CISCO SYSTEMS

White Paper

Parallel Express Forwarding on the Cisco 10000 Series

Parallel Express Forwarding (PXF) is a powerful, adaptive network-processing technology that maximizes both forwarding performance and services to build more powerful, scalable networks. PXF on the Cisco® 10000 Series Routers helps enable multiple millions of packets per second (mpps) forwarding rates while allowing customers to continually add features to their solution using existing network hardware.

This paper describes the architecture and benefits of the Cisco Parallel Express Forwarding technology

Summary

PXF is implemented in a programmable application-specific integrated circuit (ASIC) that manages forwarding decisions and per-subscriber packet processing including access control lists (ACLs), quality of service (QoS), flow accounting, and traffic shaping. The PXF ASIC allows the addition of new features without an incremental penalty in packet performance. Each packet is processed by the PXF engine in a deterministic fashion, unlike CPU forwarding platforms where the number of CPU cycles required depends largely on the features applied.

PXF makes use of the expedited IP lookup and forwarding algorithms introduced with Cisco Express Forwarding, while offering expanded functionality and accelerated performance through the implementation of a parallel architecture. The PXF forwarding engine applies the combination of parallel processing and pipelining techniques to the Cisco Express Forwarding algorithms to efficiently manage a variety of complex services and operations.

The Cisco 10000 Series Performance Routing Engine 2 (PRE2) uses a Versatile Time Management Scheduler (VTMS) scheduling algorithm to provide a single level of queuing. The Cisco 10000 Series PRE3 uses the latest PXF architecture for improved forwarding and feature scalability. The PRE3 implements the Hierarchal Queuing Framework (HQF) with up to three levels of hierarchy to deliver the most demanding triple-play services.

Challenge

Service providers require scalable networks to profitably meet their customers' requirements for better performance, more services, and higher reliability. Edge routers, such as the Cisco 10000 Series, will be needed to manage higher-bandwidth requirements, more subscribers, and multiple service levels ranging from best-effort consumer Internet data services to high-priority business applications, voice, and video.

Historically, routers have been optimized for feature flexibility at the edge and fast packet forwarding at the core of networks. Edge router packet forwarding architectures based on general purpose CPU components have offered the highest level of flexibility to quickly program new features, but often with the penalty of degrading performance and scalability. Core routing designs have been optimized for fast and consistent performance with a fixed feature set by using ASIC components. To address the needs of growing service providers and their customers, a solution is needed that blends the best of programmable general purpose CPU designs and high speed ASIC technology.

Solution

Designed to meet new requirements from service providers for high-capacity aggregation with sophisticated IP services, the Cisco 10000 Series uses the Cisco patented Parallel Express Forwarding (PXF) technology. PXF is a parallel multiprocessor architecture that helps enable deployment of multiple IP services while maintaining peak performance throughput.

PXF Performance Benefits

PXF performance is achieved through an architecture that uses the following components and methods:

- *QoS* – PXF processes packets and applies QoS at line rate. PXF performs classification, queuing, shaping, and policing on a per-subscriber basis. Up to three levels of QoS are available (starting with the Cisco 10000 Series PRE3).

- *Separation of control plane and forwarding plane functions* – This separation helps achieve high availability.

- *Use of parallel processors* – PXF takes advantage of a parallel array of processors for an accelerated switching path. Processor-intensive tasks such as policy routing, QoS, and statistics collection are segmented and distributed to columns of multiple processors.

- *Distributed memory access* – The PXF architecture allocates independent memory to each processor, as well as assigning memory to each column of processors. Use of multiple memory banks within a per-column array provides memory access optimization.

- *Modularization of functionality* – PXF distributes data structures across multiprocessor arrays. Distributed, task-specific memory resources allow parallel processing of complex tasks. Unlike uniprocessor systems where changes to code may have unforeseen results, including a negative impact on performance in other areas, the PXF architecture allows independent processing of tasks. Multiple tasks can execute efficiently without degrading performance.

- *Virtual interface indexing* – PXF switching mechanisms employ virtual interface indexing for the communication of connection and session information from one processor to another. This creates a streamlined pipeline through the processor array. The initial index is calculated based on the ingress of a frame into the platform. These indices subsequently direct the switching of all packet contexts through the forwarding path. Significantly, while the route processor may inject packets into the forwarding path that it has created, it is not necessary for it to manage the initial packets of a stream or flow, as in a route-caching model. As with Cisco Express Forwarding, PXF switching is topology-based. However, PXF is an evolution of the Cisco Express Forwarding model; it implements switching at a more fundamental level of the packet-management process, starting with the classification of the MAC header information, to further streamline forwarding.

Scalability and Versatility Benefits

The multiprocessing architecture of PXF makes it inherently more scalable than single processor-based forwarding schemes. New features can be implemented without degrading performance of existing features.

In addition, PXF technology allows platforms to scale to manage tens of thousands of interfaces. Routing and control functions are separated on the PRE to maintain performance as the network grows.

PXF also greatly simplifies the addition of new features. Unlike strictly hardware- or ASIC-based switching models, PXF minimizes the cost and time to market associated with feature improvements or enhancements.

Packet Forwarding with PXF

PXF enhances the forwarding information base (FIB) model by separating control-plane functions from forwarding-plane functions. In the PXF architecture, the route processor manages control-plane functions, including routing protocols, chassis and network management, system configuration, error management, and packets addressed to the router. Forwarding-plane functions are handled by PXF technology.

Each PXF network processor provides a packet-processing pipeline consisting of 16 microcoded processors arranged as multiple pipelines. Each of the coprocessors in a PXF network processor is an independent, high-performance processor, customized for packet processing. Each processor, called an Express Micro Controller (XMC), provides a sophisticated dual-instruction-issue execution unit, with a variety of special instructions designed to execute packet-processing tasks efficiently.

Within a single PXF network processor, the 16 XMCs are linked together in eight parallel pipelines. Each pipeline comprises four microcontrollers arranged as a systolic array, where each processor can efficiently pass its results to its neighboring downstream processor. Eight parallel pipelines are used, further increasing throughput.

Four PXF network processor ASICs are used in each PRE, yielding eight parallel-processing pipelines, each containing eight processors in a row (Figure 1).

Figure 1
PRE Forwarding-Path Processor Array
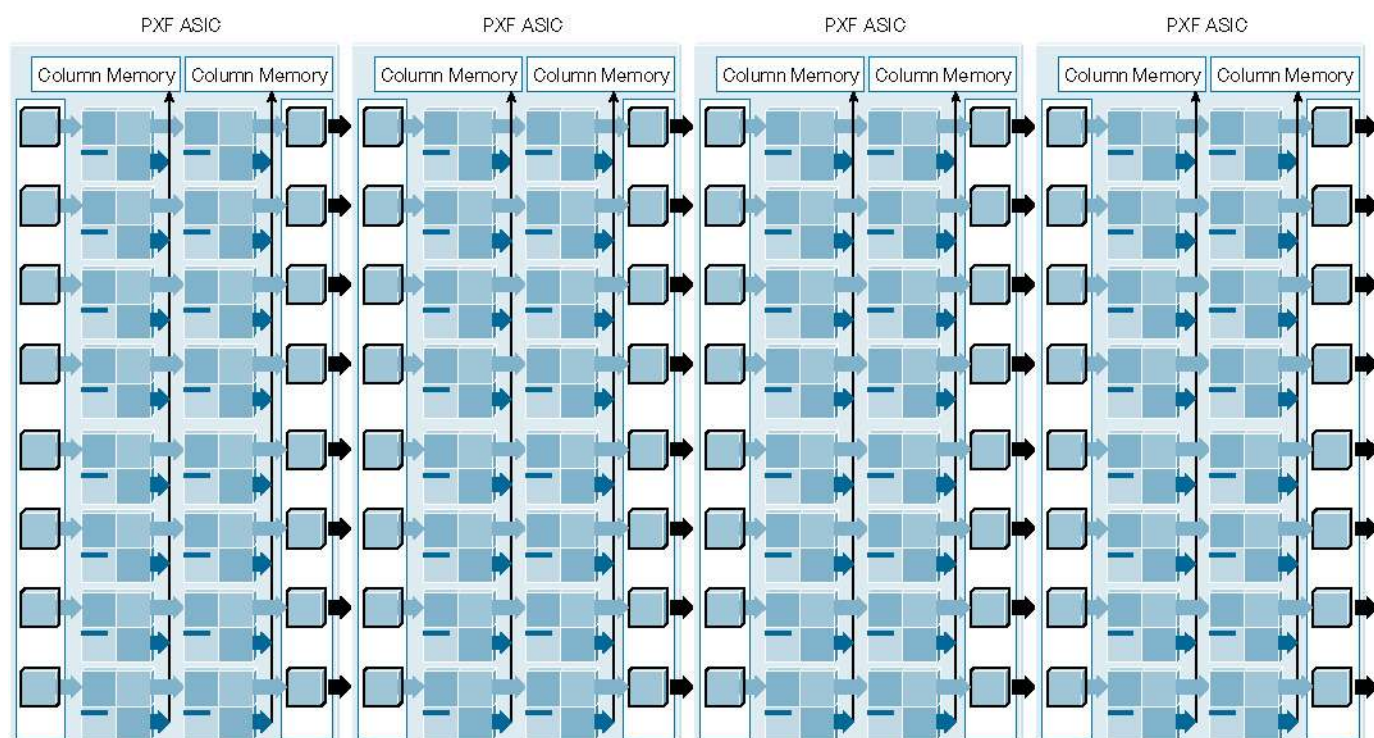
In the array of processors, microcode and Cisco IOS® Software are combined to provide feature processing. The exact allocation of features to microcontrollers in the processor pipeline is completely flexible and can change as new features are added.

The PXF network-processor architecture allows all processors to work efficiently on per-packet feature processing, yielding high throughput while still allowing substantial feature processing.

Figure 2

PXF Packet Processing on the Cisco 10000 Series Router
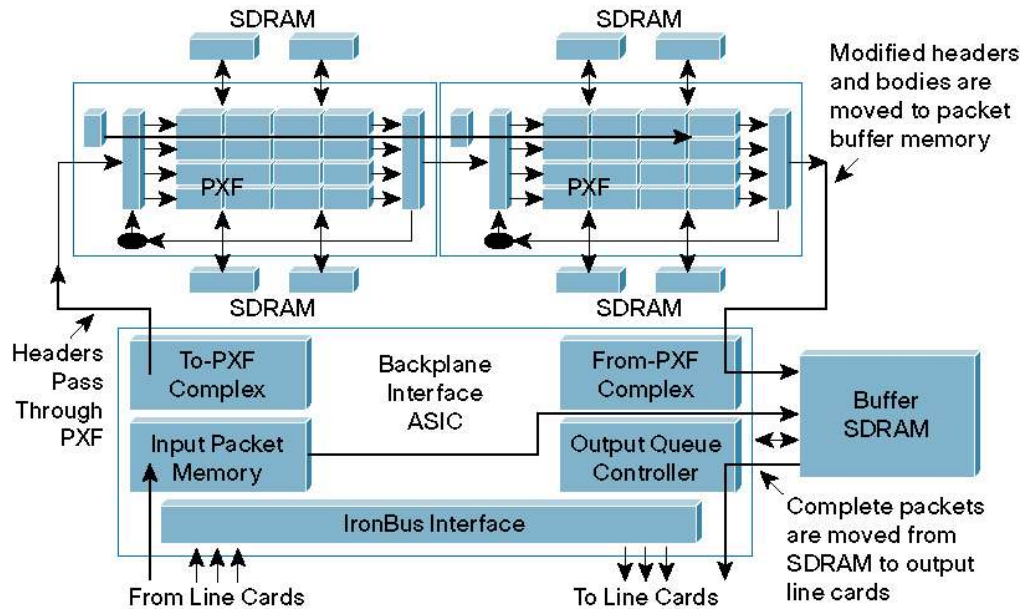
Cisco 10000 Series PXF Packet Processing

Figure 2 shows the process of a packet transiting the Cisco 10000 Series. A packet entering the Cisco 10000 Series from an interface card is sent over the backplane by the interface card. When it reaches the PRE, the packet is stored in buffer memory internal to the backplane interface ASIC. The backplane interface ASIC then passes the packet header to the PXF pipeline, which classifies the packet, modifies the packet header, and may modify the packet data. As part of this processing, PXF selects the output interface on which to forward the packet. In the simplest packet-routing cases, PXF then commands the backplane interface ASIC to store the packet in the backplane interface ASIC packet-buffer memory, in one of possibly several software output queues associated with the output interface. Some complex routing functions such as tunnels may require a packet to be processed twice (feedback) by PXF before it is placed in the output queue. Subsequently, the PXF output scheduling function applies various algorithms on these output queues to select which packet to send next. When a packet is scheduled for actual transmission on the output interface, PXF directs the backplane interface ASIC to copy the packet to the hardware queue associated with the output interface. The backplane interface ASIC will then transfer the packet across the backplane, and the packet will be sent on the output link by the line-card hardware.

Packets addressed to the router (for example, control plane or routing updates) are also processed by PXF. PXF will forward these packets to the route processor instead of an interface card. The output interface selected by the routing algorithm corresponds to a software queue associated with the backplane interface ASIC to route processor direct memory access (to-RP DMA) interface. Packets are queued in the backplane interface ASIC packet-buffer memory on a to-RP DMA software queue and scheduled for output as described earlier. When the packet is actually scheduled for output, however, PXF directs the backplane interface ASIC to move the packet from the backplane interface ASIC packet-buffer memory to the to-RP DMA interface output queue. The to-RP DMA engine copies the packet to the route-processor system memory, specifically to a buffer previously passed to the backplane interface ASIC by the route-processor Cisco IOS Software backplane interface ASIC driver. The backplane interface ASIC then announces the arrival of a packet to the route processor via a completion interrupt. In response to the interrupt, the backplane interface ASIC driver removes the packet from the to-RP DMA interface, performs preliminary classification of the packet, and passes the packet to upper-layer code within Cisco IOS Software for processing. As part of this processing, the driver corresponding to the network interface on which the packet was received may be invoked.

CISCO SYSTEMS

Corporate Headquarters
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA
www.cisco.com
Tel: 408 526-4000
      800 553-NETS (6387)
Fax: 408 526-4100

European Headquarters
Cisco Systems International BV
Haarlerbergpark
Haarlerbergweg 13-19
1101 CH Amsterdam
The Netherlands
www-europe.cisco.com
Tel: 31 0 20 357 1000
Fax: 31 0 20 357 1100

Americas Headquarters
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA
www.cisco.com
Tel: 408 526-7660
Fax: 408 527-0883

Asia Pacific Headquarters
Cisco Systems, Inc.
168 Robinson Road
#28-01 Capital Tower
Singapore 068912
www.cisco.com
Tel: +65 6317 7777
Fax: +65 6317 7799

Cisco Systems has more than 200 offices in the following countries and regions. Addresses, phone numbers, and fax numbers are listed on the
Cisco Website at www.cisco.com/go/offices.

Argentina • Australia • Austria • Belgium • Brazil • Bulgaria • Canada • Chile • China PRC • Colombia • Costa Rica
Croatia • Cyprus • Czech Republic • Denmark • Dubai, UAE • Finland • France • Germany • Greece • Hong Kong SAR
Hungary • India • Indonesia • Ireland • Israel • Italy • Japan • Korea • Luxembourg • Malaysia • Mexico
The Netherlands • New Zealand • Norway • Peru • Philippines • Poland • Portugal • Puerto Rico • Romania • Russia
Saudi Arabia • Scotland • Singapore • Slovakia • Slovenia • South Africa • Spain • Sweden • Switzerland • Taiwan
Thailand • Turkey • Ukraine • United Kingdom • United States • Venezuela • Vietnam • Zimbabwe

Printed in the USA   C11-349557-00 05/06