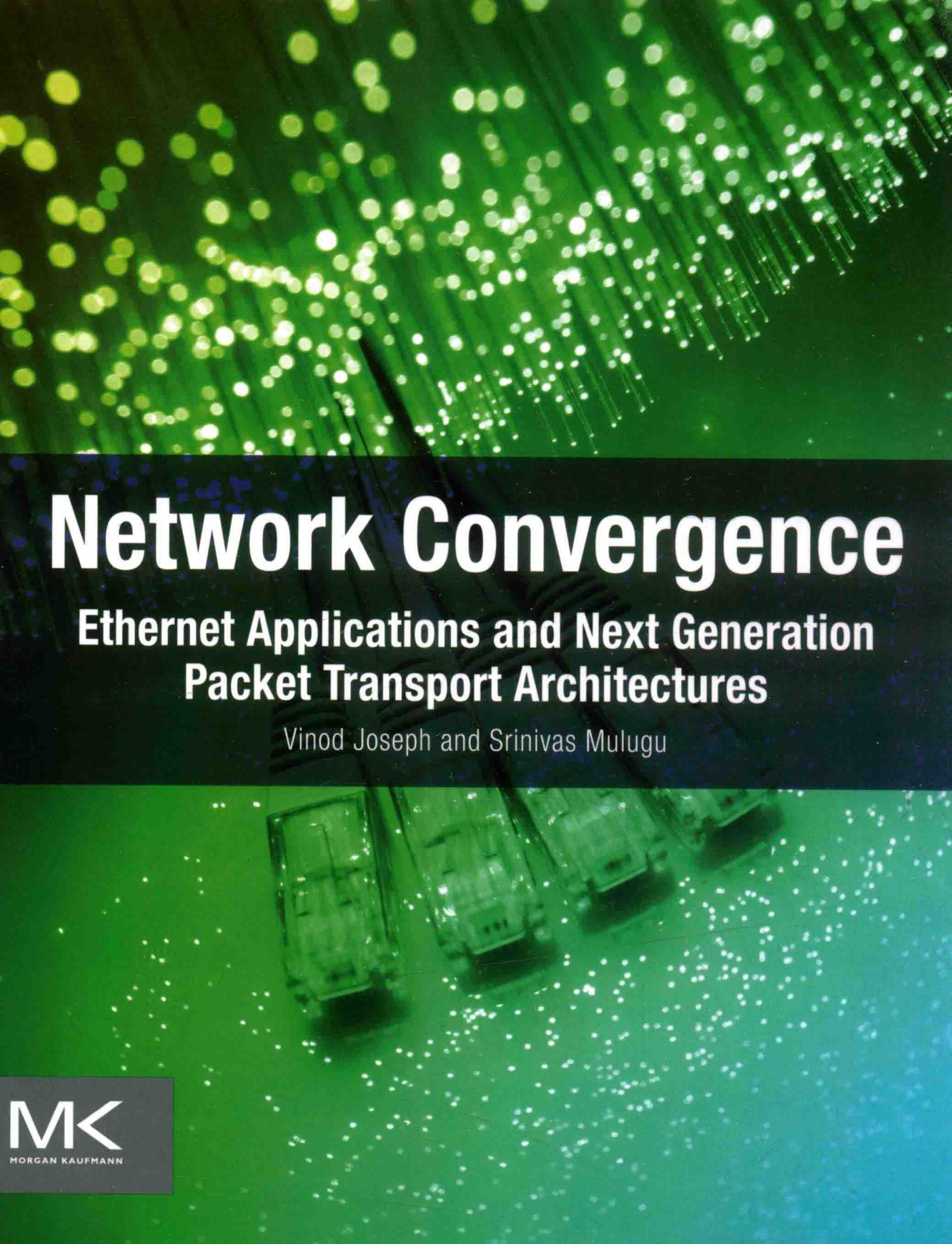

Network Convergence
Ethernet Applications and Next Generation Packet Transport Architectures
Vinod Joseph and Srinivas Mulugu
MK
MORGAN KAUFMANN

NETWORK CONVERGENCE

Ethernet Applications and Next Generation Packet Transport Architectures

VINOD JOSEPH and
SRINIVAS MULUGU

ELSEVIER
AMSTERDAM • BOSTON • HEIDELBERG • LONDON
NEW YORK • OXFORD • PARIS • SAN DIEGO
SAN FRANCISCO • SINGAPORE • SYDNEY • TOKYO
Morgan Kaufmann is an imprint of Elsevier
MK MORGAN KAUFMANN

Publisher: *Steve Elliot*
Editorial Project Manager: *Kaitlin Herbert*
Project Manager: *Malathi Samayan*
Designer: *Mark Rogers*

Morgan Kaufmann is an imprint of Elsevier
225 Wyman Street, Waltham, MA 02451, USA

Notices
Knowledge and best practice in this field are constantly changing. As new research and experience broaden our understanding, changes in research methods or professional practices, may become necessary. Practitioners and researchers must always rely on their own experience and knowledge in evaluating and using any information or methods described herein. In using such information or methods they should be mindful of their own safety and the safety of others, including parties for whom they have a professional responsibility.

To the fullest extent of the law, neither the Publisher nor the authors, contributors, or editors, assume any liability for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions, or ideas contained in the material herein.

Library of Congress Cataloging-in-Publication Data
Joseph, Vinod.
Network convergence : Ethernet applications and next generation packet transport architectures / Vinod Joseph, Srinivas Mulugu.
pages cm
Includes bibliographical references and index.
ISBN 978-0-12-397877-6 (pbk.)
1. Ethernet (Local area network system) 2. Packet transport networks. 3. Computer network architectures. 4. Convergence (Telecommunication) 5. Internetworking (Telecommunication) I. Mulugu, Srinivas. II. Title.
TK5105.383.J68 2013
004.6–dc23

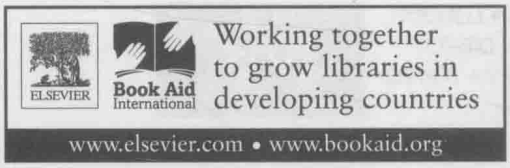
2013025197

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library

ISBN: 978-0-12-397877-6

For information on all MK publications visit our website at http://store.elsevier.com

Printed and bound in USA
14 15 16 13 12 11 10 9 8 7 6 5 4 3 2 1



NETWORK CONVERGENCE

INTRODUCTION

Over the years, Ethernet has become the de facto vehicle for deploying Internet communication transport infrastructures at the access, aggregation, and even the core aspects. Because of the simplicity, capacity for scalability, availability, and levels of integration that Ethernet offers across the various networking layers, it has been adopted widely across the industry. The objective of the book is to highlight the convergence of new developments, applications, and services that are emerging in Ethernet transport.

The book discusses various applications and services that can be deployed using Ethernet as a converged infrastructure linking multiple carrier and/or enterprise infrastructures. In the book we examine several services, such as MPLS Layer 3 VPNs, Point-to-Point and Multi-Point Ethernet over MPLS PWs, and provider backbone bridging, which is an option available for scaling Ethernet layer 2 services. We then move on to look at how MPLS can be used in all Ethernet access, aggregation, and core aspects to offer services such as mobility and still retain operational scale and control. We also examine MPLS-TP, a trend that is applicable in certain Ethernet access environments, before moving on to discuss how packet and optical layers can be integrated.

Please note that among all the graphics and figures appearing in this book, all symbols of routers and switches are purely generic to illustrate a device or concept. None of them represents any actual vendor. Some of the configuration templates provided are from actual vendors, such as Juniper, Cisco, and Alcatel-Lucent. This is to provide diversity and also help the reader relate to specific topics. It is by no means an endorsement of any vendors or their respective technologies.

Finally, this book is written by the two authors in their own capacities. It has no affiliation to any organizations they are directly or indirectly involved with.

CONTENTS

Introduction ix

Chapter 1 Deploying Ethernet Multi-Point Services Using VPLS 1

Introduction 1
Virtual Private LAN Service (VPLS) 1
Inter-Provider VPLS 66
Conclusion 70

Chapter 2 Understanding Advanced MPLS Layer 3 VPN Services 73

Introduction 73
MPLS Layer 3 VPNs 73
MPLS Layer 3 Inter-AS VPNs 103
Advanced Next Generation Multicast over Layer 3 VPNs 176
Migration from Draft-Rosen to NG-MVPNs 295
Summary 323

Chapter 3 Provider Backbone Bridging with VPLS 325

Introduction 325
Operational Examples 332
Conclusion 392

Chapter 4 Scaling Packet Ethernet Services Using Seamless MPLS 393

Introduction 393
State of the Mobile Transport Industry 393
Next Generation Mobile Transport Characteristics 396
System Concept 401
Transport Models 404
Service Models 412
Large Network, Inter-AS Design with IP/MPLS Access 419

Large Network Inter-AS Design with Non-IP/MPLS Access 424
Small Network, Integrated Core and Aggregation with IP/MPLS Access 427
Small Network, Integrated Core, and Aggregation with Non-IP/MPLS Access 429
Service Architecture 430
Inter-Domain Hierarchical LSPs 439
Inter-Domain LSPs for Multi-Area IGP Design 439
Inter-Domain LSPs for Inter-AS Design 444
Inter-Domain LSPs for Integrated Core and Aggregation Design 450
Transport and Service Control Plane 451
BGP Control Plane for Multi-Area IGP Design 452
BGP Control Plane for Inter-AS Design 454
BGP Control Plane for Integrated Core and Aggregation Design 457
Quality of Service 458
Synchronization Distribution 461
Conclusion 464

Chapter 5 QoS in Mobile IP Networks 465

Introduction 465
The First Generation Mobile Network 465
The Second Generation Mobile Network 465
The 2.5G Mobile Network 468
The Third Generation Mobile Vision 469
The Evolution toward 3G 471
The Third Generation Evolution for GSM 471
Converged NGN Infrastructure for Mobile Transport 474
Third Generation Release 99 Overview 476
Third Generation Release 4 Overview 480
Third Generation Release 5/6 Overview 483
Mobile Network RAN Access 487
IPV6 in 3G Networks 490
Overview of QoS and the 3GPP Standards 493

3G Traffic Classes 496
Bearer Service Attributes 498
Mobile Network Protocol Stacks 498
The Iu Interface in the Circuit-Switched Domain 508
Mobile Network Transport Framework 509
Quality of Service Framework 512
Conclusion 524

Chapter 6 Emerging Trends in Packet Optical Convergence 525
Introduction 525
Basics about Optical Networks 526
Optical Transmission Principles 527
DWDM Transponders 528
Multiplexers and De-Multiplexers 529
Optical Add/Drop Multiplexers 529
Overview of Fixed OADM & ROADM Technology 530
The Optical Transport Network (OTN) 536
Design Constraints and Initial Design Considerations for Fiber Optic Networks 542
IPoDWDM 546

Chapter 7 MPLS Transport Profile 553
Introduction 553
Why MPLS-TP? Why Now? 553
IP/MPLS and MPLS-TP Inter-operability 559
Overlay Model 561
Peer Model 568
Application Scenario: The Overlapping Interconnection Scenario of the L2/L3 VPN Bridging in the LTE Environment 572
Application Scenario: The Peer-to-Peer Interconnection Scenario in the MS-PW Environment 573

Index 577

1

DEPLOYING ETHERNET MULTI-POINT SERVICES USING VPLS

Introduction

In this chapter we take a look at virtual private LAN services (VPLS) and the various building blocks of deploying multipoint Ethernet services using VPLS.

Virtual Private LAN Service (VPLS)

Although our topic is VPLS, let us begin by taking a quick look at MPLS Layer 2 VPNs, also referred to as point-to-point services.

A point-to-point L2 VPN circuit, as defined by the pseudowire encapsulation edge to edge working group (PWE3) of the Internet Engineering Task Force (IETF), is a provider service that offers a point-to-point service infrastructure over an IP/MPLS packet switched network. The PWE3 working group describes mechanisms for delivering L2 VPN services across this kind of network. The basic reference model is shown in Figure 1.1.

A pseudowire (PW) is a connection between two provider edge (PE) devices, which connects two attachment circuits (ACs). An AC can be a Frame Relay DLCI, an ATM VPI/VCI, an Ethernet port, a VLAN, a HDLC, a PPP connection on a physical interface, a PPP session from an L2TP tunnel, an MPLS LSP, etc. During the setup of a PW, the two PE routers are configured or automatically exchange information about the service to be emulated so that later they know how to process packets coming from the other end. The PE routers use Targeted LDP (T-LDP) sessions for setting the PW. After a PW is set up between two PE routers, frames received by one PE from an AC are encapsulated and sent over the PW to the remote PE, where native frames are re-constructed and forwarded to the other CE.

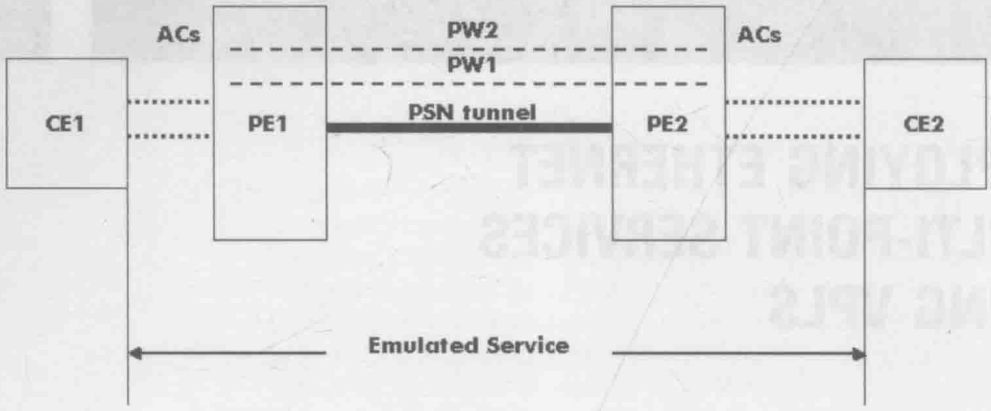


Figure 1.1

From a data-plane perspective, different PWs in the same packet-switched network (PSN) tunnel are identified using a multiplexing field. This multiplexing field is an MPLS label, and the encapsulation of the customer frames over these (MPLS) connections or PWs is defined by the PWE3 working group. PSN tunnels are implemented in the provider's network as MPLS LSPs (RSVP, LDP), or using IP-in-IP (GRE). Figure 1.2 shows the protocol stack in the core of the provider's network for Ethernet frames.

Ethernet is particularly appealing to enterprise networkers: It is mature, reliable, cheap, scalable, and well understood. Common networking practice is to connect local sites (subnets, floors, or buildings of a campus) with an Ethernet backbone switch, managing and scoping the network with layer 2 VLANs. So it comes as no surprise that such network operators would like to be able to connect sites across a wider area using the same Ethernet backbones. Nor is this interest new; as much as 15 years ago many local providers were offering metropolitan-area Ethernet services such as Transparent LAN Service (TLS), based on proprietary technologies, and LAN Emulation (LANE), based on ATM backbones. But such service offerings were not ideal for the provider due to factors such as dependency on a single vendor, for a proprietary TLS solution, and prohibitive complexity, for a LANE solution. As Ethernet technology itself advanced, permitting greater speeds at greater transmission distances, more recent metropolitan Ethernet offerings have been built around Ethernet switches. But these switch-based infrastructures have their own limitations, primarily lack of scalability due to the numeric limitations on VLAN IDs.

In recent years, VPLS has arisen as a practical, economical, and scalable alternative for creating metro Ethernet services. VPLS, in

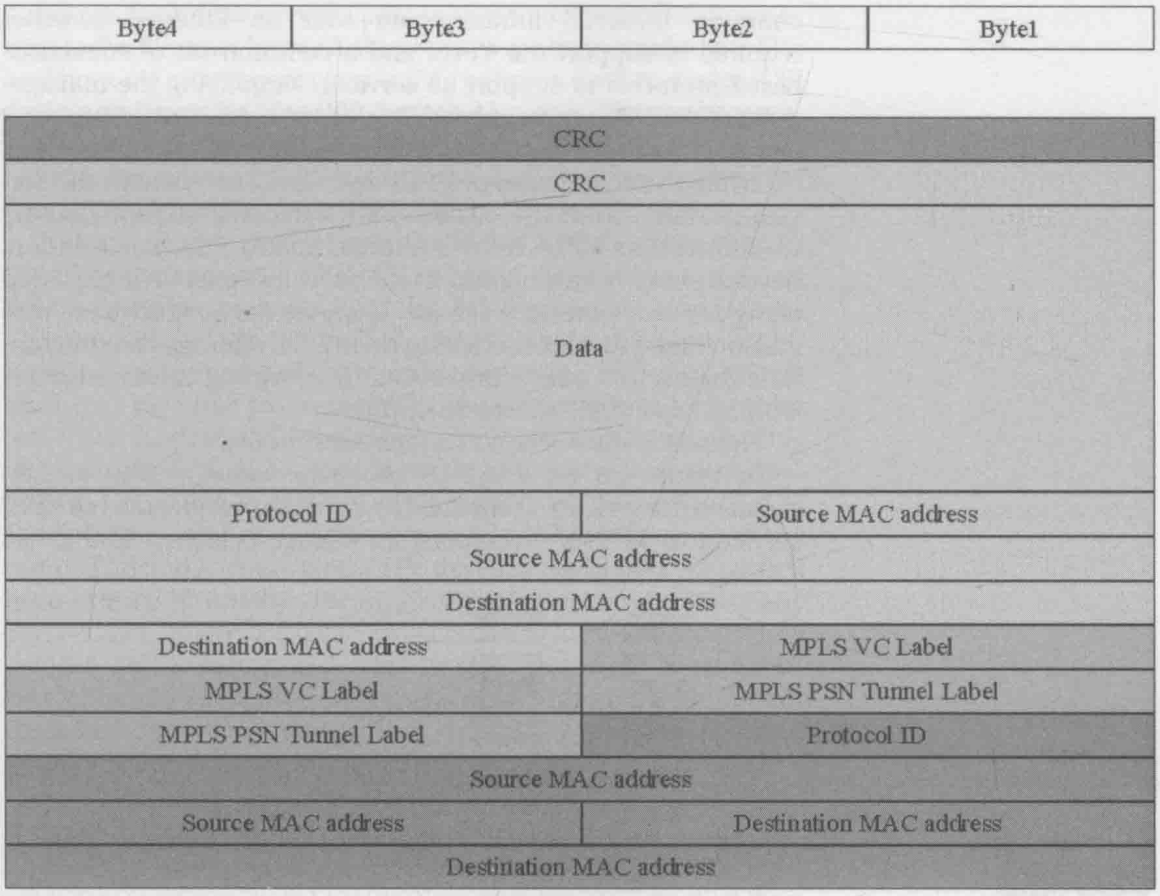


Figure 1.2

turn, has been made possible by the advent of MPLS, which has seen accelerating deployment in carrier and service provider networks beginning in the late 1990s. MPLS provides a means of creating virtual circuits, similar to Frame Relay DLCIs and ATM VCI/VPIs, over IP networks. Its appeal is its ability to eliminate Frame Relay and ATM infrastructures while moving the services provided by those infrastructures to an IP network, thereby reducing the overall capital and operational costs of the network. These MPLS virtual circuits—called label-switched paths (LSPs)—have for many years been used to provide Layer 3 IPv4 VPNs and Layer 2 point-to-point VPNs. More recently, the technology has been extended to support Layer 3 IPv6 VPNs, Layer 3 Multicast VPNs, and VPLS.

The advantage of VPLS for the service provider is in building on the capital and operational cost savings of an MPLS VPN network: a

common IP/MPLS infrastructure with no Ethernet switches required to support the VPLS, and a common set of standards-based protocols to support all services, simplifying the management of the network. Supplying the desired service to the customer is a simple matter of installing and configuring the correct interface.

While the advantages of VPLS described here benefit the service provider, from the customer's perspective there is nothing to differentiate VPLS from any other metro Ethernet solution, beyond possibly having some of the provider's cost savings passed along as a less expensive service. However, service providers who add an inter-provider element to their VPLS offering, can differentiate themselves from competitors by providing their customers with an expanded "service footprint."

Figure 1.3 shows the VPLS reference model.

In Figure 1.3 an IP/MPLS backbone network (the packet-switched network, PSN) operated by a service provider offers a VPLS service to two VPN customers: an Orange customer and a Red

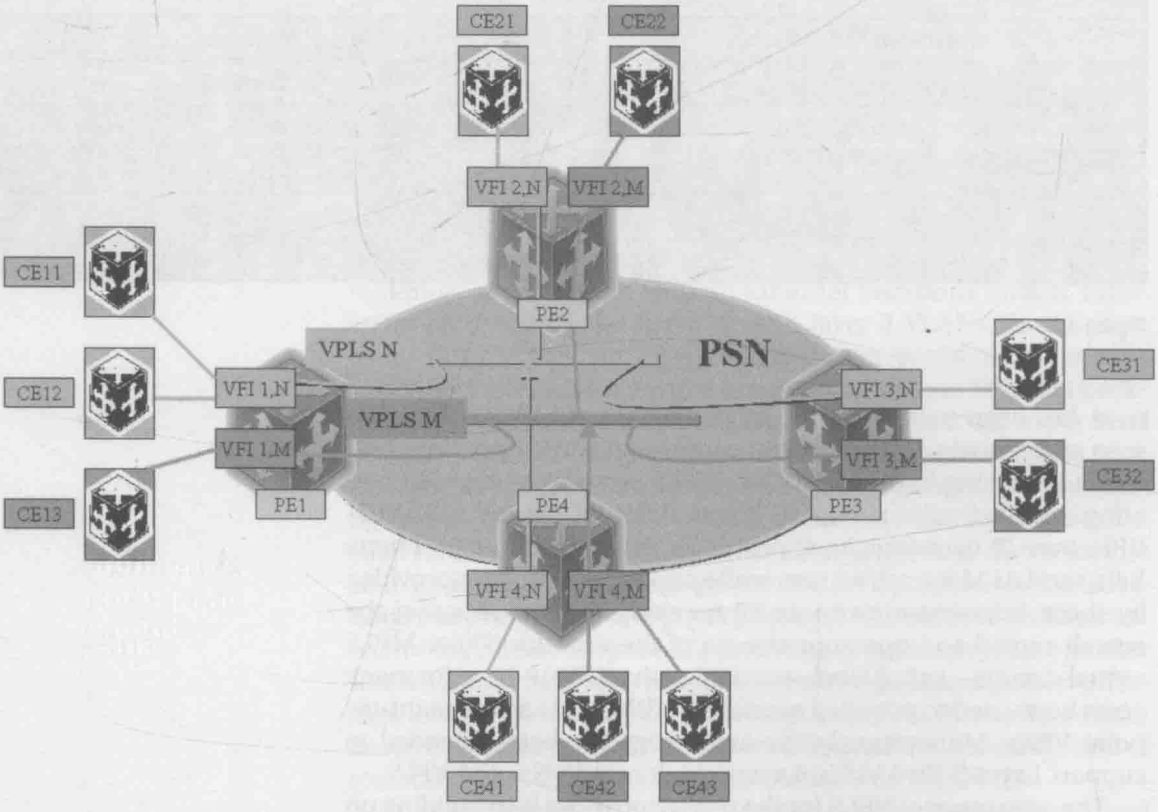


Figure 1.3

customer. Each customer has private sites that it wants to interconnect at the Ethernet layer. Customer sites are connected to the SP's backbone via attachment circuits (AC) between customer edge (CE) devices and provider edge (PE) devices. As such, a VPN can be represented by a collection of CE devices. In this illustration, the Orange L2VPN N consists of <CE11, CE12, CE21, CE31, CE41> while the Red L2VPN M consists of <CE22, CE31, CE32, CE42, CE43>.

As with all PE-based VPNs, with VPLS, the CE devices are unaffected by the service: a VPLS CE can be a standard router, or an Ethernet bridge or host. It is the PE device that implements VPLS-specific functions. Indeed, the PE device needs to implement a separate virtual forwarding instance (VFI)–also known as virtual switched instance (VSI), the equivalent of VRF tables for MPLS Layer 3 VPNs)–for every VPLS it is attached to. This VFI has physical direct interfaces to attached CE devices that belong to the VPLS, and virtual interfaces or pseudowires that are point-to-point connections to remote VFIs belonging to the same VPLS and located in other PE devices. These PWs are carried from one PE to another PE via PSN tunnels. From a data-plane perspective, different PWs in the same PSN tunnel are identified using a multiplexing field. This multiplexing field is an MPLS label. The encapsulation of the customer Ethernet frames over these MPLS connections or PWs is defined by the PWE3 working group. PSN tunnels are implemented in the provider's network as MPLS LSPs (RSVP, LDP) or using IP-in-IP (GRE). Figure 1.4 shows the protocol stack in the core of the provider's network.

A Draft-Rosen MVPN represents itself as an emulated LAN. Each MVPN has a logical PIM interface and will form an adjacency to every other PIM interface across PE routers within the same MVPN. This is illustrated in Figure 1.__.

Note that with VPLS, a full mesh of PSN tunnels between the network's PE devices is assumed, and for every VPLS instance there is a full mesh of pseudowires between the VFIs belonging to that VPLS. The IETF Layer 2 VPN working group has produced two separate VPLS standards,0 documented in RFC 4761 and RFC 4762 (see Kompella and Rekhter, Jan. 2007, and Lasserre and Kompella, Jan. 2007). These two RFCs define almost identical approaches with respect to the VPLS data plane, but they specify significantly different approaches to implementing the VPLS control planes.

VPLS Control Plane

The VPLS control plane has two primary functions: autodiscovery and signaling. Discovery refers to the process of finding all PE routers that participate in a given VPLS instance. A PE router can be

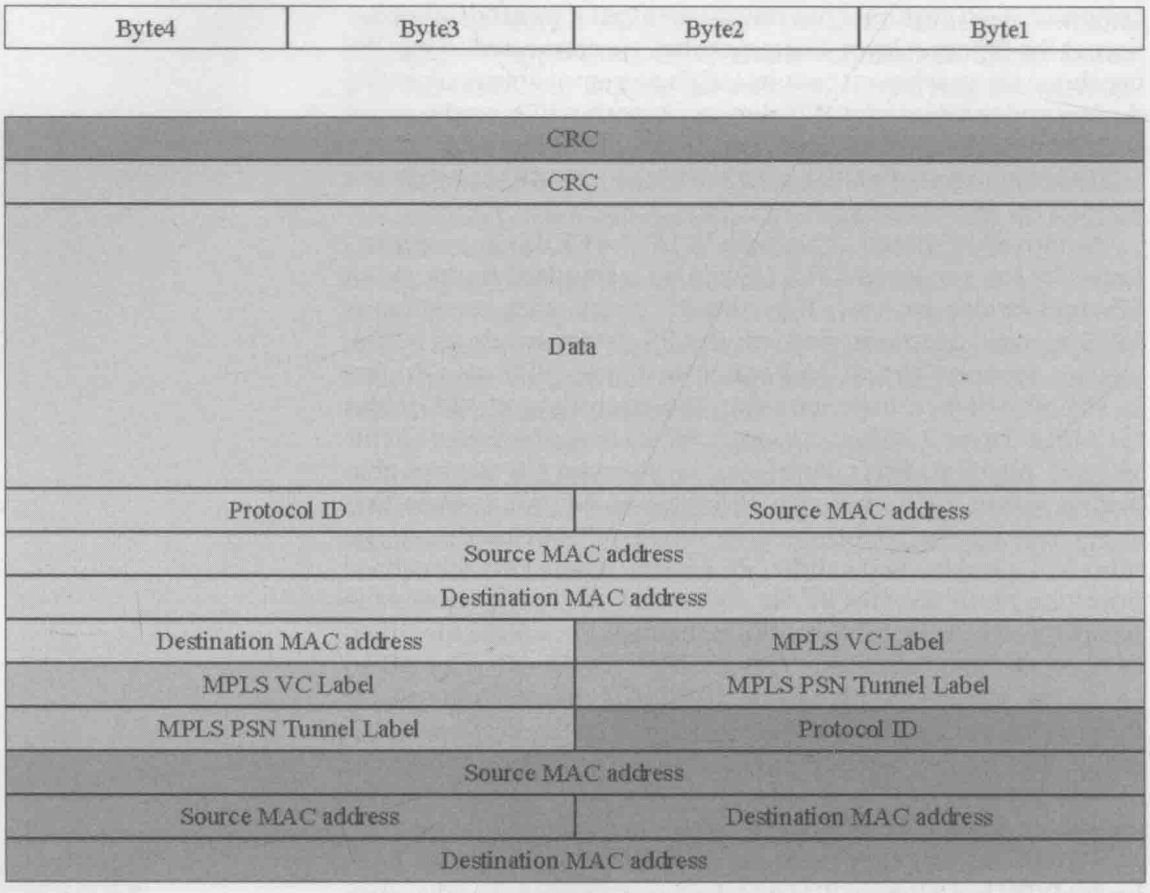


Figure 1.4

configured with the identities of all the other PE routers in a given VPLS instance, or the PE router can use a protocol to discover the other PE routers. The latter method is called autodiscovery. After discovery occurs, each pair of PE routers in a VPLS network must be able to establish pseudowires to each other, and in the event of membership change, the PE router must be able to tear down the established pseudowires. This process is known as signaling. Signaling is also used to transmit certain characteristics of the pseudowire that a PE router sets up for a given VPLS.

BGP-VPLS Control Plane

The BGP-VPLS control plane, as defined by RFC 4761, is similar to that for Layer 2 and Layer 3 (see Kompella, Jan. 2006, and Rosen and Rekhter, Feb. 2006). It defines a means for a PE router to

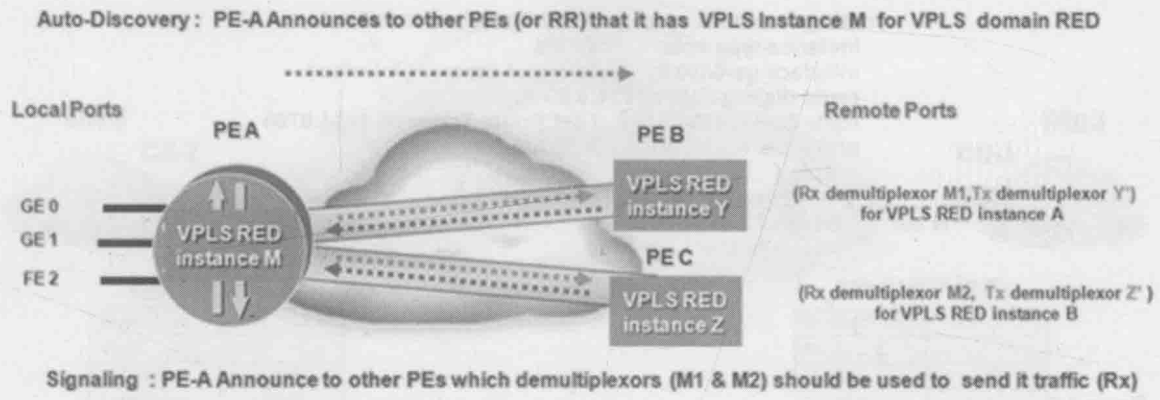


Figure 1.5

discover which remote PE routers are members of a given VPLS (autodiscovery), and for a PE router to know which pseudowire label a given remote PE router will use when sending the data to the local PE router (signaling). With the BGP-VPLS control plane, BGP carries enough information to provide the autodiscovery and signaling functions simultaneously. See Figure 1.5.

The details for de-multiplexer fields will be discussed in the following sections. As in the BGP scheme for Layer 2 and Layer 3 VPNs, a route target is configured on each PE router for each VPLS present on the PE router. The route target is the same for a particular VPLS across all PE routers and is used to identify the VPLS instance to which an incoming BGP message pertains. For each VPLS on each PE router, an identifier, known as a site identifier, is configured. Each PE router involved in a particular VPLS must be configured with a unique site identifier. The site identifier is the same as the virtual edge identifier (VE ID) referred to in RFC 4761, which prescribes one VE ID per PE for each VPLS instance, irrespective of how many local ports belong to that VPLS. A label block is a set of de-multiplexer labels used to reach a given VPLS site within a set of remote sites. The PE router uses a label block to send a single common update message to establish a pseudowire with multiple PE routers, instead of having to send an individual

```
routing-instances vpnA {  // Configuration for VPN A
  instance-type vpls;       // vpls
  interface ge-0/0/0.0;  // multipoint Ethernet interface
  route-distinguisher 1234:5.6.7.8;
  route-target 1234:8765; // set Route Target to 1234:8765
  protocols {               // PE-CE protocol
    vpls {
      site-range 20;
      site CE-A3 {
        site-identifier 3;
      }
    }
  }
}
```

Figure 1.6

message to each PE router. A number of illustrations in the following sections elaborate on this in greater detail.

Note: Each PE router creates a virtual connection table (VCT) per VPLS instance. The VCT is similar to the virtual forwarding instance (VFI) referred to earlier in this chapter). Hence the terms VCT and VFI are used interchangeably in this chapter.

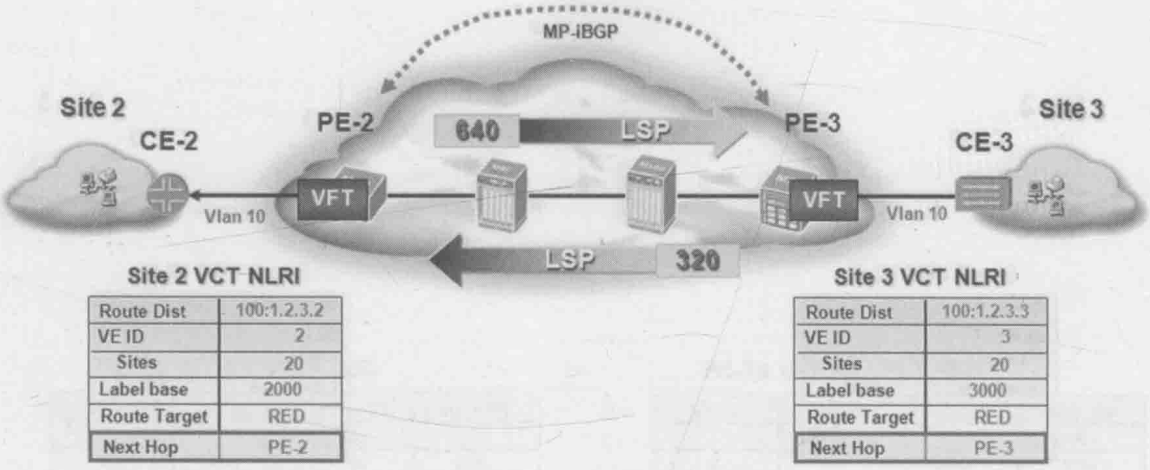
Figure 1.6 shows a JUNOS Configuration snippet describing the basic setup of a BGP-based VPLS instance. The configuration here is exactly the same as the configuration for BGP/MPLS Layer 3 VPNs, with the exception of the keyword "VPLS" defined under the protocols hierarchy, which in the case of BGP/MPLS VPNs would BGP or OSPF or RIP, etc.

Figure 1.7 illustrates a two-site VPLS instance created between two PE routers clarifies the configuration statements provided above, and their relevance.

In Figure 1.7, PE2 allocates a label base of "2000" for a given VPLS instance: VPLS RED. PE3 uses label base "3000" for the same VPLS instance. The illustration that follows shows the role of the label base.

In Figure 1.8, PE2 has been allotted 3002 by PE3, as the inner label to be used to reach Site 3 on PE3. Similarly PE3 will use 2003 as the Inner label for reaching Site 2 on PE2. Another label would be used by each of the PE routers, if they needed to connect to another site within the same VPLS instance (VPLS RED) on another PE. The MPLS outer labels are also displayed in PE2's VFT (Label 640).

If more PE routers are added to VPLS instance RED, each of them uses different label. This is illustrated in Figure 1.9.



- **PE-PE VCT distribution using Multi-Protocol BGP (RFC 2858)**
  - Requires full-mesh MP-iBGP or Route Reflectors
  - Route Distinguisher: "uniquifies" VCT information
  - Route Target: determines VPN topology
  - Analogous to CE-PE routes advertisements in RFC2547 VPNs
  - One single LNRI advertisement per VPLS instance per PE is sufficient

Figure 1.7

In PE2's VFT, site-id "1" and site-id "15" have different MPLS outer and inner labels. This indicates that those sites belong to different PE routers. The same is the case with PE3.

LDP-VPLS Control Plane

In contrast to the BGP-VPLS control plane, the LDP-VPLS control plane provides only signaling but no autodiscovery (more on this in the following sections). In this control plane, LDP is used to signal the pseudowires that interconnect the VPLS instances of a given customer on the PE routers. The LDP signaling scheme for VPLS is similar to the LDP scheme for point-to-point Layer 2 connections (see Martini et al., Apr. 2006). In the absence of an autodiscovery mechanism, the identities of all the remote PE routers that are part of a VPLS instance must be configured on each PE router manually.

The virtual circuit identifier (VCID), which is in the point-to-point Layer 2 connection used to identify a specific pseudowire, is configured to be the same for a particular VPLS instance on all PE routers. Hence, the VCID enables a PE router to identify