS packets

For egress processing, the LIF table can help with the following:

- Helps derive per logical interface configured parameters such as ACL labels, VPN, and more (some of this information is used as part of the lookup key in other tables)
- Performs interface check for routed packets
- Provides source filtering for multicast packets
- Performs scope enforcement for IPv6 packets
- Supports MPLS Fast Reroute for FRR-TE tunnels

RPF Processing and the RPF Map Table

The reverse path forwarding check is used to confirm that the source IP address associated with a frame is received on the interface that the FIB table lists as the correct source or RPF interface. For unicast forwarding, RPF check is performed to stop IP address spoofing through malformed or forged source IP addresses.

The PFC4 supports a maximum of 16 RPF interfaces for both IPv4 and IPv6. Both PFC3B/XL and PFC3C/XL supported looking up two interfaces during the RPF lookup, and none of the PFC3x forwarding engines supported IPv6 RPF lookups.

The following is an example of how RPF is used. A packet arrives on interface 3/1 and has a source address that is part of the subnet 193.10.1.x. The RPF process performs a reverse lookup on the forwarding table. Rather than looking at the destination address, it uses the source address for the lookup. The lookup determines that packets from network 193.10.1.x should arrive on interface 3/5. In this case, as the packet arrived on interface 3/1, it is deemed to be a spoofed packet, and thus, is dropped in hardware.

The RPF processing block responsible for the RPF check is also responsible for a number of other processing checks as well. These additional processing checks include the following:

- IP Multicast forwarding requires an RPF check, to build its distribution tree. PIM uses the RPF information to determine which interface to send Join and Prune messages, for a given IP source
- To support MPLS aggregate VPN support, a lookup using the MPLS label can be performed to determine the MPLS VPN id associated with it
- For IPv4 and IPv6 packets, a source MAC to IP binding check can be performed
- For IPv4 and IPv6 packets, source AS mapping check can be performed which facilitates a later lookup in the ACL TCAM
- For IPv4 and IPv6 packets, a source group tag (SGT) lookup can be performed
- VPN and QoS mapping for MPLS packets is supported

Classification Processing and the Classification Memory (ACLs)

Two TCAM banks providing a total of up to 256 K access control entries are available for classification ACLs. By default, the PFC4 reserves 16 K entries for QoS ACEs and 48 K entries for Security entries. The PFC4XL reserves 64 K entries for QoS ACEs and 192 K entries for Security ACEs. The first ACL TCAM bank is used for standard QoS and Security ACL entries, and the second TCAM bank is used for Security and CTS (RBACL) entries, which require label-based functions. This combined TCAM design can be used to store Security ACLs, QoS ACLs, RBACLs, or accounting results.

The Layer 3 forwarding engine provides for a dual lookup into each bank, allowing for four lookups to be performed simultaneously. This means that for each input packet, up to four classification rules can be matched during IFE (ingress) processing, and up to four classification rules can be matched during OFE (egress) processing.

Each classification TCAM entry is 144 bits wide, and uses a lookup key to initiate a lookup into the TCAM. This lookup key uses input fields such as the ACL label (which is obtained from the LIF table), packet type (IPv4, IPv6, and more) and other fields to generate the key. The result of the TCAM lookup provides a pointer into the classification SRAM that holds the actual ACE entry.

The Layer 3 forwarding engine works in conjunction with the Layer 2 forwarding engine, which holds and maintains the hit counters for each access control entry. Every time a classification ACL is matched during IFE or OFE processing, the ACL counters in the Layer 2 forwarding engine are updated to reflect the hit.

The classification TCAMs have the following capabilities:

- Ingress DSCP mutation, using one of 14 ingress DSCP mutation maps selected by ingress LIF map table lookup
- Address compression on IPv6 addresses, prior to lookup into the TCAM
- Security ACLs (ingress/egress classification) returns a permit/deny for the packet, based on configured ACL policies
- ACL redirects (ingress/egress classification), which define which packets are to be redirected (this mechanism can be used for features such as policy-based routing through input classification)
- RPF+ (ingress classification only), which provides the ability to ignore RPF result for certain classified flows
- RBACL support. The generation of SGT and DGT based on IP classification. Also, classification of flows based on SGT (input and output classification) or DGT (only output classification) of the packet is enabled through these TCAMs.
- Ability to generate ACL-based VPNs (ingress classification only), useful for implementing VRF select or newer routing mechanisms like MTR (multi-topology routing)

- Service cards virtualization, which provides the ability to generate a virtual ID, based on ingress or egress classification
- Ability to specify the QoS for a classified flow, and a policer index for aggregate policing, based on input or output classification
- ACL-based accounting (ingress and egress classification), plus ability to provide drop counters on each ACEs to implement ACL-based accounting
- Generation of fields required for NetFlow lookup (ingress and egress classification)
- Code logic that interfaces with the classification TCAMs also interfaces with the Layer 2 forwarding engine to maintain ACE statistics

NetFlow Processing and the NetFlow Hash and Data Tables

NetFlow is a hardware-enabled process that collects statistics on packet flows that traverse the switch. A flow is identified by a flow mask, which uses fields from the packet header to determine what constitutes a flow. The default flow mask uses the source and destination IP address, the source and destination port number, and the IP protocol to determine a flow.

Here is an example of how this is used. A user initiates three sessions: one email, one web, and one print. Each packet flow uses the same source IP address, but has a different destination IP address and port number pair. Using the default full flow mask, each session will be viewed as a separate flow, and statistics for each flow would be counted individually (a full flow mask uses the IP protocol field, src/dest ip address and src/dest port number to identify a unique flow).

However, if the administrator was to use, for example, a source-only flow mask (each flow identified by source IP address only), then the statistics count result would be different. A source-only flow mask identifies all packets associated with the same source IP address as part of the same flow. In our example above, the same user who initiated three sessions (email, web, and print) would now have all of that session data collected under one flow record, rather than three individual flow records when using the other flow mask. There are a number of flow masks that can be chosen by the user, and these can be configured as needed by the administrator.

At this stage of the processing flow, NetFlow processing is implemented across two banks of Reduced Latency DRAM (RLDRAM). One bank (NetFlow hash table) acts as a hash table holding pointers into the second bank (NetFlow data table), where the actual NetFlow data is held. The NetFlow data table is 288 bits wide, of which 160 bits represents the NetFlow key, and the other 128 bits are used for holding NetFlow data. A total of 1 M NetFlow entries can be stored (in the PFC4XL), and this is split evenly between IFE and OFE. For ingress (IFE) NetFlow, the Layer 3 forwarding engine maintains 512 K entries. Likewise, for egress (OFE) NetFlow, additional 512 K entries are also offered. For the non-XL version of the PFC4, the NetFlow table can hold up to 512 K entries, and those can be shared between both IFE and OFE.

NetFlow flow records can be created for IPv4 flows, IPv6 flows, and VPN flows. The IPv4 and VPN consume one entry, while IPv6 flows consume two entries. NetFlow statistics are maintained for each flow.

When a packet arrives for NetFlow processing, the flow mask is used to determine which fields in the packet header will be used to build a lookup key, which is used to search the NetFlow hash table for an existing entry. If an entry is found, the lookup will return a pointer into the NetFlow table and the NetFlow statistics table. If no entry is found, then a pointer entry is created and a flow record will be created in the NetFlow table.

In parallel with the increase in NetFlow entries is a corresponding increase in the number of flow masks that are supported. The flow mask is an important element in the NetFlow process, and warrants additional attention. As described above, the flow mask defines what constitutes a flow, and has an impact on how statistics are collected for different packet streams. With the previous PFC3x forwarding engines, there were six flow masks, of which two were reserved for system use. This left two flow masks that could be used by the user. The main use for the two different flow masks was with user-based rate limiting (UBRL), which allowed the user to configure different policers with different flow masks. With the PFC4, there are now 80 flow masks available for use. Of the 80, there are 32 flow masks for IPv4, 32 flow masks for IPv6, eight flow masks for MPLS, and eight flow masks for Layer 2 packet flows.

Microflow policing can be performed on every NetFlow entry in the NetFlow data table. This means you could potentially have 512 K ingress microflow policers and 512 K egress microflow policers in action at the same time when the system operates in PFC4XL mode.

While sampled NetFlow was available on the PFC3x forwarding engines, it was performed in software by the control plane. Sampled NetFlow on the PFC4 is a new feature now supported by the Layer 3 forwarding engine as a hardware process. It offers a way to reduce the amount of information that is collected about a given flow. Supervisor 2T supports 1:N Sampling, a method which allows one packet to be chosen out of N packets (e.g. 1 per 1000). Sampling operates in a random mode, meaning any packet within sample N will be randomly selected. Samplers are global, and a total of 1 K NetFlow samplers are supported.

A third piece of memory (ICAM) is also used for NetFlow. This piece of memory is used to store flows that have a conflict, due to a hash collision or page full condition in the primary NetFlow table. The ICAM can support 16 NetFlow entries that are shared between the IFE and OFE processes. Given the 99 percent efficiency of the NetFlow hashing algorithm and the aging mechanisms to age out flows that have ended, the likelihood of needing this ICAM is greatly decreased.

Layer 3 Processing and the Forwarding Information Base/Adjacency Tables

The Forwarding Information Base (FIB) table contains a hardware representation of the Layer 3 forwarding tables found in the control plane. The Catalyst 6500 uses Cisco Express Forwarding (CEF) architecture to enable hardware forwarding. The routing protocols collect next-hop information about the routing topology, and maintain this in their respective protocol tables in the control plane. This forwarding information is consolidated into the global routing table called the RIB (Routing Information Base). CEF then takes this global routing table (RIB) and creates a FIB table that is pushed down into the PFC4/DFC4 FIB TCAM. All Layer 3 forwarding in hardware uses this FIB table to perform Layer 3 forwarding lookups.

The FIB in the PFC4 contains 256 K entries, while the FIB in the PFC4XL contains 1 million entries. These are the same as their PFC3x forwarding engine counterparts. The FIB in the PFC4 contains prefix entries for IPv4 and IPv6 global address, IPv4 and IPv6 multicast addresses and MPLS label entries. There is a level of partitioning that exists to ensure there is always some space available for different types of forwarding entries. There is some flexibility from a user configuration standpoint that allows these partition boundaries to be changed to accommodate more of one type of forwarding entry. For example, in the PFC4XL, the default setting provides for 512 K IPv4 entries, and this can be increased through configuration control to support up to 1 M entries if required.

The PFC4 FIB is actually comprised of two memory blocks: one is TCAM-based and the other is RLDRAM-based. When performing a Layer 3 lookup, the FIB TCAM lookup is performed first. To execute the lookup, a FIB TCAM lookup key is derived, based on incoming packet type and other fields, to perform a FIB TCAM lookup. The result of the FIB lookup returns a pointer into the FIB RLDRAM, which will hold a pointer into the adjacency table for normal forwarded packets. If the adjacency pointer indicates the destination can be reached through multiple paths, it computes a unique adjacency pointer for each path.

With the earlier PFC3x on the Sup 720, the adjacency table also contained rewrite information for rewriting the Layer 2 header, when the packet left the switch. With the PFC4, this information has now been moved to the rewrite table (discussed later). The adjacency table contains a pointer to the LTL (Local Target Logic) index. The LTL index is an internal pointer that represents an interface in the switch, which essentially identifies the egress interface out of which the packet is going to be sent. The adjacency table also contains a flood index should the packet need to be sent out multiple interfaces.

The FIB is also now used to perform a RPF check for IP multicast packets. This is a new feature of PFC4 not found in earlier PFC3x forwarding engines.

Note that for OFE processing, the FIB is bypassed and does not participate in packet processing for that pipeline.

Policer Processing and the NetFlow Statistics Table

The policing logic implements the following functionality:

- Performs the IFE and OFE TTL check
- Performs an MTU check prior to egress policing
- Performs distributed and non-distributed three-color aggregate policing
- Performs two-color shared and non-shared NetFlow (microflow) policing
- Maintains NetFlow and aggregate policing statistics
- Supports both packet and byte based policing for aggregate and NetFlow policing

The NetFlow statistics table is maintained by the policing processing logic and consists of 1 M entries. It consists of three banks of memory, and is used to maintain a NetFlow statistics entry for every active flow in the system. A NetFlow statistics entry consists of multiple fields, including the packet timestamp, byte count for NetFlow policed packets, forwarded packet and byte count, last used timestamp, TCP flag, and more.

Rewrite Processing and the Rewrite Info Table

The rewrite process engine on the Layer 3 forwarding engine is designed to perform the tasks associated with generating next-hop rewrite instructions for outbound packets. The rewrite process will generate a new source or destination MAC address, and provide a rewrite pointer for the multicast expansion table (MET). Other important processing elements of the rewrite engine include:

- The rewrite information for source and destination MAC address
- The Multicast Expansion Table (MET) pointer that is used to specify all Outgoing Interfaces (or OIFs that a multicast packet needs to be replicated to
- Can initiate packet recirculation, for special processes that require it such as VPLS, GRE tunnelling, and IPv6 tunnelling operations
- MPLS operations for pushing, popping, and swapping labels
  - Push up to five labels
  - Pop one or two labels
  - Swap one label
  - Swap one label + push up to four labels
  - Swap two labels
- EoMPLS encapsulation for up to five labels
- EoMPLS de-encapsulation with one non-null label or two labels with top label as null label
- IPv4 Network Address Translation (NAT)
- Fast Re-Route (FRR) support
- Provides TOS, traffic class, and EXP rewrites
- Provides TTL rewrites

The rewrite info table contains 1 M entries, matching what the LIF table supports. It will generate a rewrite instruction set providing instructions for what the outgoing packet's Layer 2 frame addressing will use.

Packet Processing and the Adjacency Stats Table

The main objective of this component is to build and pass the result of the processing of all processing blocks above back to the Layer 2 forwarding engine. The packet processor also serves as a conduit between IFE and OFE processing, taking the results of IFE processing and building the input required for OFE processing.

The adjacency statistics table contains 512 K entries and maintains a packet and byte counter for original packets (and not copy packets). In other words, a packet that has SPAN processing applied to it would not be counted by the adjacency statistics.

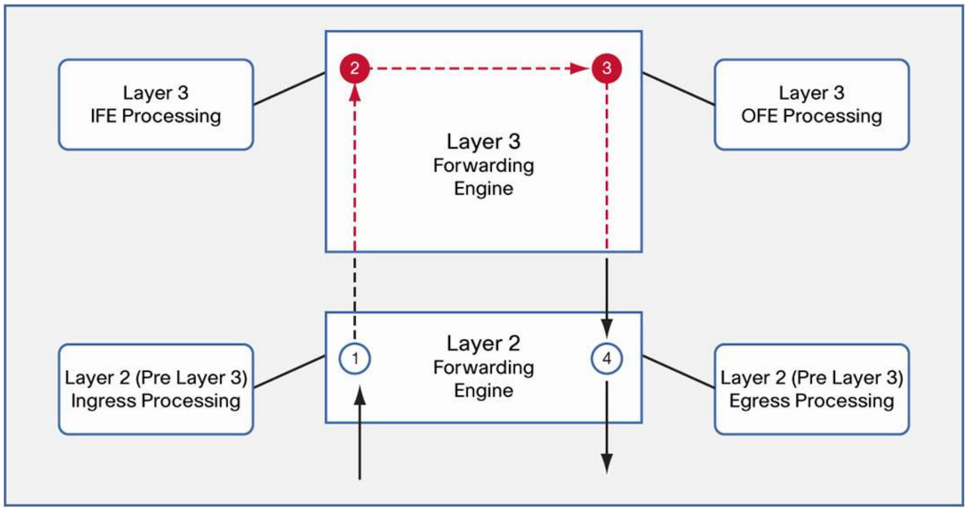
This processing block also maintains a table that holds accounting statistics (4 K entries).

PFC4 (and PFC4XL) Packet Walk

The PFC4 (and PFC4XL) is capable of processing up to 60 Mpps for both Layer 2 and Layer 3 processed packets. There are four stages to the packet walk. These stages are listed below and are shown in the following diagram:

1. Layer 2 (pre Layer 3) ingress processing
2. Layer 3 IFE processing
3. Layer 3 OFE processing
4. Layer 2 (post Layer 3) egress processing

Figure 13. PFC4 Layer 3 Forwarding Engine Packet Walk



This section will explore in detail the packet walk through the PFC4 for an IPv4 unicast packet.

Layer 2 (Pre Layer 3) Ingress Processing

The entire process for the packet walk begins with the arrival of a frame over the DBUS destined to the Layer 2 forwarding engine ASIC. The processing steps associated with this stage are as follows:

1. Packet arrives over the DBUS or the fabric replication ASIC.
2. A CRC check is performed by the Layer 2 forwarding engine.
3. A LIF database lookup is performed which returns the ingress LIF, bridge domain and lookup index for later Layer 3 forwarding engine lookups.

4. A Result Bundle Hash (RBH) is generated, indicating which link of an EtherChannel bundle is to be used (if applicable).
5. Perform static MAC address match for control packets (CDP, BPDU, and more).
6. Perform Layer 2 table MAC address lookup.
7. Merge result of above processing (L2 Lookup, RBH, and more) and create a frame which is forwarded to Layer 3 forwarding engine for processing.

Layer 3 IFE Processing

The IFE processing steps for the Layer 3 forwarding engine are detailed below:

1. The Layer 3 forwarding engine receives a frame from the Layer 2 forwarding engine, which it checks for CRC errors, then forwards to the IFE pipeline.
2. A LIF map table lookup is performed to collect information regarding the interface on which the packet arrived (for example, the ACL label for input classification, RPF mode, VPN number, and more).
3. An RPF check is performed on the Source address.
4. Packet classification is performed on the packet, with a lookup into ACL TCAMs.
5. TCAM lookup result is merged into an ACL result, a QoS result, an accounting result, and a CTS result.
6. Notify the Layer 2 engine to update ACL statistics (using the ACL hit counter, located on Layer 2 engine).
7. Ingress NetFlow lookup is performed, using results from Step 5.
   a. If it is a hit, NetFlow statistics are updated and Microflow policing is performed (if applicable).
   b. If no hit is found, a new NetFlow entry is created.
8. The Layer 3 FIB lookup is performed, which returns an adjacency pointer and a link count, which determines how many equal cost paths exist for this prefix.
9. Ingress aggregate and NetFlow policing is then performed.
10. All the ingress lookups are now complete and a result can be generated.

Layer 3 OFE Processing

Once the IFE pipeline processing is finished, the packet is then passed to the OFE pipeline for further processing. The OFE steps are detailed below.

1. The adjacency pointer retrieved by the FIB lookup during IFE processing is used to return an egress LIF index and a rewrite pointer. The rewrite pointer is held for later in the OFE process when rewrite information required for this packet needs to be retrieved from the rewrite table.
2. The egress LIF lookup is performed to retrieve information about the egress interface out of which the packet is to be sent.
3. An egress classification lookup is performed into the ACL TCAM for Security and QoS ACLs. If a redirect is found (such as PBR, VACL redirect, and more), then a new rewrite pointer is received for later lookup into the rewrite table.
4. An egress NetFlow lookup is performed.
5. Egress aggregate and microflow policing is applied (if applicable).
6. A lookup into the RIT (rewrite information table) is performed, using the rewrite pointer retrieved earlier.
7. A final result is generated containing the information retrieved from the OFE lookups.
8. The result frame is passed back to the Layer 2 engine containing information such as the destination VLAN, egress LIF and rewrite information.

Layer 2 (Post Layer 3) Egress Processing

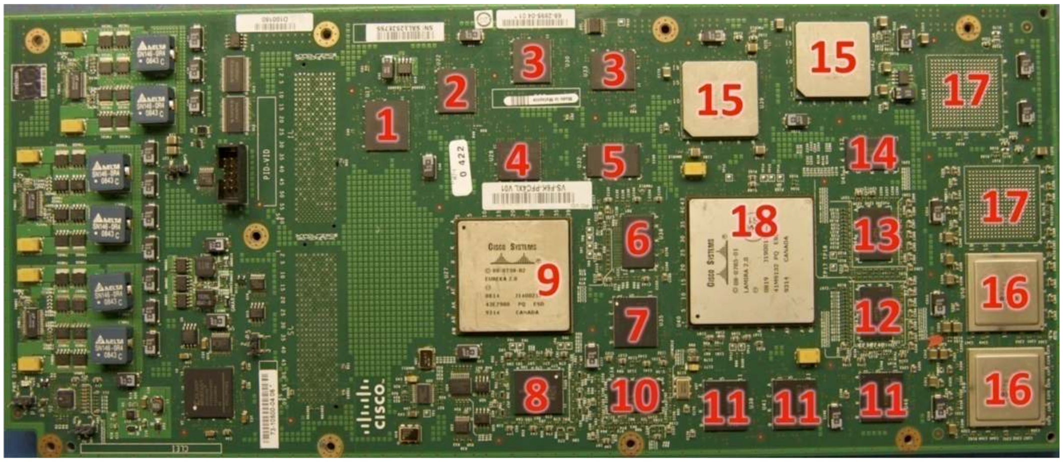
Subsequent to the Layer 3 engine finishing processing, it hands the result of its operation back to the Layer 2 engine for the final processing stage to commence. These final steps are detailed here:

1. Checks the Layer 3 engine result for CRC errors.
2. Merges its L2 result with the L3 result from Layer 3 engine.
3. Inspects results of Layer 3 engine processing for post L2 table lookups, such as LIF.
4. Performs CPU rate limiting function (if applicable).
5. Sends result to outgoing interface through BUS.
6. LIF statistics are updated to reflect the forwarding of this packet.

This completes the processing for a single packet passing through the PFC4 complex. Up to 60 million of these packet lookups can be processed in one second. The same processing sequence and performance metrics apply for the DFC4 complex, independent of the PFC4. This allows an aggregate lookup rate of 720 Mpps for a 6513-E system.

PFC4 Board Layout

The PFC4 board is shown below:



Each of the numbered components is listed below:

1. L2 forwarding engine - LIF (logical Interface) Table
2. L2 forwarding engine - Adjacency Statistics (1/2)
3. L3 forwarding engine - Classification Table
4. L2 forwarding engine - Adjacency Statistics (2/2)
5. L3 forwarding engine - RPF Map Table
6. L3 forwarding engine - LIF Map Table
7. L3 forwarding engine - Adjacency Table
8. L2 forwarding engine - LIF Statistics
9. **Layer 2 forwarding engine ASIC**
10. L3 forwarding engine - Rewrite Info Table
11. L3 forwarding engine - NetFlow Statistics
12. L3 forwarding engine - NetFlow Data Table
13. L3 forwarding engine - NetFlow Hash
14. L3 forwarding engine - FIB Table

15. L3 forwarding engine - Classification TCAM
16. L3 forwarding engine - FIB TCAM
17. L3 forwarding engine - FIB TCAM for XL PFC (+ 16)
18. **Layer 3 forwarding engine ASIC**

Switch Fabric and Fabric Connection Daughter Card

The following section provides more details about the new switch fabric design and Fabric Connection Daughter Card (FCDC) on the Supervisor 2T.

Supervisor Engine 2T Switch Fabric

LAN switches predominantly use either a shared memory switch fabric or a crossbar switch fabric for their switch backplane. The switch fabric implemented on the Supervisor 2T uses the crossbar architecture, which is the same backplane architecture used on the Supervisor 720. A crossbar architecture allows multiple simultaneous data transfers to occur at the same time between different linecards.

Each linecard slot in a chassis has its own set of dedicated channels over which to send data into the switch fabric. Twenty-six 40 Gbps fabric channels are provided by the Supervisor 2T Switch Fabric, which distributes two fabric channels to each linecard slot in any given chassis.

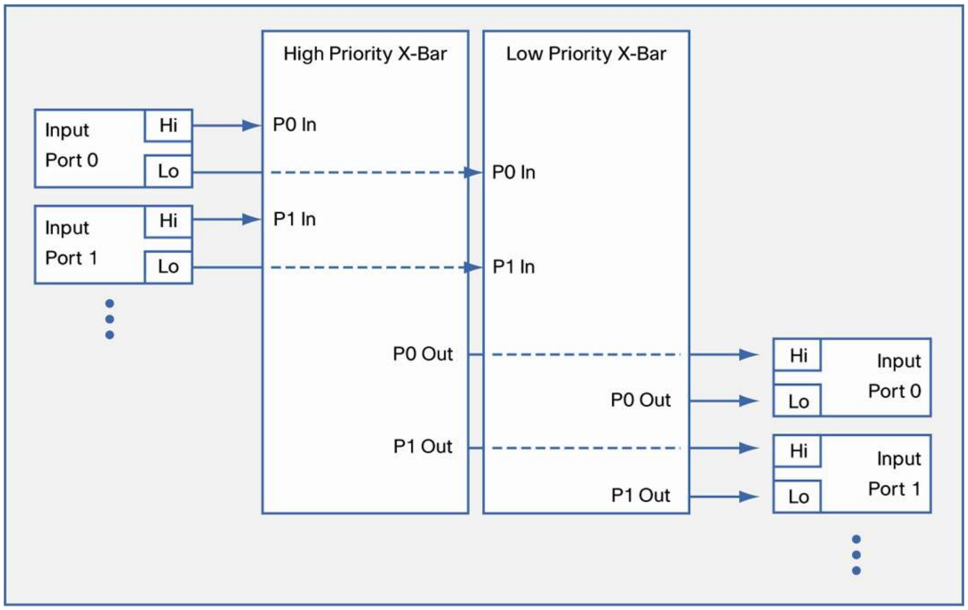
The fabric ASIC used in the Supervisor 2T represents a major upgrade from the fabric ASIC used in Supervisor 720. Some of the major enhancements include:

- Buffered crossbar design
- 26 fabric channels (compared with 20 fabric channels on the Sup 720-10G)
- Each channel can operate at either 40 Gbps to support the new WS-X69xx linecards, or 20 Gbps to provide backward compatibility for WS-X67xx and WS-X68xx linecards
- Enhanced multi-destination arbitration
- Support for two-priority level data path through two dedicated input queues and two output queues
- Packet counters per queue, so packet history is visible for debugging purposes
- Separate non-shared input and output queuing on a per-port-per-queue basis
- Multiple modes of backpressure supported
    - Per-queue flow control from input buffer to downlink linecard
    - Per-queue flow control from output buffer to input buffer (internal to fabric ASIC)
    - Per-queue flow control from linecard to output buffer
- Support for Jumbo frames up to 9248 bytes in size

Switch Fabric Architecture

The Supervisor 2T Switch Fabric implements two identical crossbar switch fabrics. Each crossbar is used to handle either high priority or low priority traffic. This is shown in the following diagram.

Figure 14. Supervisor 2T Crossbar Architecture



When a linecard forwards a packet over its fabric port, it is received by the fabric and stored in the fabric ports input buffer. The input buffer on the ingress fabric port contains a high priority queue and a low priority queue. When a packet arrives on the ingress fabric port, it is demultiplexed into one of the two queues, depending on the priority assigned to it. The assignment of the packet into one of the two queues provides the determination of which fabric the packet will traverse (the high priority or the low priority fabric).

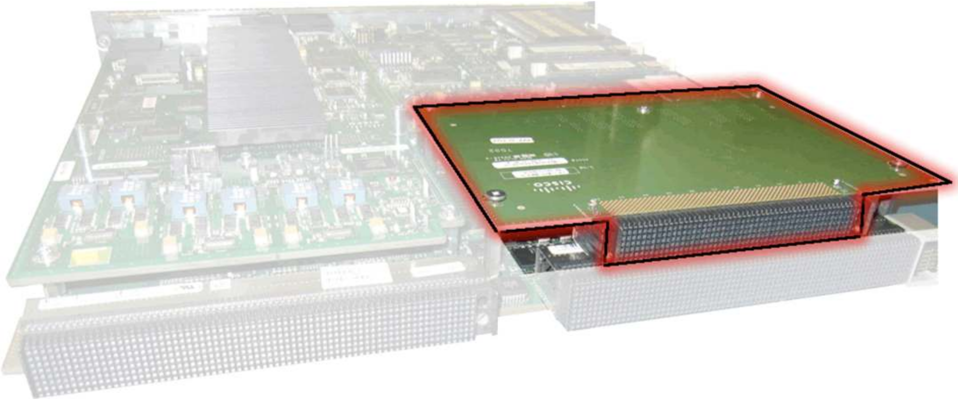
Within the packet header is a Fabric Port of Exit (FPOE) index field, which indicates the switch fabric port of exit. The switch fabric will arbitrate for output buffers associated with that egress fabric port. Once the arbitration is successful, the packet data will flow from the ingress packet buffer on the ingress fabric port to the egress packet buffers on the egress fabric port. During this transmission, a three-times-over-speed is used to minimize switching latency and reduce Head of Line Blocking (HOLB) contention. Once the entire data packet has been received by the egress fabric interface ASIC and is fully stored in the egress fabric port buffers, it is then transmitted by the egress linecard.

Fabric Connection Daughter Card

The Supervisor 720 supported only 20 fabric channels (18 for linecards, two for Supervisor uplinks). When was installed in the 6513, it limited the top eight slots to a single channel only. This limited linecard options for Slots 1 through 8, restricting dual fabric linecards (WS-X67xx) to being inserted in Slots 9 through 13 only. The crossbar switch fabric on the Supervisor 2T and the 6513-E chassis address this limitation.

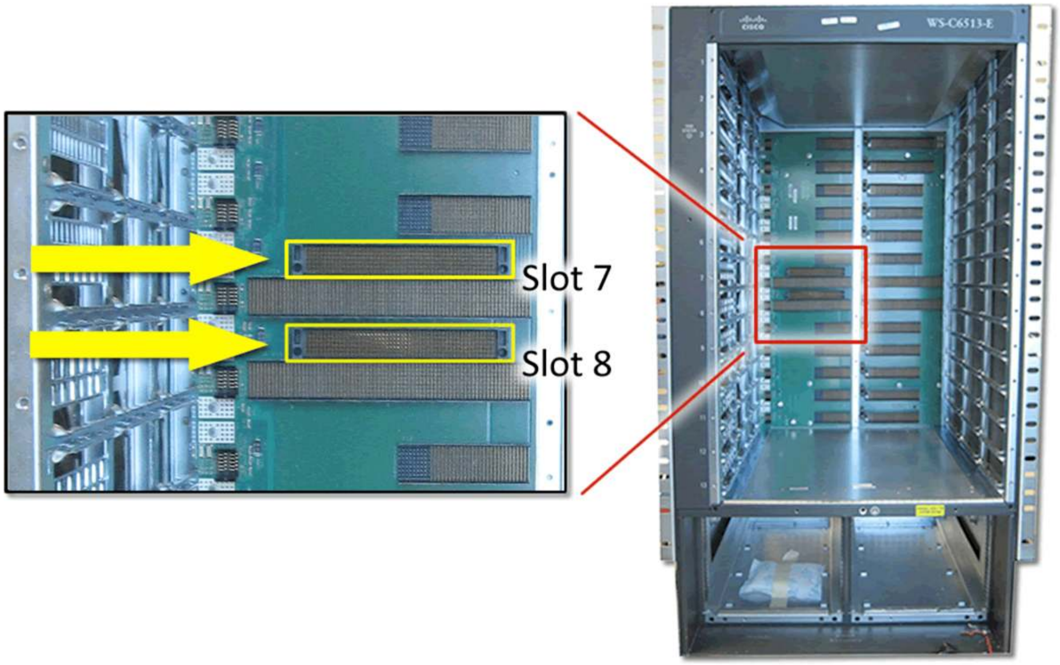
With this chassis and Supervisor combination, all linecard slots (Slots 1 through 6 and Slots 9 through 13) have dual fabric channels built into the chassis, allowing dual fabric linecards (WS-X67xx, WS-X68xx and WS-69xx) to operate in all slots. The fabric connection daughter card is a new addition on the Supervisor baseboard. This daughter card provides connection for 6 additional channels across Slots 1 through 6 that were missing in the original 6513. This can be seen in the following diagram.

Figure 15. Fabric Connection Daughter Card (FCDC)



The diagram above shows the FCDC from the rear of the Supervisor 2T. It provides a second connector that sits just above the original fabric connector found on earlier Supervisor 720 models. For all chassis models (except the 6513-E), this connector does not play a part in provisioning fabric channels to any of the linecard slots. In the 6513-E, there is an additional connector found for Slots 7 and 8 on the chassis backplane that is used by the FCDC. This additional backplane connector can be seen in the following diagram of the 6513-E.

Figure 16. Catalyst 6513-E Backplane Connectors



This additional slot connector allows the 6 additional fabric channels to be relayed to the first six linecard slots in the 6513-E, thus providing dual fabric channel connections to all linecard slots.

Glossary

The following section provides brief explanations for the major acronyms used throughout this document.

Table 16. Definition of Acronyms

Acronym	Description
ACE	Access Control Entry
ACL	Access Control List
ASIC	Application Specific Integrated Circuit
BD	Bridge Domain
CMP	Connectivity Management Processor (Port)
CTS	Cisco TrustSec
DFC	Distributed Forwarding Card
DSCP	Differentiated Services Code Point
EoMPLS	Ethernet over MPLS
FCDC	Fabric Connection Daughter Card
FNF	Flexible NetFlow
GRE	Generic Routing Encapsulation
LIF	Logical Interface
MSFC	Multi-layer Switch Feature Card
MPLS	Multi-Protocol Label Switching
PFC	Policy Feature Card
QoS	Quality of Service
TCAM	Tertiary Content Addressable Memory
VLAN	Virtual Local Area Network
VPLS/H-VPLS	Virtual Private LAN Services (or Hierarchical VPLS)

Americas Headquarters
Cisco Systems, Inc.
San Jose, CA

Asia Pacific Headquarters
Cisco Systems (USA) Pte. Ltd.
Singapore

Europe Headquarters
Cisco Systems International BV Amsterdam,
The Netherlands

Cisco has more than 200 offices worldwide. Addresses, phone numbers, and fax numbers are listed on the Cisco Website at **www.cisco.com/go/offices.**

Printed in USA C11-676346-04 01/13