Internet Draft
Document: draft-lasserre-tls-mpls-00.txt

Marc Lasserre
Nick Slabakov
Rob Nath
Riverstone Networks

Pascal Menezes
Terabeam Networks

Loa Andersson
Utfors

Andrew Smith
Consultant

Shahid Akhtar
Ciena

Tissa Sevenirathne
Force10 Networks

Pierre Lin
Yipes Communication

Lewis Eatherton
Excite@Home

Vasile Radoaca
Nortel Networks

Ivy Hsu
Foundry Networks

Expires: February 2002 August 2001

Transparent VLAN Services over MPLS

Status of this Memo

This document is an Internet-Draft and is in full conformance with all provisions of Section 10 of RFC2026.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at
http://www.ietf.org/ietf/1id-abstracts.txt
The list of Internet-Draft Shadow Directories can be accessed at
http://www.ietf.org/shadow.html.

Abstract

This document describes a transparent virtual LAN service (VLS) solution over MPLS, sometimes known as Transparent LAN Services (TLS). VLS simulates an Ethernet virtual 802.1d bridge [6] [7] for a given set of users. It delivers a layer 2 broadcast domain that is fully capable of learning and forwarding on Ethernet MAC addresses that is closed to a given set of users. Many VLS services can be supported from a single PE node.

Conventions

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119

Placement of this Memo in Sub-IP Area

RELATED DOCUMENTS

http:// search.ietf.org/internet-drafts/draft-martini-l2circuit-trans-mpls-06.txt

http://search.ietf.org/internet-drafts/draft-martini-l2circuit-encap-mpls-02.txt

WHERE DOES THIS FIT IN THE PICTURE OF THE SUB-IP WORK

PPVPN

WHY IS IT TARGETTED AT THIS WG

The charter of the PPVPN WG includes L2 VPN services and this draft specifies a model for Ethernet L2 VPN services over MPLS.

JUSTIFICATION

Existing Internet drafts specify how to provide point-to-point Ethernet L2 VPN services over MPLS. This draft defines how multipoint Ethernet services can be provided.

Table of Contents

Status of this Memo....1
Abstract....2
Conventions....2
Table of Contents....3

1. Overview....4
2. Bridging Model for MPLS....4
2.1 Flooding and Forwarding....5
2.2 Address Learning....6
2.3 LSP Topology....6
2.4 Loop free L2 VPN....6
2.5 L2 VPN Provisioning....7
2.6 LDP Based Discovery....7
3. Security Considerations....8
4. References....9
5. Author's Addresses....10

1. Overview

The following discussion applies to devices that serve as Label Edge Routers (LERs) on a MPLS network that is VLS capable. It will not discuss the behavior of transit Label Switch Routers (LSRs) that are considered a part of MPLS network. The MPLS network provides a number of Label Switch Paths (LSPs) that form the basis for connections between LERs attached to the same MPLS network. The resulting set of interconnected LERs forms a private MPLS VPN where each LSP is uniquely identified at each MPLS interface by a label.

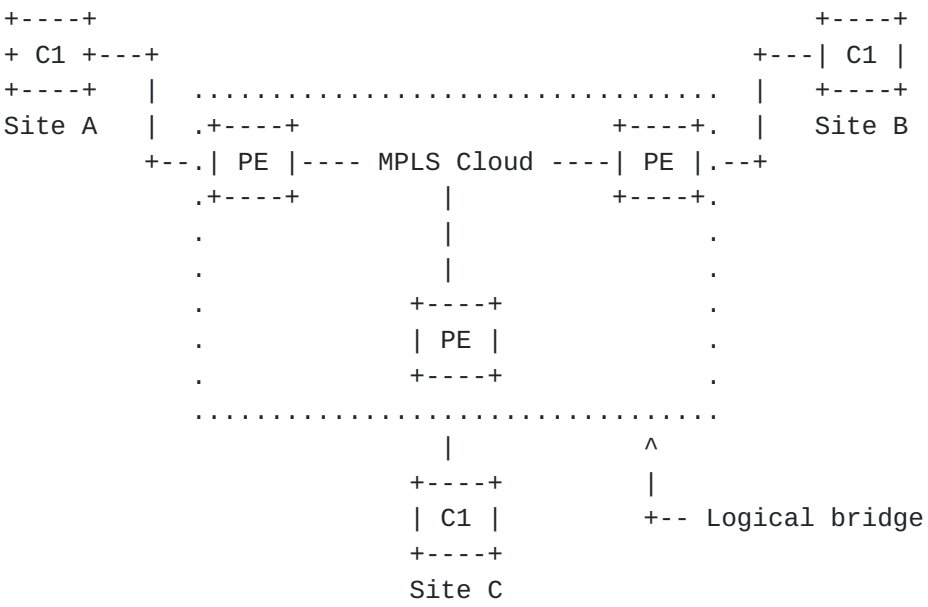
Ethernet has become a predominant technology initially for Local Area Networks (LANs) and now as an access technology, specifically in metropolitan networks. Ethernet ports or IEEE VLANs are dedicated to customers on Provider Edge (PE) routers acting as LERs. Customer traffic gets mapped to a specific MPLS L2 VPN by configuring L2 FECs based upon the input port and/or VLAN.

Broadcast and multicast services are available over traditional LANs. MPLS does not support such services currently. Sites that belong to the same broadcast domain and that are connected via an MPLS network expect broadcast, multicast and unicast traffic to be forwarded to the proper location(s). This requires MAC address learning/aging on a per LSP basis, packet replication across LSPs for multicast/broadcast traffic and for flooding of unknown unicast destination traffic.

[1] defines how to carry L2 PDUs over point-to-point MPLS LSPs. This document describes extensions to [1] for transporting Ethernet/802.3 and VLAN [8] traffic across multiple sites that belong to the same L2 broadcast domain. Note that the same model can be applied to other 802.1 technologies. It describes a simple and scalable way to offer Virtual LAN services, including the appropriate flooding of Broadcast, Multicast and unknown unicast destination traffic over MPLS, without the need for address resolution servers or other external servers, as discussed in [10].

2. Bridging Model for MPLS

An MPLS interface acting as a bridge must be able to flood, forward, and filter bridged frames.



The set of PE devices interconnected via MPLS appears as a single 802.1d bridge/switch to customer C1. Each PE device will learn remote MAC addresses on LSPs (and keeps learning directly attached MAC addresses on customer facing ports).

2.1 Flooding and Forwarding

Flooding is performed by sending unknown unicast and multicast frames to all possible appropriate destinations. In the MPLS environment this means sending the PDU through each relevant VC LSP. This is accomplished by explicitly copying it to each VC LSP that is part of the corresponding VPN.

Note that multicast frames do not necessarily have to be sent to all VPN members. For simplicity, the default approach of broadcasting multicast frames can be used. Extensions explaining how to interact with 802.1 GMRP protocol, IGMP snooping and static MAC multicast filters will be discussed in a future revision.

To forward a frame, a bridge must be able to associate a destination MAC address with a VC LSP. It is unreasonable and perhaps impossible to require bridges to statically configure an association of every possible destination MAC address with a VC LSP. Therefore, MPLS bridges must provide enough information to allow an MPLS interface to dynamically learn about foreign destinations beyond the set of LSRs. To accomplish dynamic learning, a bridged PDU MUST conform to the encapsulation described within [1].

2.2 Address Learning

Unlike BGP VPNs [8], reachability information does not need to be advertised and distributed via a control plane. Reachability is obtained by standard learning bridge functions in the data plane.

Since VC LSPs are uni-directional, two LSPs of opposite directions are required to form a logical bi-directional link. When a new MAC address is learned on an inbound LSP, it needs to be associated with the outbound LSP that is part of the same pair. The state of this logical link can be considered as up as soon as both incoming and outgoing LSPs are established. Similarly, it can be considered as down as soon as one of these two LSPs is torn down.

2.3 LSP Topology

PE routers run either an IGP or an EGP between them. Tunnel LSPs between PE routers are therefore established along routed paths. Note that tunnel LSPs can also be explicitly routed. Such tunnels form a full mesh. Partial mesh of tunnel LSPs will be discussed in a future revision. VC LSPs are then mapped onto these tunnel LSPs. The resulting VC LSP mesh for the corresponding VPN instance has to be loop free.

In a Ethernet based topology, since VC LSPs are not terminated at the CE boundary, unlike Frame Relay or ATM, it is the responsibility of the Service Providers to offer a loop free topology. With Frame Relay or ATM, it is the customers' responsibility to run STP or a routing protocol to prevent loops. With Ethernet as the access medium, a port and/or a VLAN is used per customer. Customer facing ports can be used to tunnel untagged or 802.1q tagged traffic. The VC LSPs connected to each site in the corresponding VPN are only visible to the PE device.

2.4 Loop free L2 VPN

In order to avoid running a STP instance per VPN, which would not scale, partial mesh configurations of VC LSPs are not allowed. Note that customers are allowed to run STP such as when a customer has a back door link used for backup. In such a case STP BPDUs are simply tunneled through the MPLS cloud.

Each PE MUST create a rooted tree to every other PE router that serve the same L2 VPN. Each PE MUST support a "split-horizon" scheme in order to prevent loops, that is, a PE MUST NOT forward traffic from one VC LSP to another in the same VPN (since each PE has direct connectivity to all other PEs in the same VPN).

2.5 L2 VPN Provisioning

Provisioning of a full mesh can be automatic with a discovery protocol where each PE advertizes which VPN(s) it serves to other PEs in the same domain, reducing the provisioning complexity to O(1) PE to configure when a new site is added, and O(n) new LSPs to be set up.

The number of sites per VPN and the number of VPNs per PE router define the total number of VC LSPs to be managed. Let's consider for example that each VPN has an average of 4 sites and that 1000 customers are supported in a PE, 4000 VC LSPs need to be established and maintained. Since VC LSPs are set up via LDP, there is no need to refresh LSP states like in the RSVP case. Tunnel LSPs can be either be established via LDP or via RSVP.

Since LDP is required for establishing VC LSPs, LDP is a logical choice to exchange VPN membership between PE routers. BGP offers the advantage of being able to set up inter-provider VPNs. However, BGP is typically not enabled on PE routers, but only on core routers.

The signaling of VPN membership via BGP will be discussed in a future revision. A possible method for exchanging VPN membership via BGP can be based on [5].

2.6 LDP Based Discovery

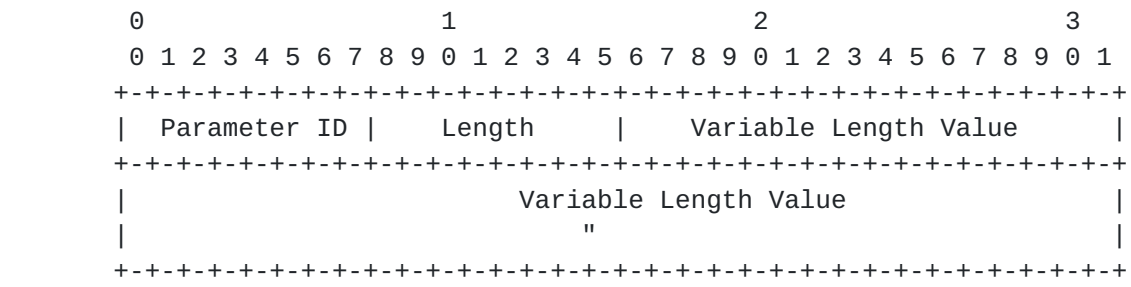
Once an LDP adjacency has been formed between two PEs, all VC LSPs get established over this single TCP session.

Directly connected PEs multicast LDP hellos periodically. Non-directly connected PEs exchanged unicast targeted hellos to each desired peer. Once a hello adjacency is formed, a LDP TCP connection is established. A LDP session is established over this connection by exchanging LDP initialization messages.

IGP extensions to automatically discover VLS capable PEs can be used, such as defined in [4]. This will allow a TLS capable PE node to automatically setup a LDP adjacency to the discovered node and automate provisioning for VC LSPs. If LDP is used to create tunnel LSPs, then this will be automated once the newly discovered TLS PE node has an IP entry in the IGP (this assumes that a loopback address is used to represent the newly discovered TLS capable PE node). However if RSVP-TE is used for the tunnel LSP, then it is important that the two PE nodes have a tunnel LSP between them (the means of doing this is beyond the scope of this document).

In [2], the L2 VPN information is carried in a Label Mapping message sent in downstream unsolicited mode. This document defines a new

parameter id, a 7-byte VPN Id as defined in [9], to the interface parameters field in the VC FEC described in [2].



The parameter ID is defined as follows:

Parameter ID	Length	Description
0x01	4	Interface MTU in octets.
0x02	4	Maximum Number of concatenated ATM cells.
0x03	up to 82	Optional Interface Description string.
0x04	4	CEM [8] Payload Bytes.
0x05	4	CEM options.
0x06	7	VPN Id.

The first five parameters are as described in [2].

The VPN Id field is a unique 7-byte number that identifies a specific L2 VPN instance in a service provider network. All VLS capable PE nodes MUST use the same VPN ID for a given L2 VPN. PE nodes belonging to the same VLS must be capable of mapping Ethernet ports and/or VLANs to the corresponding VPN Id.

3. Security Considerations

This document does not affect the underlying security issues of MPLS.

4. References

[1] "Encapsulation Methods for Transport of Layer 2 Frames Over MPLS", draft-martini-l2circuit-encap-mpls-02.txt (Work in progress)

[2] "Transport of Layer 2 Frames Over MPLS", draft-martini-l2circuit-trans-mpls-06.txt (Work in progress)

[3] "LAN Emulation over ATM version 1.0", af-lane-0021.0000 (ATM Forum)

[4] "Distribution of 802.1Q VLAN information using Opaque LSA", draft-tsenevir-8021qospf-00.txt (Work in progress)

[5] "Use of BGP-MP for Layer 2 VPN Membership discovery", draft-tsenevir-bgp-l2vpn-00.txt (Work in progress)

[6] Original 802.1D - ISO/IEC 10038, ANSI/IEEE Std 802.1D-1993 "MAC Bridges".

[7] 802.1D - "Information technology - Telecommunications and information exchange between systems - Local and metropolitan area networks - Common specifications - Part 3: Media Access Control (MAC) Bridges: Revision. This is a revision of ISO/IEC 10038: 1993, 802.1j-1992 and 802.6k-1992. It incorporates P802.11c, P802.1p and P802.12e." ISO/IEC 15802-3: 1998.

[8] 802.1Q - ANSI/IEEE Draft Standard P802.1Q/D11, "IEEE Standards for Local and Metropolitan Area Networks: Virtual Bridged Local Area Networks", July 1998.

[9] Fox, Gleeson, "Virtual Private Networks Identifier", RFC2685

[10] "Requirements for Network Based Layer 2 VPN", draft-tsenevir-l2-req-00.txt (Work in progress)

5. Author's Addresses

Marc Lasserre
Riverstone Networks
5200 Great America Pkwy Phone: 1-408-878-6550
Santa Clara, CA 95054 Email: marc@riverstonenet.com

Nick Slabakov
Riverstone Networks
5200 Great America Pkwy Phone: 1-303-471-6926
Santa Clara, CA 95054 Email: nslabakov@riverstonenet.com

Rob Nath
Riverstone Networks
5200 Great America Pkwy Phone: 1-408-878-6742
Santa Clara, CA 95054 Email: rnath@riverstonenet.com

Loa Andersson
Utfors Bredband AB Phone: +46 8 5270 50 38
Rasundavagen 12 169 29 Solna Email: loa.andersson@utfors.se

Pascal Menezes
TeraBeam Networks
2300 Seventh Ave
Seattle, WA 98121 Email: Pascal.Menezes@Terabeam.com

Andrew Smith Fax: +1 415 345 1827
Consultant Email: ah_smith@pacbell.net

Tissa Senevirathne
Force10 Networks
1440 McCarthy Blvd Phone: 408-865-5103
Milpitas, CA Email: tsenevir@hotmail.com

Pierre Lin
Yipes Communication
114 Sansome St Phone: 415-218-9520
San Francisco, CA 94104 Email: pierre.lin@yipes.com

Lewis Eatherton
Excite@Home
450 Broadway Street Phone: 650-556-5022
Redwood City, CA 94063 Email: leathert@excitehome.net

Vasile Radoaca
Nortel Networks
600 Technology Park Phone: 978-288-6097
Billerica MA 01821 Email: vasile@nortelnetworks.com

Foundry Networks
2100 Gold Street
PO Box 649100 Phone: 408-586-1795
San Jose, CA 95164 Email: ihsu@foundrynet.com