

The Cisco Flow Processor: Cisco's Next Generation Network Processor

Executive Summary

Cisco® realized five years ago that the network edge was changing and that the existing network-processor design would not be sufficient to meet the needs of exploding video traffic growth and emerging collaborative applications, and the need for borderless networking where any service can be offered securely anywhere in the network. General-purpose processors enabled flexible network services and rapid feature development but lacked traffic prioritization mechanisms and suffered from slow forwarding performance. Over the last few years, faster network-processor technology is now a mainstream technology that is growing at more than 100 percent per year (Network Processors: A Heavy Reading Competitive Analysis – Jan 2005). Even in the past few years however, many of the multi-core network processors available offered high throughputs but limited service flexibility with fixed feature pipeline processing and inflexible classification and packet-processing programming languages. The Cisco Flow Processor takes network-processor technology to the next level by moving away from other rigid forms of processing technology to one of true massively parallel and flexible flow processing for classification, services integration, and traffic management.

A world-class development team at Cisco has invested more than 600 person-years to create a new line of advanced network processors, the Cisco Flow Processor. These network processors combine the best aspects of a high-performance forwarding engine with the service flexibility of the general-purpose processor. It is the industry's first fully integrated and programmable flow processor designed to unify massive parallel processing, integrated quality of service (QoS), and advanced memory management while at the same time offering integral service delivery and programmability. The four primary goals for the processor, ideal for both service provider and enterprise applications, follow:

Scale

The Cisco Flow Processor strikes the right balance between fast network processing and scalability for virtually any current network service requirement. The processor powers this advanced feature set with hundreds of customized network-processor resources, each capable of flexible flow processing for many applications into more than 100,000 queues anywhere in the network.

Performance

Cisco Flow Processors can provide a range (5 to 100+ Gbps) of packet-processing bandwidth inside the chip itself, meaning that the entire payload and headers of frames are available for packet processing -- accelerating Deep Packet Inspection and application-specific processing.

Feature Velocity and Versatility

The Cisco Flow Processor takes advantage of a massively parallel CPU architecture while incorporating a software architecture that allows fast delivery of new data-path features and services. The processor represents a new hardware and software architecture that can be used across many Cisco product lines. Cisco offers true consistency in forwarding architecture across product lines in the midrange and high-end routing space.

Multigenerational

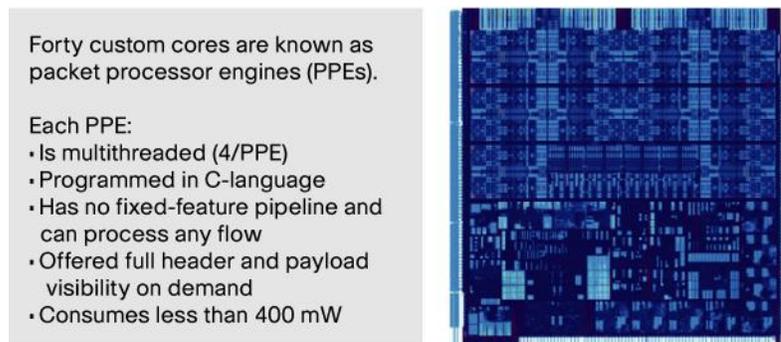
The unique software architecture of the Cisco Flow Processor will allow Cisco to evolve this network processor over time and use the same software across generations of hardware.

Introduction

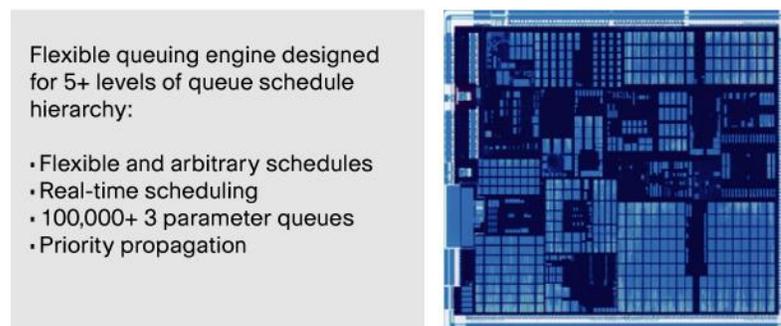
A multigenerational family of processors, the Cisco Flow Processor has been designed by Cisco as both a hardware and software architecture. The first generation resides on two pieces of silicon; later generations may be single-chip solutions that adhere to the same software architecture described herein. The term “Cisco Flow Processor” alone refers to the overall hardware and software architecture of the network processor. The first generation of the Cisco Flow Processor uses industry-leading 90-nm technology and converges three principal components:

- It consolidates 40 customized packet-processor cores (900 MHz to 1.2 GHz) into a single piece of silicon. This massive amount of parallel processing reduces any requirements for external service blades inside the router. All processing is performed on the chip (Figure 1)

Figure 1. Cisco Flow Processor Engine (First Generation)



- This dedicated flexible-queuing engine offers a dramatic offload of buffer, queue, and scheduling (BQS) processing to address very complex subscriber- and interface-level queuing requirements of both enterprise and carrier networks today (Figure 2)

Figure 2. Cisco Flow Processor Traffic Manager (First Generation)

- And finally, the Cisco Flow Processor uses a software architecture based on a full ANSI-C development environment implemented in a true parallel processing environment. Some traditional network processors rely upon difficult-to-implement microcode, making it difficult and time-consuming to add new capabilities. Other network processors offer higher-level language development but into a feature pipelined architecture. With the Cisco Flow Processor, new features can be added quickly as customer requirements evolve by taking advantage of industry-standard tools and languages built upon a powerful parallel processing architecture. This architecture represents a paradigm shift and evolution in the software architectures associated with network processing today

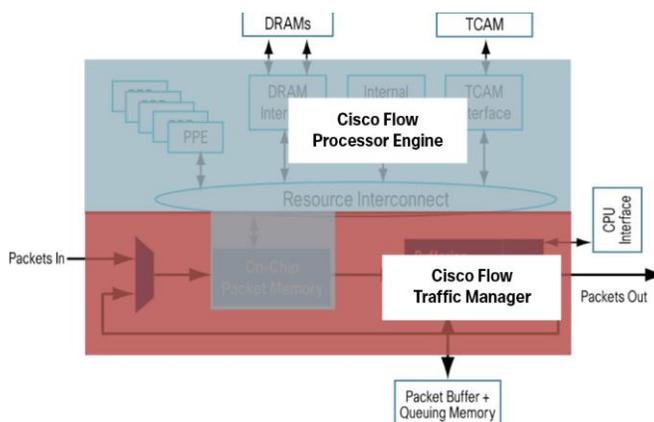
After extensive analysis of current network-processor technology and the Cisco IOS[®] Software, Cisco has produced a flexible processor that can address numerous key requirements in terms of combining forwarding with services:

Solution Requirement	Cisco Flow Processor Innovation
Carrier aggregation points in the network must be able to do more than simply forward Multiprotocol Label Switching (MPLS) or IP packets; they must be able to intelligently implement any number of services related to voice and video, security, and subscriber management	With a hardware architecture that guarantees timely and ordered access to internal databases, the Cisco Flow Processor can deliver scalable multiprocessor software for a wide range of networking applications and services.
Multiple enterprise segments and industries need a processor that can offer very-high-bandwidth, low-latency processing while at the same time implementing any number of segment-specific feature sets, whether they are government-, financial, insurance-, or health-related.	In the Cisco Flow Processor, the queuing work is off-loaded from the PPEs to a dedicated traffic manager. In addition, the PPEs are supplemented with dedicated hardware assists for normal packet processing such as lookups, policing, and classification. The Cisco Flow Processor helps ensure that critical flows are forwarded with lowest latency and no resource bottlenecks.
Networks today need flexibility and the capability of timely insertion of new high-speed network services and easily deployable follow-on features for applications in areas such as voice and video, address and subscriber management, routing, and security	The Cisco Flow Processor is not hindered by a fixed feature pipeline that reduced the application flexibility on previous generations of network processors. Each multithreaded PPE uses a full multilevel instruction caching architecture, and is fed enough instructions and fast access to various service databases to keep overall processing usage optimal across hundreds of thousands of flows.

The Cisco Flow Processor not only represents a significant evolutionary innovation in network processing, but also offers a delivery vehicle for business and network solutions for years to come.

Cisco Flow Processor Architecture

The Cisco Flow Processor is not just a chip or hardware solution; it offers a next-generation hardware and software architecture that will be used in future Cisco products (Figure 3).

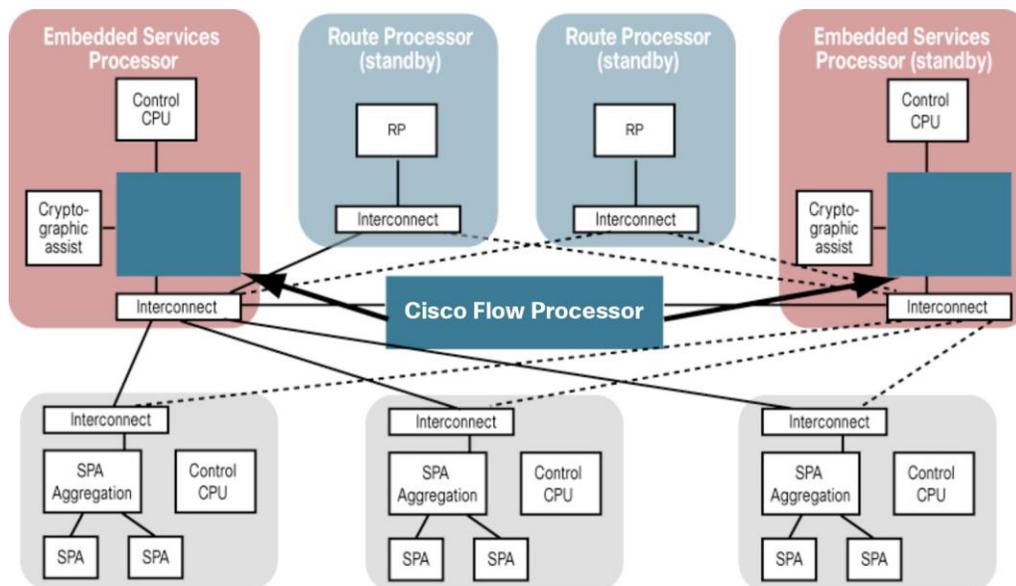
Figure 3. Cisco Flow Processor Major Block Diagram

The Cisco Flow Processor architecture is non-pipelined, parallel processing with centralized shared memory. The processor engine is responsible for all processing of all flows in terms of running through forwarding-path software, and the traffic manager function is responsible for queuing and scheduling functions for both the system and I/O interfaces. The traffic manager can forward packets to other parts of the system, I/O interfaces, and back to the packet processor itself through a recycle path. The Cisco Flow Processor has numerous large-bandwidth external interfaces (not all shown in the figure):

- This family of network processors supports a range of 5 to more than 100 Gbps of bidirectional external (to the chip) bandwidth across multiple interfaces. These interfaces are used for both internal and external traffic destined for actual shared-port-adapter (SPA) interfaces, fabric interfaces on distributed systems, as well as traffic to system-specific resources such a route processor for punted traffic
- A separate high-speed interface is used to access the control processors residing on the host system that integrates the Cisco Flow Processor. This high-speed interface is used to program the processor and to distribute tables and databases both to and from the processor
- Numerous high-speed interfaces to external reduced-latency memory, ternary content addressable memory [TCAM], and static RAM (SRAM) are used for storing large software libraries, packet buffers, and queuing state

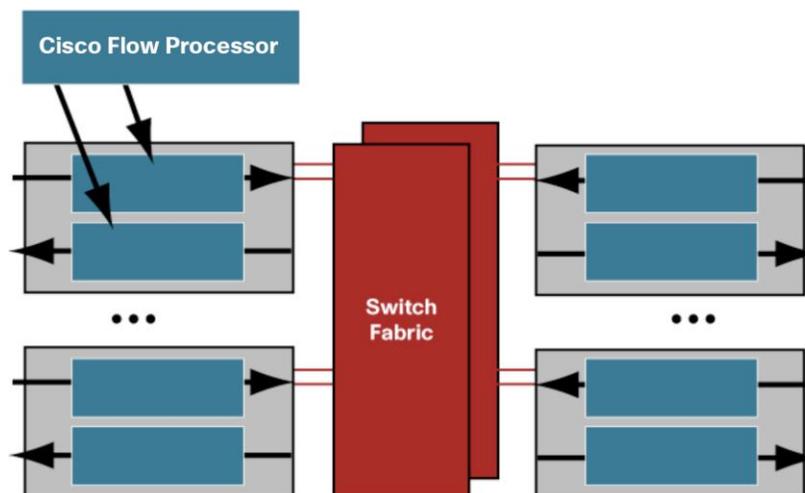
An important aspect of the architecture of the Cisco Flow Processor is that it can be used across many different types of systems. It will be embedded in the new Cisco ASR 1000 Series Aggregation Services Routers. On the Cisco ASR 1000 platform, this chipset is so powerful it constitutes the entire forwarding engine of the centralized system (shown in Figure 4 with a second processor for *active/standby redundancy*).

Figure 4. Cisco Flow Processor Residing in a Cisco ASR 1000 Series Router (6-rack-unit [6RU] centralized forwarding plane shown)



On distributed systems such as that shown in Figure 5, the Cisco Flow Processor will be used on future line cards and constitute part or all of the processing of the ingress or egress data path of a line card:

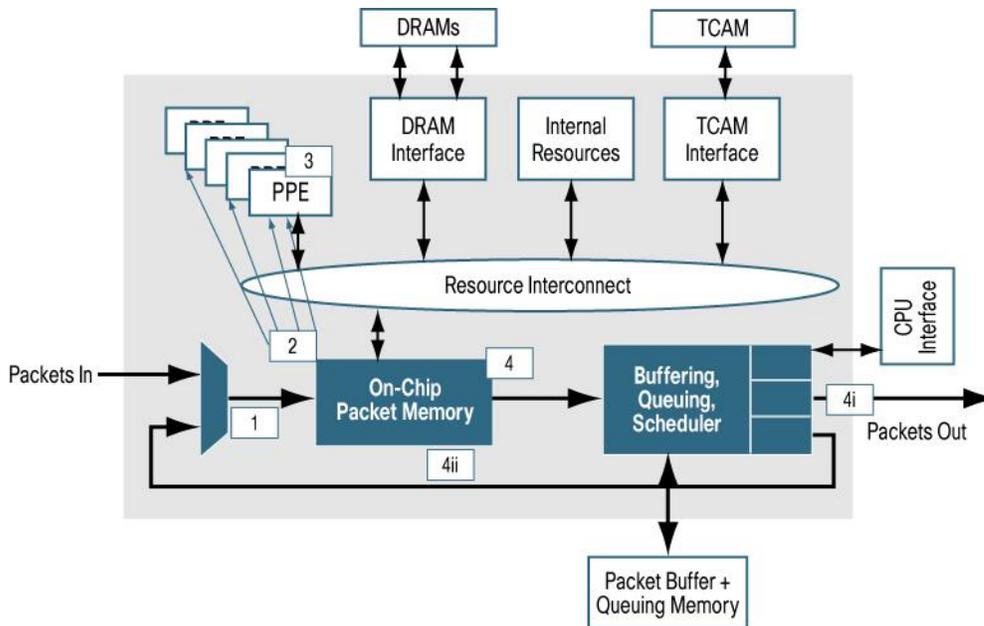
Figure 5. Cisco Flow Processor Residing on Distributed Line Cards



High-Level Packet Flow

Figure 6 shows the basic architecture of the system.

Figure 6. Cisco Flow Processor Basic Architecture



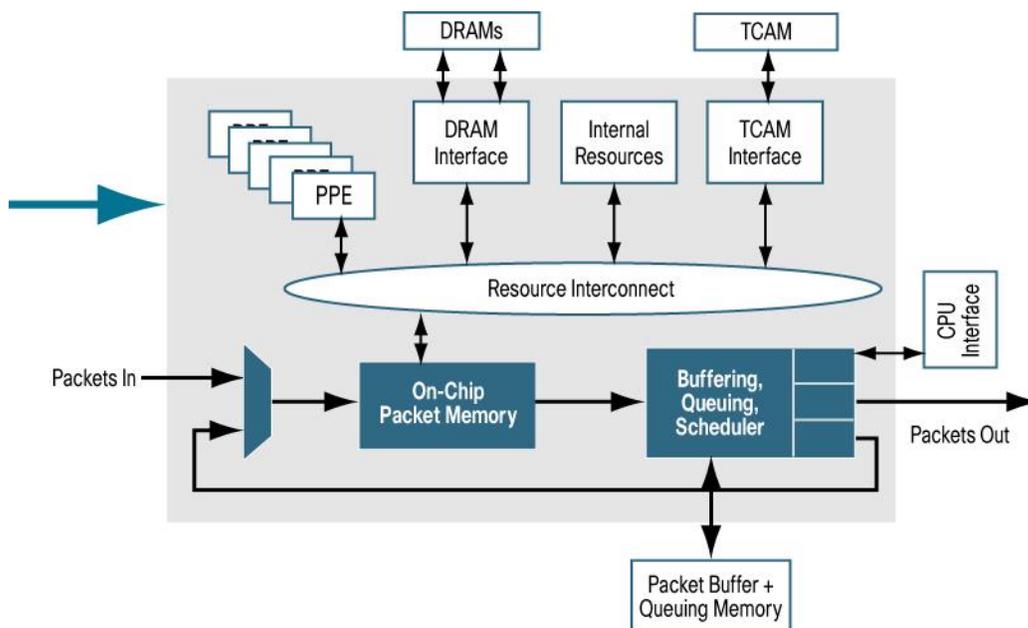
1. The entire Layer 2 frame arrives on the Cisco Flow Processor and is received into on-chip packet memory on the packet processor. The processor can at its discretion process and analyze the entire packet, not just the headers. It implements, for example, Network Based Application Recognition (NBAR) and Flexible Packet Matching (FPM) without any other external processors necessary.
2. The Cisco Flow Processor then distributes the packet context to a free PPE thread. The processor has 160 such threads available for processing at any given time; later implementations will expand this number. The PPEs themselves are described in more detail in the next section.
3. The PPE thread runs through a feature chain in software, which processes the packet. While processing the packet, the PPE thread uses various hardware assists (described later) on the packet-processor chip.
4. When processed, the PPE thread releases the packet to the traffic manager and its own packet buffer for placement into an output queue for scheduling and transmission. At this point the traffic manager can do one of the following:
 - a. Schedule the packet for transmission out a physical interface adhering to very flexible, arbitrary, and hierarchical queuing configurations
 - b. Schedule the packet for transmission back to another PPE thread (for example, ingress shaping or multicast processing) or even off-packet processor resources such as cryptographic devices or route processors in cases where the system itself needs to receive traffic

The last PPE thread to process the packet determines whether or not a packet needs multipass processing. State is carried between passes such that subsequent PPE processing does not repeat already-completed packet processing.

Packet Processor Engines

Each PPE is a 32-bit Reduced Instruction Set Computer (RISC) core. Each PPE can process 4 threads, which corresponds to the ability of processing 4 packets independently and in parallel. In other words, the Cisco Flow Processor implements a potential well of 40 x 4, or 160 separate packet-processing resources. The software model used in packet processing allows for each PPE thread to process a packet independently and does not enforce a tight coupling between processing packets and usage of resources inside the chip. As a thread waits for various hardware assists (described in the next section), the PPE can work on the other three and help ensure that PPE processor usage always stays optimal and latencies for packet processing remain low (Figure 7).

Figure 7. Packet Processor Engines



When a packet arrives into on-chip memory, a distributor function assigns one of the threads of the PPE to a particular context (or packet). The assigned PPE is responsible for the packet for its entire life in on-chip memory before it is sent to the traffic manager for scheduling. Each PPE thread runs through embedded software written in ANSI-C. The actual instructions performed and the speed and latency at which packets are processed by each PPE depend on many factors:

- The number of features enabled for this particular packet (prefix lookups, Weighted Random Early Detection [WRED], Traffic Policing, Cisco IOS Firewall and Network Address Translation [NAT] processing, broadband access [BBA], etc.) is one factor
- Another factor is the reliance on internal and external resources (described in the next section)
- PPE usage of multipass processing is a third factor. In general each PPE tries to completely process the packet in one pass through the Cisco Flow Processor, but in cases in which optimizing packet latency is a goal it has to recycle packets through the processor a second or third time:
 - Ingress shaping

- Multicast (The Cisco Flow Processor can hold state between passes; it is extremely efficient with fan-out style multicast processing because it can process packets in stages.)

The Cisco Flow Processor also helps ensure that priority treatment of packets exists such that high-priority traffic being recycled has priority over newly arrived lower-priority traffic in terms of accessing PPE resources. Likewise, recycled lower-priority traffic does not have priority over newly arrived high-priority traffic.

In summary, the following is true of each PPE:

- It performs run to completion only the instructions necessary for complete packet processing of a particular packet. Each PPE thread can process packets for any flows that are found in large per service databases available to each PPE. The PPE thread is not dedicated to specific flow processing or specific functions and can run through any software implemented by the Cisco Flow Processor
- It is not responsible for queuing or scheduling, but can use hardware assists to perform operations such as advanced classifications, NBAR, FPM, Traffic Policing, and WRED
- During the life of a flow, the Cisco Flow Processor can maintain state of each flow such that as packets arrive any PPE thread can process regardless of whether the flow is new or existing
- The Cisco Flow Processor can collect flow information for applications and services such as video distribution, voice terminations, NetFlow data collection, and security flows associated with Cisco IOS Firewall and NAT. It can collect detailed data concerning each flow:
 - Duration, packet, and byte counts
 - Packet fragments and copies for applications such as Encapsulated (generic routing encapsulation [GRE]) Remote Switched Port Analyzer (ERSPAN) and Flexible NetFlow and Lawful Intercepts
 - Detailed session setup and teardown logs for Cisco IOS Firewall and NAT: The Cisco Flow Processor does not have to punt session creation or teardown of packets, reducing CPU usage of control processors in the routing platform itself
 - Call quality for session-border-controller (SBC) applications
 - Flexible flow monitoring for NetFlow collection: The Cisco Flow Processor can export huge amounts of flow cache data directly off the processor, further reducing CPU usage of control processors in the routing platform. Using Sampled NetFlow and various well known aggregation schemes results in optimized use of this cache
 - Auditing: The Cisco Flow Processor can be audited for information concerning all active flows and also signal the control plane when some policy function needs to be made aware of the end of life of a flow

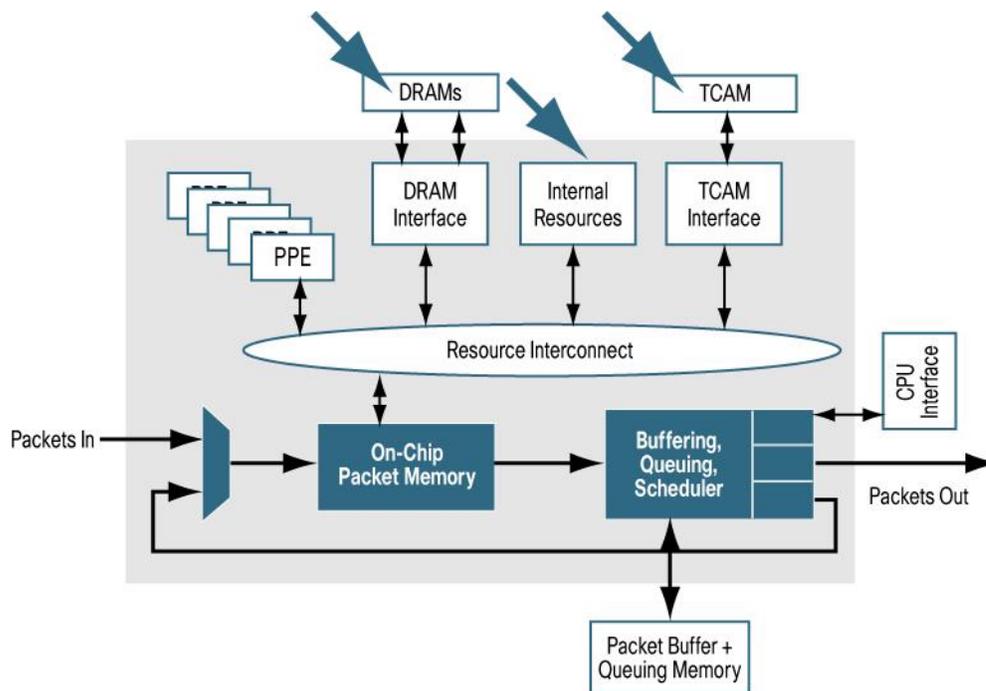
Each PPE on the Cisco Flow Processor integrated into the Cisco ASR 1000 platform can execute 1200 millions of instructions per second (MIPS), translating into *single-pass packet latency in as little as 10's of microseconds* with features enabled.

Packet Processor Engine Resources

Each PPE has access to an array of hardware-assist functions on and off chip:

- The Layer 1 cache (16 KB) is shared across threads from the same PPE, and the Layer 2 cache (dual 256 KB) is shared across threads from all PPEs. Layer 1 and Layer 2 instruction caches provide extremely fast access to instructions for each of the PPEs inside the Cisco Flow Processor. In general, if PPEs are processing packets that all require the same forwarding-path features, it is highly probable that each PPE thread can use the fastest memory and process packets so that the PPEs are used to best advantage
- Each thread has access to its own Layer 1 data cache and accesses a large external reduced-latency memory, which is memory mapped directly into the address space of each PPE thread
- Further, each PPE can access hardware feature acceleration of network address and prefix lookups, hash lookups, WRED, Traffic Policers, range lookups, and TCAM for advanced classification and access-control-list (ACL) acceleration as it processes packets (Figure 8)

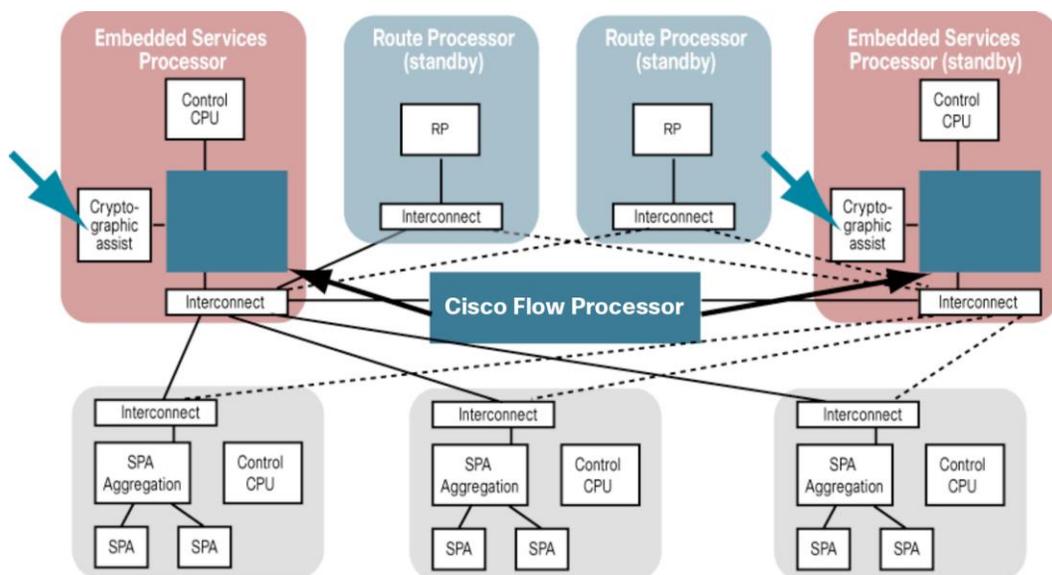
Figure 8. Hardware Assists



- A key internal resource for each PPE is the hardware lock manager. Individual threads and PPEs do not communicate directly with each other on the Cisco Flow Processor, helping ensure that flows that need less processing are not held up waiting for other flows that may take more time to process. A flow in the Cisco Flow Processor is defined more by what processing and advanced classification is required rather than arbitrary hashes of network addresses and ports. In situations where some cooperation in terms of when a packet leaves on-chip memory for egress scheduling is required, the lock manager assures the proper packet ordering for flows (at the control of software)

- Another key resource for the Cisco Flow Processor on the Cisco ASR 1000 is the off-chip cryptographic engine. Cryptographic functions are accessed from the PPE by sending the packet to the traffic manager, where it is buffered, queued, and scheduled to the cryptographic engine. When the cryptographic operation is completed, the packet returns to a PPE for additional packet processing. This example describes multipass processing on the Cisco Flow Processor (Figure 9)

Figure 9. Cisco ASR 1000 Cryptographic Assists

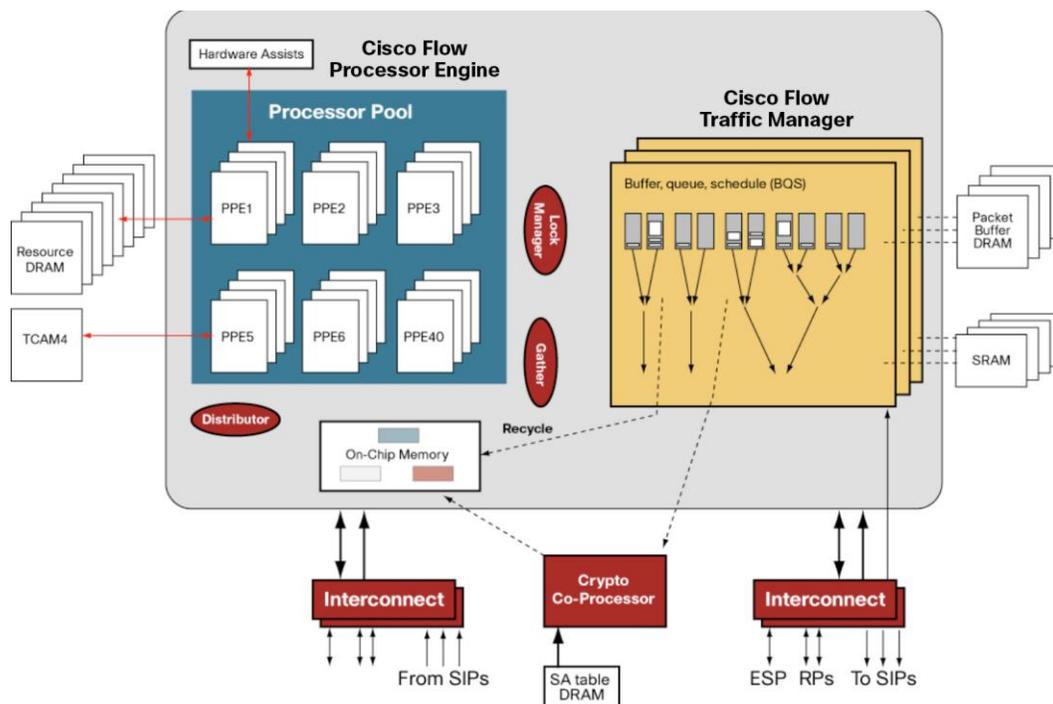


The net result of having hardware resources available to each PPE is that in general they do not have to perform computationally heavy lifting, as would be the case with long ACLs and cryptographic functions, for example. And while the hardware resources are being used, the use of each PPE stays at an optimal level because the PPEs can continue to process packets of flows in other threads. The overall goal of providing an efficient processing path through the Cisco Flow Processor is maintained. For either centralized forwarding engines or distributed processing, the processor complex can provide for rich feature processing and at the same time make sure that all system resources are efficiently used and bottlenecks are avoided.

Cisco Flow Processor Buffer, Queue, and Scheduling: The Traffic Manager

The application of the Cisco Flow Processor integrated in the Cisco ASR 1000 Series Embedded Services Processor (ESP) is shown in Figure 10.

Figure 10. Cisco Flow Processor Used on Cisco ASR 1000 Embedded Services Processor



When packet processing is completed, the PPE thread releases the packet to the traffic manager. It is here that any operations related to actual queue scheduling occur. The traffic manager (shown in the figure as the BQS block) implements an advanced real-time flexible queuing hardware:

- One hundred twenty-eight thousand queues and the ability to set three parameters for each queue:
 - Maximum bandwidth (“shape” rate in Cisco IOS Software Modular QoS CLI (MQC))
 - Minimum bandwidth (“bandwidth” rate in Cisco IOS Software MQC)
 - Excess bandwidth (“bandwidth-remaining ratio | percent” in MQC)
 - In addition, two Priority Queues can be enabled for each QoS policy applied.
- The number and makeup of layers inside the QoS policy are flexible. The possible hierarchies are not tied to any existing hierarchies used in networks today. The traffic manager can implement almost any hierarchy under the control of software
- The traffic manager can schedule multiple levels (no limit) of hierarchy in one pass through the chip. Further, queuing operations can be split across multiple passes through the chipset (for example, ingress shaping or queuing followed by egress queuing). With the traffic manager found in the Cisco Flow Processor, separate QoS policies can be applied to physical interfaces and logical interfaces simultaneously and then linked together into the same queuing hierarchy, allowing for some of the most advanced queue hierarchies available in the industry today, all at the control of software

- The priority of a packet is preserved and propagated down through the hierarchy such that, for example, a packet has suitable priority treatment at the VLAN level, at the physical Gigabit Ethernet level, and on the Cisco ASR 1000; even the bandwidth of the SPA interface processors (SIPs) is accounted for by the schedulers residing on the traffic manager
- External resources such as cryptographic devices and route processors are also scheduled appropriately on the Cisco Flow Processor such that they are never overwhelmed. Again, priority treatment is applied such that network control packets are punted accordingly and high-priority encrypted traffic is encrypted before lower-priority traffic that may also require cryptographic processing

The key advantage of real-time scheduling is that the schedulers on the Cisco Flow Processor are constantly monitoring the status of shallow egress buffers found on SIPs and SPAs. The traffic manager can monitor millions of events (egress queue status information from transmit buffers found on SIPs and SPAs) per second across 64,000 ports or virtual circuits, making it one of the most accurate scheduling engines found in the industry today. The Cisco Flow Processor is expected to deliver accuracies and granularity in the range of 0.1 to 1 percent relative to the interfaces in question. Egress interfaces are used at maximum rates but are never oversubscribed such that other parts of the platform or router (for example, SIPs on the Cisco ASR 1000) would have to drop outbound traffic.

Summary

In a world where all services and applications are converging onto the same network, it is quintessential that the elements in the network actually delivering the services be able to keep up with ever-increasing demands. The Cisco Flow Processor will help enable next-generation networks of the future with the innovations listed in Table 1.

Table 1. Cisco Flow Processor Innovations Compared with Other Vendors

Cisco Flow Processor Innovation	Cisco Flow Processor Found in all ASR 1000 series routers	Vendor X	Vendor Y
Hardware application capability	Forwarding, traffic management, and services	Forwarding and traffic management only	Forwarding and traffic management only
Architecture	Non-pipelined, parallel processing with centralized shared memory, full visibility of entire layer 2 frame including payload	Fixed hardware feature pipeline; limited visibility	Fixed hardware feature pipeline; limited visibility
Programmability for feature velocity	Fast feature development with ANSI-C software development	Slow	Slow
Voice features	SBC, terminating, and interconnecting media terminations with full accounting and flow control	None	None
Video features	Multicast Visual Quality Experience (VQE) and Video Call Admission Control (CAC)	Sub-line-rate IP Multicast	Sub-line-rate IP Multicast
Combined data features: IPv4 and IPv6 Multipath Forwarding, IP Multicast, ACL, Traffic Policing and QoS, NetFlow, and Unicast Reverse Path Forwarding (URPF)	Low-latency packet processing with large set of combined baseline features enabled First generation of Cisco Flow Processor is capable of 10 Gbps and 16 Mpps with all data features combined	2 to 3 Mpps	650,000 pps (tested triple-play [data, voice, and video] application)
Security	Deep Packet Inspection, Cisco IOS Zone-Based Firewall, intrusion detection and prevention, and control-plane protection	None	None

Traffic management	128,000 queues in any arbitrary hierarchy Flexible three-parameter (maximum, minimum, and excess) hierarchies Dual Low Latency Queuing (LLQ) for each QoS policy	512 queues	16,000 queues
---------------------------	--	------------	---------------

Moreover, this processor is just the beginning. The Cisco Flow Processor represents a multigenerational program within Cisco that will allow portability of software from generation to generation. The Cisco Flow Processor on the Cisco ASR 1000 and future Cisco platforms provides a vehicle for delivering integration and service models that have not even been designed yet. Carrier and enterprise networks with the Cisco Flow Processor are well-positioned for the future because flexible upgrades are possible without compromising quality or performance.



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.

CCDE, CCVP, Cisco Eos, Cisco StadiumVision, the Cisco logo, DCE, and Welcome to the Human Network are trademarks; Changing the Way We Work, Live, Play, and Learn is a service mark; and Access Registrar, Aironet, AsyncOS, Bringing the Meeting To You, Catalyst, CCDA, CCDP, CCIE, CCIIP, CCNA, CCNP, CCSP, Cisco, the Cisco Certified Internetwork Expert logo, Cisco IOS, Cisco Press, Cisco Systems, Cisco Systems Capital, the Cisco Systems logo, Cisco Unity, Collaboration Without Limitation, Enterprise/Solver, EtherChannel, EtherFast, EtherSwitch, Event Center, Fast Step, Follow Me Browsing, FormShare, GigaDrive, HomeLink, Internet Quotient, IOS, iPhone, IP/TV, IQ Expertise, the IQ logo, IQ Net Readiness Scorecard, iQuick Study, IronPort, the IronPort logo, LightStream, Linksys, MediaTone, MeetingPlace, MGX, Networkers, Networking Academy, Network Registrar, PCNow, PIX, PowerPanels, ProConnect, ScriptShare, SenderBase, SMARTnet, Spectrum Expert, StackWise, The Fastest Way to Increase Your Internet Quotient, TransPath, WebEx, and the WebEx logo are registered trademarks of Cisco Systems, Inc. and/or its affiliates in the United States and certain other countries.

All other trademarks mentioned in this document or Website are the property of their respective owners. The use of the word partner does not imply a partnership relationship between Cisco and any other company. (0801R)