Cisco Systems

MPLS and VPN Architectures

Volume II

Master the latest MPLS VPN solutions to design, deploy, and troubleshoot advanced or large-scale networks

Ivan Pepelnjak, CCIE® No. 1354
Jim Guichard, CCIE No. 2069
Jeff Apcar

ciscopress.com

MPLS and VPN Architectures, Volume II

Jim Guichard, CCIE No. 2069
Ivan Pepelnjak, CCIE No. 1354
Jeff Apcar

Cisco Press
201 West 103rd Street
Indianapolis, IN 46290 USA

MPLS and VPN Architectures, Volume II

Jim Guichard, CCIE No. 2069

Ivan Pepelnjak, CCIE No. 1354

Jeff Apcar

Published by:
Cisco Press
201 West 103rd Street
Indianapolis, IN 46290 USA

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

Library of Congress Cataloging-in-Publication Number: 619472051122

ISBN: 1-58705-112-5

Warning and Disclaimer

This book is designed to provide information about MPLS and VPN architectures. 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 authors, 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 discs or programs that may accompany it.

The opinions expressed in this book belong to the authors 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.

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 through e-mail at feedback@ciscopress.com. Please make sure to include the book title and ISBN in your message.

We greatly appreciate your assistance.

Publisher	John Wait
Editor-In-Chief	John Kane
Cisco Representative	Anthony Wolfenden
Cisco Press Program Manager	Sonia Torres Chavez
Manager, Marketing Communications, Cisco Systems	Scott Miller
Cisco Marketing Program Manager	Edie Quiroz
Acquisitions Editor	Amy Moss
Production Manager	Patrick Kanouse
Development Editor	Grant Munroe
Project Editor	Lori Lyons
Copy Editor	Karen A. Gill
Technical Editors	Matt Birkner, Dan Tappan
Content Editor	Monique Morrow
Team Coordinator	Tammi Ross
Book Designer	Gina Rexrode
Cover Designer	Louisa Adair
Production Team	Mark Shirar
Indexer	Tim Wright

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.
Capital Tower
168 Robinson Road
#22-01 to #29-01
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.com Web site at www.cisco.com/go/offices.**

Argentina • Australia • Austria • Belgium • Brazil • Bulgaria • Canada • Chile • China PRC • Colombia • Costa Rica • Croatia • 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

About the Authors

Jim Guichard, CCIE No. 2069, is a Technical Leader II within the Internet Technologies Division (ITD) at Cisco Systems. During the past six years at Cisco and previously at IBM, Jim has been involved in the design, implementation, and planning of many large-scale WAN and LAN networks. His breadth of industry knowledge, hands-on experience, and understanding of complex internetworking architectures have enabled him to provide valued assistance to many of Cisco's larger service provider customers. His previous publications include *MPLS and VPN Architectures* , by Cisco Press.

Ivan Pepelnjak, CCIE No. 1354, is the Chief Technology Advisor and member of the board with NIL Data Communications (www.NIL.si), a high-tech data communications company that focuses on providing high-value services in new-world service provider technologies.

Ivan has more than 10 years of experience in designing, installing, troubleshooting, and operating large corporate and service provider WAN and LAN networks, several of them already deploying MPLS-based virtual private networks (VPNs). He is the author or lead developer of a number of highly successful advanced IP courses covering MPLS/VPN, BGP, OSPF, and IP QoS, and he is the architect of NIL's remote lab solution. Ivan's previous publications include *MPLS and VPN Architectures* and *EIGRP Network Design Solutions* , by Cisco Press.

Jeff Apcar is a Senior Design Consulting Engineer in the Asia Pacific Advanced Services group at Cisco Systems. He is one of the Cisco lead consultants on MPLS in the region and has designed MPLS networks for many service providers in AsiaPac using packet-based and cell-based MPLS. Jeff has also designed and maintained large IP router networks (500+ nodes) and has a broad and deep range of skills covering many facets of networking communications.

Jeff has more than 24 years of experience in data communications and holds Dip. Tech (Information Processing) and B.App.Sc (Computing Science) (Hons) from the University of Technology, Sydney, Australia.

About the Technical Reviewers

Matthew H. Birkner, CCIE No. 3719, is a Technical Leader at Cisco Systems, specializing in IP and MPLS network design. He has influenced multiple large carrier and enterprise designs worldwide. Matt has spoken at Cisco Networkers on MPLS VPN technologies in both the U.S. and EMEA over the past few years. A "double CCIE", he has published the Cisco Press book, Cisco Internetwork Design. Matt holds a BSEE from Tufts University, where he majored in electrical engineering.

Dan Tappan is a distinguished engineer at Cisco Systems. He has 20 years of experience with internetworking, having worked on the ARPANET transition from NCP to TCP at Bolt, Beranek, and Newman. For the past several years, Dan has been the technical lead for Cisco's implementation of MPLS (tag switching) and MPLS/VPNs.

About the Content Reviewer

Monique Morrow is currently CTO Consulting Engineer at Cisco Systems, Inc. She has 20 years of experience in IP internetworking that includes design, implementation of complex customer projects, and service development for service providers. Monique has been involved in developing managed network services such as remote access and LAN switching in a service provider environment. She has worked for both enterprise and service provider companies in the United States and in Europe. She led the Engineering Project team for one of the first European MPLS-VPN deployments in 1999 for a European service provider.

Dedications

To my wife Sadie, for putting up with me writing another book and the long lonely nights associated with such an undertaking. To my children Aimee and Thomas, who always help to keep me smiling.—Jim

To my wife Karmen, who was always there when I needed encouragement or support. To my children Maja and Monika, who waited patiently for my attention on too many occasions.—Ivan

To my wife Anne, who is an exceptional person in every way. To my children Caitlin, Conor, and especially Ronan: Despite his constant efforts to reboot my PC, I managed to lose a draft only once.—Jeff

Acknowledgments

Every major project is a result of teamwork, and this book is no exception. We'd like to thank everyone who helped us in the long writing process: our development editor, Grant Munroe, who helped us with the intricacies of writing a book; the rest of the editorial team from Cisco Press; and especially our reviewers, Dan Tappan, Matt Birkner, and Monique Morrow. They not only corrected our errors and omissions, but they also included several useful suggestions to improve the quality of this publication.

Jeff would like to thank his management team Tony Simonsen, Michael Lim, and Steve Smith, for providing the time and encouragement to do the book. Also special thanks to the guys in the AsiaPac Lab Group, Nick Stathakis, Ron Masson, and George Lerantges, who let him hog lots of gear. Last, Jeff would like to thank Jim and Ivan for inviting him to collaborate with them.

Finally, this book would never have been written without the continuous support and patience of our families, especially our wives, Sadie, Karmen, and Anne.

Contents at a Glance

Table of Contents

Introduction

Since our first MPLS book (*MPLS and VPN Architectures*) was published by Cisco Press a few years ago, MPLS has matured from a hot leading-edge technology—supporting Internet services and leased-line–based VPN solution—to a set of solutions that are successfully deployed in large-scale service provider networks worldwide. A number of additional solutions had to be developed to support the needs of these networks, and many additional IOS services were made VPN-aware to enable the service providers to deploy the services they were already offering within the new architectural framework. Therefore, it was a natural step to continue on the path we charted with the first book and describe the enhancements made to MPLS architecture or its implementation in Cisco IOS in *MPLS and VPN Architectures: Volume II*.

Who Should Read This Book?

This book is not designed to be an introduction to Multiprotocol Label Switching (MPLS) or virtual private networks (VPNs); Volume I (*MPLS and VPN Architectures*) provides you with that knowledge. This book is intended to tremendously increase your knowledge of advanced MPLS VPN deployment scenarios and enable you to deploy MPLS and MPLS VPN solutions in a variety of complex designs. Anyone who is involved in design, deployment, or troubleshooting of advanced or large-scale MPLS or MPLS VPN networks should read it.

How This Book Is Organized

Although this book could be read cover-to-cover, it is designed to be flexible and allow you to easily move between chapters and sections of chapters to cover just the material that you need more information on. If you do intend to read them all, the order in the book is an excellent sequence to use.

Part I: Introduction

Chapter 1, "MPLS VPN Architecture Overview," serves as a refresher to the information contained within MPLS and VPN Architectures. It does not describe the MPLS or MPLS VPN technology in detail; if you need baseline MPLS or MPLS VPN knowledge, read *MPLS and VPN Architectures*: Volume I first.

Part II: Advanced PE-CE Connectivity

Chapter 2, "Remote Access to an MPLS VPN," discusses integration of access technologies such as dial, DSL, and cable into an MPLS VPN backbone. This chapter shows how you can integrate various access technologies into the backbone, thereby providing VPN service to many types of customers.

Chapter 3, "PE-CE Routing Protocol Enhancements and Advanced Features," builds on Volume 1 of the *MPLS and VPN Architectures* book and introduces more advanced options/features for OSPF connectivity as well as support for IS-IS and EIGRP routing protocols.

Chapter 4, "Virtual Router Connectivity," discusses the use of the VRF constructs to build virtual router type connectivity, extending the VRF concept to the CE router. This chapter also discusses new VRF-related features, including VRF-lite and PE-based network address translation (PE-NAT).

Part III: Advanced Deployment Scenarios

Chapter 5, "Protecting the MPLS-VPN Backbone," looks at various security issues within the backbone and describes the necessary steps that a service provider must take to protect the backbone and any attached VPN sites.

Chapter 6, "Large-Scale Routing and Multiple Service Provider Connectivity," describes the advanced features, designs, and topologies that were made possible with the enhancements to Cisco IOS since the first MPLS and VPN Architectures book was written.

Chapter 7, "Multicast VPN," discusses the deployment of IP multicast between VPN client sites.

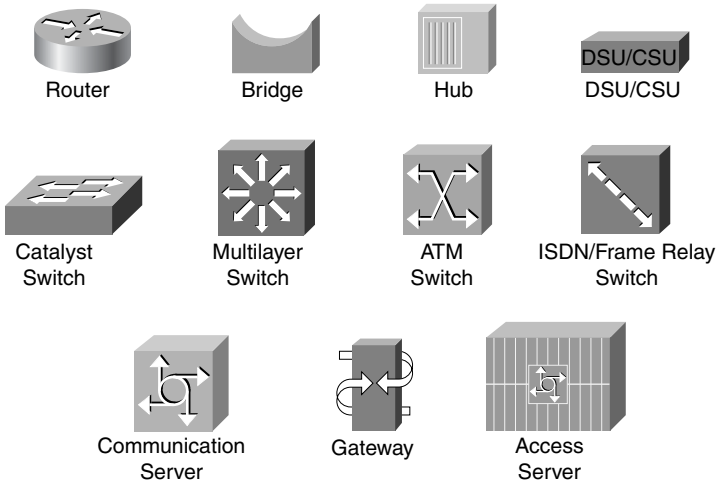
Chapter 8, "IP Version 6 Across an MPLS Backbone," discusses a model (6PE) that gives the service providers an option to provide IPv6 connectivity across an MPLS-enabled IPv4 backbone.

Part IV: Troubleshooting

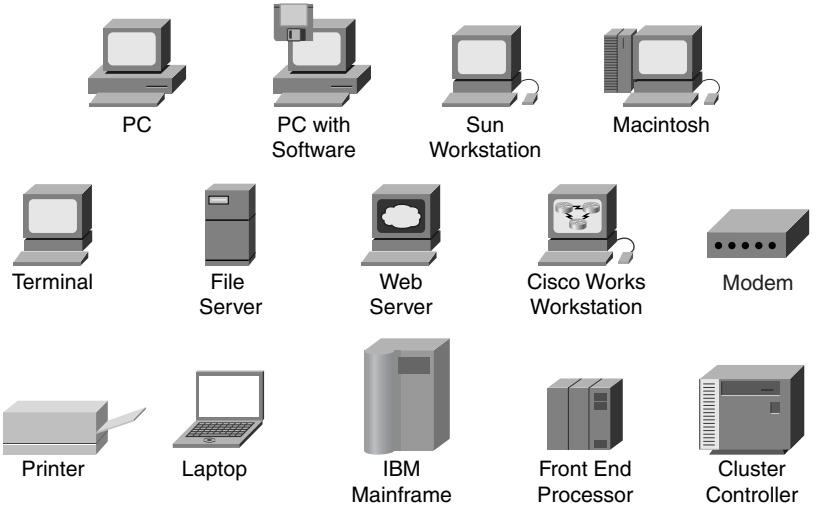
Chapter 9, "Troubleshooting of MPLS-Based Solutions," provides a streamlined methodology for identifying faults in MPLS solutions and troubleshooting an MPLS VPN backbone.

Icons Used in This Book

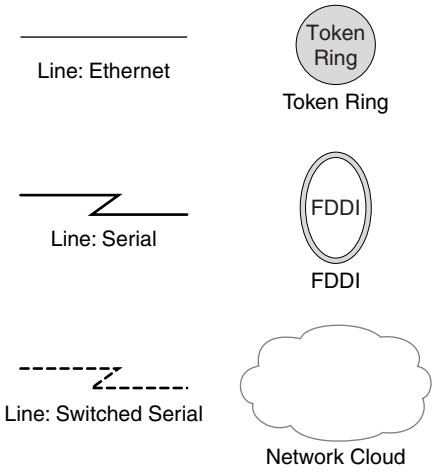
Throughout this book, you will see the following icons used for networking devices:



The following icons are used for peripherals and other devices:



The following icons are used for networks and network connections:



Command Syntax Conventions

The conventions used to present command syntax in this book are the same conventions used in the IOS Command Reference. The Command Reference describes these conventions as follows:

- Vertical bars (|) separate alternative, mutually exclusive elements.
- Square brackets [ ] indicate optional elements.
- Braces { } indicate a required choice.
- Braces within brackets [{ }] indicate a required choice within an optional element.
- Boldface indicates commands and keywords that are entered literally as shown. In actual configuration examples and output (not general command syntax), boldface indicates commands that are manually input by the user (such as a show command).

Italics indicate arguments for which you supply actual values.

This page intentionally left blank

PART

Introduction

Chapter 1 MPLS VPN Architecture Overview

CHAPTER 1

MPLS VPN Architecture Overview

Virtual private networks (VPNs) have recently received a lot of attention from equipment manufacturers, consultants, network designers, service providers, large enterprises, and end users due to their cost advantages over traditional enterprise networks. As with most technologies, the foundation for today's VPN networks and underlying technologies was created more than 20 years ago. During its development, end users discovered that it made financial sense to replace links between sites in their own private network with virtual connections across a shared infrastructure. The assumption for doing this was that a shared environment (or VPN) is equivalent in terms of security and privacy to the network (links) it was replacing.

This chapter reviews the basic Multiprotocol Label Switching (MPLS) and MPLS-based VPN concepts and terminologies to ensure an understanding of the terms used in this book. It also covers the latest developments in the MPLS VPN arena and how they enable the service provider to offer new MPLS-based services, such as remote access into an MPLS-based VPN or Internet Protocol (IP) multicast within a VPN. These developments are also described in depth in later chapters.

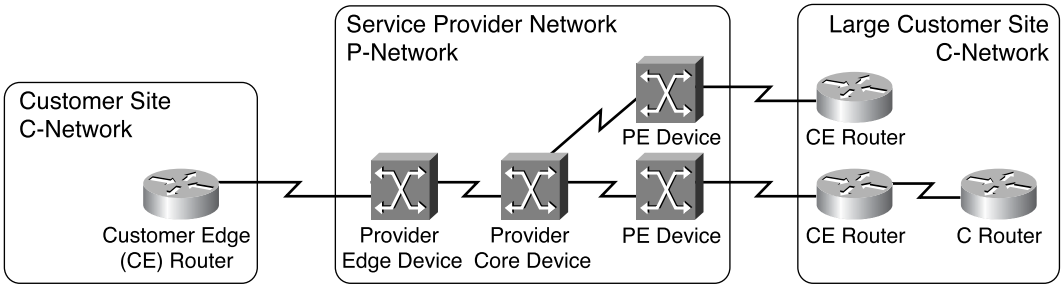
NOTE You can find more in-depth descriptions of these concepts and additional MPLS or VPN background information in Ivan Pepelnjak and Jim Guichard's *MPLS and VPN Architectures* (Volume I), published by Cisco Press, which is a prerequisite to understanding this book.

MPLS VPN Terminology

Since the early days of X.25 and Frame Relay (the two technologies initially used to deploy VPN services), many different technologies have been proposed as the basis to enable a VPN infrastructure. These ranged from Layer 2 technologies (X.25, Frame Relay, and Asynchronous Transfer Mode [ATM]) to Layer 3 technologies (primarily IP) or even Layer 7 technologies. IBM once had a product that transported IP datagrams over Systems Network Architecture (SNA) application sessions, and TGV (a company later acquired by Cisco Systems) had implemented IP transport over DECnet sessions. Not surprisingly, with such a variety of implementation proposals, the overall terminology in the field has changed dramatically. This book uses the terminology introduced with the MPLS-based VPN.

MPLS VPN-based terminology is based on a clear distinction between the service provider network (P-network) and the customer network (C-network), as shown in Figure 1-1.

Figure 1-1 *MPLS VPN-Based Terminology*



The P-network is always topologically contiguous, whereas the C-network is usually clearly delineated into a number of *sites* (contiguous parts of the customer network that are connected in some way other than through the VPN service). Note that a site does not need to be geographically contained; if the customer is using a VPN service for its international connectivity only, a site could span a whole country.

The devices that link the customer sites to the P-network are called *customer edge* (CE) *devices*, whereas the service provider devices to which the CE routers connect are called *provider edge* (PE) *devices*. In most cases, the P-network is made up of more than just the PE routers. These other devices are called P devices (or, if the P-network is implemented with Layer 3 technology, P routers). Similarly, the additional Layer 3 devices in the customer sites that have no direct connectivity to the P-network are called C routers.

VPN technologies have evolved into two major approaches toward implementing VPN services:

- **Connection-oriented VPN**—The PE devices provide virtual leased lines between the CE devices. These virtual leased lines are called *virtual circuits* (VCs). The VCs can be permanent, established out-of-band by the service provider network management team (called *permanent virtual circuits,* or *PVCs*). They can also be temporary, established on demand by the CE devices through a signaling protocol that the PE devices understand. (These VCs are called *switched virtual circuits,* or *SVCs*).
- **Connectionless VPN**—The PE devices participate in the connectionless data transport between CE devices. It is unnecessary for the service provider or the customer to establish VCs in these VPNs, except perhaps between the PE and CE routers if the service provider uses switched WAN as its access network technology.

Connection-Oriented VPNs

Connection-oriented VPNs were the first ones to be introduced. They offer a number of clear advantages, including the following:

- The service provider does not need to understand the customer's network; the service provider just provides virtual circuits between the customer sites.
- The service provider is not involved in the customer's routing (as shown in Figures 1-2 and 1-3), and it doesn't need to know which Layer 3 protocols the customer is deploying. Consider, for example, the network shown in Figure 1-2. The VPN network is implemented with Frame Relay VCs; therefore, the service provider is unaware of the routing protocols that the customer is using. From the customer's routing perspective, the customer routers are directly adjacent (linked with virtual point-to-point links), as shown in Figure 1-3.

Figure 1-2 *Connection-Oriented VPN: Physical Topology*

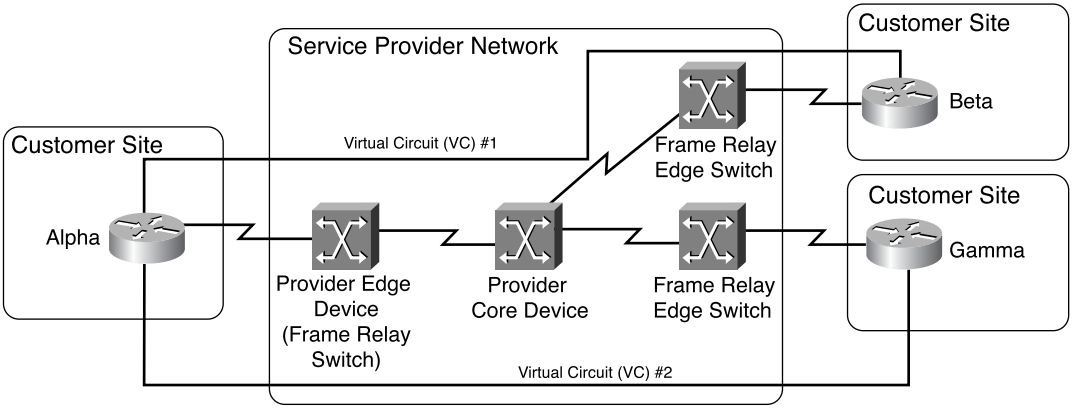
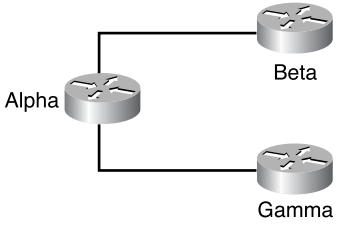


Figure 1-3 *Connection-Oriented VPN: Customer Routing Perspective*



Connection-oriented VPNs also have several obvious disadvantages:

- All VCs between the customer sites have to be provisioned, either manually by the service provider network management team or by the CE devices. Even if the VCs are established automatically by the CE devices, these devices need to be configured with enough information to establish the links through the signaling protocol of choice.
- The CE routers must exchange the routing information with other CE routers, resulting in more router adjacencies, slower convergence, and generally more complex routing setups.

NOTE If you are interested in more of the advantages and disadvantages of connection-oriented or connectionless VPNs, you can find them in Chapter 8, "Virtual Private Network (VPN) Implementation Options," of Jim Guichard and Ivan Pepelnjak's *MPLS and VPN Architectures* (Volume I), published by Cisco Press, 2002.

Modern connection-oriented VPNs are implemented with a variety of different technologies, including the following:

- They can be implemented with traditional connection-oriented Layer 2 technologies (X.25, Frame Relay, or ATM) or with connectionless Layer 2 technologies, such as virtual LANs (VLANs).
- They can also be implemented with *tunnels* that are established over public Layer 3 infrastructure (usually over public IP infrastructure—most commonly the Internet). These VPNs can use Layer 3 over Layer 3 tunnels, such as generic routing encapsulation (GRE), which is described in RFC 2784, or tunnels based on IP security (IPSec) technology. These VPNs can also use Layer 2 over Layer 3 tunnels, which are most commonly found in dial-up access networks to implement virtual private dialup networks (VPDNs).

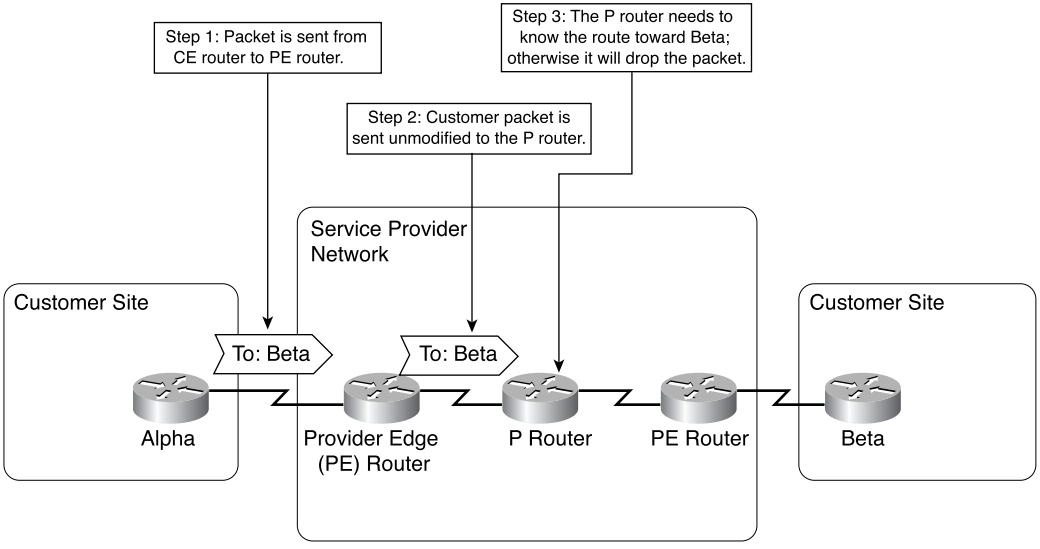
Connectionless VPNs

Contrary to connection-oriented VPNs, connectionless VPNs propagate individual datagrams that the CE devices send across the P-network. This approach, although highly scalable as proven by today's Internet, does impose a number of limitations on the customers:

- The customers can use only the Layer 3 protocol that the service provider supports. This was a serious drawback a few years ago, but it is quickly becoming a moot issue because most networking devices now support IPv4.

- The customers must use addresses coordinated with the service provider. In a connectionless network, every P device must be able to forward every individual datagram to its final destination; therefore, each datagram must have a unique destination address, known to every P device, as shown in Figure 1-4.

Figure 1-4 *Packet Propagation on Connectionless VPNs*



The simplicity of CE router configuration in a connectionless VPN world, as well as the capability to support IP-based VPN services together with public IP services on the common infrastructure, prompted many service providers to consider the rollout of connectionless VPN services. However, the acceptance of these services was initially quite low because the customers were unwilling to renumber their existing network infrastructure to comply with the service provider's addressing requirement. Clearly, a different VPN technology was needed that would combine the benefits of a connectionless VPN (simple CE router configuration and lack of explicit provisioning of the virtual circuits) with the benefits of a connection-oriented VPN (such as the support of overlapping address spaces and the simplicity of data forwarding in the P devices).

MPLS-Based VPNs

MPLS-based VPN technology uses a combination of connection-oriented and connectionless VPN technologies, including the following features:

- The interface between the CE routers and the PE routers is connectionless. No additional configuration is needed on the CE devices.

- The PE routers use a modified IP forwarding paradigm; a distinct IP routing and forwarding table (called *virtual routing and forwarding* table, or *VRF*) is created for each customer.
- The customer's addresses are extended with 64-bit *route distinguishers* to make nonunique 32-bit IP addresses globally unique within the service providers' backbone. The resulting 96-bit addresses are called *VPNv4* addresses.
- A single routing protocol is run between the PE routers for all VPN customers. Modified *Border Gateway Protocol* (BGP) with multiprotocol extensions is used in this function.
- The PE routers use MPLS-based VCs (called *label-switched paths,* or *LSPs*) to transport the customer's datagrams between PE routers. Additional MPLS labels are inserted in front of the customer's IP datagrams to ensure their proper forwarding from ingress PE routers toward the destination CE router.
- The LSPs between all PE routers are established automatically based on the IP topology of the P-network. It is unnecessary to configure or manually establish these paths.
- The mapping between the customer's destination addresses and LSPs leading toward the egress PE routers is performed automatically based on the BGP next-hops.

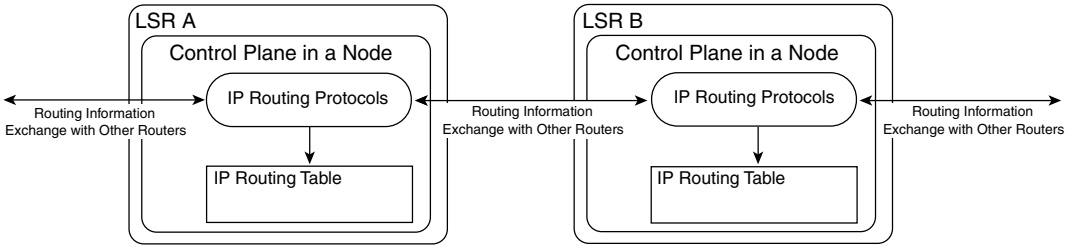
The following sections will briefly refresh your MPLS and MPLS VPN knowledge. For more in-depth discussion of the MPLS and MPLS VPN technology, please refer to Cisco Press's *MPLS and VPN Architectures* (Volume I). For more details on ATM-based MPLS implementations, refer to *Advanced MPLS Design and Implementation*, published by Cisco Press.

The MPLS Technology

In essence, the MPLS technology combines the richness of IP routing and the simplicity of hop-by-hop label switching of Frame Relay or ATM to provide the seamless integration of the connection-oriented forwarding with the IP world. Due to their dual nature (they operate on both the IP layer as well as the label-switching layer), the MPLS devices are called *label switch routers* (LSRs). This section describes the typical operation of MPLS devices, focusing on the simplest MPLS application: forwarding of IP datagrams across an MPLS network.

All devices in an MPLS network run IP routing protocols on their *control plane* to build IP routing tables. In MPLS devices that support IP forwarding, the IP routing tables are used to build IP forwarding tables, also called *forwarding information base* (FIB). In MPLS devices that support only label forwarding (such as the ATM switches with MPLS functionality), the IP routing FIB does not exist. The IP routing operation of the MPLS control plane is shown in Figure 1-5.

Figure 1-5 *LSRs Build the IP Routing Table*

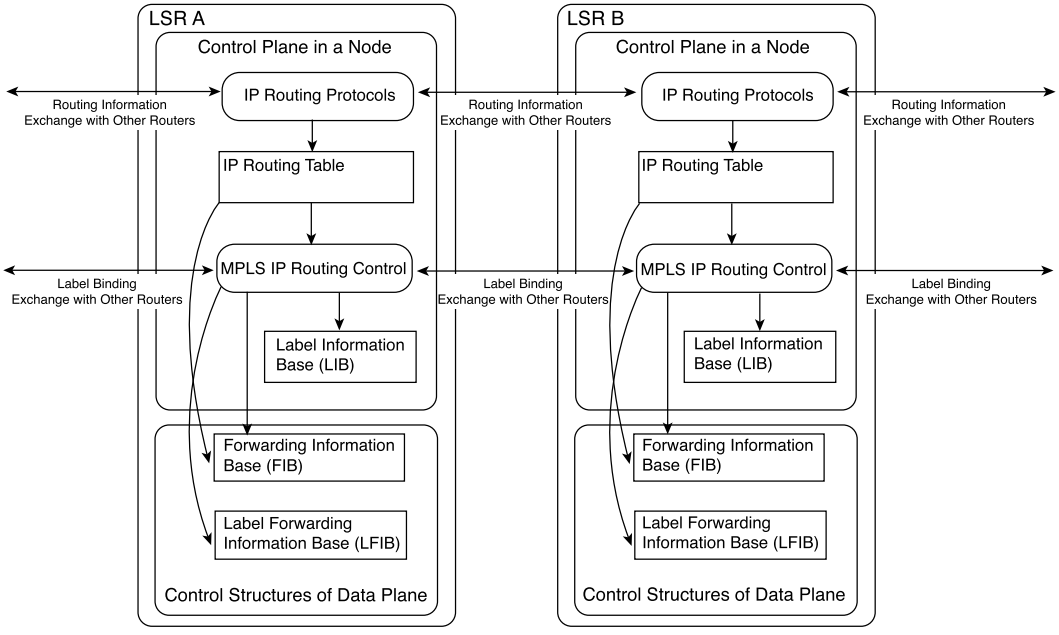


After the IP routing tables have been built, MPLS labels are assigned to individual entries in the IP routing table (individual IP prefixes) and propagated to adjacent MPLS devices through a *Label Distribution Protocol* (LDP).

NOTE In usual MPLS operation, labels are not assigned to BGP destinations because the router always reaches BGP destinations through recursive lookup on BGP next-hop. Therefore, BGP destinations can be reached through the label that is associated with the BGP next-hop for those destinations.

Each MPLS device uses its own local label space; globally unique labels or centralized label assignment is unnecessary, making MPLS extremely robust and scalable. Every label assigned by an MPLS device is entered as an input label in its *label forwarding information base* (LFIB), which is the forwarding table used for label switching. The label assignment and distribution of an MPLS device are illustrated in Figure 1-6.

Figure 1-6 *Control Plane Operations in an LSR*

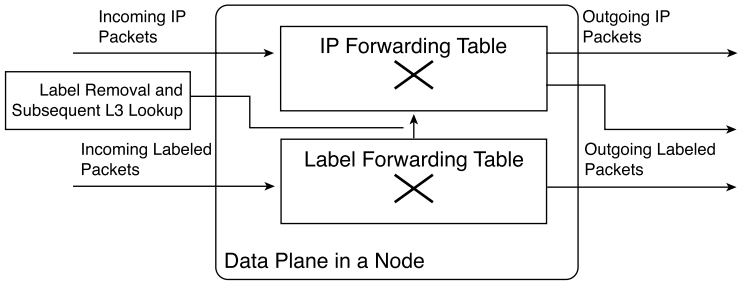


Most label assignments, both local as well as those made by adjacent devices, are entered into a table called the *label information base* (LIB). The label that the IP next-hop assigns for a particular IP prefix is entered as an output label in the local LFIB to enable pure label forwarding. In devices that support IP forwarding, such a label is also entered into the FIB to support IP-to-label forwarding.

After the IP routing tables, IP forwarding tables, and label forwarding tables have been built, the MPLS devices can start to forward IP traffic. All MPLS devices must support label forwarding; whenever they receive a labeled packet, they perform a label lookup in the LFIB, replace the input label with the output label, and forward the labeled packet to the next-hop LSR. Some MPLS devices (ingress LSRs) can receive IP datagrams, perform a lookup in the FIB, insert an MPLS label stack in front of the IP datagram based on information stored in the FIB, and forward the labeled packet to the next-hop LSR. The PE router within the MPLS VPN architecture is an example of such a device.

Other MPLS devices (egress LSR) can receive labeled packets, perform an LFIB lookup, and (based on the absence of an output label in the LFIB) remove the label from the ingress labeled datagram and forward the IP datagram to the next-hop IP router. In most cases, all LSRs in an MPLS network can act as both ingress and egress LSRs, the notable exception being ATM switches acting as LSRs. The various paths that an IP datagram or a labeled datagram can take through an LSR are displayed in Figure 1-7.

Figure 1-7 *Packet Forwarding in an LSR*



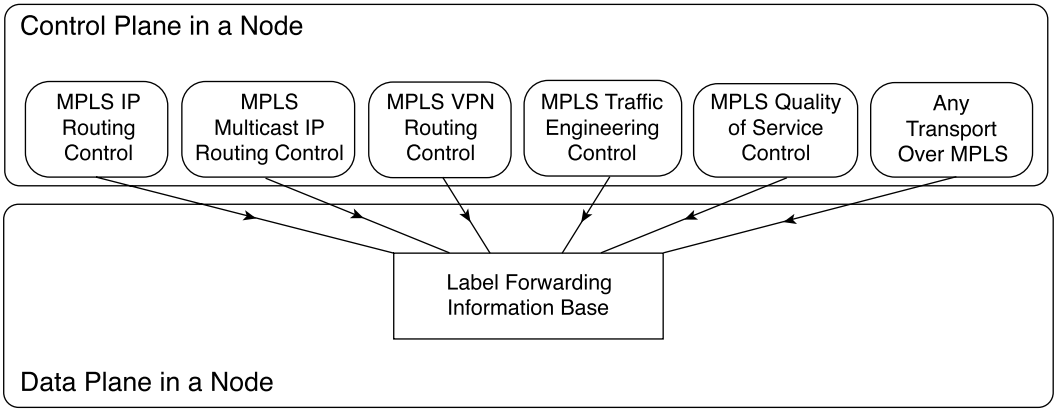
The basic principle of MPLS has been extended to a variety of other applications, including these:

- **MPLS traffic engineering (TE)**—The modified link-state routing protocols (OSPF and ISIS) are used to discover free resources in the network, labels are assigned through the *Resource Reservation Protocol* (RSVP), and the global FIB is modified based on MPLS TE labels.
- **MPLS VPNs**—Many FIBs are created (one or more per VPN customer), and Multiprotocol BGP is used to distribute the customer routing information and MPLS labels across the network.
- **MPLS quality of service (QoS) in ATM environments**—The standard LDP is modified to assign up to four labels for each IP prefix, with each label serving a different QoS class.

New MPLS applications are constantly emerging. For example, one of the new MPLS applications (also covered in this book) enables IPv6 transport across an MPLS network; IPv6 routing protocols are used to build IPv6 routing tables, which are then used as the basis for label assignment and distribution.

The large variety of different MPLS applications still adhere to the common framework. Each application might have its own "routing protocol," its own LDP, and its own forwarding database. However, the MPLS applications all share a common LFIB, enabling the LSRs to transparently integrate new MPLS applications without affecting the existing services, as shown in Figure 1-8.

Figure 1-8 *Multiple MPLS Applications in a Single LSR*



The MPLS VPN Technology

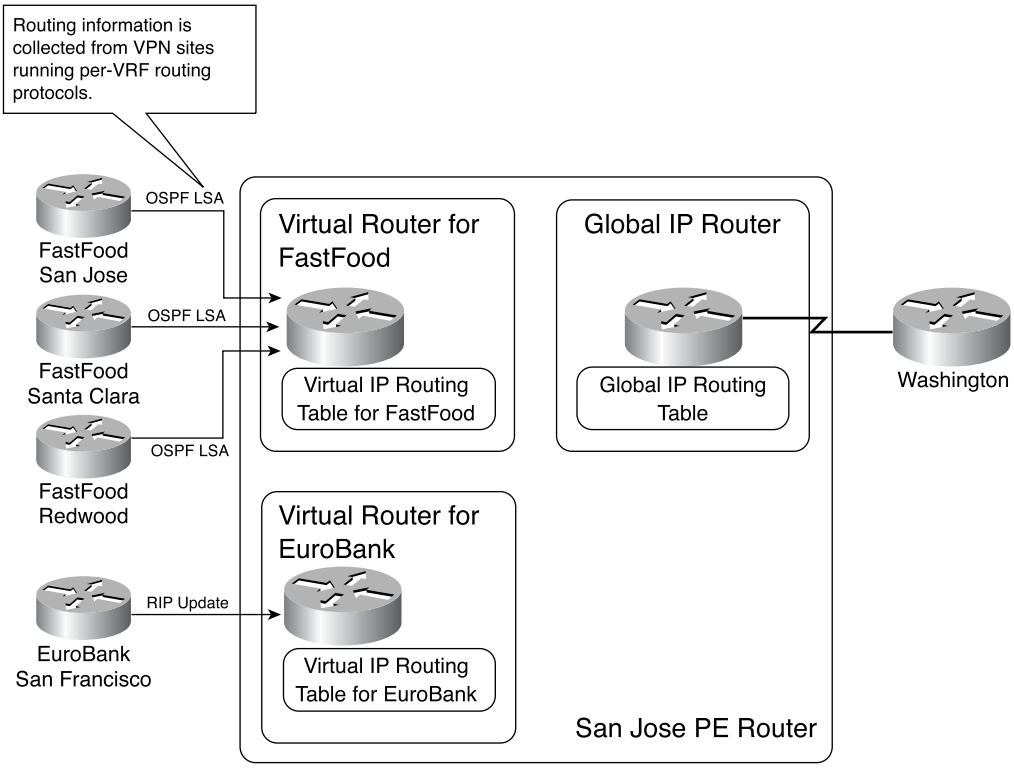
As discussed previously, MPLS-based VPNs use a combination of connectionless VPNs between the customers and service providers (thus minimizing the provisioning complexity and cost) with connection-oriented VPNs in the network core (reducing the overhead on the P devices). Furthermore, several additional mechanisms have been implemented to allow the customers to use overlapping address spaces.

In a typical MPLS-VPN network, the CE routers and PE routers exchange the customer routes using any suitable IP routing protocol. These routes are inserted into VRFs on the PE routers, which guarantees the perfect isolation between customers. This process is illustrated in Figure 1-9, which details the internal structure of a PE router (San Jose) to which two VPN customers are connected (FastFood and EuroBank) and which also connects to a P router (Washington).

When customer routes are placed into VRFs, the PE routers allocate a separate MPLS label that will be needed for VPN data forwarding to each customer route. The customer routes and associated MPLS labels are transported across the P-network using multiprotocol BGP. The customer IP addresses are augmented with a 64-bit route distinguisher before being inserted into the provider's BGP to ensure global uniqueness of potentially nonunique customer addresses. Additional BGP attributes (extended BGP communities) are used to control the exchange of routes between VRFs to allow the service providers to build VPN topologies that are almost impossible to build with any other VPN technology.

NOTE You can find detailed descriptions of these topologies and implementation guidelines in the *MPLS and VPN Architectures* (Volume I) book.

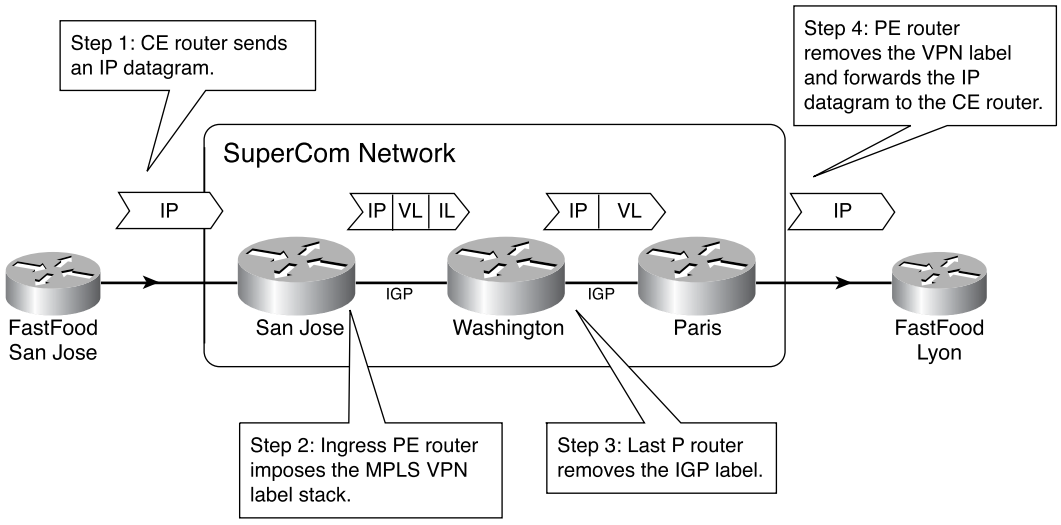
Figure 1-9 *Virtual Routing Tables in a PE Router*



The extended BGP communities are also used to implement additional MPLS VPN features, including *automatic route filtering* with the site-of-origin (SOO) community or automatic propagation of Open Shortest Path First (OSPF) route attributes across the BGP backbone. (OSPF support is described in more detail in Chapter 3, "PE-CE Routing Protocol Enhancements and Advanced Features.")

VPN packet forwarding across the MPLS VPN backbone is implemented with MPLS forwarding using an MPLS label stack imposed in the IP datagram by the ingress PE router. The first label in the stack is the label assigned to the IP address of the egress PE router (BGP next-hop) in the service provider core. The second label is the label assigned to the customer route by the egress PE router. The first label is usually removed one hop before the egress PE router through a process called *penultimate hop popping*. The egress PE router then performs label lookup on the VPN label, removes the VPN label, and forwards the packet to the CE router. The whole process is illustrated in Figure 1-10.

Figure 1-10 *VPN Packet Propagation in an MPLS VPN Network*



An IP datagram, sent from San Jose to Lyon, is forwarded across the service provider backbone in a number of steps:

1. An IP datagram is sent from the CE router to the PE router.
2. The PE router performs an IP lookup and prepends an MPLS header consisting of two labels: a label assigned via LDP (also known as IGP label, or IL), identifying the path toward the egress PE router (Paris); and a VPN label (VL) assigned by the Paris PE router.
3. The penultimate router in the service provider network removes the IGP label, leaving only the VPN label in the MPLS header.
4. The egress PE router performs label lookup on the VPN label, removes the MPLS header, and forwards the IP datagram to the Lyon CE router.

New MPLS VPN Developments

Many service providers worldwide have enthusiastically embraced the MPLS and MPLS VPN technologies as they enable the service providers to deploy the two most common applications—Internet access and VPN services—on a common network infrastructure. The diversity of their infrastructures, access layer technologies, and IP routing setups, as well as the new services these service providers would like to deploy, have triggered the development of several new MPLS-related features, including these:

- Tight integration of access technologies such as dial-up, digital subscriber line (DSL), and cable with MPLS VPN

- New routing protocol options and support for additional VPN routing protocols
- Transport of additional Layer 3 protocols over MPLS

Each of these is discussed in the following sections.

Access Technology Integration with MPLS VPN

The initial implementation of MPLS VPN technology supported customer sites that were connected primarily to the service provider backbone through a permanent connection. These connections were implemented with Layer 2 technology, which was well established in the IOS code base. Although you could, with skill, support other access technologies (most notably, dial-up users), a number of supporting technologies were not MPLS VPN-enabled, forcing the service providers to accept compromises they would rather avoid.

Tighter integration of MPLS VPN with access technologies was implemented by making several additional Cisco IOS services VPN-aware:

- Virtual-Profile Cisco Express Forwarding (CEF)
- Overlapping address pools
- On-demand address pools (ODAP)
- Framed Route VRF Aware
- Per VRF authentication, authorization, and accounting (AAA)
- VRF-aware large-scale dial out (LSDO)
- VPN-ID
- DHCP relay—MPLS VPN support

All these features and the access technology integration with MPLS VPN is described in detail Chapter 2, "Remote Access to an MPLS VPN."

New Routing Protocol Options

New Cisco IOS releases extend the range of IP routing protocols that are supported between the PE routers and the CE routers. Enhanced IGRP (EIGRP) and Integrated Intermediate System-to-Intermediate System (Integrated IS-IS) are supported, as well as additional OSPF connectivity options, including virtual OSPF links between PE routers (*sham links*). Furthermore, Cisco IOS supports IP Multicast inside the MPLS VPN and per-VRF network address translation (NAT) on the PE router. These new features are described in Chapters 3, "PE-CE Routing Protocol Enhancements and Advanced Features," 4, "Virtual Router Connectivity," and 7, "Multicast VPN."

New Layer-3 Protocols Transported Over MPLS

IP version 6 (IPv6), also known as IP: The Next Generation (IPng), has joined IPv4 as another Layer 3 protocol that can be transported across an MPLS backbone. MPLS support for globally routed IPv6 is described in Chapter 8, "IPv6 Across an MPLS Backbone."

Summary

Many service providers that wanted to minimize their costs of provisioning and operations by offering all their services (VPN and public Internet) over a common infrastructure have enthusiastically embraced MPLS-based VPN networks. Furthermore, these service providers have achieved significant cost savings due to the provisioning simplicity offered by MPLS VPN's integration with the benefits of both connectionless and connection-oriented VPN approaches.

An end-to-end MPLS VPN solution is, like any other VPN solution, divided into the central P-network to which a large number of customer sites (sites in the C-network) are attached. The customer sites are attached to the PE devices (PE routers) through CE devices (CE routers). Each PE router contains several virtual routing and forwarding tables (VRFs)—at least one per VPN customer. These tables are used together with Multiprotocol BGP run between the PE routers to exchange customer routes and to propagate customer datagrams across the MPLS VPN network. The PE routers perform the label imposition (ingress PE router) and removal (egress PE router). The central devices in the MPLS VPN network (P routers) perform simple label switching.

MPLS-based VPNs have been significantly enhanced since their initial rollout. The new MPLS VPN features allow better integration of access technologies, support of additional PE-CE routing protocols, as well as support of new transport options across MPLS backbones (transport of IPv6 and legacy Layer 2 technologies).

This page intentionally left blank

PART II

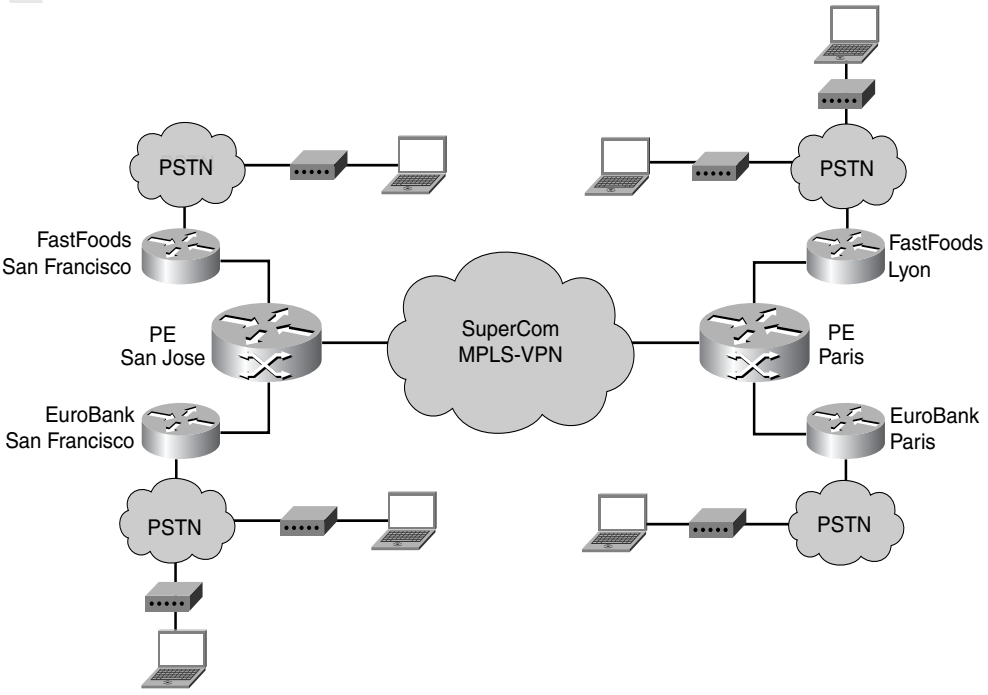
Advanced PE-CE Connectivity

CHAPTER 2

Remote Access to an MPLS VPN

The initial service offerings for Multiprotocol Label Switching (MPLS) virtual private networks (VPNs) were provided to customers through fixed connections to the provider edge (PE) router by using technologies such as leased line, Frame Relay, Asynchronous Transfer Mode (ATM) permanent virtual circuits (PVCs), or last mile Ethernet. The provision of remote or off-net access to the MPLS VPN was incumbent upon the customer having the appropriate access infrastructure in place to cater to his mobile or remote workforce. Therefore, the ability for an MPLS VPN service provider to supply MPLS VPN value-added services (which, in turn, generates more revenue) to remote users was completely dependent on the customer's remote access network and the geographic coverage that the network provided. This is illustrated in Figure 2-1.

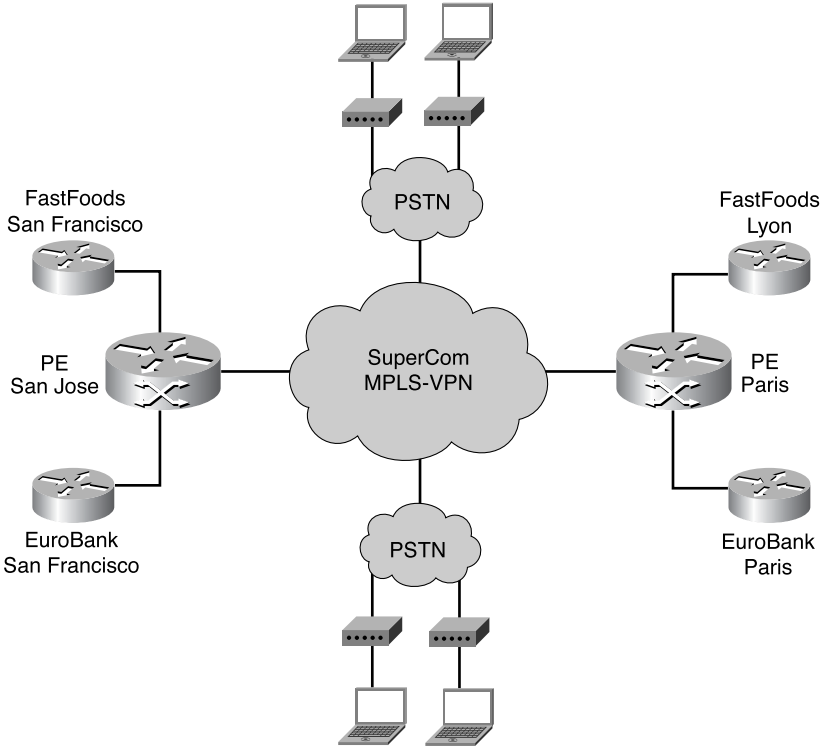
Figure 2-1 *Remote Access Provided by Customer*



In this scenario, the SuperCom network provides only fixed-line access to the EuroBank and FastFoods customer edge (CE) routers. Remote access is provided by using EuroBank and FastFoods hardware at their remote locations.

To provide a scalable and complete end-to-end VPN service, the service provider must have a network infrastructure that is capable of integrating remote access directly into an MPLS VPN network. Such an infrastructure can enable remote users to seamlessly access their corporate VPNs through a service provider point of presence (POP), not a customer POP. The advantage of this is that a service provider can offer a value-add service by leasing wholesale dial access to many VPN customers. The VPN customers can be ISPs or large enterprises that want to provide access to remote users but avoid the need for maintaining their own separate and expensive access network. The same service provider remote access network can be sold as a unique service to many VPN customers (build once, sell many), which decreases the customer's operating costs and increases the revenue of the service provider. This is illustrated in Figure 2-2.

Figure 2-2 *Remote Access Provided by a Service Provider*



In this scenario, SuperCom provides remote access services terminating into the MPLS VPN network. This remote access network allows any EuroBank or FastFoods remote user direct access to his VPNs, which alleviates the need for EuroBank and FastFoods to provide a separate remote access infrastructure.

Service providers will invariably use one or more of the following access technologies to provide remote access to an MPLS VPN:

- Public Switched Telephone Network (PSTN)
- Integrated Services Digital Network (ISDN)
- Asymmetric digital subscriber line (ADSL)
- Data-over Cable Service Interface Specifications (DOCSIS), or simply called cable

These access technologies are used in conjunction with various protocols and procedures to provide the remote access service. The protocols and procedures include the following:

- Point-to-Point Protocol (PPP)
- Layer 2 Tunneling Protocol (L2TP)
- Virtual private dialup network (VPDN)
- Remote Authentication Dial-In User Service (RADIUS)
- Dynamic Host Configuration Protocol (DHCP)

The first part of this chapter provides an overview of each of these protocols and procedures to provide you with a foundation for understanding how remote access is provided to an MPLS VPN. The second part of this chapter covers the following remote access scenarios and features:

- Dial-in access to an MPLS VPN via VPDN (L2TP) or direct ISDN
- Large-scale dial-out access from an MPLS VPN via L2TP or direct ISDN
- Dial backup to an MPLS VPN
- Digital subscriber line (DSL) access to an MPLS VPN by using various encapsulation methods
- Cable access to an MPLS VPN
- Advanced features, such as on-demand address pools, per-VRF AAA, and VRF-aware DHCP relay

Feature Enhancements for MPLS VPN Remote Access

Several new features and enhancements were made to Cisco IOS so that MPLS VPN services could be provisioned over various remote access technologies. Most of these features are incorporated into the detailed examples provided throughout this chapter or are

addressed in the later section, "Advanced Features for MPLS VPN Remote Access." The features can be summarized as follows:

- **Virtual-profile Cisco Express Forwarding (CEF)**—PPP sessions that terminate on a Cisco router through an L2TP tunnel or direct ISDN interface do so via a virtual-access interface. The virtual-access interface is an instance of a virtual-profile or a virtual-template. Each system has a maximum of 25 virtual-templates; virtual-profiles do not have this limitation; therefore, they are preferred because they are more scalable and flexible. The virtual-profile CEF feature allows these interfaces to be CEF switched, which is a prerequisite for MPLS.
- **Overlapping address pools**—Previously, per-router local address pools could only be specified in the global IP routing instance. This meant that all VRFs as well as all global interfaces shared a single local pool to provide interface addresses for PPP sessions. The overlapping pool feature allows the same IP address range to be used concurrently in different VRFs, thereby providing better utilization of the IP address space.
- **On-demand address pools (ODAP)**—Instead of configuring pool address ranges locally, the ODAP feature allows a central RADIUS server to provide VRF-aware pool addresses as required. In this way, the local pool can expand and contract based on usage, and the RADIUS server can provide better address management by allocating subnets where they are needed.
- **Framed Route VRF aware**—When a remote CE router dials into a PE router via a PPP session, there must be a mechanism to allow the remote subnet to be injected into the VRF for the duration of the call. This is done through the Framed-Route RADIUS attribute or the corresponding cisco-avpair "ip:route" attribute. This attribute usually applies to the global routing table; however, enhancements have been made so that Cisco IOS can determine whether it should be applied to a VRF.
- **Per VRF authentication, authorization, and accounting (AAA)**—This feature allows RADIUS information to be sent directly to a customer RADIUS server that is located within the VRF. Previously, the only way to get to a customer RADIUS server was to use a proxy via the service provider RADIUS server reachable in the global routing table.
- **VRF-aware large-scale dial out (LSDO)**—This feature allows the LSDO solution to operate within the context of a VRF. VRF-aware LSDO allows multiple VRFs to use the same dialer interface on a router with individual profiles downloaded from an AAA server.
- **VPN-ID**—This feature allows remote access applications such as a RADIUS or DHCP server to identify the VPN that originates a RADIUS or DHCP request. The VPN-ID feature is based on RFC 2685.
- **DHCP Relay—MPLS VPN Support**—This feature allows a single DHCP server to identify and service many VRFs by supplying addresses from distinct IP address pools. Creating different namespaces within the server separates address pools. Either the VRF name or the VPN ID identifies these namespaces. The DHCP server can reside in the global routing table or in any customer or shared services VRF.

Overview of Access Protocols and Procedures

This section briefly describes the typical protocols that are used in remote access technologies. It serves as a refresher or an introduction to those of you who are not intimately familiar with these protocols. For a more in-depth description of remote access protocols and Cisco IOS configuration guidelines, please refer to Cisco Connect Online (www.cisco.com) under the Technologies section.

PPP

PPP is fundamental to the deployment of nearly all the remote access scenarios discussed in this chapter. PPP provides a link layer service (Layer 2 of the OSI model) between two devices (in this case, the customer device and the PE router), and it can operate over a variety of physical media such as ISDN, ADSL, leased line, and virtual circuits such as ATM PVCs and L2TP tunnels. PPP provides a datagram service only; reliable transport is the responsibility of the higher layers in the protocol stack. The connection that PPP operates over can be either fixed or switched (dial-up) and running in asynchronous or synchronous bit serial mode. The only requirement for PPP is that the circuit provided be full duplex. An advantage of PPP is that it can support many different network protocols (Layer 3 of the OSI hierarchy), such as IP, DECnet, AppleTalk, and OSI simultaneously over the same link.

PPP is a layered protocol that has three components:

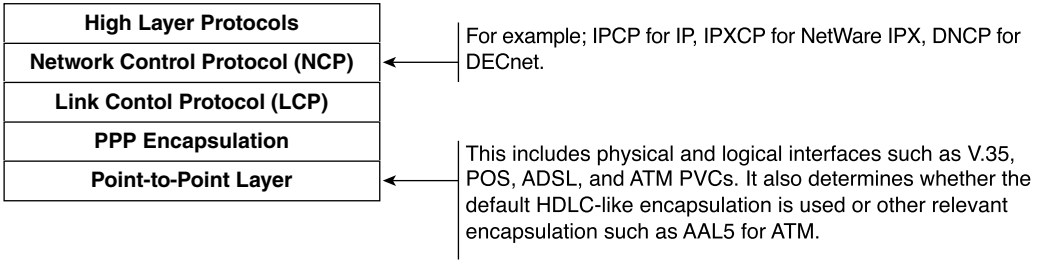
- An encapsulation component that is used to transmit datagrams over the specified physical layer.
- A Link Control Protocol (LCP) to establish, configure, and test the link as well as negotiate capabilities.
- One or more NCPs used to negotiate optional configuration parameters and facilities for the network layer. There is one Network Control Protocol (NCP) for each protocol supported by PPP.

NOTE The device that terminates PPP sessions in a service provider network is called a *network access server (NAS).* A NAS is capable of terminating many connections over a variety of physical media. Among other examples, a NAS could be a Cisco Systems 7200 acting as a PE router with switched ISDN connections or a Cisco Systems AS5300 universal access concentrator terminating dial-in ISDN or analog modem calls.

To establish a link for point-to-point communication, each endpoint uses LCP to open the connection, negotiate capabilities, and configure the link appropriately. Examples of capabilities that can be negotiated are the maximum receive unit (MRU), compression of certain PPP fields, and Password Authentication Protocol (PAP) or Challenge Handshake Authentication Protocol (CHAP).

Optionally, you can assess the link quality to determine whether the network protocols can be activated. If the link quality is not of acceptable quality, then LCP can hold off passing to the NCP phase. When the LCP phase is completed, the relevant NCP for that protocol must separately negotiate each network layer protocol. For example, the NCP for IP called Internet Protocol Control Protocol (IPCP) can negotiate options such as IP addresses to be used at each end of the link, DNS server addresses, and the compression protocol. LCP and NCP are both extensible protocols; therefore, new features and options can be easily added when required. Figure 2-3 shows where LCP and NCP fit in the PPP model.

Figure 2-3 *PPP Model*



The LCP layer also provides the optional authentication function, which is a fundamental requirement when providing remote access services. Authentication takes place after the link has been established and prior to the NCP negotiation phase.

As previously mentioned, LCP has two authentication protocols available: PAP and CHAP. PAP is a simple two-way handshake protocol. The username/password is repeatedly sent across the link from the originating end until an acknowledgement is received. PAP sends passwords in clear text; there is no protection from playback or trial and error attacks (such as trying to guess passwords from the outside).

CHAP is a more robust authentication protocol that uses a three-way handshake to verify the identity of the remote end. The authentication is done initially when the link is established and might be periodically repeated. CHAP is the preferred authentication method and will be used in examples throughout this chapter. The three-way handshake operates as follows:

- The local peer sends a challenge message to the remote peer
- The remote peer combines the challenge with a shared secret key and responds with a value calculated by using a one-way hash function (such as a message-digest algorithm MD5).
- The local peer then compares the returned hash value with what it expected to receive. (It calculates its own value by using the hash function.)
- If the hash values match, the authentication is acknowledged; otherwise, the connection is terminated.

NOTE The password, or "secret key" as it is referred to, is never sent across the link. Only the hashed response of the secret is transmitted. Because CHAP can be used to authenticate many different remote systems, the challenge/response packet can also contain a name (usually the hostname) that will be used to index a list of secret keys or passwords.

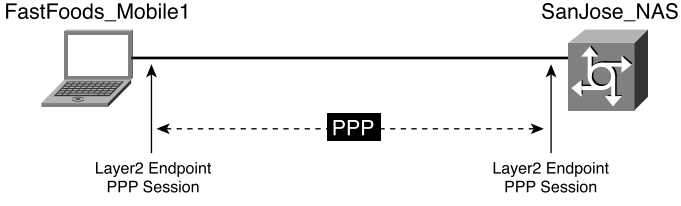
Figure 2-4 illustrates CHAP in operation. A remote FastFoods user has dialed into the San Jose NAS. SanJose_NAS will send a challenge message to the FastFoods_Mobile1 PC asking for its secret. FastFoods_Mobile1 will use information in the challenge message as well as the secret that is locally stored to send a response back. The response message will contain the name of the FastFoods remote user (FastFoods_Mobile1) as well as the encrypted secret (*whatsthebuzz*). The SanJose_NAS will then compare the response received from FastFoods_Mobile1 with the name/secret pair stored either locally on the NAS server or on a RADIUS/AAA server. If the encrypted versions of the secrets match, then an accept message is sent back and the NCP layer can proceed. This handshake can be periodically repeated during the call.

Figure 2-4 *CHAP Three-Way Handshake*

L2TP

In a typical PPP connection, the Layer 2 termination point and the PPP session endpoint reside on the same physical device. For example, a user could obtain a connection to the NAS by way of an analog dial-up or ISDN connection and then run PPP over that connection. In this case, the Layer 2 and PPP session would terminate on the NAS as shown in Figure 2-5.

Figure 2-5 *PPP Endpoints*

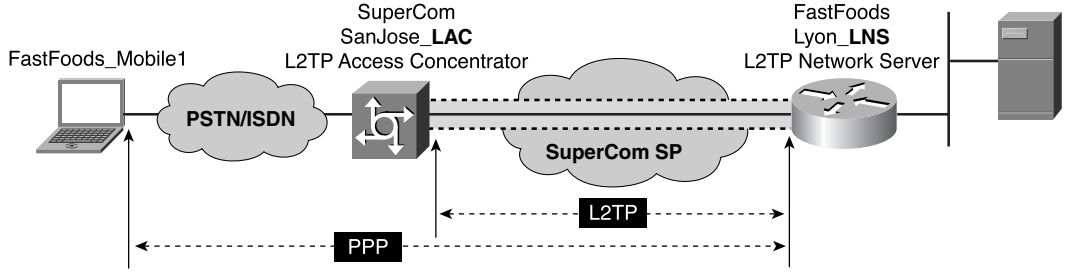


L2TP allows the PPP session endpoint to be divorced from the Layer 2 termination point. This means that a PPP session can be extended across the Internet or an ISP network. While traversing an IP backbone, the PPP session is carried inside an L2TP tunnel. The PPP session can pass through many intermediate nodes before terminating on the target remote access server. L2TP allows the remote client to communicate with the remote server by using PPP as if the two were directly connected. The network infrastructure is transparent to either end of the PPP session. The device that terminates the Layer 2 connection and originates the L2TP tunnel is called the *L2TP Access Concentrator (LAC)*. The device that terminates the L2TP tunnel and the original PPP session from the remote client is called the *L2TP Network Server (LNS)*. The LAC passes packets between the remote client and the LNS.

NOTE L2TP allows the creation of a virtual private dialup network (VPDN) to connect a remote client to its corporate network by using a shared infrastructure, which could be the Internet or a service provider's network. VPDNs are described in the following section.

Figure 2-6 illustrates the basic concept of an L2TP tunnel.

Figure 2-6 *PPP Session Through an L2TP Connection*



In this scenario, FastFoods has a remote client called FastFoods_Mobile1 that needs to communicate directly with a server that is located at the FastFoods Lyon site. The nearest dial-in POP to the FastFoods mobile user is provided by SuperCom in San Jose. The Lyon server is reachable through a FastFoods router that is connected directly to the SuperCom network in Paris. Therefore, when FastFoods_Mobile1 calls into the SuperCom LAC in San Jose, the San Jose LAC will exchange PPP messages with FastFoods_Mobile1 and communicate by way of L2TP requests and responses with FastFood's Lyon_LNS to set up an L2TP tunnel. The PPP session will be established between FastFoods_Mobile1 and the Lyon_LNS.

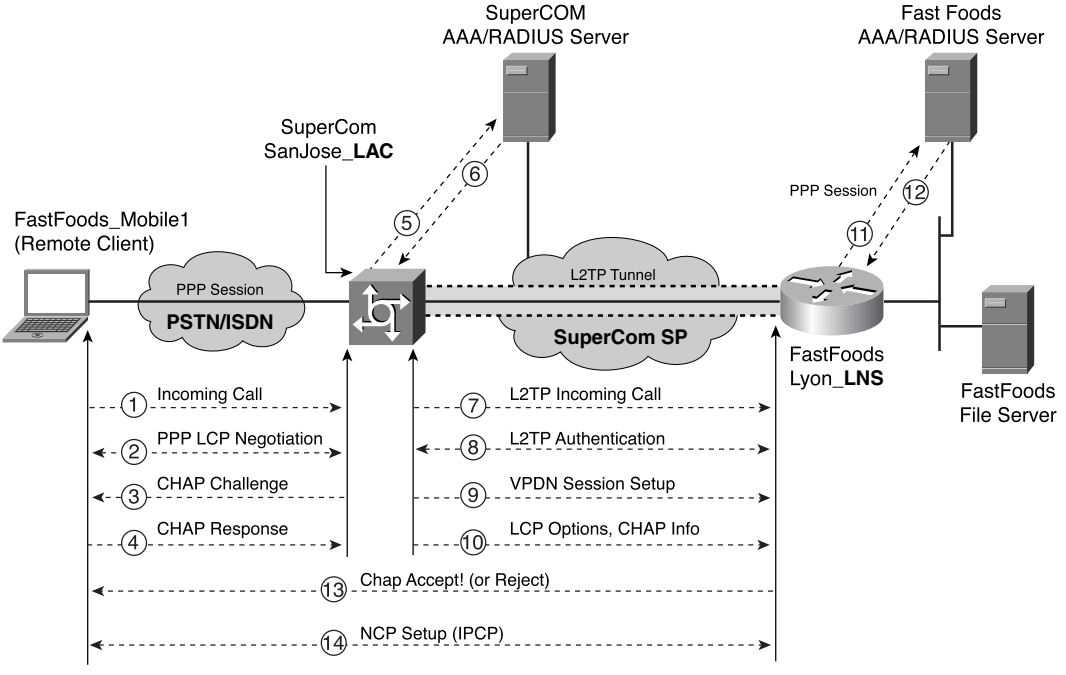
PPP frames from FastFoods_Mobile1 will be accepted by the SanJose_LAC, stripped of any linked framing or transparency bytes, encapsulated in L2TP, and forwarded over the appropriate tunnel toward Lyon_LNS. The LNS will accept these L2TP frames, strip the L2TP encapsulation, and process the incoming PPP frames.

VPDN

A *VPDN* is a network that connects a remote access client to a private network by using a shared or public IP infrastructure. A VPDN uses a tunnel protocol, such as L2TP, Point-to-Point Tunneling Protocol (PPTP), or Layer 2 Forwarding (L2F) to extend the Layer 2 and higher parts of the network connection from a remote user across an ISP network to a private network. VPDNs allow a service provider to share its common remote access infrastructure among many remote clients. Each client can dial in to a service provider NAS/LAC and be connected to the private corporate network based on the logon domain name or the number that was dialed (by using the dialed number identification service, or DNIS).

Figure 2-7 describes the VPDN process. It is essentially the same scenario as described in Figure 2-6, except that the protocol exchanges are fully detailed. It uses a combination of PPP, L2TP, and RADIUS to provide the virtual private dial-in service.

Figure 2-7 *VPDN Process*



The following steps outline what happens during the VPDN process:

Step 1 The FastFoods remote client initiates a PPP call to the SuperCom San Jose LAC via PSTN or ISDN.

Step 2 The remote client and the LAC begin to negotiate PPP options by using LCP. This covers elements such as the authentication method (CHAP or PAP), compression, and the PPP multilink.

Step 3 Assuming that CHAP was selected, the LAC sends a challenge message.

Step 4 The FastFoods remote client responds with its username (assume it is mobile1@fastfoods.com) and password. The LAC partially authenticates the user by using the information it has received in the CHAP response.

Step 5 The LAC checks whether the FastFoods remote client is a VPDN user. It determines this by examining the username (mobile1), domain name (fastfoods.com), or called number (DNIS). This information can either be stored locally (configured statically) on the LAC or it can be retrieved from the SuperCom RADIUS server. In our example, the information is forwarded via a RADIUS request to the SuperCom RADIUS server.

Step 6 The RADIUS server has an entry for the domain name of the FastFoods remote client; therefore, the client is a VPDN user. The RADIUS server replies to the LAC with a message containing the IP address of the FastFoods LNS and other information to allow the LAC to create an L2TP tunnel to the specific LNS.

NOTE If the remote client were determined not to be a VPDN client, then authentication would continue on the LAC. In this case, it would be likely that this customer would be subscribing to Internet access or some other SuperCom common service and would be connected directly to the global routing space of SuperCom.

Step 7 If the L2TP tunnel does not already exist, the SanJose_LAC builds a tunnel to the FastFoods Lyon_LNS by using L2TP control messages. Only one tunnel is built for each domain. For example, all fastfoods.com that subsequently dial in use the same tunnel.

Step 8 L2TP provides an optional CHAP-like authentication mechanism during tunnel establishment. The LNS can check to see if the LAC can open a tunnel (via local configuration) to it and both the LAC and LNS can authenticate each other using a shared secret configured locally or on a RADIUS server. Alternatively, the LNS can accept the tunnel without any authentication.

Step 9 After the tunnel is created, a VPDN session is created over the L2TP tunnel for the FastFoods remote client. Each remote client is associated with a unique VPDN session on an L2TP tunnel.