



Next-Generation Network Services

A guide to building service-oriented networks
to differentiate and grow your business

Next-Generation Network Services

Robert Wood

Copyright © 2006 Cisco Systems, Inc.

Published by:

Cisco Press

800 East 96th Street

Indianapolis, IN 46240 USA

All rights reserved. No part of this book may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or by any information storage and retrieval system, without written permission from the publisher, except for the inclusion of brief quotations in a review.

Printed in the United States of America 1 2 3 4 5 6 7 8 9 0

First Printing November 2005

Library of Congress Cataloging-in-Publication Number: 2003107970

ISBN: 1587051591

Warning and Disclaimer

This book is designed to provide information about Cisco next-generation network services. Every effort has been made to make this book as complete and as accurate as possible, but no warranty or fitness is implied.

The information is provided on an “as is” basis. The author, Cisco Press, and Cisco Systems, Inc., shall have neither liability nor responsibility to any person or entity with respect to any loss or damages arising from the information contained in this book or from the use of the disks or programs that might accompany it.

The opinions expressed in this book belong to the author and are not necessarily those of Cisco Systems, Inc.

Trademark Acknowledgments

All terms mentioned in this book that are known to be trademarks or service marks have been appropriately capitalized. Cisco Press or Cisco Systems, Inc., cannot attest to the accuracy of this information. Use of a term in this book should not be regarded as affecting the validity of any trademark or service mark.

Corporate and Government Sales

Cisco Press offers excellent discounts on this book when ordered in quantity for bulk purchases or special sales.

For more information please contact:

U.S. Corporate and Government Sales 1-800-382-3419 corpsales@pearsontechgroup.com

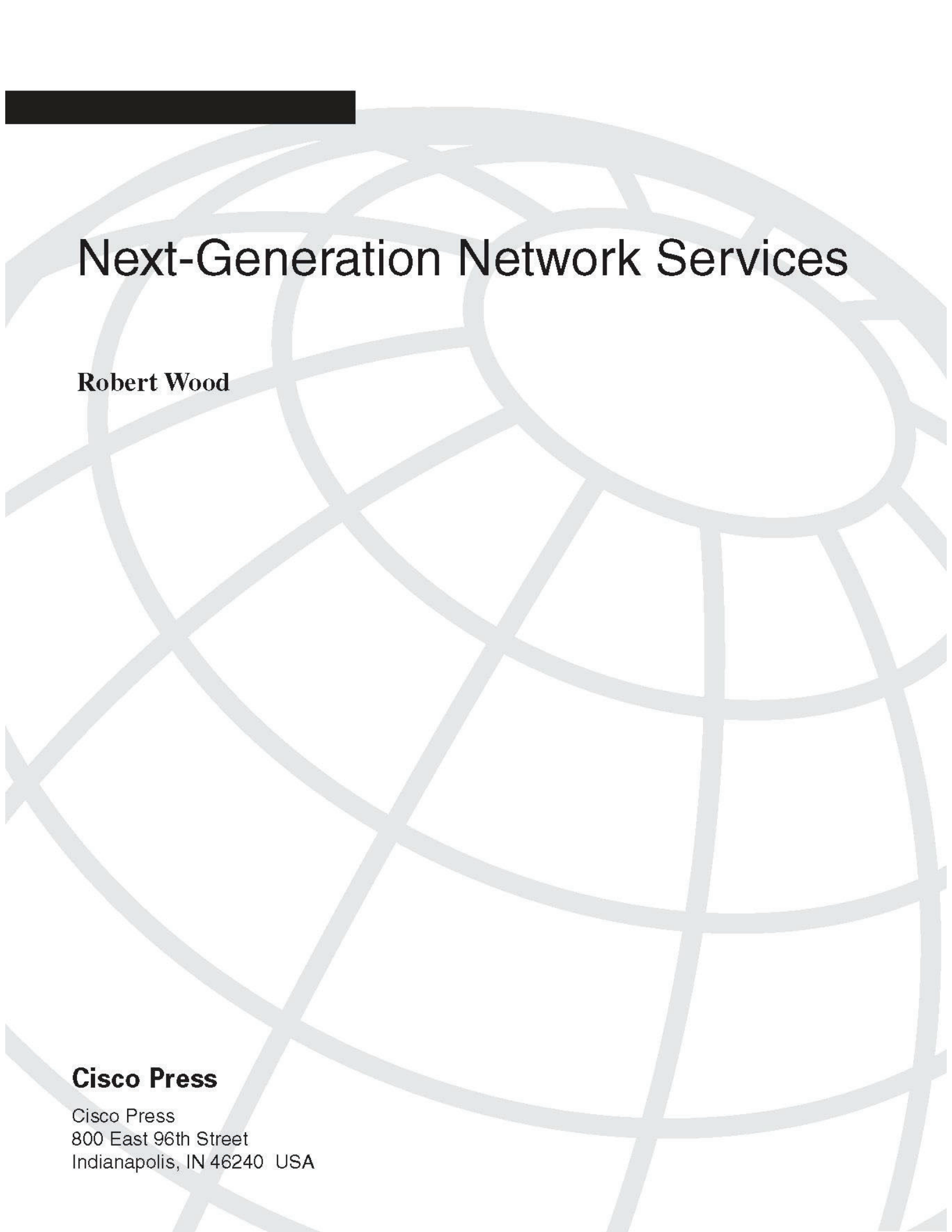
For sales outside the U.S. please contact: **International Sales** international@pearsoned.com

Feedback Information

At Cisco Press, our goal is to create in-depth technical books of the highest quality and value. Each book is crafted with care and precision, undergoing rigorous development that involves the unique expertise of members from the professional technical community.

Readers’ feedback is a natural continuation of this process. If you have any comments regarding how we could improve the quality of this book, or otherwise alter it to better suit your needs, you can contact us by e-mail at feedback@ciscopress.com. Please make sure to include the book’s title and ISBN in your message.

We greatly appreciate your assistance.



Next-Generation Network Services

Robert Wood

Cisco Press

Cisco Press
800 East 96th Street
Indianapolis, IN 46240 USA

Table 3-3 introduces a general application of MPLS technology.

Table 3-3 *MPLS Technology Application*

MPLS Characteristics	MPLS Features and Solutions
Requirements	Consolidated packet-based core Migrate Layer 2 customers to consolidated core Migrate Layer 2 services to Layer 3 services Multiservice provisioning platforms Transfer of complex routing tasks by enterprises to service providers Rapid IP service creation Ease of accounting and billing
Technology options	RFC 3031, "Multiprotocol Label Switching Architecture" MPLS Layer 3 VPNS (IETF 2547bis) MPLS TE Any Transport over ATM
Design options	Frame-based MPLS (IP) Cell-based MPLS (IP+ATM)
MPLS services	Layer 2 VPN services Layer 3 VPN services VPLS QoS Traffic Engineering

Cisco Next-Generation Multiservice Routers

For next-generation multiservice networks, routing platforms born and bred on the service pull of IP networking have the advantage. The greatest customer demand is for IP services. Networks built on IP are naturally multiservice-capable, given IP embellished data, VoIP, and video over IP convergence capabilities.

IP routing architecture has reached the hallowed five 9s of availability status, and representative platforms are faster, more scalable, and more service rich than any networking technology that has come before. Innovations such as MPLS have created the flexibility to combine both conventional and contemporary networking approaches, achieving more customer-service granularity in the process. The combination of distributed processing architectures, IP and hardware acceleration in programmable silicon, virtualization architecture, and continuous system software operations now deliver high-end, service provider IP routing platforms that are constant, flexible, affordable, and secure.

For high-end service provider multiservice routing, the notable products are the Cisco CRS-1 Carrier Routing System, the Cisco IOS XR Software, and the Cisco XR 12000/12000 Series Routers.

Cisco CRS-1 Carrier Routing System

Once you turn on the Cisco CRS-1 Carrier Routing System, you might never turn it off. Unlike many routers that have preceded the CRS-1 design, the CRS-1 is scalable and simple, continuous and adaptable, flexible and high performance. None of these individual characteristics of the CRS-1 compromises another, leading to new achievements in nondisruptive scalability, availability, and flexibility of the system. Using the CRS-1 Carrier Routing System, providers can visualize one network with many services and limitless possibilities.

Using a new Cisco IOS XR Software operating system that is also modular and distributed, the CRS-1 is the first carrier IP routing platform that can support thousands of interfaces and millions of IP routes, using a pay-as-you-grow architectural strategy. The CRS-1 blends some of the best of computing, routing, and programmable semiconductor and software architectures for a new, high-end routing system that you can use for decade-plus lifecycles.

Using the CRS-1's concurrent scalability, availability, and performance, you can use the CRS-1 to consolidate service provider point-of-presence (POP) designs, collapsing core, peering, and aggregation layers inside the covers of one system. Previous routing platforms had limitations in the number of peers, interfaces, or processing cycles, leading to network POP designs that layered functionality based on the performance constraints of the routing platforms. With the CRS-1, these limitations are removed—hardware works in concert with software for extensible convergence of network infrastructure and services. The CRS-1 represents the next-generation IP network core and is the foundation for IP/MPLS provider core consolidation.

CRS-1 Hardware Design

The CRS-1 hardware system design uses the primary elements of line card and fabric card shelves. Each type of shelf comprises a standard telecommunications rack from a dimensional footprint.

Line Card Shelf

Line card shelves support the routing processors, integrated fabric cards and the line card slots, each of which is capable of 40 Gbps performance. Known collectively as a *Line Card Chassis*, the chassis comes in either an 8-slot version or a 16-slot version.

NOTE

Cisco uses *shelf* as marketing terminology and the term *chassis* as engineering terminology; both terms are interchangeable.

Two Route Processors are installed per chassis, one for active and one for hot standby. The Route Processors have their own dedicated slots and don't subtract from the 8 or 16 potential line cards slots of either chassis. Each Line Card Chassis contains up to 8 fabric cards in the rear of the chassis to support the Benes switching fabric for a single shelf system configurations. Each line card is composed of rear-facing Interface Modules and front-facing Modular Services Cards connected together via a midplane design. The Line Card Chassis is where route processing, forwarding, and control-plane intelligence of the system resides.

Within each Line Card Chassis are 2 Route Processors, up to 16 Interface Modules, each pairing with 16 Modular Services Cards, and 8 fabric cards. Redundant fan trays, power supplies, and cable management complete the distinctive elements within Line Card Chassis.

Each Route Processor is made up of a symmetrical multiprocessing architecture based on a Dual PowerPC CPU complex with at least 4 GB of DRAM, 2 GB of Flash memory, and a 40 GB micro hard drive. One of the Route Processors operates in active mode with the other in hot standby. The Route Processors, along with system software, can provide non-stop forwarding (NSF) and stateful switchover (SSO) functions without losing packets. Another plus of the CRS-1 architecture is that any Route Processor can control any line card slot, on any Line Card Chassis in a multishelf system. Using features of the Cisco IOS XR Software operating system, Route Processors and line cards can be formed across the system chassis to create logical routers within the physical CRS-1 overall system. Any time that supplementary processing power is needed, the architecture supports the addition of distributed Route Processors, providing two additional Dual PowerPC CPU complexes with their associated DRAM, Flash, and hard drive.

To create a line card, a combination of Interface Modules and Modular Services Cards are used. The Interface Modules, also referred to as *Physical Layer Interface Modules (PLIMs)* contain the physical interface ports and hardware interface-specific logic. Interface Modules for the CRS-1 exist for OC-768c/STM-256x, OC-192c/STM-64c, OC-48c/STM-16c, and 10 Gigabit Ethernet. The Interface Modules, installed in the rear card cage of the Line Card Chassis, connect through the midplane to Modular Services Cards in the front card cage of the chassis.

The Cisco Modular Services Cards are made up of a pair of Cisco Silicon Packet Processors (SPPs), each of which is an array of 188 programmable Reduced Instruction Set Computer (RISC) processors. These SPPs are deployed two per Modular Services Card, with one for the input direction and one for output packet processing. The SPP is another key innovation, as the SPP architecture achieves 40 Gbps line rates with multiple services, offering new features through in-service software upgrades to the SPP. The Interface Module and the

Modular Services Card work together as a pair to form a complete line card slot. The Modular Services Card interfaces with the fabric cards, using the switching fabric to reach other line cards or the Route Processor memory.

Fabric Chassis

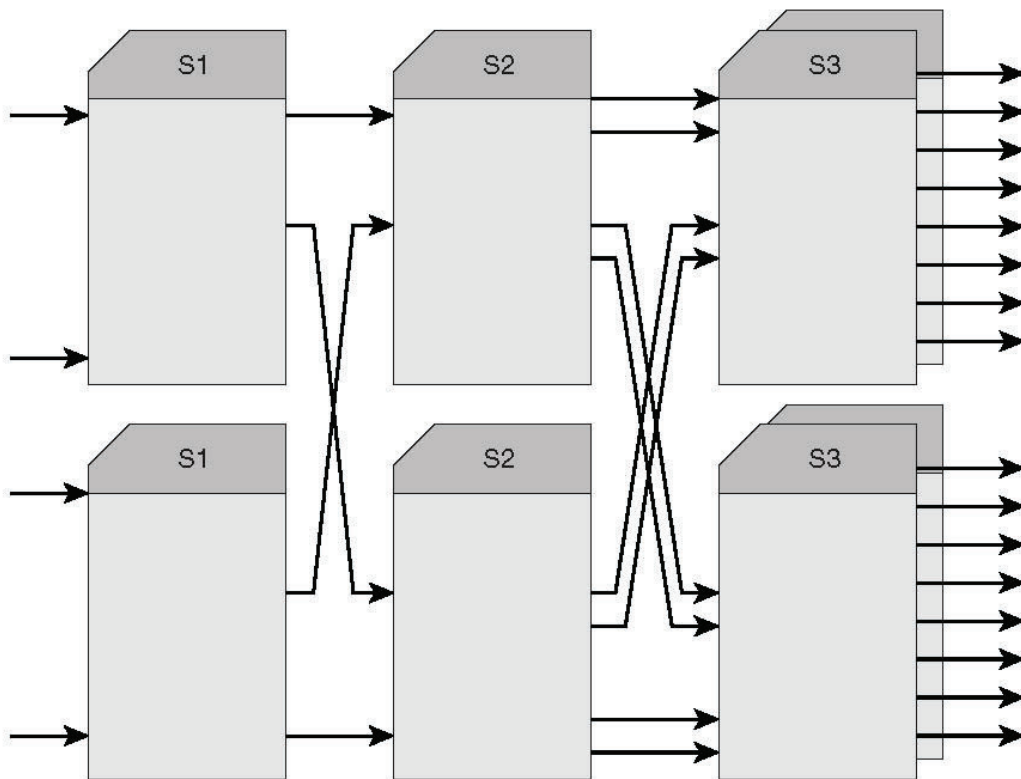
The Fabric Chassis is used to extend the CRS-1 into a CRS-1 Multishelf System. Up to 8 Fabric Chassis can interconnect as many as 72 Line Card Chassis to create the maximum CRS-1 Multishelf System. The Fabric Chassis is used as a massively scalable stage 2 of the three-stage Benes switching fabric in a multishelf system configuration.

A switching fabric is a switch backplane, and many of the Cisco products use various types of switching fabrics to move packets between ingress interfaces and Route Processor memory and out to egress interfaces. For example, a crossbar fabric is a popular fabric used in many Cisco products, such as the 12000 series and the 7600 series. For hundreds or even thousands of interface ports, a crossbar switching mechanism becomes too expensive and scheduling mechanisms too complex.

Therefore, the CRS-1 implements a three-stage, dynamically self-routed, Benes topology cell-switching fabric. This fabric is a multistage buffered switching fabric that represents the lowest-cost $N \times N$ cell-switching matrix that avoids internal blocking. The use of a backpressure mechanism within the fabric limits the use of expensive off-chip buffer memory, instead making use of virtual output queues in front of the input stage. Packets are converted to cells, and these cells are used for balanced load distribution through the switch fabric. The cells are multipath routed between stages 1 and 2 and again between stages 2 and 3 to assist with the overall goal of a nonblocking switching architecture. The cells exit stage 3 into their destination line card slots where the Modular Services Cards reassemble these cells into the proper order, forming properly sequenced packets. The Benes topology switching fabric is implemented in integrated fabric cards for single shelf systems and additionally implemented as standalone Fabric Chassis in a multishelf system configuration. Each standalone Fabric Chassis can contain up to 24 fabric cards for stage 2 operation.

A CRS-1 Single-Shelf system will use integrated fabric cards within the Line Card Chassis that include all three stages within the card. In a CRS-1 Multishelf System, from one to eight CRS-1 Fabric Chassis are used to form stage 2 of the switching fabric, with stage 1 operating on the fabric card of the egress line card shelf and stage 3 operating on the ingress line card shelf across the fabric.

Figure 3-8 shows a conceptual diagram of the CRS-1 switching fabric. Physically, the Cisco CRS-1 fabric is divided into eight planes over which packets are divided into fixed-length cells and then evenly distributed. Within the planes, the three fabric stages—S1, S2, and S3—dynamically route cells to their destination slots, where the Modular Services Cards reassemble cells in the proper order to form properly sequenced packets.

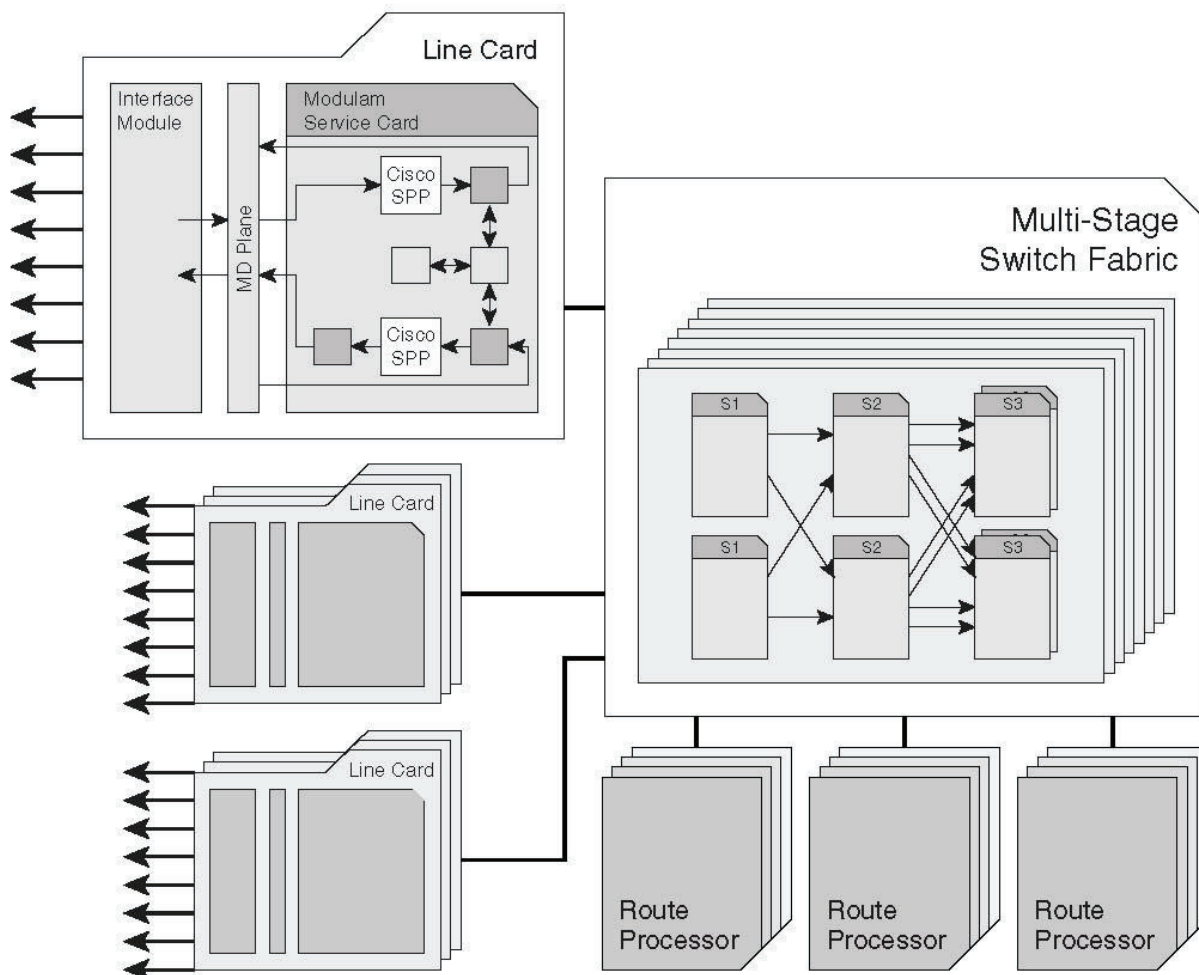
Figure 3-8 *One Plane of the Eight-Plane Cisco CRS-1 Switching Fabric*

Source: Cisco Systems, Inc.

Together the Route Processors, fabric cards, Interface Modules, and Modular Services Cards work with the IOS XR operating system to create a routing architecture that is scalable from 640 Gbps to 92 Tbps (terabits per second) of performance. These capacities are accomplished through various configurations of a CRS-1 Multishelf System or a CRS-1 Single-Shelf System. The overall CRS-1 architectural design is conceptualized in Figure 3-9.

Cisco CRS-1 Multishelf System

The Cisco CRS-1 Multishelf Systems are constructed using a combination of Line Card Chassis and Fabric Chassis. Up to 72 Line Card Chassis can be interconnected with 8 Fabric Chassis to create a multishelf system with as many as 1,152 line card slots, each capable of 40 Gbps, yielding approximately 92 Tbps (full duplex) of aggregate performance capacity. Cisco CRS-1 Multishelf Systems can start with as few as 2 Line Card Chassis and 1 Fabric Chassis and grow as demand occurs.

Figure 3-9 *Cisco CRS-1 Hardware Architecture*

Source: Cisco Systems, Inc.

Within a multiservice system, any Route Processor can control any line card on any Line Card Chassis in the system. For example, a Route Processor in Line Card Chassis number 1 can be configured to control a line card in Line Card Chassis number 72 using the Fabric Chassis as an internal connectivity path. Route Processors and distributed Route Processors are responsible for distributing control plane functions and processing for separation, performance, or logical routing needs.

Using a Cisco CRS-1 Multishelf System, providers can achieve the following configurations:

- 2 to 72 Line Card Chassis
- 1 to 8 Fabric Chassis
- Switching capacity from 640 Gbps to 92 Tbps (full duplex)

- Support for up to 1,152 line cards at 40 Gbps each
 - 1,152 OC-768c/STM-256c POS ports
 - 4,608 OC-192c/STM-64c POS/DPT ports
 - 9,216 10 Gigabit Ethernet ports
 - 18,432 OC-48c/STM-16c POS/DPT ports

Cisco CRS-1 16-Slot Single-Shelf System

The CRS-1 Single-Shelf Systems come as either a 16-slot or an 8-slot Line Card Chassis. Single-shelf systems use integrated Switch Fabric Cards (SFCs), installed in the rear card cage of the Line Card Chassis rather than using a standalone Fabric Chassis. In a single-shelf system configuration, the integrated SFCs perform all three stages of the Benes topology switching fabric operation. Using a Cisco CRS-1 16-Slot Single-Shelf System, providers can achieve the following configurations:

- 16-slot Line Card Chassis with integrated fabric cards
- Switching capacity to 1.28 Tbps (full duplex)
- Support for up to 16 line cards at 40 Gbps each
 - 16 OC-768c/STM-256c POS ports
 - 64 OC-192c/STM-64c POS/DPT ports
 - 128 10 Gigabit Ethernet ports
 - 256 OC-48c/STM-16c POS/DPT ports

Cisco CRS-1 8-Slot Single-Shelf System

The CRS-1 Single-Shelf Systems also come in an 8-slot Line Card Chassis. The 8-slot Line Card Chassis is one half as tall as a 16-slot Line Card Chassis. As previously mentioned, single-shelf systems use the integrated SFCs, installed in the rear card cage of the Line Card Chassis, performing all three stages of the Benes topology switching fabric operation. Using a Cisco CRS-1 8-Slot Single-Shelf System, providers can achieve the following configurations:

- 8-slot Line Card Chassis with integrated fabric cards
- Switching capacity to 640 Gbps (full duplex)
- Support for up to 8 line cards at 40 Gbps each
 - 8 OC-768c/STM-256c POS ports
 - 32 OC-192c/STM-64c POS/DPT ports
 - 64 10 Gigabit Ethernet ports
 - 128 OC-48c/STM-16c POS/DPT ports

medium adaptability of Ethernet all lend to the descending cost curve and the ascending adoptability of Ethernet. Outpacing its rivals by greater than 20 to 1, Ethernet has moved from LAN technology to the MAN and toward WAN consideration. People are taking Ethernet home with them. The number of home-based Ethernet networks is soaring for users of broadband Internet connections worldwide. To link consumer Ethernet with corporate Ethernets for communications, caching, and commerce, optical Ethernet is increasingly the conduit.

Media-transparent—whether cable, DSL, optical, or wireless—Ethernet is surfing LAN-switching capability, rocketing on computer gigahertz improvement, and riding optical fiber—breaking speed and distance barriers to 10 Gbps and beyond. Using optical fiber, Ethernet has moved out of the Campus LAN and into the MANs and WANs. By using optical Ethernet in provider networks, it becomes possible to set and forget the customer interface equipment, enjoying linear scalability of bandwidth through software control rather than truck rolls and circuit-dependent interfaces. Ethernet in the provider network is a significant contributor to OpEx savings, as it is a cost-effective way to match interfaces, speeds, and protocols with customers. For these reasons, Ethernet carries a lot of value with the customer.

Through the benefits of multimode and single-mode optical fiber and continuing advancements in Ethernet technology, Gigabit Ethernet is now a familiar tenant in enterprises and service provider metropolitan offerings. To aggregate the bandwidth of Gigabit Ethernet from desktops, servers, and mainframes, 10 Gigabit Ethernet (10GE) moves the decimal point for a 10x improvement in backbone capacity. At 10 Gbps, 10GE is the current champ of the Layer 2 speed race. Perhaps the most significant of all, 10GE in provider networks enables physical layer convergence of the LAN, MAN, and WAN.

Both Gigabit Ethernet and 10GE become customary transports for all enterprise protocols such as Fiber Connection (FICON), Enterprise System Connection (ESCON), Fibre Channel, and IP. Gigabit Ethernet is most often used as an edge connectivity service and will appeal to a large variety of customers. Sporting a ten-fold performance increase over Gigabit Ethernet at about three to four times the cost of GE, 10GE will find many applications as a core and aggregation technology in MANs and long-haul networks using DWDM. It can also be used as an edge service for very high-bandwidth applications. At this line rate, 10GE can be used to link together supercomputers for collaborative, shared, or grid computing; medical imaging; remote medical telesurgery with robotics; geographically synchronized storage farms; and so on.

Gigabit Ethernet and 10GE over Optical Networks

Gigabit Ethernet and 10GE are becoming the technologies of choice within enterprise, metropolitan, wide area, and perhaps even residential networks. With the ability to directly modulate Ethernet over optical fiber at distances up to 40 km or more, providers, operators,

companies, and users can converge LANs with MANs, and WANs that use the same Layer 2 transport end to end. This enables lower-cost MANs using Layer 3 and Layer 4 Ethernet switching and 10GE backbones. With Ethernet over optical, you can do the following:

- Market profitable Ethernet services such as
 - Ethernet LAN extension
 - Ethernet-based Internet service
 - Storage area network services
 - Disaster recovery services
 - Ethernet home networking
- Use Ethernet as a simple layer to leverage less expensive point-to-point services equivalent to an OC-3 service.
- Use CWDM and DWDM wavelengths to provision multiple virtual Ethernet services.
- Sell optical wavelengths to other service providers who want to offer Ethernet services to businesses.
- Provide routing services that easily extend intelligent IP services between customer buildings.
- Use Ethernet as an effective transport for video on demand (VOD).

Gigabit Ethernet over optical fiber was first standardized in 1998 (IEEE 802.3z). 10GE was standardized in 2002 as IEEE 802.3ae. Additional Ethernet over optical standardization activity continues as shown in Table 5-8.

Table 5-8 *Gigabit Ethernet and 10GE Standards*

Standard	IEEE 802.3z	IEEE 802.3ae	IEEE 802.3ah	IEEE 802.3aq
Year ratified	1998	2002	2004	Estimated 2006
Description	Gigabit Ethernet 1000BASE-X 1000BASE-T	10GE 10GBASE-xR 10GBASE-xW 10GBASE-LX4	Ethernet in the First Mile (EFM) 100BASE-FX 100BASE-LX10 1000BASE-SX 1000BASE-LX 1000BASE-LX10 1000BASE-BX-D 1000BASE-BX-U	10GE over multimode fiber

Table 5-8 *Gigabit Ethernet and 10GE Standards (Continued)*

Standard	IEEE 802.3z	IEEE 802.3ae	IEEE 802.3ah	IEEE 802.3aq
Fiber support	Multimode and single-mode fiber	Multimode and single-mode fiber	100/1000BASE unidirectional and bidirectional optics Multimode and 10 km single-mode fiber	>220 meters of FDDI-grade multimode fiber
Standard optics	1000BASE-SX 850 nm for MMF only 1000BASE-LX 1310 nm for MMF and SMF	10GBASE-SR and SW using 850 nm for MMF only 10GBASE-LR and LW using 1310 nm for SMF 10GBASE-LX4 using four lanes of 1310 nm over MMF and SMF at 2.5 Gbps each 10GBASE-ER and EW using 1550 nm over SMF	1000BASE-LX and 1000BASE-BX for SMF	Under definition

NOTE

The IEEE specifications describe the port types for each of the IEEE 802.3 interfaces. While 1000BASE represents that Gigabit Ethernet (GE) is a 1000 Mbps baseband transmission, the T and the X represent GE over category 5 twisted pair (T for twisted pair) and GE over optical fiber (X for wavelength). Therefore, there are four varieties of Gigabit Ethernet port types that are common in the market

- 1000BASE-T for 1000 Mbps operation over category 5 copper
- 1000BASE-SX for 1000 Mbps operation over multimode fiber with a short-wavelength 850 nm laser
- 1000BASE-LX for 1000 Mbps operation over single-mode fiber with a long wavelength 1310 nm laser and 8B/10B encoding on MMF or SMF

Gigabit Ethernet for Optical Networks

Gigabit Ethernet is a 1000 Mbps (1 Gbps) Ethernet service that is very popular in the access and distribution layers of MANs. Many enterprises need internetworking between campuses, and Ethernet is a logical choice. Because many of these enterprises use Gigabit Ethernet in their LAN backbones, a Gigabit Ethernet offering is usually a minimum requirement. Gigabit Ethernet finds wide application in the metropolitan area. Metro Ethernet networks can use Gigabit Ethernet to deliver both full- and fractional-speed point-to-point and multipoint Ethernet services. Many providers with a SONET/SDH infrastructure add Gigabit Ethernet connectivity to their offerings.

Gigabit Ethernet can be scaled through the use of the Cisco EtherChannel feature. Based on the IEEE 802.3ad Link Aggregation standard, Cisco products can bond multiple Gigabit Ethernet facilities into a Gigabit EtherChannel to provide increments of 1000 Mbps (1 Gbps) connectivity up to 8 Gbps. With Gigabit EtherChannel, multiple links become an application-transparent bandwidth pool. If one of the links within the channel were to fail, that link traffic is redirected to another link within one second, much less than the trip point of any protocol timers that could cause or report an application session error. Gigabit EtherChannel is a customary high-availability interconnection method between IP routers. A two-link Gigabit EtherChannel between a pair of routers provides not only 2 Gbps of bandwidth capacity but ensures that the connection survives in the event of a link or port failure.

Gigabit Ethernet can be provisioned over a metro access ring via single-mode fiber using either CWDM or DWDM. Gigabit Ethernet is also a prime technology for cable service operators to deliver VOD to subscriber households, again generally using DWDM wavelengths to increase the service density and efficiency of their deployed optical fiber backbones that serve their metropolitan subscriber base.

Gigabit Ethernet is also a key technology for use with the Ethernet in the First Mile (EFM) standard 802.3ah. Using the 1000BASE-LX and 1000BASE-BX transceivers over single-mode fiber, Gigabit Ethernet connectivity can be extended to business and residential access areas. The EFM standard is rated for the increased fiber distances that are needed to reach these suburban areas, up to 10 km. EFM provides for use of single- or dual-fiber operation and allows for extended temperature support for the optic transceivers. The vision of the EFM standard is the capability to deliver Gigabit Ethernet to each user, business or residential.

10 Gigabit Ethernet for Optical Networks

10GE is the next stride for scaling the performance of both enterprise and service provider networks, combining multigigabit bandwidth and intelligent services end to end. The 10GE-emerging applications are campus extension, disaster recovery, and data center extensions such as a standard transport for data center protocols including Fibre Channel—

for example, nine Fibre Channels into one 10GE lambda or ESCON with 40 ESCON channels per 10GE lambda. 10GE is a common requirement to facilitate grid computing and supercomputing.

10GE is based on Ethernet, using the same Ethernet MAC protocol, the Ethernet frame format, and the Ethernet frame size. Interestingly, the specification supports full duplex as the only mode of bidirectional communication. Because it is full duplex only, 10GE is unlimited in reach, limited in distance only by the physics and impairments of the optical media transmission that is used. In a full-duplex link, there are no Ethernet packet collisions, so link distances are determined by the limitations of optics and cost and not by the diameter of an Ethernet collision domain.

10GE is more popular in the core portion of MANs. Roughly equivalent in bandwidth to a SONET/SDH OC-192c, 10GE possesses the capacity to scale core networks even further using 10GE EtherChannel.

Many enterprises are considering a move to Ethernet for WAN connectivity, especially in the larger metropolitan statistical areas (MSAs). With an abundance of metropolitan fiber and multigigabit switching products, enterprises can extend fractional or full Fast Ethernet and Gigabit Ethernet across owned or leased fiber wavelengths to create enterprise WAN connections up to 40 km or more. Because 10GE requires no change to the Ethernet MAC protocol or packet format, 10GE also supports all upper-layer services.

Large enterprises with high-bandwidth requirements between data centers might elect to acquire a 10GE connection across a leased fiber wavelength or perhaps across a leased OC-192/STM-64 circuit from a provider.

Ethernet is an asynchronous link protocol with its network timing and data synchronization maintained within each character in the bit stream. As a router, switch, or hub receives the Ethernet bit stream, it has the opportunity to resynchronize at the beginning of a character and to retiming the data transfer. This avoids the use of a network-wide synchronous clock source that must be hierarchically extended end to end within an enterprise network. As you will see in the next section, "Ethernet over Next-Generation SONET/SDH," a SONET network must share the same stratum level clock (usually Stratum 1) across all SONET/SDH nodes, an underlying clock source which is used to avoid timing drift between transmission and receipt of a datagram or TDM voice sample.

The reasons for Gigabit Ethernet's and 10GE's ascendancy into the provider and operator domains are multifold, beyond their inherent simplicity. First, the technologies leverage the vast installed base of Ethernet worldwide, from enterprises to consumers. Second, the design philosophy of the Ethernet manufacturing industry consistently uses high-volume manufacturing and low-cost design to create the appropriate price/performance metrics to ensure the product's market success. Additionally, the open standard specifications of the 802.3 framework enable mass-market competition at every stage of the Ethernet value chain, lowering the cost of components, subsystems, and final products.

Most of the complexities of encoding schemes, electronic circuitry, and optical interfaces are embedded in merchant silicon, enabling a wide choice of Ethernet subsystems at competitive costs. 10GE is following the same price/performance curves brought about by such innovation principles—for example, allowing 10GE to be transported up to 40 km using lower-cost, uncooled pluggable optics. As mentioned earlier in the section “Light Emitters,” the use of vertical cavity surface-emitting lasers (VCSELs) is likely to lower these costs even further.

Using Ethernet over optical, many choices are now possible, such as Ethernet over SONET/SDH, Ethernet over RPR/DPT, Ethernet over optical dark fiber using pluggable optic transceivers, and optical transponders on metro and long-haul platforms.

Ethernet over Next-Generation SONET/SDH

The success of Ethernet has created a groundswell, charging out-of-enterprise LANs and creating demand for Ethernet connectivity options in the provider domains. Many service providers want to offer Ethernet services and overlay them onto their embedded SONET/SDH network infrastructure. Next-generation SONET/SDH features have been developed to optimize Ethernet over SONET/SDH bandwidth efficiencies.

Many of the metro provider products use Ethernet cards that interface and map Ethernet ports into SONET STS payloads by aggregating at the STS-1 level. With an STS-1 payload equal to 49.5 Mbps, you can provision various-speed Ethernet services by using STS bandwidth scaling. For example, the following can be provisioned:

- Ethernet (10BASE-T) can be transported within an STS-1 circuit.
- Fast Ethernet (100BASE-T) can be transported within an STS-3c circuit.
- Gigabit Ethernet (1000BASE-X) can be transported within an STS-24c circuit.
- Subrate Gigabit Ethernet can be transported in STS-6c (297 Mbps), STS-9c (445 Mbps), and STS-12c (594 Mbps) circuits.

While STS bandwidth scaling is an effective way to map Ethernet into SONET frames, it is bandwidth inefficient.

Delivering Ethernet over SONET/SDH is not without challenge, because Ethernet is an asynchronous form of data transmission, whereas SONET/SDH uses a synchronous payload. Ethernet is a byte-oriented, self-clocking data stream; SONET is a bit-oriented routing system. Ethernet creates dynamic bandwidth, whereas SONET/SDH expects constant bandwidth. Ethernet is connectionless, whereas SONET/SDH is connection oriented. Merging the pair, as is, causes inefficiencies of transmission, in effect wasting bandwidth. Wasting bandwidth on SONET/SDH platforms tends to go against the basic covenants of telecommunications.

The move to next-generation SONET/SDH is desired, because several component technologies and standards-based developments are layered on SONET/SDH to increase