MPLS based Virtual Private Network Services

An MFA Forum Sponsored Tutorial

Matt Kolon
MFA Forum Ambassador
Senior Technical Solutions Manager
Juniper Networks

MPLS VPN Tutorial Agenda

• Introduction to the MFA Forum
• Introduction to MPLS and MPLS VPNs
  ▪ Defining Layer 2 and 3 VPNs
• Layer 3 MPLS VPN
  ▪ BGP Review
  ▪ RFC 2547bis Key Characteristics
  ▪ BGP/MPLS VPN Architecture Overview
    ▪ VPN Routing and Forwarding (VRF) Tables
    ▪ Overlapping VPNs
    ▪ VPN Route Distribution
    ▪ VPN Packet Forwarding
    ▪ Scaling L3VPNs and Route Reflectors
MPLS VPN Tutorial Agenda

- Layer 2 VPNs
  - IETF PWE3 and L2VPN WG update
  - Encapsulation and Label Stacking
  - Virtual Private Wire Services – VPWS
    - Pt-to-pt Ethernet, Pt-to-pt ATM, Pt-to-pt Frame Relay
  - Virtual Private LAN Services – VPLS
- Introduction to Multi-Service Interworking
  - Carrier Challenges at the Edge
  - Interworking History and Definition
  - Network and Service Interworking (FRF.5 and FRF.8.1)
  - MPLS FR Alliance Multi-Service Interworking Work Actions

Introduction to the MFA Forum
Mission Statement

The MFA Forum is an international, industry-wide, nonprofit association of telecommunications, networking, and other companies focused on advancing the deployment of multi-vendor, multi-service packet-based networks, associated applications, and interworking solutions.

MFA Forum

- Formed in July 2005 by merging the ATM Forum and the MPLS & Frame Relay Alliance
- 39 member companies
- Three primary committees
  - Technical Committee
    - Applications and Deployment Working Group
    - ATM Architecture Working Group
    - ATM Signaling Working Group
    - Interoperability Working Group
    - Interworking and Frame Relay Working Group
  - Marketing Awareness and Education Committee
  - Service Provider Council
- MPLS User Group – Enterprises, Carriers
Technical Committee
Major Work Items

- MPLS Inter-Carrier Interface
- Packet-Based GMPLS Client to Network Interconnect (CNI)
- ATM and Frame Relay to MPLS Control Plane interworking (in Final Ballot)
- Fault Management Interworking (in Final Ballot)
- ATM, Ethernet, and Frame Relay Interworking over MPLS (in Final Ballot)
- Performance Monitoring Across Multiservice Networks (in Straw Ballot)
- Layer 2 Service Mediation
- AAL1 and AAL2 Voice Trunking over MPLS

MFA Forum

- Market Awareness & Education
  - Tutorials
    - Introduction to MPLS ½ day and full day
    - MPLS Virtual Private Networks ½ day and full day
    - MPLS VPN Security ½ day
    - Traffic Engineering ½ day
    - GMPLS ½ day
    - Migrating Legacy Services to MPLS ½ day
    - MPLS OAM ½ day
    - Voice over MPLS ½ day
  - New tutorials based upon demand
- Conferences and exhibitions
  - MFA Forum speaker at almost every MPLS conference globally
- Website and Newsletter
- Public message board
- Next meeting: June 27-29 in Vancouver, BC, Canada
- Please join us!
  - Subscribe to information mail list info@mfaforum.org
  - To join the Forum contact Alexa Morris, Executive Director
    E-Mail: amorris@mfaforum.org  Phone: 510 608-5914

Copyright © 2006 MFA Forum
Section 1
Introduction to MPLS and MPLS VPNs

Why MPLS?
A Common Control Plane

Best of the packet-switched and circuit-switched worlds
**MPLS: Addresses many network needs**

- End Users
- Applications
- IP
- SONET/SDH
- Photonics
- Access
- MPLS L3 VPN & L2 VPN
- MPLS TE + CoS/QoS
- MPLS Protection & rerouting
- GMPLS

**Virtual Private Networks**

- Provide private line and private LAN connections between multiple sites
- Leverage public network to provide competitive service pricing and reduce service operating cost

**Diagram:**

- Corporate Headquarters
- Storage Facility
- Home office
- Remote office
- Public Network
Virtual Private Networks

- VPN (Virtual Private Network) is simply a way of using a public network for private communications, among a set of users and/or sites
- Remote Access: Most common form of VPN is dial-up remote access to corporate database - for example, road warriors connecting from laptops
- Site-to-Site: Connecting two local networks (may be with authentication and encryption) - for example, a Service Provider connecting two sites of the same company over its shared network

MPLS, VPNs, and Standards

A lot of confusion

- L2TP
- VPWS
- Point to multipoint
- Layer 2
- IP VPNs
- Lasserre
- Kompella
- Tunneling
- VPLS
- Martini
- Vkompella
- PWE3
- Point to point
- Layer 3
- IPLS
- BGP / MPLS VPNs
- RFC 2547 bis
- IPsec
- PPVPN
### VPNs

**Types, Layers, and Implementations**

<table>
<thead>
<tr>
<th>VPN Type</th>
<th>Layer</th>
<th>Implementation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leased Line</td>
<td>1</td>
<td>TDM/SDH/SONET</td>
</tr>
<tr>
<td>Frame Relay</td>
<td>2</td>
<td>DLCI</td>
</tr>
<tr>
<td>ATM</td>
<td>2</td>
<td>VC</td>
</tr>
<tr>
<td>GRE/UTI/L2TPv3</td>
<td>3</td>
<td>IP Tunnel</td>
</tr>
<tr>
<td>Ethernet</td>
<td>2</td>
<td>VLAN / VPWS / VPLS</td>
</tr>
<tr>
<td>IP</td>
<td>3</td>
<td>RFC2547bis / VR</td>
</tr>
<tr>
<td>IP</td>
<td>3</td>
<td>IPsec</td>
</tr>
</tbody>
</table>

### VPNs

**How do they compare?**

<table>
<thead>
<tr>
<th></th>
<th>FR or ATM</th>
<th>IPsec</th>
<th>L3 MPLS</th>
<th>L2 MPLS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Point-to-multipoint</td>
<td>×</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Multi-protocol</td>
<td>✓</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>QoS and CoS</td>
<td>✓</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Low latency</td>
<td>✓</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Security</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>SLAs</td>
<td>✓</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>
MPLS VPNs in the IETF

- **IETF Areas**
  - Application
  - General
  - Internet
  - Op and Man
  - Routing
  - Security
  - Sub-IP
  - Transport
  - User Services

- **IAB**
- **ISOC**
- **IANA**

**Layer 2 VPNs**
- **L2VPN**
- **L3VPN**

**Layer 3 VPLS**
- **PWE3**

- **Internet**
- **Transport**

- **Encapsulations**
  - ATM
  - FR
  - Ethernet
  - PPP/HDLC
  - SONET/SDH
What are Layer 2, Layer 3 VPNs

- VPNs based on a Layer 2 (Data Link Layer) technology and managed at that layer are defined as Layer 2 VPNs (MPLS, ATM, Frame Relay)
- VPNs based on tunneling at Layer 3 (Network Layer) are Layer 3 VPNs, (BGP/MPLS, VR, IPSec)

Visually - Layer 2 VPN
Visually - Layer 3 VPN

In a Layer 3 VPN, CE Device and PE Device are IGP peers. VPN Tunneling Protocols are L2TP, IPSec, and MP-iBGP.

VPN Tunnels:
- VPN A
- VPN B

Section 2

Layer 3 MPLS VPN
Layer 3 MPLS VPN
- BGP Review
- RFC 2547bis Key Characteristics
- BGP/MPLS VPN Architecture Overview
  - VPN Routing and Forwarding (VRF) Tables
  - Overlapping VPNs
  - VPN Route Distribution
  - VPN Packet Forwarding
  - Scaling L3VPNs and Route Reflectors

What is BGP?
- BGP is an exterior gateway protocol that allows IP routers to exchange network reachability information.
- BGP published as RFC 1105 in 1989 and after several updates, the current version, BGP-4 was published in 1995 as RFC 1771.
- Numerous other RFCs and Internet Drafts focus on various aspects and extensions including multiprotocol extensions, extended communities, carrying label information in BGP, etc
IGP vs. EGP

- Interior Gateway Protocols
  - RIP, OSPF, IS-IS
  - Dynamic, some more than others
  - Define the routing needed to pass data **within** a network
- Exterior Gateway Protocol
  - BGP
  - Less Dynamic than IGPs
  - Defines the routing needed to pass data **between** networks

Internal Border Gateway Protocol

**iBGP** - BGP between border routers in the same AS.

Provides a consistent view within the AS of the routes exterior to the AS.
**External Border Gateway Protocol**

eBGP - BGP between border routers in two different AS’s.

---

**BGP/MPLS IP VPN (RFC 2547bis)**

*Key Characteristics*

- **Requirements:**
  - Support for overlapping, private IP address space
  - Different customers run different IGPs (i.e. RIP, OSPF, IS-IS)
- **Solution:**
  - VPN network layer is terminated at the edge (PE)
    - PE routers use plain IP with CE routers
**BGP/MPLS IP VPN**

**Key Characteristics**

- P routers (LSRs) are in the core of the MPLS cloud
- P and PE (LERs) routers run an IGP and a label distribution protocol
  - Labelled VPN packets are transported over MPLS core
- PE routers are MP-iBGP fully meshed
  - for dissemination of VPN membership and reachability information between PEs

---

**Virtual Routing and Forwarding (VRF) Tables**

- Each VPN needs a separate **Virtual routing and forwarding instance (VRF)** in each PE router to
  - Provides VPN isolation
  - Allows overlapping, private IP address space by different organizations
Virtual Routing and Forwarding (VRF) PE to CE Router Connectivity

- Protocols used between CE and PE routers to populate VRFs with customer routes
  - BGP-4
    - useful in stub VPNs and transit VPNs
  - RIPv2
  - OSPF
  - static routing
    - particularly useful in stub VPNs
- Note:
  - Customer routes need to be advertised between PE routers
  - Customer routes are not leaked into backbone IGP

Virtual Routing and Forwarding (VRF) Overlapping VPNs

- A VPN is a collection of sites sharing a common routing information (routing table)
- A VPN can be viewed as a community of interest (or Closed User Group)

Examples:
- Extranet
- VoIP Gateway
Virtual Routing and Forwarding (VRF) Overlapping VPNs

- A site can be part of different VPNs
- A site belonging to different VPNs may or may not be used as a transit point between VPNs
- If two or more VPNs have a common site, address space must be unique among these VPNs

Examples:
- Extranet
- VoIP Gateway

VRFs and Route Distribution

- Multiple VRFs are used on PE routers
- The PE learns customer routes from attached CEs
- Customer routes are distributed to other PEs with MP-BGP
- Different IGPs or eBGP supported between PE and CE peers
• **Route Target attributes**
  - “Export” Route Target: Every VPN route is tagged with one or more route targets when it is exported from a VRF (to be offered to other VRFs)
  - “Import” Route Target: A set of routes targets can be associated with a VRF, and all routes tagged with at least one of those route targets will be inserted into the VRF
VPN Route Distribution

• How will the PE routers exchange information about VPN customers and VPN routes between themselves?

Option #1: PE routers run a different routing algorithm for each VPN
  • Scalability problems in networks with a large number of VPNs
  • Difficult to support overlapping VPNs

VPN Route Distribution

• How will the PE routers exchange information about VPN customers and VPN routes between themselves?

Option #2: BGP/MPLS IP VPN - PE routers run a single routing protocol to exchange all VPN routes
  • Problem: Non-unique IP addresses of VPN customers. BGP always propagates one route per destination not allowing address overlap.
**VPN Route Distribution**

**VPN-IPv4 Addresses**

- **VPN-IPv4 Address**
  - VPN-IPv4 is a globally unique, 96bit routing prefix

<table>
<thead>
<tr>
<th>Route Distinguisher (RD)</th>
<th>IPv4 Address</th>
</tr>
</thead>
<tbody>
<tr>
<td>64 bits</td>
<td>32 bits</td>
</tr>
<tr>
<td>Creates a VPN-IPv4 address that is globally unique, RD is configured in the PE for each VRF, RD may or may not be related to a site or a VPN</td>
<td>IP subnets advertised by the CE routers to the PE routers</td>
</tr>
</tbody>
</table>

- **Route Distinguisher format**
  - **ASN:nn**
    - Autonomous System Number (ASN) assigned by Internet Assigned Number Authority (IANA)
  - **IP-address:nn**
    - use only if the MPLS/VPN network uses a private AS number
  - **BGP-AS4:nn**
    - 4-byte Autonomous System Number (BGP-AS4)
VPN Route Distribution
BGP with Multiprotocol Extensions

• How are 96-bit VPN-IPv4 routes exchanged between PE routers?

• BGP with Multiprotocol Extensions (MP-BGP) was designed to carry such routing information between peer routers (PE)
  ▪ propagates VPN-IPv4 addresses
  ▪ carries additional BGP route attributes (e.g. route target) called extended communities

VPN Route Distribution
BGP with Multiprotocol Extensions

• A BGP route is described by
  ▪ standard BGP Communities attributes (e.g. Local Preference, MED, Next-hop, AS_PATH, Standard Community, etc.)
  ▪ extended BGP Communities attributes

• Extended Communities
  ▪ Route Target (RT)
    • identifies the set of sites the route has to be advertised to
  ▪ Route Origin (RO)/Site of Origin
    • identifies the originating site
    • to prevent routing loops with multi-homed customer sites
IGP Label Distribution

- All routers (P and PE) run an IGP and a label distribution protocol
- Each P and PE router has routes for the backbone nodes and a label is associated to each route
- MPLS forwarding is used within the backbone

MP-BGP Route Distribution

“Net1” is the provider’s autonomous system
MP-BGP Route Distribution

Summary

• VPN Routing and Forwarding (VRF) Table
  ▪ Multiple routing tables (VRFs) are used on PEs
    • VPNs are isolated
  ▪ Customer addresses can overlap
    ▪ Need for unique VPN route prefix
    ▪ PE routers use MP-BGP to distribute VPN routes to each other
    ▪ For security and scalability, MP-BGP only propagates information about a VPN to other routers that have interfaces with the same Route Target value.

  $MP\text{-}BGP = BGP \text{ with Multiprotocol Extensions}$

VPN Packet Forwarding

PE-to-PE connectivity via LSPs
  • All routers (P and PE) run an IGP and a label distribution protocol
  • Each P and PE router has routes for the backbone nodes and a label is associated to each route
  • MPLS forwarding is used within the backbone
VPN Packet Forwarding
Label Stacking

Ingress PE router uses two-level label stack
- VPN label (inner label) assigned by the egress PE router
- IGP label (top label) identifying the PE router

Label stack is attached in front of the IP packet that belongs to a VPN
The MPLS packet is forwarded across the P routers in the backbone network

VPN Packet Forwarding
Label Stacking

P routers switch the packet based on the IGP Label (top label)

Egress PE router removes top label, uses inner label to select which VPN/CE to forward the packet to. Inner label is removed and packet sent to CE2 router
VPN Packet Forwarding

Penultimate Hop Popping

P routers switch the packet based on the IGP Label (top label)

PE1 router receives normal IP packet from CE1 router. PE router does “IP Longest Match” from VRF, finds iBGP next hop PE2 and imposes a stack of labels

PE1

IP Packet

IGP Label(PE2)

VPN Label

PE2

IP Packet

VPN Label

IGP Label(PE2)

Penultimate Hop Popping

P2 is the penultimate hop for the BGP next-hop. P2 removes the top label. This has been requested through LDP by PE2

PE2 receives packet with the label corresponding to the outgoing VRF. One single lookup. Label is popped and packet sent to CE2 router

Core Routers (P Routers)

• Not involved in MP-BGP
• Does not make routing decision based on VPN addresses
• Forwards packet based on the top label value
• P routers do not need to carry VPN routing information or Internet routing information, thus providing better network scalability
Scaling BGP/MPLS VPNs

- Scalability of BGP/MPLS VPNs
  - Expanding the MPLS core network
    - without impact on the VPN services, e.g. adding P routers (LSRs), new or faster links
  - Label stacking
    - allows reducing the number of LSPs in the network core and avoiding LSP exhaustion
  - VPN Route Distribution
    - Route Reflectors

Scaling BGP/MPLS VPNs

Route Reflectors

- BGP Route Reflectors
  - Existing BGP technique, can be used to scale VPN route distribution
    - PEs don’t need full mesh of BGP connections, only connect to RRs
    - By using multiple RRs, no one box needs to have all VPN routes
  - Each edge router needs only the information for the VPNs it supports
    - directly connected VPNs
Reference Material

Books:
"BGP4 Inter-Domain Routing in the Internet" by John Stewart ISBN 0-201-37951-1
"Internet Routing Architectures" by Bassam Halabi ISBN 1-56205-652-2
"Interconnections: Bridges and Routers" by Radia Perlman ISBN 0-13-468505-9
"Routing in the Internet" by Christian Huitema ISBN 0-13-132192-7

Mail Lists:
SSR mailinglist - majordomo@cabletron.com
GateD mailinglists - See www.gated.org
North American Network Operators Group (NANOG) mailist - See www.merit.org

Reference Material

Request For Comments - RFCs

08/98 - RFC2385PS "Protection of BGP Sessions via the TCP MD5 Signature Option"
02/98 - RFC 2283PS "Multiprotocol Extensions for BGP-4"
01/97 - RFC 2042 "Registering New BGP Attribute Types"
08/96 - RFC 1998 "An Application of the BGP Community Attribute in Multi-home Routing"
08/96 - RFC 1997 "BGP Communities Attribute"
06/96 - RFC 1966 "BGP Route Reflection An alternative to full mesh"
06/96 - RFC 1965 "Autonomous System Confederations for BGP"
10/95 - RFC 1863 "A BGP/IDRP Route Server alternative to a full mesh routing"
08/95 - RFC 1817 "CIDR and Classful Routing"
03/95 - RFC 1774 "BGP-4 Protocol Analysis"
03/95 - RFC 1773 "Experience with the BGP-4 protocol"
03/95 - RFC 1772 "Application of the Border Gateway Protocol in the Internet"
03/95 - RFC 1771 "A Border Gateway Protocol 4 (BGP-4)"
12/94 - RFC 1745 "BGP4/IDRP for IP---OSPF Interaction"
07/94 - RFC 1657 "Definitions of Managed Objects for BGP-4 using SMIv2"
09/93 - RFC 1520 "Exchanging Routing Information Across Provider Boundaries in CIDR"
09/93 - RFC 1519 "CIDR; an Address Assignment and Aggregation Strategy"
09/93 - RFC 1518 "An Architecture for IP Address Allocation with CIDR"
Reference Material

Internet Drafts

08/98 "LDP Specification"
08/98 "Border Gateway Multicast Protocol (BGMP): Protocol Specification"
08/98 "A Framework for Inter-Domain Route Aggregation"
08/98 "Routing Policy Configuration Language (RPCL)"
08/98 "Carrying Label Information in BGP-4"
08/98 "Capabilities Negotiation with BGP-4"
08/98 "BGP Security Analysis"
08/98 "A Border Gateway Protocol 4 (BGP-4)"
07/98 "Using RPSL in Practice"
07/98 "Multiprotocol Label Switching Architecture"
06/98 "NHRP for Destinations off the NBMA Subnetwork"
05/98 "BGP Route Flap Damping"
04/98 "BGP-4 Capabilities Negotiation for BGP Multiprotocol Extensions"
03/98 "To Be Multihomed: Requirements & Definitions"
03/98 "BGP-4 over ATM and Proxy PAR"
02/98 "Use of BGP-4 Multiprotocol Extensions for IPv6 Inter-Domain Routing"
02/98 "Carrying Label Information in BGP-4"
01/98 "DNS-base NLRI origin AS verification in BGP"

Section 3

Layer 2 VPNs
MPLS VPN Tutorial Agenda

Layer 2 VPNs
- IETF PWE3 and L2VPN WG update
- Encapsulation and Label Stacking
- Virtual Private Wire Services – VPWS
  - Pt-to-pt Ethernet, Pt-to-pt ATM, Pt-to-pt Frame Relay
  - Virtual Private LAN Services – VPLS

MPLS L2 VPN Market Drivers

What can we conclude?

- Layer 3 IP is not the only traffic
  - Still a lot of legacy SNA, IPX etc
  - Large enterprises have legacy protocols
- Layer 3 IP VPNs are not the whole answer
  - IP VPNs cannot handle legacy traffic
- Layer 2 legacy traffic widely deployed

Carriers need to support
Layer 2 and Layer 3 VPNs
MPLS Layer 2 VPNs

- **Point-to-point layer 2 solutions**
  - Virtual Private Wire Services - VPWS
  - Similar to ATM / FR services, uses tunnels and connections (LSPs)
  - Customer gets connectivity only from provider
  - Ongoing work to encapsulate Ethernet, ATM, FR, TDM, SONET, etc

- **Multi-point layer 2 solutions**
  - Virtual Private LAN Services - VPLS
  - Virtual Private LAN Services aka Transparent LAN Service (TLS)
  - Ethernet Metro VLANs / TLS over MPLS
  - Independent of underlying core transport
  - All drafts “currently” support PWE3 (Martini) Ethernet encapsulation
  - Differences in drafts for discovery and signaling

MPLS Point-to-Point Service

- Tunnel Label determines path through network
- VC Label identifies VLAN, VPN, or connection at the end point
MPLS Point-to-Point Services

Label Stacking

Three Layers of Encapsulation

1) Tunnel Header: Contains information needed to transport the PDU across the IP or MPLS network
2) Demultiplexer Field: Used to distinguish individual emulated VCs within a single tunnel
3) Emulated VC Encapsulation: Contains the information about the enclosed PDU (known as Control Word)

Tunnel Header determines path through network
Demultiplexer Field identifies VLAN, VPN, or connection at the end point
All services look like a Virtual Circuit to MPLS network

Tunnel Header | Demux Field | VC Encaps Information | Layer 2 payload
---|---|---|---
1 | 2 | 3 |

Encaps Information Field

Layer 2 header fields may be discarded at ingress
Control word carries “flag” bits depending on encapsulation
- (FR: FECN, BECN, C/R, DE, ATM: CLP, EFCI, C/R, etc)
Length required when padding small frames on links which have a minimum frame size
Sequence number is optional. It is used to detect out of order delivery of frames.

Control Word for PW Associated Channel

Draft-ietf-pwe3-cw-06.txt
### LDP - Label Mapping Message

<table>
<thead>
<tr>
<th>Label Mapping</th>
<th>Message Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>Message ID</td>
<td></td>
</tr>
<tr>
<td><strong>FEC TLV</strong></td>
<td></td>
</tr>
<tr>
<td>Label TLV</td>
<td></td>
</tr>
<tr>
<td>Label Request Message ID TLV</td>
<td></td>
</tr>
<tr>
<td>LSPID TLV (optional)</td>
<td></td>
</tr>
<tr>
<td>Traffic TLV (optional)</td>
<td></td>
</tr>
</tbody>
</table>

### New VC FEC Element Defined

<table>
<thead>
<tr>
<th>VC TLV</th>
<th>C</th>
<th>VC Type</th>
<th>VC Info Length</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Group ID</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>VC ID</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Interface Parameters</td>
<td></td>
</tr>
</tbody>
</table>

- **Virtual Circuit FEC Element**
  - C - Control Word present
  - VC Type - FR, ATM, Ethernet, HDLC, PPP, ATM cell
  - VC Info Length - length of VCID field
  - Group ID - user configured - group of VCs representing port or tunnel index
  - VC ID - used with VC type to identify unique VC
  - Interface Parameters - Specific I/O parameters
Layer 2 Encapsulation
Ongoing work in PWE3

- RFC 3916: Requirements for PWE3
  - "This document describes base requirements for the Pseudo-Wire Emulation Edge to Edge Working Group (PWE3 WG). It provides guidelines for other working group documents that will define mechanisms for providing pseudo-wire emulation of Ethernet, ATM, Frame Relay."

- RFC 3985: PWE3 Architecture
  - "This document describes an architecture for Pseudo Wire Emulation Edge-to-Edge (PWE3). It discusses the emulation of services (such as Frame Relay, ATM, Ethernet TDM and SONET/SDH) over packet switched networks (PSNs) using IP or MPLS. It presents the architectural framework for pseudo wires (PWs), defines terminology, specifies the various protocol elements and their functions."

Layer 2 Encapsulation PWE3
WG documents (original Martini work)

- Pseudowire Set-up and Maintenance using LDP
  - draft-ietf-pwe3-control-protocol-17.txt – June 05

- ATM AAL5 and ATM cell
  - draft-ietf-pwe3-atm-encap-10.txt – Sept 05

- Frame Relay
  - draft-ietf-pwe3-frame-relay-06.txt – June 05

- Ethernet / 802.1q VLAN
  - draft-ietf-pwe3-ethernet-encap-11.txt – June 05

- PPP/HDLC
  - draft-martini-ppp-hdlc-encap-mpls-00.txt
MPLS Ethernet Encapsulation

draft-ietf-pwe3-ethernet-encap-11.txt

Original Ethernet frame

<table>
<thead>
<tr>
<th>Preamble</th>
<th>DA</th>
<th>SA</th>
<th>802.1q</th>
<th>L</th>
<th>payload</th>
<th>FCS</th>
</tr>
</thead>
</table>

Encapsulated Ethernet over MPLS over Ethernet Transport

- Ingress device strips the Ethernet preamble and FCS
- Ethernet header becomes “control word”
- New MPLS Ethernet header (type 0x8847) and new FCS is added to MPLS Ethernet packet

Martini VC Encaps field is normally 32 bits
- Ethernet VC Encaps field equals
  - DA and SA at 6 bytes each
  - Length at 2 bytes
  - 802.1q VLAN at 4 bytes
  - Total at 18 bytes (144 bits vs 32 bits)
Life of a Frame
Ethernet over Ethernet MPLS

- Last Mile
- Provider's MPLS Backbone
- CPE
- Penultimate Hop LSR
- PE
- payload
- FCS
- DA
- SA
- 802.1q
- Tunnel Label
- VC Label
- DA
- SA
- 802.1q
- payload
- FCS

ATM and Frame Relay Service
Reference Model

- PSEUDO WIRE
- PSN TUNNEL
- ATM or FR Service
- CE-1
- CE-2
- UNI or NNI
- PE1
- PE2
- Emulated Service
- PE = Provider Edge
- CE = Customer Edge

Requirements for Pseudo Wire Emulation
Edge-to Edge (PWE3)
ATM AAL5 Encapsulation

draft-ietf-pwe3-atm-encap-10.txt

- Ingress reassembles AAL5 frames and strips 8 octet AAL5 trailer
- Required control word includes:
  - T = Transport type bit
  - Common Part Convergence Sublayer-Protocol Data Unit (AAL5 CPCS-PDU)
  - Or ATM Cell
  - E = EFCl bit - Efficient Forward Congestion
  - L = CLP bit - Cell Loss Priority
  - C = Command / Response bit

ATM Control Word

<table>
<thead>
<tr>
<th>Tunnel Header</th>
<th>Demux Field</th>
<th>Control word</th>
<th>AAL5 CPCS-PDU</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

ATM Cell Mode Encapsulation

draft-ietf-pwe3-atm-encap-10.txt

- Ingress performs no reassembly
- Control word is optional:
  - Length may be used to infer number of cells
  - Flags set to zero

ATM Control Word

<table>
<thead>
<tr>
<th>Tunnel Header</th>
<th>Demux Field</th>
<th>Control word</th>
<th>ATM cell #1 minus FCS</th>
<th>ATM cell #2 minus FCS</th>
<th>...</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Tunnel Header</th>
<th>Demux Field</th>
<th>Control word</th>
<th>ATM cell #1 minus FCS</th>
<th>ATM cell #2 minus FCS</th>
<th>...</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
MPLS PWE3 FR Encapsulation

draft-ietf-pwe3-frame-relay-06.txt

- Main Functions: FR over Pseudo Wire - FRoPW
  - Encapsulation of FR specific information in a suitable FRoPW packet (ingress function)
  - Transfer of a FRoPW packet through IP / MPLS network
  - Extraction of FR specific information from a FRoPW packet (egress function)
  - Generation of native FR frames at egress
  - Other operations to support FR services

- Two Mapping modes defined between FR VCs and FR PWs
  - One-to-one mapping
  - One FR VC mapped to a pair of unidirectional PWs
MPLS PWE3 FR Encapsulation

draft-ietf-pwe3-frame-relay-06.txt

- Many Bi-directional FR VC
- End-to-end FR VCs
- Pair of Uni-directional PW LSPs
- Many Bi-directional FR VC

Pseudo Wire Emulated Service

- Two Mapping modes defined between FR VCs and FR PWs
  - Many-to-one or port mode mapping (Optional – w / header)
  - Many FR VCs mapped to a pair of Unidirectional PWs

MPLS Frame Relay Encapsulation

draft-ietf-pwe3-frame-relay-06.txt

- F = FECN (Forward Explicit Congestion Notification)
- B = BECN (Backward Explicit Congestion Notification)
- D = DE (Discard Eligibility Indicator)
- C = C/R (Command / Response Field)
MPLS VPN Tutorial Agenda

Layer 2 VPNs
- IETF PWE3 and L2VPN WG update
- Encapsulation and Label Stacking
- Virtual Private Wire Services – VPWS
  - Pt-to-pt Ethernet, Pt-to-pt ATM, Pt-to-pt Frame Relay
  - Virtual Private LAN Services – VPLS

IETF Layer 2 VPNs

- draft-ietf-l2vpn-requirements-01.txt – Nov 05
  - Provides requirements for Layer 2 Provider Provisioned Virtual Private Networks (L2VPNs) Provides taxonomy and terminology and states generic and general services requirements. It covers point-to-point VPNs referred to as Virtual Private Wire Services (VPWS), as well as multipoint-to-multipoint VPNs as known as Virtual Private LAN services (VPLS)
  - This document provides a framework for Layer 2 Provider Provisioned Virtual Private Networks (L2VPNs). This framework is intended to aid in standardizing protocols and mechanisms to support interoperable L2VPNs.
### MPLS VPLS Architecture

Distributed PE functions
- PE-POP = PE at SP POP
- PE-CLE = PE at customer site

### MPLS VPLS VPLS Reference Model

Multi Tenant Unit(s) (MTU-s)
**MPLS VPLS**

**VPLS Reference Model**

- Tunnel LSPs are established between PEs.
- Customer Virtual Private LANs are tunneled through MPLS network.

**Virtual Private LAN Services**

*draft-ietf-l2vpn-vpls-ldp-08.txt*

- Updated November 2005
- Defines an Ethernet (IEEE 802.1D) learning bridge model over MPLS Martini Ethernet circuits
- Defines the LER (PE) function for an MPLS VPLS network
- Creates a layer 2 broadcast domain for a closed group of users
- MAC address learning and aging on a per LSP basis
- Packet replication across LSPs for multicast, broadcast, and unknown unicast traffic
- Includes Hierarchical VPLS
  - formerly draft-khandekar-ppvnp-hvpls-mpls-00.txt
Virtual Private LAN Services

draft-ietf-l2vpn-vpls-ldp-08.txt

- Tunnel LSPs are established between PEs
- Customers designated C1 and C2 are part of two independent Virtual Private LANs
- Layer 2 VC LSPs are set up in Tunnel LSPs
- Core MPLS network acts as a LAN switch

Virtual Private LAN Services

draft-ietf-l2vpn-vpls-ldp-08.txt

- Reduces signaling and packet replication to allow large scale deployment of VPLS
- Uses Martini VC / LSPs between edge MTU and VPLS aware PE devices

VC-1 = Single pt-to-pt Martini Tunnel LSP
MTU-s = Bridging Capable MTU (Multi Tenant Unit)
PE-rs = VPLS Capable PE
        = Virtual VPLS (Bridge) Instance
VPLS Internal PE Architecture

Attachment circuit

CE

PE

IEEE 802.1D bridging code
IETF VPLS code

VPLS Code

- VPLS Forwarding
  - Learns MAC addresses per pseudo-wire (VC LSP)
  - Forwarding based on MAC addresses
  - Replicates multicast & broadcast frames
  - Floods unknown frames
  - Split-horizon for loop prevention
- VPLS Signaling
  - Establishes pseudo-wires per VPLS between relevant PEs
- VPLS Discovery (Manual, LDP, BGP, DNS)
Bridging Code

- Standard IEEE 802.1D code
  - Used to interface with customer facing ports
  - Might run STP with CE
  - Used to interface with VPLS
  - Might run STP between PEs

VPLS Scalability

Parameters

- Number of MAC Addresses
- Number of replications
- Number of LSPs
- Number of VPLS instances
- Number of LDP peers
- Number of PEs
VPLS Scalability
Signaling Overhead – Flat Topology

- Architecture has a direct impact on the Signaling Overhead (control plane)

VPLS Scalability
Signaling Overhead – Hierarchical Topology

- Architecture has a direct impact on the Signaling Overhead (control plane)
VPLS Scalability

Replication Overhead – **Flat Topology**

- Architecture has a direct impact on Replication Overhead (forwarding plane)

VPLS Scalability

Replication Overhead – **Hierarchical Topology**

- Architecture has a direct impact on Replication Overhead (forwarding plane)
VPLS Scalability
Adding a New Site – Flat Topology

- Architecture affects Provisioning & Signaling between all nodes

VPLS Scalability
Adding a New Site – Hierarchical Topology

- Architecture affects Provisioning & Signaling between all nodes
• Architecture has a direct impact on ability to offer Inter-Metro Service
VPLS Scalability

FIB Size

- VPLS FIB size depends on the type of Service Offering:
  - Multi-protocol Inter-connect service
    - Mimics the DSL Tariff Model
    - Customers are charged per site per block of MAC addresses
  - Router Inter-connect
    - One MAC address per site
- Same Network Design principles apply for
  - MAC FIB Size of VPLS Service and,
  - Route Table Size of Virtual Private Routed Network (VPRN) Service

MPLS VPNs Summary

- Layer 2 versus Layer 3
  - Apples and Oranges
- Layer 3 MPLS VPNs
  - Deployed with Internet Draft 2547bis
- Layer 2 MPLS VPNs
  - Lots of Interest from Carriers and Vendors
  - Many new drafts – lots of consolidation
  - We are in “concept” stage
  - Solutions available
Section 4

Introduction to Multi-Service Interworking

Why Interwork?

- Carriers want a common edge infrastructure to support and “Interwork” with legacy and new services
- Carriers want to support all legacy transports technologies and services
- Carriers are planning to converge on an IP / MPLS core
- Carriers want to seamlessly introduce Metro Ethernet services and IP VPNs
**Interworking**

**History**

- The Frame Relay Forum defined the **Network Interworking** function between Frame Relay and ATM in the FRF.5 document finalized in 1994.
- The Frame Relay Forum defined the **Service interworking** function between Frame Relay and ATM in the FRF.8.1 document finalized in 2000.
- **Why define FR and ATM interworking?**
  - ATM cores with FR access services deployed
  - ATM and Frame Relay circuits are point-to-point
  - Both data links have services that are somewhat similar in nature even though the signaling is different

---

**InterWorking Function - IWF**

*Network vs Service IWF*

**Network Interworking**

- **Network Interworking** is used when one protocol is “tunneled” across another “intermediary” network / protocol
- The **Network Interworking** function “terminates” and “encapsulates” the protocol over a Pt-to-Pt connection

**Service Interworking**

- **Service Interworking** is required to “translate” one protocol to another protocol – used between two unlike protocols
- The **Service Interworking** function “translates” the control protocol information transparently by an interworking function (IWF)
**Network Interworking FRF.5**

*Reference Model*

- Frame Relay Service - FRS
- ATM
- Emulated FR Service

- *Network* Interworking *encapsulates* the L2 Service
- FRS is *encapsulated* and sent across ATM network
- Service at end points *has to be* the same

---

**Service Interworking FRF.8.1**

*Reference Model*

- Frame Relay Service
- ATM Service
- IWF
- ATM
- IWF

- *Service* Interworking *translates* the L2 Service
- FR service is *translated* into ATM service
- Services at the end points *are not* the same
Why not continue with ATM IW?

- ATM is optimized for voice transport – cell overhead etc
- Cells are simply fixed length packets and can be carried unchanged across an MPLS network
- Packets are not cells and must be adapted to be carried across ATM
- MPLS is optimized for packet transport
- Carriers want to converge on IP/MPLS cores supporting both new and legacy services

Why Migrate to MPLS?

- MPLS allows service providers to converge onto a single infrastructure while offering existing services
- MPLS enables new service offerings and simplifies service provisioning
- MPLS supports rapid growth in IP applications and services
- MPLS allows the integration of services management into a common OSS strategy
- MPLS supports the integration of packet technologies and optical cores
MPLS Multi-Service Interworking

MPLS Connects Services at the Edge

- ATM
- FR
- TDM
- Ethernet
- IP / MPLS

MPLS Network Interworking

IETF PWE3 Pt-to-Pt Encapsulation

Point-to-point tunnels

- ATM to ATM
- FR to FR
- Ethernet to Ethernet

Service has to be pt-to-pt between like services: ATM to ATM, FR to FR, Enet to Enet, etc
**MPLS Multi-Service Interworking**

**Reference Model**

- **PE** = Provider Edge
- **CE** = Customer Edge
- **PSN** = Packet Switched Network
- **IWF** = InterWorking Function
- **Multi-Service**: Services equal to FR, ATM, Ethernet

![Diagram](image)

**MPLS as a Service Enabler**

- **SERVICE**: VPLS, VPWS, FR / ATM, IP VPNs, TDM / Voice
- **SERVICE ENABLER**: MPLS (QoS, TE, FRR)
- **TRANSMISSION**: SDH / SONET, Other

VPLS = Virtual Private LAN Services
VPWS = Virtual Private Wire Services
L3 IP VPN = RFC2547-bis
For More Information. . .

- http://www.mfaforum.org
- http://www.ietf.org
- http://www.itu.int
- http://www.mplsrc.com

For questions, utilize the MFA Forum Message Board
Website: http://www.mfaforum.org/board/

Thank you for attending the

MPLS based Virtual Private Network Services Tutorial