Network Working Group
Request for Comments: 4667
Category: Standards Track

W. Luo
Cisco Systems, Inc.

September 2006

Layer 2 Virtual Private Network (L2VPN) Extensions for Layer 2 Tunneling Protocol (L2TP)

Status of This Memo

Copyright Notice

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 L2TP Access Concentrators (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 L2VPNs in a unified fashion.

Table of Contents

1. Introduction

[RFC3931] defines a dynamic tunneling mechanism to carry multiple Layer 2 protocols besides Point-to-Point Protocol (PPP), the only protocol supported in [RFC2661], over a packet-based network. The baseline protocol supports various types of applications, which have been highlighted in the different Layer 2 Tunneling Protocol (L2TP) reference models in [RFC3931]. An L2TP Access Concentrator (LAC) is an L2TP Control Connection Endpoint (LCCE) that cross-connects attachment circuits and L2TP sessions. Layer 2 Virtual Private Network (L2VPN) applications are typically in the scope of the LAC-LAC reference model.

This document discusses the commonalities and differences among L2VPN applications with respect to using L2TPv3 as the signaling protocol. In this document, the acronym "L2TP" refers to L2TPv3 or L2TP in general.

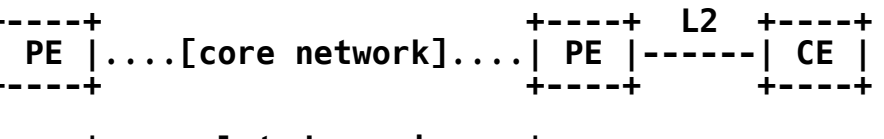
1.1. Specification of Requirements

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 [RFC2119].

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 [L2VPNFW], 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.



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 VSIs 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 is connected through a network of point-to-point cross-connects. 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 purposes. An AII is used to distinguish forwarders within a group. AII can be unique on 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, the source forwarder and target forwarder can be denoted as <AGI, SAII> and <AGI, TAII>, 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 for signaling L2VPNs, the incoming call procedures have 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 Control Connection Management and the Incoming Call procedures, defined in Sections 3.3 and 3.4.1 of [RFC3931], respectively. L2TP control message types are defined in Section 3.1 of [RFC3931]. This document references the following L2TP control messages:

```
Start-Control-Connection-Request (SCCRQ)
Start-Control-Connection-Reply   (SCCRP)
Incoming-Call-Request            (ICRQ)
Incoming-Call-Reply              (ICRP)
Incoming-Call-Connected          (ICCN)
Set-Link-Info                    (SLI)
```

4.2. Existing AVPs for L2VPN

The following Attribute Value Pairs (AVPs), defined in Sections 5.4.3, 5.4.4, and 5.4.5 of [RFC3931], are used for signaling L2VPNs.

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 Call-Disconnect-Notify (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 TAII 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.

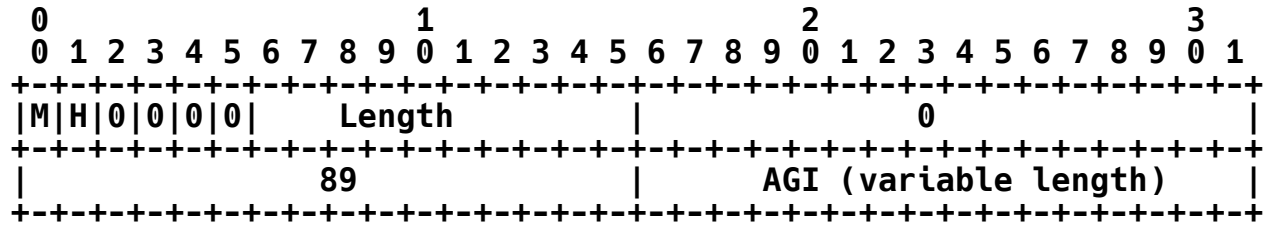
4.3. New AVPs for L2VPN

Attachment Group Identifier

The AGI AVP, Attribute Type 89, 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 SAII and TAII, 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 to use

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:



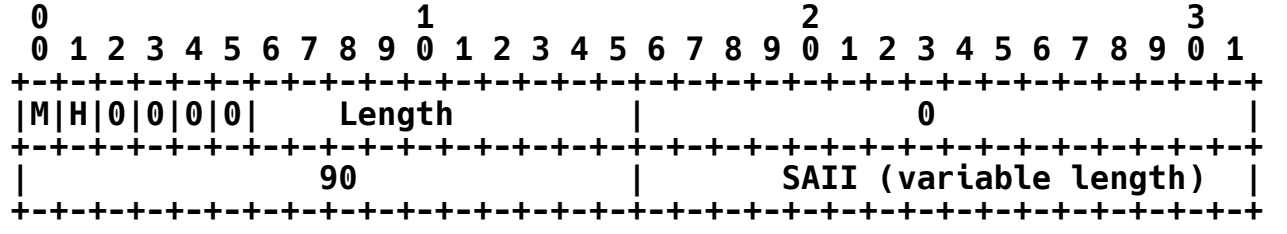
The AGI field is a variable-length field. This AVP MAY be present in ICRQ.

This AVP MAY be hidden (the H bit MAY be 0 or 1). The hiding of AVP attribute values is defined in Section 5.3 of [RFC3931]. The M bit for this AVP SHOULD be set to 0. The Length (before hiding) of this AVP is 6 octets plus the length of the AGI field.

Local End ID

The Local End ID AVP, Attribute Type 90, encodes the SAII value. The SAII may also be used in conjunction with the TAII to detect pseudowire ties. When it is omitted in the control messages, it is assumed that it has the same value as the TAII.

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



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

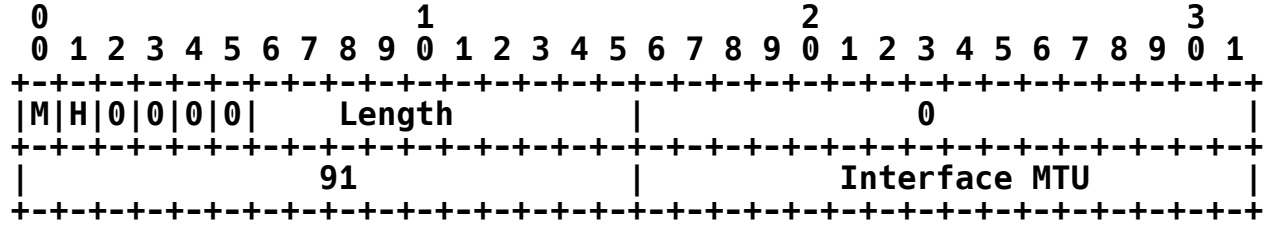
This AVP MAY be hidden (the H bit MAY be 0 or 1). The M bit for this AVP SHOULD be set to 0. The Length (before hiding) of this AVP is 6 octets plus the length of the SAII field.

Interface Maximum Transmission Unit

The Interface Maximum Transmission Unit (MTU) AVP, Attribute Type

91, 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:



The Interface MTU field is a 2-octet integer 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.

23 - 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 octets.

4.4. AVP Interoperability

To ensure interoperability, the mandatory (M) bit settings of the existing AVPs used in L2VPN applications should be the same as those specified in [RFC3931]. The generic M-bit processing is described in Section 5.2 of [RFC3931]. Setting the M-bit of the new AVPs to 1 will impair interoperability.

5. Signaling Procedures

5.1. Overview

Assume that a PE assigns a forwarder identifier to one of its local forwarders and that it 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, as defined in [PWE3IANA], 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 should it initiate a pseudowire connection request.

When the local PE receives an ICRQ for a pseudowire connection, it examines the forwarder identifiers 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

24 - Attempt to connect to non-existent forwarder
25 - 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 TAII and SAII are present in the ICRQ, the receiving PE compares the TAII and SAII against the SAII and TAII previously sent to the sending PE. If the received TAII matches the sent SAII and the received SAII matches the sent TAII, a tie is detected.

If only the TAII is present in the ICRQ, the SAII is assumed to have the same value as the TAII. The receiving PE compares the received TAII with the SAII that it previously sent to the sending PE. If the

SAII in that ICRQ is also omitted, then the value encoded in the sent TAII is used for comparison. If they match, a tie is detected.

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

Once a tie is discovered, the PE uses the standard L2TP tie breaking procedure, as described in Section 5.4.4 of [RFC3931], 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 or some auto-discovery procedure.

The algorithm is used to set up a full mesh of interconnections between SOURCE_FORWARDERS and TARGET_FORWARDERS. An L2VPN is formed when the algorithm is finished 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 whether 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, TAII, and SAII 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.

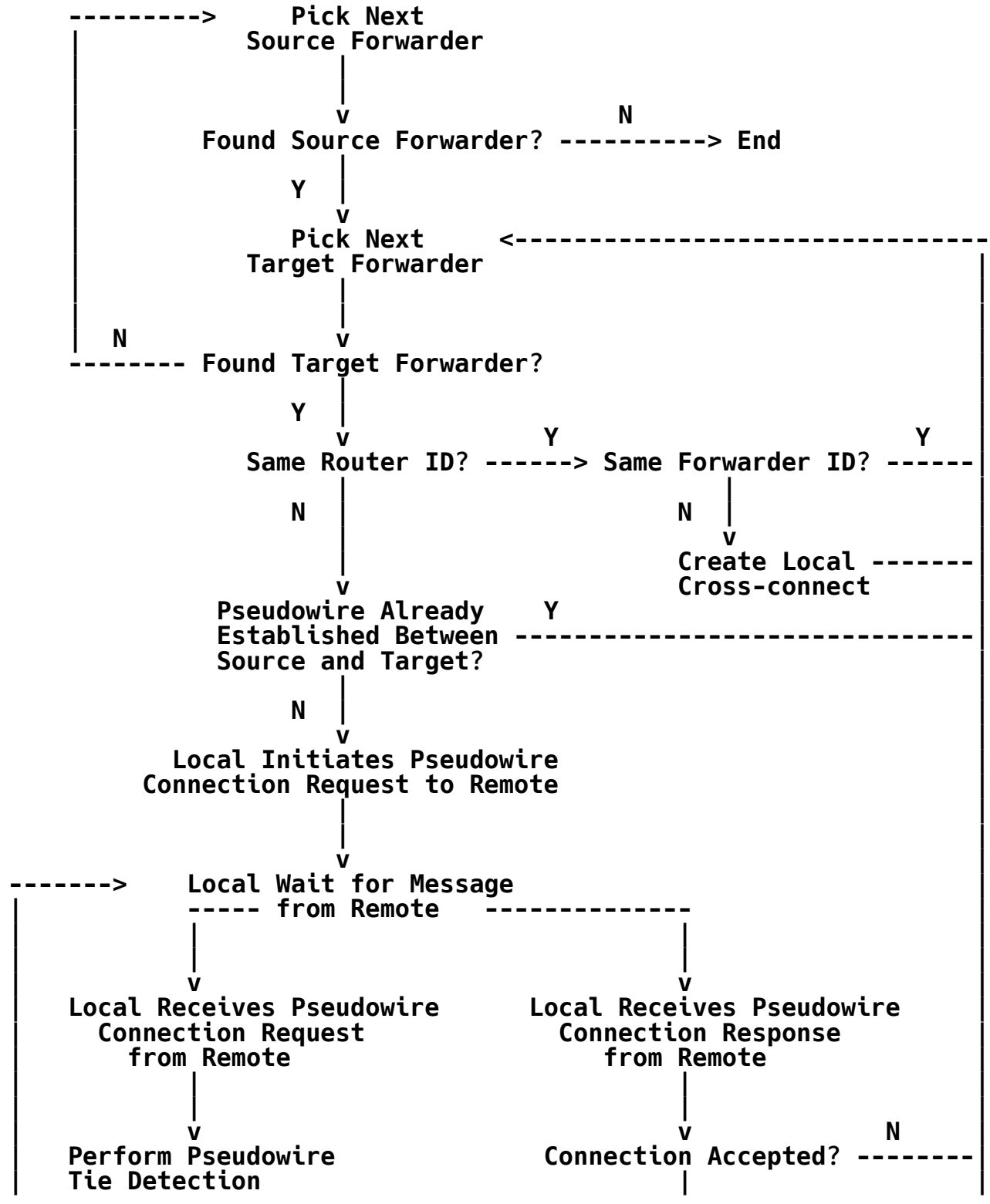
12. If the local PE loses the tie breaker, it sends a CDN with the result code that indicates that the disconnection is due to losing the tie breaker. Proceed to step 14.

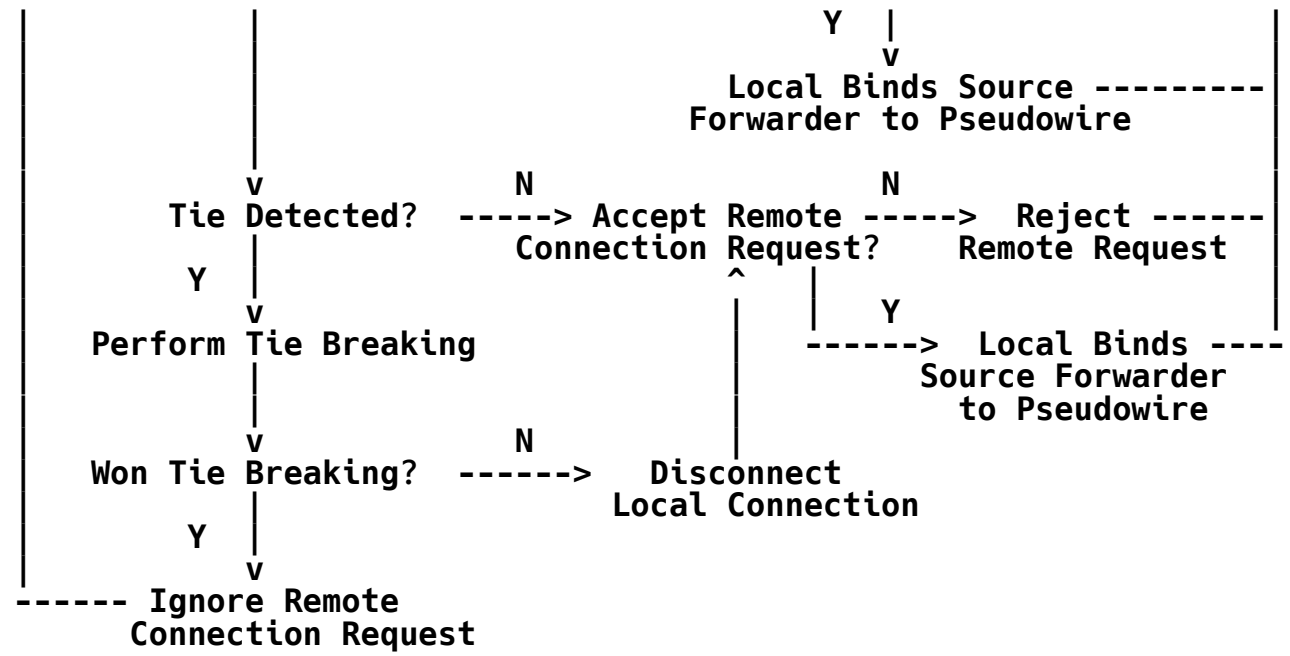
13. If the local PE wins the tie breaker, it ignores the remote PE's ICRQ, but acknowledges receipt of the control message and continues waiting for the response from the remote PE. Go to step 10.

14. The local PE determines whether it should accept the connection request, as described in 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:


Pick Next
Source Forwarder
Found Source Forwarder?
N
End
Y
Pick Next
Target Forwarder
N
Found Target Forwarder?
Y
Same Router ID?
Y
Same Forwarder ID?
Y
N
N
Create Local
Cross-connect
Pseudowire Already
Established Between
Source and Target?
Y
N
Local Initiates Pseudowire
Connection Request to Remote
Local Wait for Message
from Remote
Local Receives Pseudowire
Connection Request
from Remote
Local Receives Pseudowire
Connection Response
from Remote
Perform Pseudowire
Tie Detection
Connection Accepted?
N



6. IANA Considerations

The IANA registry procedure in this document follows that in Section 10 of [RFC3931]. The IANA has assigned the following new values for existing registries managed by IANA.

This document defines three new L2TP control message Attribute Value Pairs (AVPs) that have been assigned by the IANA. These are described in Section 4.3 and are summarized below:

- 89 - Attachment Group Identifier
- 90 - Local End Identifier
- 91 - Interface Maximum Transmission Unit

Sections 4.3 and 5.1 define three new result codes for the CDN message that have been assigned by the IANA:

- 23 - Mismatching interface MTU
- 24 - Attempt to connect to non-existent forwarder
- 25 - Attempt to connect to unauthorized forwarder

7. Security Considerations

This specification does not introduce any additional security considerations beyond those discussed in [RFC3931] and [L2VPNFW].

8. Acknowledgement

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

9. References

9.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC3931] Lau, J., Townsley, M., and I. Goyret, "Layer Two Tunneling Protocol - Version 3 (L2TPv3)", RFC 3931, March 2005.

9.2. Informative References

[PWE3IANA] Martini, L., "IANA Allocations for Pseudowire Edge to Edge Emulation (PWE3)", BCP 116, RFC 4446, April 2006.

[L2VPNFW] Andersson L., Ed. and E. Rosen, Ed., "Framework for Layer 2 Virtual Private Networks (L2VPNs)", RFC 4664, September 2006.

[PWE3L2TP] W. Townsley, "Pseudowires and L2TPv3", Work in Progress.

[RFC2661] Townsley, W., Valencia, A., Rubens, A., Pall, G., Zorn, G., and B. Palter, "Layer Two Tunneling Protocol "L2TP"", RFC 2661, August 1999.

Author's Address

Wei Luo
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134

EMail: luo@cisco.com

Full Copyright Statement

Intellectual Property

Acknowledgement

Funding for the RFC Editor function is provided by the IETF Administrative Support Activity (IASA).