L2VPN Extensions for L2TP

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.

Abstract

The Layer 2 Tunneling Protocol (L2TP) provides a standard method for setting up and managing L2TP sessions to tunnel a variety of L2 protocols. One of the reference models supported by L2TP describes the use of an L2TP session to connect two Layer 2 circuits attached to a pair of peering LACs, which is a basic form of Layer 2 Virtual Private Network (L2VPN). This document defines the protocol extensions for L2TP to set up different types of L2VPN in a unified fashion.

Specification of Requirements

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].

Luo
# Table of Contents

Status of this Memo................................................. 1  
1. Introduction.................................................... 3  
2. Network Reference Model........................................ 3  
3. Forwarder Identifier............................................ 4  
4. Protocol Components............................................ 5  
   4.1 Control Messages........................................... 5  
   4.2 Existing AVPs for L2VPN.................................. 5  
   4.3 New AVPs for L2VPN........................................ 6  
5. Signaling Procedures............................................. 8  
   5.1 Overview................................................... 8  
   5.2 Pseudowire Tie Detection.................................. 9  
   5.3 Generic Algorithm......................................... 9  
6. IANA Considerations.............................................. 13  
7. Security Considerations......................................... 14  
8. Intellectual Property Statement............................... 14  
9. Copyright Notice................................................ 14  
10. Acknowledgement................................................ 15  
11. References..................................................... 15  
   11.1 Normative References.................................... 15  
   11.2 Informative References.................................. 15  
12. Authors’ Address............................................... 15
1. Introduction

[L2TPv3] defines a dynamic tunneling mechanism to carry multiple L2 protocols besides PPP (as originally defined in [RFC 2661]) over a packet-based network. The baseline protocol supports various types of applications, which have been highlighted in the different L2TP reference models in [L2TPv3]. L2VPN applications are typically in the scope of the LAC-LAC reference model.

This document discusses the commonalities as well as differences among L2VPN applications with respect to utilizing L2TPv3 as the signaling protocol.

The acronym "L2TP" refers to L2TPv3 or L2TP in general in this document.

2. Network Reference Model

In the LAC-LAC reference model, a LAC serves as a cross-connect between attachment circuits and L2TP sessions. Each L2TP session acts as an emulated circuit, also known as pseudowire. A pseudowire is used to bind two "forwarders" together. For different L2VPN applications, different types of forwarders are defined.

In the L2VPN framework [L2VPN FW], a LAC is a Provider Edge (PE) device. LAC and PE are interchangeable terms in the context of this document. Remote systems in the LAC-LAC reference model are Customer Edge (CE) devices.

```
+----+    L2   +----+    L2   +----+
 | CE |------| PE |....[core network]....| PE |------| CE |
+----+    ++++  +----+    ++++  +----+

|<- emulated service ->|
|----------------- L2 service -------------->
```

L2VPN Network Reference Model

In a simple cross-connect application, an attachment circuit is a forwarder directly bound to a pseudowire. It is a one-to-one mapping. Traffic received from the attachment circuit on a local PE is forwarded to the remote PE through the pseudowire. When the remote PE receives traffic from the pseudowire, it forwards the traffic to the corresponding attachment circuit on its end. The forwarding decision is based on the attachment circuit or pseudowire demultiplexing identifier.
With Virtual Private LAN Service (VPLS), a Virtual Switching Instance (VSI) is a forwarder connected to one or more attachment circuits and pseudowires. A single pseudowire is used to connect a pair of VSI on two peering PEs. Traffic received from an attachment circuit or a pseudowire is first forwarded to the corresponding VSI based on the attachment circuit or pseudowire demultiplexing identifier. The VSI performs additional lookup to determine where to further forward the traffic.

With Virtual Private Wire Service (VPWS), attachment circuits are grouped into "colored pools". Each pool is a forwarder and connected through a network of point-to-point cross-connect. The data forwarding perspective is identical to the cross-connect application. However, constructing colored pools involves more complicated signaling procedures.

3. Forwarder Identifier

A forwarder identifier is assigned to each forwarder on a given PE and is unique in the context of the PE. It is defined as the concatenation of an Attachment Group Identifier (AGI) and an Attachment Individual Identifier (AII), denoted as <AGI, AII>. The AGI is used to group a set of forwarders together for signaling purpose. An AII is used to distinguish forwarders within a group. AII can be unique at a per platform or per group basis.

As far as the signaling procedures are concerned, a forwarder identifier is an arbitrary string of bytes. It is up to implementations to decide the values for AGI and AII.

When connecting two forwarders together, both MUST have the same AGI as part of their forwarder identifiers. The AII of the source forwarder is known as the Source AII (SAII), and the AII of the target forwarder is known as the Target AII (TAII). Therefore, source forwarder and target forwarder can be denoted as <AGI, SAII> and <AGI, TAI> respectively.
4. Protocol Components

4.1 Control Messages

L2TP defines two sets of session management procedures: incoming call and outgoing call. Even though it is entirely possible to use the outgoing call procedures to signaling L2VPNs, the incoming call procedures has some advantages in terms of the relevance of the semantics. [PWE3L2TP] gives more details on why the incoming call procedures are more appropriate for setting up pseudowires.

The signaling procedures for L2VPNs described in the following sections are based on the Incoming Call procedures.

4.2 Existing AVPs for L2VPN

Router ID

The Router ID sent in SCCRQ and SCCRP during control connection setup establishes the unique identity of each PE.

Pseudowire Capabilities List

The Pseudowire Capabilities List sent in the SCCRQ and SCCRP indicates the pseudowire types supported by the sending PE. It merely serves as an advertisement to the receiving PE. Its content should not affect the control connection setup.

Before a local PE initiates a session of a particular pseudowire type to a remote PE, it MUST examine whether the remote PE has advertised this pseudowire type in this AVP, and SHOULD NOT attempt to initiate the session if the intended pseudowire type is not supported by the remote PE.

Pseudowire Type

The Pseudowire Type sent in ICRQ signals the intended pseudowire type to the receiving PE. The receiving PE checks it against its local pseudowire capabilities list. If it finds a match, it responds with an ICRP without a Pseudowire Type AVP, which implicitly acknowledges its acceptance of the intended pseudowire. If it does not find a match, it MUST respond with a CDN with an "unsupported pseudowire type" result code.

L2-Specific Sublayer

The L2-Specific Sublayer can be sent in ICRQ, ICRP, and ICCN. If the receiving PE supports the specified L2-Specific Sublayer, it MUST
include the identified L2-Specific Sublayer in its data packets sent to the sending PE. Otherwise, it MUST reject the connection by sending a CDN to the sending PE.

Circuit Status

The Circuit Status is sent in both ICRQ and ICRP to inform the receiving PE about the circuit status on the sending PE. It can also be sent in ICCN and SLI to update the status.

Remote End Identifier

The TAI value is encoded in the Remote End ID AVP and sent in ICRQ along with the optional AGI to instruct the receiving PE to bind the proposed pseudowire to the forwarder that matches the specified forwarder identifier.

4.3 New AVPs for L2VPN

Attachment Group Identifier

The AGI AVP, Attribute Type TBA, is an identifier used to associate a forwarder to a logical group. The AGI AVP is used in conjunction with the Local End ID AVP and Remote End ID AVP, which encode the SAI and TAI respectively, to identify a specific forwarder. When the AGI AVP is omitted in the control messages or contains a zero-length value, the forwarders are considered using the default AGI. Note that there is only one designated default AGI value for all forwarders.

The Attribute Value field for this AVP has the following format:

```
0                   1                   2                   3
+-------------------+-------------------+-------------------+-------------------+
| M | H | 0 | 0 | 0 | 0 | Length | 0 |
+-------------------+-------------------+-------------------+-------------------+
| TBA | AGI (variable length) |
+-------------------+-------------------+-------------------+-------------------+
```

The AGI field is a variable-length field. This AVP MAY be present in ICRQ.

This AVP MAY be hidden (the H bit may be 0 or 1). The M bit for this AVP SHOULD be set to 0. The Length (before hiding) of this AVP is 6 plus the length of the AGI field.
The Local End ID AVP, Attribute Type TBA, encodes the SAIII value. The SAIII may also be used in conjunction with the TAI to detect pseudowire ties. When it is omitted in the control messages, it is assumed that it has the same value as the TAI.

The Attribute Value field for this AVP has the following format:

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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|H|0|0|0|0|    Length         |              0                |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|              TBA              |       SAIII (variable length)  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

The SAIII field is a variable-length field. This AVP MAY be present in ICRQ.

This AVP MAY be hidden (the H bit MAY be 0 or 1). The M bit for this AVP SHOULD be set to 0. The Length (before hiding) of this AVP is 6 plus the length of the SAIII field.

Interface Maximum Transmission Unit

The Interface Maximum Transmission Unit (MTU) AVP, Attribute Type TBA, indicates the MTU in octets of a packet that can be sent out from the CE-facing interface. The MTU values of a given pseudowire, if advertised in both directions, MUST be identical. If they do not match, the pseudowire SHOULD NOT be established. When this AVP is omitted in the control messages in either direction, it is assumed that the remote PE has the same interface MTU as the local PE for the pseudowire being signaled.

The Attribute Value field for this AVP has the following format:

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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|H|0|0|0|0|    Length         |              0                |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|              TBA              |          Interface MTU        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

The Interface MTU is a 2-octet value. This AVP MAY be present in ICRQ and ICRP. When a PE receives an Interface MTU AVP with an MTU value different from its own, it MAY respond with a CDN with a new result code indicating the disconnect cause.
RC-TBA-1 - Mismatching interface MTU

This AVP MAY be hidden (the H bit MAY be 0 or 1). The M bit for this AVP SHOULD be set to 0. The Length (before hiding) of this AVP is 8.

5. Signaling Procedures

5.1 Overview

Assume a PE assigns a forwarder identifier to one of its local forwarders, and knows it needs to set up a pseudowire to a remote forwarder on a remote PE that has a certain Forwarder ID. This knowledge can be obtained either through manual configuration or some auto-discovery procedure.

Before establishing the intended pseudowire, each pair of peering PEs exchanges control connection messages to establish a control connection. Each advertises its supported pseudowire types in the Pseudowire Capabilities List AVP.

After the control connection is established, the local PE examines whether the remote PE supports the pseudowire type it intends to set up. Only if the remote PE supports the intended pseudowire type, it should initiate a pseudowire connection request.

When the local PE receives an ICRQ for a pseudowire connection, it examines the forwarder identifier encoded in the AGI, SAII, and TAII in order to determine the following:

- whether it has a local forwarder with the forwarder identifier value specified in the ICRQ,
- whether the remote forwarder with the forwarder identifier specified in the ICRQ is allowed to connect with this local forwarder.

If both conditions are met, it sends an ICRP to the remote PE to accept the connection request. If either of the two conditions fails, it sends a CDN to the remote PE to reject the connection request.

The local PE can optionally include a result code in the CDN to indicate the disconnect cause. The possible result codes are:

RC-TBA-2 - Attempt to connect to non-existent forwarder
RC-TBA-3 - Attempt to connect to unauthorized forwarder
5.2 Pseudowire Tie Detection

Conceivably in the network reference models, as either PE may initiate a pseudowire to another PE at any time, the PEs could end up initiating a pseudowire to each other simultaneously. In order to avoid setting up duplicated pseudowires between two forwarders, each PE must be able to independently detect such a pseudowire tie. The following procedures need to be followed to detect a tie:

If both TAI and SAI are present in the ICRQ, the receiving PE compares the TAI and SAI against the SAI and TAI itself previously sent to the sending PE. If the received TAI matches the sent SAI and the received SAI matches the sent TAI, a tie is detected.

If only the TAI is present in the ICRQ, the SAI is assumed to have the same value as the TAI. The receiving PE compares the received TAI with the SAI that it previously sent to the sending PE. If the SAI in that ICRQ is also omitted, then the value encoded in the sent TAI is used for comparison. If they match, a tie is detected.

If the AGI is present, it is first prepended to the TAI and SAI values before the tie detection occurs.

Once a tie has been discovered, the standard L2TP tie breaking procedure is used to disconnect the duplicated pseudowire.

5.3 Generic Algorithm

The following uses a generic algorithm to illustrate the protocol interactions when constructing an L2VPN using L2TP signaling.

Each PE first forms a list, SOURCE_FORWARDERS, consisting of all local forwarders of a given VPN. Then it puts all local forwarders that need to be interconnected and all remote forwarders of the same VPN into another list, TARGET_FORWARDERS. The formation of the network topology depends on the content in the SOURCE_FORWARDERS and TARGET_FORWARDERS lists. These two lists can be constructed by manual configuration and/or some auto-discovery procedure.

The algorithm is used to set up pseudowires among all the forwarders that intend to be interconnected by iterating through each source and target forwarder. An L2VPN is formed upon finishing the algorithm in every participating PE of this L2VPN.

1. Pick the next forwarder, from SOURCE_FORWARDERS. If no forwarder is available for processing, the processing is complete.
2. Pick the next forwarder, from TARGET_FORWARDERS. If no forwarder is available for processing, go back to step 1.

3. If the two forwarders are associated with different Router IDs, a pseudowire must be established between them. Proceed to step 6.

4. Compare the <AGI, AII> values of the two forwarders. If they match, the source and target forwarders are the same, so no more action is necessary. Go back to step 2.

5. As the source and target forwarders both reside on the local PE, no pseudowire is needed. The PE simply creates a local cross-connect between the two forwarders. Go back to step 2.

6. As the source and target forwarders reside on different PEs, a pseudowire must be established between them. The PE first examines if the source forwarder has already established a pseudowire to the target forwarder. If so, go back to step 2.

7. If no pseudowire is already established between the source and target forwarders, the local PE obtains the address of the remote PE, and establishes a control connection to the remote PE if one does not already exist.

8. The local PE sends an ICRQ to the remote PE. The AGI, TAI, and SAI values are encoded in the AGI AVP, the Remote End ID AVP, and the Local End ID AVP respectively.

9. If the local PE receives a response corresponding to the ICRQ it just sent, proceed to step 10. Otherwise, if the local PE receives an ICRQ from the same remote PE, proceed to step 11.

10. The local PE receives a response from the remote PE. If it is a CDN, go back to step 2. If it’s an ICRP, the local PE binds the source forwarder to the pseudowire and sends an ICCN to the remote PE. Go back to step 2.

11. If the local PE receives an ICRQ from the same remote PE, it needs to perform session tie detection, as described in Section 5.2. If a session tie is detected, the PE performs tie breaking.

12. If the local PE loses the tie breaker, it sends a CDN with the result code that indicates the disconnection is due to losing the tie breaker. Proceed to step 14.
13. If the local PE wins the tie breaker, it ignores the remote PE’s ICRQ, but acknowledges receipt of the control message and continues waiting for the response from the remote PE. Go to step 10.

14. The local PE determines whether it should accept the connection request, as described in the Section 5.1. If it accepts the ICRQ, it sends an ICRP to the remote PE.

15. The local PE receives a response from the remote PE. If it is a CDN, go back to step 2. If it is an ICCN, the local PE binds the source forwarder to the pseudowire, go back to step 2.

The following diagram illustrates the above procedure:
6. IANA Considerations

This document defines three new L2TP control message Attribute Value Pairs (AVPs) to be assigned by the IANA. These are described in Section 4.3 and Section 5.1 define three new result codes for the CDN message are to be assigned by the IANA:

AVP-TBA-1 - Attachment Group Identifier
AVP-TBA-2 - Local End Identifier
AVP-TBA-3 - Interface Maximum Transmission Unit

Section 4.3 and Section 5.1 define three new result codes for the CDN message are to be assigned by the IANA:

RC-TBA-1 - Mismatching interface MTU
RC-TBA-2 - Attempt to connect to non-existent forwarder
RC-TBA-3 - Attempt to connect to unauthorized forwarder
7. Security Considerations

The signaling procedures described in this document does not incur additional security considerations that L2TP already provisions.

8. Intellectual Property Statement

The IETF takes no position regarding the validity or scope of any Intellectual Property Rights or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; nor does it represent that it has made any independent effort to identify any such rights. Information on the procedures with respect to rights in RFC documents can be found in BCP 78 and BCP 79.

Copies of IPR disclosures made to the IETF Secretariat and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the IETF on-line IPR repository at http://www.ietf.org/ipr.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement this standard. Please address the information to the IETF at ietf-ipr@ietf.org.

9. Copyright Notice

Copyright (C) The Internet Society (2005).

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.

This document and the information contained herein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY AND THE INTERNET ENGINEERING TASK FORCE DISCLAIM ALL WARRANTIES, EXPRESS OR 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.
10. Acknowledgement

The author would like to thank Mark Townsley and Carlos Pignataro for their valuable input.

11. References

11.1 Normative References


11.2 Informative References


12. Authors’ Address

Wei Luo
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134
Email: luo@cisco.com