IP-Only LAN Service (IPLS)

Status of this Memo

By submitting this Internet-Draft, each author represents that any applicable patent or other IPR claims of which he or she is aware have been or will be disclosed, and any of which he or she becomes aware will be disclosed, in accordance with Section 6 of BCP 79.

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
This Internet-Draft will expire on January 2006.

Copyright Notice
Copyright (C) The Internet Society (2005). All Rights Reserved.
For potential updates to the above required-text see: http://www.ietf.org/ietf/1id-guidelines.txt

Abstract

A Virtual Private LAN Service (VPLS) [VPLS] is used to interconnect systems across a wide-area or metropolitan-area network, making it
appear that they are on a private LAN. The systems which are interconnected may themselves be LAN switches. If, however, they are IP hosts or IP routers, certain simplifications to the operation of the VPLS are possible. We call this simplified type of VPLS an "IP-only LAN Service" (IPLS). In an IPLS, as in a VPLS, LAN interfaces are run in promiscuous mode, and frames are forwarded based on their destination MAC addresses. However, the maintenance of the MAC forwarding tables is done via signaling, rather than via the MAC address learning procedures specified in [IEEE 802.1D]. This draft specifies the protocol extensions and procedures for support of the IPLS service.

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

Table of Contents

Internet Draft draft-ietf-l2vpn-ipls-04.txt

Status of this Memo..................................................1
Abstract............................................................................1
Table of Contents......................................................2
1.0 Contributing Authors.............................................3
2.0 Overview.....................................................................3
2.1 Terminology.............................................................6
3.0 Topology.....................................................................7
4.0 Configuration...........................................................8
5.0 Discovery....................................................................9
5.1 CE discovery............................................................9
6.0 Pseudowire Creation................................................9
6.1 Receive Unicast Multipoint-to-point Pseudowire................9
6.3 Send Multicast Replication tree..................................10
7.0 Signaling....................................................................11
7.1 IPLS PW Signaling..................................................11
7.2 Signaling Advertisement Processing............................13
7.3 IANA Considerations for LDP Status Code......................13
8.0 Forwarding............................................................13
8.1 Non-IP or non-ARP traffic........................................14
8.2 Unicast IP Traffic....................................................14
8.3 Broadcasts and Multicast IP Traffic.........................14
8.4 ARP Traffic..........................................................14
8.5 Encapsulation........................................................15
9.0 Attaching to IPLS via ATM or FR............................15
10.0 VPLS vs IPLS......................................................16
11.0 IP Protocols........................................................16
12.0 Dual Homing with IPLS........................................17
13.0 Acknowledgements................................................17
14.0 Security Considerations..........................................17
15.0 References..........................................................17
15.1 Normative References...........................................17
15.2 Informative References.........................................18
16.0 Author’s Address................................................18
Intellectual Property Statement......................................19
1.0 Contributing Authors
This document is the combined effort of the following individuals and many others who have carefully reviewed this document and provided the technical clarifications

K. Arvind                    Enterasys Networks
Vach Kompella                Alcatel
Vasille Radoaca              Consultant

2.0 Overview

As emphasized in [VPLS], Ethernet has become popular as an access technology in Metropolitan and Wide Area Networks. [VPLS] describes how geographically dispersed customer LANs can be interconnected over a service provider’s network. The VPLS service is provided by Provider Edge (PE) devices that connect Customer Edge (CE) devices.

The VPLS architecture provides this service by incorporating bridging functions such as MAC address learning in the PE devices.

Provider Edge platforms are designed primarily to be IP routers, rather than to be LAN switches. To add VPLS capability to a PE router, one has to add MAC address learning capabilities, along with aging and other mechanisms native to ethernet switches. This may be fairly complex to add to the forwarding plane architecture of an IP router. As discussed in [L2VPN-FWK], in scenarios where the CE devices are NOT LAN switches, but rather are IP hosts or IP routers, it is possible to provide the VPLS service without requiring MAC address learning and aging on the PE. Instead, a PE router has to have the capability to match the destination MAC address in a packet received from a CE to an outbound pseudowire. The requirements for the IPLS service are described in [L2VPN-REQTS]. The purpose of this document is to specify a solution optimized for IPLS.

IPLS provides a VPLS-like service using PE routers that are not designed to perform general LAN bridging functions. One must be willing to accept the restriction that an IPLS be used for IP traffic only, and not used to interconnect CE devices that are themselves LAN switches. This is an acceptable restriction in many environments, given that IP is the predominant type of traffic in today’s networks.

In IPLS, a PE device implements multi-point LAN connectivity for IP traffic using the following key functions:

1. CE Address Discovery: Each Provider Edge (PE) device discovers IP/MAC address associations for the locally attached Customer Edge (CE) devices, for each IPLS instance configured on the PE device.

2. Pseudowire (PW) for Unicast Traffic: For each locally attached CE device in a given IPLS instance, a PE device sets up a pseudo-wire (PW-LSP) to each of the other PEs that supports the same IPLS instance.
For instance, if PEx and PEy both support IPLS I, and PEy is locally attached to CEa and CEb, PEy will initiate the setup of two pseudowires between itself and PEy. One of these will be used to carry unicast traffic from any of PEy's CE devices to CEa. The other will be used to carry unicast traffic from any of PEy's CE devices to CEb.

Note that these pseudowires carry traffic only in one direction. Further, while the pseudowire implicitly identifies the destination CE of the traffic, it does not identify the source CE; packets from different source CEs bound to the same destination CE are sent on a single pseudowire.

3. Pseudowires for Multicast Traffic: In addition, every PE supporting a given IPLS instance will set up a special "broadcast pseudowire" to every other PE in that IPLS instance. If, in the above example, one of PEy's CE devices sends a multicast packet, PEy would forward the multicast packet to PEy on the special broadcast pseudowire. PEy would then send a copy of that packet to CEa and a copy to CEb.

The broadcast pseudowire carries Ethernet frames of multicast/broadcast IP and ARP packets. Thus when a PE sends a multicast packet across the network, it sends one copy to each remote PE (supporting the given IPLS instance). If a particular remote PE has more than one CE device in that IPLS instance, the remote PE must replicate the packet and send one copy to each of its local CEs.

As with the pseudowires that are used for unicast traffic, packets travel in only one direction on these pseudowires, and packets from different sources may be freely intermixed.

4. Signaling: The necessary pseudowires can be set up and maintained using the LDP-based signaling procedures described in [PWE3-CONTROL].

A PE may assign the same label to each of the unicast pseudowires that lead to a given CE device, in effect creating a multipoint-to-point pseudowire.

Similarly, a PE may assign the same label to each of the broadcast pseudowires for a given IPLS instance, in effect creating a multipoint-to-point pseudowire.

When setting up a pseudowire to be used for unicast traffic, the PE must also signal the IP address and the MAC address of the corresponding CE device.

5. ARP Packet Forwarding: ARP packets are forwarded from attachment circuit to broadcast pseudowires in the Ethernet frame format as described by [PWE3-ETH]. Following rules are observed when processing ARP packets,
   a. Both broadcast (request) and unicast (response) ARP packets are sent over the broadcast pseudowire.
   b. When an ARP packet is received from an attachment circuit,
the packet is copied to control plane for learning CEÂ’s IP and MAC address

c. All Ethernet packets, including ARP packets, received from broadcast pseudowire are forwarded out to all the attachment circuits associated with the IPLS instance. These packets are not copied to control plane.

6. Multicast IP packet Forwarding: An IP Ethernet frame received from an attachment circuit is replicated to other attachment circuits and the broadcast pseudowires associated with the IPLS instance. An IP Ethernet frame received from a broadcast pseudowire is replicated to all the egress attachment circuits associated with the IPLS instance.

7. Unicast IP packet Forwarding: An IP packet received from the attachment circuit is forwarded based on the MAC DA lookup in the forwarding table. If a match is found, the packet is forwarded to the associated egress interface. If the egress interface is unicast pseudowire, the packet is sent without MAC header. If the egress interface is a local attachment circuit the Ethernet frame is forwarded as such. An IP packet received from the unicast pseudowire is forwarded to egress attachment circuit with MAC header prepended. The MAC DA is derived from the forwarding table while PEÂ’s own MAC address is used as MAC SA.

Both VPLS [VPLS] and IPLS require the ingress PE to forward a frame based on its destination MAC address. However, two key differences between VPLS and IPLS can be noted from the above description:

. In VPLS, MAC entries are placed in the FIB of the ingress PE as a result of MAC address learning (which occurs in the data plane) while in IPLS MAC entries are placed in the FIB as a result of pseudowire signaling operations (control plane).
. In VPLS, the egress PE looks up a frame’s destination MAC address to determine the egress Attachment Circuit; in IPLS, the egress Attachment Circuit is determined entirely by the ingress PW-label.

The following sections describe the details of the IPLS scheme.

2.1 Terminology

**IPLS** IP-only LAN service (a type of Virtual Private LAN Service that is restricted to IP traffic only).

**mp2p PW** Multipoint-to-Point Pseudowire. A pseudowire that carries traffic from remote PE devices to a PE device that signals the pseudowire. The signaling PE device advertises the same PW-label to all remote PE devices that participate in the IPLS service instance. In IPLS, for a given IPLS instance, an mp2p PW used for IP
unicast traffic is established by a PE for each CE device locally attached to that PE. It is a unidirectional tree whose leaves consist of the remote PE peers (which connect at least one Attachment Circuit associated with the same IPLS instance) and whose root is the signaling PE. Traffic flows from the leaves towards the root.

**Multicast PW** Multicast/broadcast Pseudowire. A special kind of mp2p PW that carries IP multicast/broadcast traffic and all ARP frames. In the IPLS architecture, for each IPLS instance supported by a PE, that PE device establishes exactly one multicast/broadcast PW. Multicast PW uses Ethernet encapsulation.

**Unicast PW** Unicast Pseudowire carries IP unicast packets. A PE creates unicast PW for each locally attached CE. The unicast PW uses IP Layer2 transport encapsulation.

**CE** Customer Edge device. In this document, a CE is any IP node (host or router) connected to the IPLS LAN service.

**Replication Tree** The collection of all multicast PWs and attachment circuits that are members of an IPLS service instance on a given PE. When a PE receives a multicast/broadcast packet from an attachment circuit, the PE device sends a copy of the packet to every broadcast pseudowire and attachment circuit of the replication tree, excluding the attachment circuit on which the packet was received. When a PE receives a packet from a multicast PW, the PE device sends a copy of the packet to all the attachment circuits of the replication tree and never to other PWs.

### 3.0 Topology

The Customer Edge (CE) devices are IP nodes (hosts or routers) that are connected to PE devices either directly, or via an Ethernet network. We assume that the PE/CE connection may be regarded by the PE as an "interface" to which one or more CEs are attached. This interface may be a physical LAN interface or a VLAN. The Provider Edge (PE) routers are MPLS Label Edge Routers (LERs) that serve as pseudowire endpoints.
In the above diagram, an IPLS instance is shown with three sites: site S1, site S2 and site S3. In site S3, the CE device is directly connected to its PE. In the other two sites, there are multiple CEs connected to a single PE. More precisely, the CEs at these sites are on an Ethernet (switched at site 1 and shared at site 2) network (or VLAN), and the PE is attached to that same Ethernet network or VLAN. We impose the following restriction: if one or more CEs attach to a PE by virtue of being on a common LAN or VLAN, there MUST NOT be more than one PE on that LAN or VLAN.

PE1, PE2 and PE3 are shown as connected via an MPLS network; however, other tunneling technologies, such as GRE, L2TPv3, etc., could also be used to carry the pseudowires.

An IPLS instance is a single broadcast domain, such that each IP end station (e.g., IPa) appears to be co-located with other IP end stations (e.g., IPb through IPf) on the same subnet. The IPLS service is transparent to the CE devices and requires no changes to them.

4.0 Configuration

Each PE router is configured with one or more IPLS service instances, and each IPLS service instance is associated with a unique VPN-Id. For a given IPLS service instance, a set of Attachment Circuits is identified. Each Attachment Circuit can be associated with only one IPLS instance. An Attachment Circuit, in this document, is either a customer-facing Ethernet port, or a particular VLAN (identified by an IEEE 802.1Q VLAN ID) on a customer-facing Ethernet port.

The PE router can optionally be configured with a local MAC address to be used as source MAC address when IP packets are forwarded from a pseudowire to an Attachment Circuit. By default, a PE uses the MAC address of the customer-facing Ethernet interface for this purpose.

5.0 Discovery
The discovery process includes:
  . Remote PE discovery
  . VPN (i.e., IPLS) membership discovery
  . IP CE end station discovery

This draft does not discuss the remote PE discovery or VPN membership discovery. This information can either be user configured or can be obtained using auto-discovery techniques described in [BGP-Discovery] or other methods. However, the discovery of the CE is an important operational step in the IPLS model and is described below.

5.1 CE discovery

Each PE actively detects the presence of local CEs by snooping IP and ARP frames received over the Attachment Circuits. During the discovery phase, the PE examines each multicast/broadcast Ethernet frame. For link-local IP frames (for example IGP discovery/multicast/broadcast packets typically 224.0.0.x addresses), the CE's (source) MAC address is extracted from the Ethernet header and the (source) IP address is obtained from the IP header. For ARP frames, the source MAC and IP address are determined from the ARP PDU.

For each CE, the PE maintains the following tuple: <Attachment Circuit identification info, VPN-Id, IP address, MAC address>.

Once a CE is discovered, its status is monitored continuously by examining the received ARP frames and by periodically generating ARP requests. The absence of an ARP response from a CE after a configurable number of ARP requests is interpreted as loss of connectivity with the CE.

6.0 Pseudowire Creation

6.1 Receive Unicast Multipoint-to-point Pseudowire

As the PE discovers each locally attached CE, a unicast multipoint-to-point pseudowire (mp2p PW) associated exclusively with that CE is created by distributing the CE's IP address and MAC address along with a PW-Label to all the remote PE peers that participate in the same IPLS instance. Note that the same value of a PW-label SHOULD be distributed to all the remote PE peers for a given CE. The mp2p PW thus created is used by remote PEs to send unicast IP traffic to a specific CE.

(The same functionality can be provided by a set of point-to-point PWs, and the PE is not required to send the same PW-label to all the other PEs. For convenience, however, we will use the term mp2p PWs, which may be implemented using a set of point-to-point PWs.)

The PE forwards a frame received over this mp2p PW to the associated Attachment Circuit.

The unicast pseudowire uses IP Layer2 Transport encapsulation as define in [PWE3-Control].
6.2 Receive Multicast Multipoint-to-point Pseudowire

When a PE is configured to participate in an IPLS instance, it advertises a "multicast" PW-label to every other PE that is a member of the same IPLS. The advertised PW-label value is the same for each PE, which creates an mp2p pseudowire for IP multicast/broadcast traffic and ARP packets. There is only one multicast mp2p PW per PE for each IPLS instance and this pseudowire is used exclusively to carry IP multicast/broadcast and ARP traffic from the remote PEs to this PE for this IPLS instance.

Note that no special functionality is expected from this pseudowire. We call it a "multicast pseudowire" because we use it to carry multicast and broadcast IP and ARP traffic. The pseudowire itself need not provide any different service than any of the unicast pseudowires.

In particular, the Receive multicast mp2p PW does not perform any replication of frames itself. Rather, it is there to signify to the PE that the PE needs to replicate a copy of a frame received over this mp2p PW onto all the attachment circuits that are associated with the IPLS instance of the mp2p PW.

The multicast mp2p pseudowire is considered the principle pseudowire in the bundle of mp2p pseudowires that consist of one multicast mp2p pseudowire and a variable number of unicast mp2p pseudowires for a given IPLS instance. In a principle role, multicast PW represents the IPLS instance. The life of all unicast PWS in the IPLS instance depends on the existence of the multicast PW. If, for some reasons, multicast PW cease to exist, all the associated unicast pseudowires in the bundle are removed.

The multicast pseudowire uses Ethernet encapsulation as defined in [PWE3-Ethernet].

The use of pseudowires which are specially optimized for multicast is for further study.

6.3 Send Multicast Replication tree

The PE creates a send replication tree for each IPLS instance, which consists of the collection of all attachment circuits and all the "multicast" pseudowires of the IPLS instance.

Any ARP or multicast IP Ethernet frame received over an attachment circuit is replicated to the other attachment circuits and to the mp2p multicast pseudowire of the send replication tree. The send replication tree deals mostly with broadcast/multicast Ethernet MAC frames. One exception to this is unicast ARP frame, the processing of which is described in the following section.

Any Ethernet frame received over the multicast PW is replicated to all the attachment circuits of the send replication tree of the IPLS instance associated with the incoming PW label. One exception is unicast ARP frame, the processing of which is described in the
following section.

7.0 Signaling

[PWE3-CONTROL] uses the Label Distribution Protocol (LDP) to exchange PW-FECs in the Label Mapping message in a downstream unsolicited mode. The PW-FEC comes in two forms; PWid and Generalized PWid FEC elements. These FEC elements define some fields that are common between them. The discussions below refer to these common fields for IPLS related extensions. Note that the use of multipoint to point and unidirectional characteristics of the PW makes BGP as the ideal candidate for PW-FEC signaling. The use of BGP for such purposes is for future study.

7.1 IPLS PW Signaling

An IPLS carries IP packets as payload over its unicast pseudowires and Ethernet packet as payload over its multicast pseudowire. The PW-type to be used for unicast pseudowire is the IP PW, defined in [PWE3-Control] as IP Layer2 Transport. The PW-type to be used for multicast pseudowire is the Ethernet PW as defined in [PWE3-ETH].

When processing a received PW FEC, the PE matches the PW Id with the locally configured PW Id. If the PW type is Ethernet, the PW-FEC is for multicast PW. If the PW type is ‘IP Layer2 transport’, the PW FEC is for unicast PW. For unicast PW, PE must check the presence of IP and MAC address TLVs in the optional parameter fields of the Label Mapping message. If these parameters are absent, a Label Release message must be issued with a Status Code meaning ‘IP and/or MAC Address of the CE is absent’ [note: Status Code 0x0000002D is pending IANA allocation], to reject the establishment of the unicast PW with the remote PE.

The IPLS uses the Address List TLV as defined in RFC 3036 to signal the IP and MAC address of the local CE. There are two TLVs defined below; IP Address TLV and MAC Address TLV. Both TLVs must be included in the optional parameter field of the Label Mapping message when establishing the unicast IP PW.

Encoding of the IP Address TLV is:

```
0                   1                   2                   3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|0|0| Address List (0x0101) |      Length                   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     Address Family            |     CE’s IP Address           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|       CE’s IP Address         |                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

Length
When Address Family is IPV4, Length is equal to 6 bytes;
2 bytes for address family and 4 bytes of IP address.
Address Family

Two octet quantity containing a value from the ADDRESS FAMILY NUMBERS from ADDRESS FAMILY NUMBERS in [RFC1700] that encodes the addresses contained in the Addresses field.

CE’s IP Address
IP address of the CE attached to the advertising PE. The encoding of the individual address depends on the Address Family.

The following address encodings are defined by this version of the protocol:

<table>
<thead>
<tr>
<th>Address Family</th>
<th>Address Encoding</th>
</tr>
</thead>
<tbody>
<tr>
<td>IPv4 (1)</td>
<td>4 octet full IPv4 address</td>
</tr>
<tr>
<td>IPv6 (2)</td>
<td>16 octet full IPv6 address</td>
</tr>
</tbody>
</table>

Encoding of the MAC Address TLV is:

```
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|0|0| Address List (0x0101) |      Length                   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     Address Family            |     CE’s MAC address          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     Address Family            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
Length
The length field is set to value 8 (2 for address family, 6 for MAC address)

Address Family

Internet Draft  draft-ietf-l2vpn-ipls-04.txt

Two octet quantity containing a value from ADDRESS FAMILY NUMBERS in [RFC1700] that encodes the address contained in the Addresses field.

CE’s MAC Address
MAC address of the CE attached to the advertising PE. The encoding of the individual address depends on the Address Family.

The following address encodings are defined by this version of the protocol:

<table>
<thead>
<tr>
<th>Address Family</th>
<th>Address Encoding</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAC (6)</td>
<td>6 octet full Ethernet MAC address</td>
</tr>
</tbody>
</table>

7.2 Signaling Advertisement Processing

A PE should process a received [PWE3-CONTROL] advertisement with PW-type of IP Layer2 transport for IPLS as follows,
- Verify the IPLS VPN membership by matching the VPN-Id signaled in the AGI field or the PW-ID field with all the VPN-Ids configured in the PE. Discard and release the PW label if VPN-Id is not found.

- Program the Forwarding Information Base (FIB) such that when a unicast IP packet is received from an attachment circuit with its destination MAC address matching the advertised MAC address, the packet is forwarded out over the tunnel to the advertising PE with the advertised PW-label as the inner label.

A PE should process a received [PWE3-CONTROL] advertisement with the PW type of Ethernet for IPLS as follows,
- Verify the IPLS VPN membership by matching the VPN-Id signaled in the AGI field or the PW-ID field with all the VPN-Ids configured in the PE. Discard and release the PW label if VPN-Id is not found.
- Add the PW-label to the send broadcast replication tree for the VPN-Id. This enables sending a copy of a multicast/broadcast IP Ethernet frame or ARP Ethernet frame from the attachment circuit to this pseudowire.

7.3 IANA Considerations for LDP Status Code

This document uses new LDP status code. IANA already maintains a registry of name "STATUS CODE NAME SPACE" defined by RFC3036. The following value is suggested for assignment:

0x0000002D "IP and/or MAC Address of CE is absent"

8.0 Forwarding

8.1 Non-IP or non-ARP traffic

In an IPLS VPN, a PE forwards only IP and ARP traffic. All other frames are dropped silently. If the CEs must pass non-IP traffic to each other, they must do so through IP tunnels that terminate at the CEs themselves.

8.2 Unicast IP Traffic

In IPLS, IP traffic is forwarded from the Attachment Circuit to the PW based on the destination MAC address of the layer 2 frame (and not based on the IP Header).

The PE identifies the FIB associated with an IPLS instance based on the Attachment Circuit or the PW label. When a frame is received from an Attachment Circuit, the PE uses the destination MAC address as the lookup key. When a frame is received from a PW, the PE uses the PW-Label as the lookup key. The frame is dropped if the lookup fails.

8.3 Broadcasts and Multicast IP Traffic

When the destination MAC address is either a broadcast or multicast, a copy of the frame is sent to the control plane for CE discovery.
purposes (see section 5.1). It is important to note that the frames sent to the control plane is applied stricter rate limiting criteria to avoid overwhelming the control plane under adverse conditions such as Denial Of Service attack. The service provider should also provide a configurable limitation to prevent overflowing of the learned source addresses in a given IFLS instance. Also, a caution must be used such that only link local multicasts and broadcast IP packets are sent to control plane.

When a multicast/broadcast IP packet is received from an Attachment Circuit, the PE replicates it onto the Send Multicast Replication Tree (See section 6.3). When a multicast/broadcast IP Ethernet frame is received from a pseudowire, the PE forwards a copy of the frame to all attachment circuits associated with the IFLS VPN instance involved. Note that multicast PW uses Ethernet encapsulation and hence does not require additional header manipulations.

8.4 ARP Traffic

When a broadcast ARP frame is received over the attachment circuit, a copy of the frame is sent to the control plane for CE discovery purposes. The PE replicates the frame onto the Send Multicast Replication Tree (see section 6.3), which results into a copy to be delivered to all the remote PEs on the broadcast PW and other local CEs through the egress attachment circuits.

When a broadcast ARP frame is received over the broadcast PW, a copy of the Ethernet ARP frame is sent to all the attachment circuits associated with the IFLS instance.

8.5 Encapsulation

The Ethernet MAC header of a unicast IP packet received from an Attachment Circuit is stripped before forwarding the frame to the unicast pseudowire. However, the MAC header is retained for the following cases,

. when a frame is unicast or broadcast IP packet that is directed to one or more local Attachment Circuit(s).
. when a frame is a broadcast IP packet
. when a frame is an ARP packet
An IP frame received over a unicast pseudowire is prepended with a MAC header before transmitting it on the appropriate Attachment Circuit(s). The fields in the MAC header are filled in as follows:
- The destination MAC address is the MAC address associated with the PW label in the FIB
- The source MAC address is the PE’s own local MAC address or a MAC address which has been specially configured on the PE for this use.
- The Ethernet Type field is 0x0800
- The frame may be IEEE802.1Q tagged based on the VLAN information associated with the Attachment Circuit.

An FCS is appended to the frame.

9.0 Attaching to IPLS via ATM or FR

In addition to (i) an Ethernet port and a (ii) combination of Ethernet port and a VLAN ID, an Attachment Circuit to IPLS may also be (iii) an ATM or FR VC carrying encapsulated bridged Ethernet frames or (iv) the combination of an ATM or FR VC and a VLAN ID.

The ATM/FR VC is just used as a way to transport Ethernet frames between a customer site and the PE. The PE terminates the ATM/FR VC and operates on the encapsulated Ethernet frames exactly as if those were received on a local Ethernet interface. When a frame is propagated from pseudowire to a ATM or FR VC the PE prepends the Ethernet frame with the appropriate bridged encapsulation header as defined in [RFC 2684] and [RFC 2427] respectively. Operation of an IPLS over ATM/FR VC is exactly as described above, with the exception that the attachment circuit is then identified via the ATM VCI/VPI or Frame Relay DLCI (instead of via a local Ethernet port ID), or a combination of those with a VLAN ID.

10.0 VPLS vs IPLS

The VPLS approach proposed in [VPLS] provides VPN services for IP as well as other protocols. The IPLS approach described in this draft is similar to VPLS in many respects:
- It provides a Provider Provisioned Virtual LAN service with multipoint capability where a CE connected via a single attachment circuit can reach many remote CEs
- It appears as a broadcast domain and a single subnet
- Forwarding is based on destination MAC addresses

However, unlike VPLS, IPLS is restricted to IP traffic only. By restricting the scope of the service to the predominant type of traffic in today’s environment, IPLS eliminates the need for service provider edge routers to implement some bridging functions such as MAC address learning in the data path (by, instead, distributing MAC information in the control plane). Thus this solution offers a number of benefits:
- Facilitates Virtual LAN services in instances where PE devices cannot or cannot efficiently (or are specifically configured not to) perform MAC address learning.
- Unknown Unicast frames are never flooded as would be the case in VPLS.
- Encapsulation is more efficient (MAC header is stripped) for unicast IP packets while traversing the backbone network.
- PE devices are not burdened with the processing overhead associated with traditional bridging (e.g., STP processing, etc.). Note however that some of these overheads (e.g., STP processing) could optionally be turned-off with a VPLS solution in the case where it is known that only IP devices are interconnected.
- Loops (perhaps through backdoor links) are minimized since a PE could easily reject (via label release) a duplicate IP to MAC address advertisement.

11.0 IP Protocols

The solution described in this document offers IPLS service for IPv4 traffic only. For this reason, the MAC Header is not carried over the unicast pseudowire. It is reconstructed by the PE when receiving a packet from a unicast pseudowire and the Ethertype 0x0800 is used in the MAC Header since IPv4 is assumed.

Internet Draft draft-ietf-l2vpn-ipls-04.txt

However, this solution may be extended to carry other types of important traffic such as ISIS and IPv6 which are not encapsulated in Ethernet with the use of Ethertype 0x0800. In order to permit the propagation of such packets correctly, one may create a separate set of pseudowires, or pass protocol information in the "control word" of a "multiprotocol" pseudowire, or encapsulate the Ethernet MAC Header in the pseudowire. The selection of appropriate multiplexing/demultiplexing scheme is the subject of future study.

The current document focuses on IPLS service for IPv4 traffic.

12.0 Dual Homing with IPLS

As stated in previous sections, IPLS prohibits connection of a common LAN or VLAN to more than one PE. Alternatively the CE device itself can connect to more than one instance of IPLS through two separate LAN or VLAN connections to separate PEs. To the CE IP device, these separate connections appear as connections to two IP subnets. The failure of reachability through one subnet is then resolved via the other subnet using IP routing protocols.

13.0 Acknowledgements

Authors would like to thank L2VPN working group members for their valuable comments.

14.0 Security Considerations

The security aspects of this solution will be discussed at a later time.

15.0 References
15.1 Normative References


[INVARP] RFC 2390, T. Bradley et al., "Inverse Address Resolution Protocol".


Internet Draft draft-ietf-l2vpn-ipls-04.txt


15.2 Informative References


[PROXY-ARP] RFC 925, J. Postel, "Multi-LAN Address Resolution".


16.0 Author’s Address

Himanshu Shah
Ciena Corp
35 Nagog Park,
Acton, MA 01720
Email: hshah@ciena.com

K.Arvid
Enterasys Networks
50 Minuteman Rd, Suite 100
Andover, MA 01810
Email: karvind@enterasys.com

Eric Rosen
IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF
THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED
WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Copyright Statement

Copyright (C) The Internet Society (2004). This document is subject
to the rights, licenses and restrictions contained in BCP 78, and
except as set forth therein, the authors retain all their rights.

Acknowledgment

Internet Draft draft-ietf-l2vpn-ipls-04.txt

Funding for the RFC Editor function is currently provided by the
Internet Society.