

Internet Engineering Task Force (IETF)
Request for Comments: 7392
Category: Standards Track
ISSN: 2070-1721

P. Dutta
M. Bocci
Alcatel-Lucent
L. Martini
Cisco Systems
December 2014

Explicit Path Routing for Dynamic Multi-Segment Pseudowires

Abstract

When set up through an explicit path, dynamic Multi-Segment Pseudowires (MS-PWs) may be required to provide a simple solution for 1:1 protection with diverse primary and backup MS-PWs for a service, or to enable controlled signaling (strict or loose) for special MS-PWs. This document specifies the extensions and procedures required to enable dynamic MS-PWs to be established along explicit paths.

Status of This Memo

This is an Internet Standards Track document.

This document is a product of the Internet Engineering Task Force (IETF). It represents the consensus of the IETF community. It has received public review and has been approved for publication by the Internet Engineering Steering Group (IESG). Further information on Internet Standards is available in Section 2 of RFC 5741.

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at <http://www.rfc-editor.org/info/rfc7392>.

Copyright Notice

Copyright (c) 2014 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction	2
2. Requirements Language and Terminology	3
3. Explicit Path in MS-PW Signaling	3
3.1. S-PE Addressing	3
3.2. Explicit Route TLV (ER-TLV)	3
3.3. Explicit Route Hop TLV (ER-Hop TLV)	4
3.4. ER-Hop Semantics	4
3.4.1. ER-Hop Type: IPv4 Prefix	4
3.4.2. ER-Hop Type: IPv6 Prefix	4
3.4.3. ER-Hop Type: L2 PW Address	5
4. Explicit Route TLV Processing	6
4.1. Next-Hop Selection	6
4.2. Adding ER Hops to the Explicit Route TLV	8
5. IANA Considerations	8
6. Security Considerations	8
7. Normative References	9
Acknowledgements	9
Authors' Addresses	10

1. Introduction

Procedures for dynamically establishing Multi-Segment Pseudowires (MS-PWs), where their paths are automatically determined using a dynamic routing protocol, are defined in [RFC7267]. For 1:1 protection of MS-PWs with primary and backup paths, MS-PWs need to be established through a diverse set of Switching Provider Edges (S-PEs) to avoid any single points of failure at the PW level. [RFC7267] allows this through BGP-based mechanisms. This document defines an additional mechanism that allows Source Terminating Provider Edges (ST-PEs) to explicitly choose the path that a PW would take through the intervening S-PEs. Explicit path routing of dynamic MS-PWs may also be required for controlled setup of dynamic MS-PWs and network resource management.

Note that in many deployments the ST-PE will not have a view of the topology of S-PEs and so the explicit route will need to be supplied from a management application. How that management application determines the explicit route is outside the scope of this document.

2. Requirements Language and Terminology

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].

This document uses the terminology defined in [RFC7267], [RFC4447], and [RFC5036].

The following additional terminology is used:

Abstract Node: A group of nodes (S-PEs) representing an explicit hop along the path of an MS-PW. An abstract node is identified by an IPv4, IPv6, or S-PE address.

3. Explicit Path in MS-PW Signaling

This section describes the Label Distribution Protocol (LDP) extensions required for signaling explicit paths in dynamic MS-PW setup messages. An explicitly routed MS-PW is set up using a Label Mapping message that carries an ordered list of the S-PEs that the MS-PW is expected to traverse. The ordered list is encoded as a series of Explicit Route Hop TLVs (ER-Hop TLVs) encoded in an ER-TLV that is carried in a Label Mapping message.

3.1. S-PE Addressing

An S-PE address is used to identify a given S-PE among the set of S-PEs belonging to the Packet Switched Networks (PSNs) that may be used by an MS-PW. Each S-PE MUST be assigned an address as specified in Section 3.2 of [RFC7267]. An S-PE that is capable of dynamic MS-PW signaling, but has not been assigned an S-PE address, and that receives a Label Mapping message for a dynamic MS-PW MUST follow the procedures in Section 3.2 of [RFC7267].

3.2. Explicit Route TLV (ER-TLV)

The ER-TLV specifies the path to be taken by the MS-PW being established. Each hop along the path is represented by an abstract node, which is a group of one or more S-PEs, identified by an IPv4, IPv6, or S-PE address. The ER-TLV format is as per Section 4.1 of [RFC3212].

The ER-TLV contains one or more ER-Hop TLVs as defined in Section 3.3.

3.3. Explicit Route Hop TLV (ER-Hop TLV)

The contents of an ER-TLV are a series of variable-length ER-Hop TLVs. Each hop contains the identification of an "abstract node" that represents the hop to be traversed. The ER-Hop TLV format is as specified in Section 4.2 of [RFC3212].

[RFC3212] defines four ER-Hop TLV Types: IPv4 Prefix, IPv6 Prefix, Autonomous System Number, and LSP-ID. This document specifies the following new ER-Hop TLV Type:

Value	Type
0x0805	L2 PW Address of Switching Point

ER-Hop TLV

Details of the ER-Hop semantics are defined in Section 3.4.

3.4. ER-Hop Semantics

This section describes the various semantics associated with the ER-Hop TLV.

3.4.1. ER-Hop Type: IPv4 Prefix

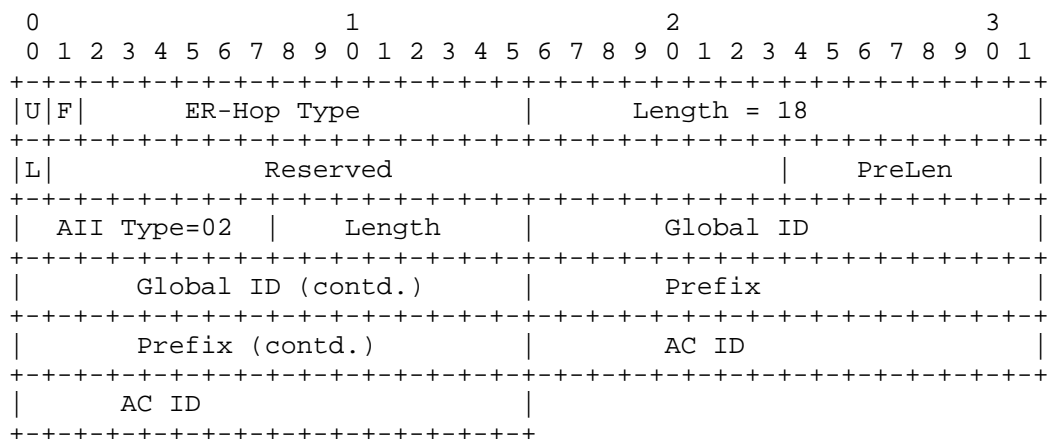
The semantics of the IPv4 ER-Hop TLV Type are specified in [RFC3212], Section 4.7.1.

3.4.2. ER-Hop Type: IPv6 Prefix

The semantics of the IPv6 ER-Hop TLV Type are specified in [RFC3212], Section 4.7.2.

3.4.3. ER-Hop Type: L2 PW Address

The semantics of the L2 PW Address ER-Hop TLV Type, which contains the L2 PW Address derived from the Generalized PWid Forwarding Equivalence Class (FEC) AII Type 2 structure as defined in [RFC5003], are as follows.



U/F

These bits MUST be set to zero and the procedures of [RFC5036] followed when the TLV is not known to the receiving node.

Type

A fourteen-bit field carrying the value of the ER-Hop 3, L2 PW Address, Value = 0x0805.

Length

Specifies the length of the value field in bytes = 18.

L Bit

Set to indicate a loose hop.
Cleared to indicate a strict hop.

Reserved

Zero on transmission. Ignored on receipt.

PreLen

Prefix Length 1-96 (including the length of the Global ID, Prefix, and AC ID fields).

All other fields (AII Type, Length, Global ID, Prefix, and AC ID) define the L2 PW Address and are to be set and interpreted as defined in Section 3.2 of [RFC5003].

4. Explicit Route TLV Processing

4.1. Next-Hop Selection

A PW Label Mapping message containing an Explicit Route TLV specifies the next hop for a given MS-PW path. Selection of this next hop by the ST-PE or S-PE inserting the ER-Hop TLV may involve a selection from a set of possible alternatives. The mechanism for making a selection from this set is implementation specific and is outside the scope of this document. The mechanism used to select a particular path is also outside the scope of this document, but each node **MUST** determine a loop-free path if it is to signal the MS-PW. [RFC6073], Section 7.6 provides a mechanism by which a node can check that the path taken by an MS-PW does not include loops.

As noted in Section 1, in many deployments the ST-PE will not have a view of the topology of S-PEs and so the path will need to be supplied from a management application.

If a loop-free path cannot be found by an ST-PE or S-PE, then a node **MUST NOT** attempt to signal the MS-PW. For an S-PE, if it cannot determine a loop-free path, then the received Label Mapping message **MUST** be released with a status code of "PW Loop Detected" as per Section 4.2.3 of [RFC7267].

To determine the next hop for the MS-PW path, a node performs the following steps. Note that these procedures assume that a valid S-PE address has been assigned to the node, as per Section 3.1, above.

1. The node receiving the Label Mapping message that contains an ER-TLV **MUST** evaluate the first ER-Hop. If the L bit is not set in the first ER-Hop and if the node is not part of the abstract node described by the first ER-Hop (i.e., it does not lie within the prefix as determined by the prefix length specified in the ER-Hop TLV), it has received the message in error. Therefore, the node **MUST** reply with a Label Release message with a "Bad Initial ER-Hop Error" (0x04000004) status code. If the L bit is set and the local node is not part of the abstract node described by the first ER-Hop, the node selects a next hop that is along the path to the abstract node described by the first ER-Hop. If there is no ER-Hop TLV contained in the ER-TLV, the message is also in error, and the node **SHOULD** return a "Bad Explicit Routing TLV Error" (0x04000001) status code in a Label Release message sent to the upstream node. Note that this statement does not

preclude a Label Mapping message with no ER-TLV. If a Label Mapping message with no ER-TLV is received, then it MUST be processed as per [RFC7267].

2. If there are no further ER-Hop TLVs following the first ER-Hop TLV, this indicates the end of the explicit route. The Explicit Route TLV MUST be removed from the Label Mapping message. This node may or may not be the end of the PW. Processing continues as per Section 4.2, where a new Explicit Route TLV MAY be added to the Label Mapping message.
3. If a second ER-Hop TLV does exist, and the node is also a part of the abstract node described by the second ER-Hop, then the node deletes the first ER-Hop and continues processing with step 2, above. Note that this makes the second ER-Hop into the first ER-Hop for the iteration for the next PW segment.
4. The node determines if it is topologically adjacent to the abstract node described by the second ER-Hop. That is, it is directly connected to the next node by a PW control-plane adjacency. If so, the node selects a particular next hop that is a member of the abstract node. The node then deletes the first ER-Hop and continues processing as per Section 4.2, below.
5. Next, the node selects a next hop within the abstract node of the first ER-Hop that is along the path to the abstract node of the second ER-Hop. If no such path exists, then there are two cases:
 - A. If the second ER-Hop is a strict ER Hop, then there is an error, and the node MUST return a Label Release message to the upstream node with a "Bad Strict Node Error" (0x04000002) status code.
 - B. Otherwise, if the second ER-Hop is a loose ER Hop, then the node selects any next hop that is along the path to the next abstract node. If no path exists within the MPLS domain, then there is an error, and the node MUST return a Label Release message to the upstream node with a "Bad Loose Node Error" (0x04000003) status code.
6. Finally, the node replaces the first ER-Hop with any ER Hop that denotes an abstract node containing the next hop. This is necessary so that when the explicit route is received by the next hop, it will be accepted.
7. Progress the Label Mapping message to the next hop.

4.2. Adding ER Hops to the Explicit Route TLV

After selecting a next hop, the node MAY alter the explicit route in the following ways.

If, as part of executing the algorithm in Section 4.1, the Explicit Route TLV is removed, then the node MAY add a new Explicit Route TLV.

Otherwise, if the node is a member of the abstract node for the first ER-Hop, then a series of ER Hops MAY be inserted before the First ER Hop or the first ER-Hop MAY be replaced. Each ER Hop in this series MUST denote an abstract node that is a subset of the current abstract node.

Alternately, if the first ER-Hop is a loose ER Hop, an arbitrary series of ER Hops MAY be inserted prior to the first ER-Hop.

5. IANA Considerations

RFC 5036 [RFC5036] defines the LDP TLV name space, which is maintained by IANA, in the LDP "TLV Type Name Space" registry. TLV types for the Explicit Route TLV, the IPv4 Prefix ER-Hop TLV, and the IPv6 Prefix ER-Hop TLV are already defined in this registry.

IANA has assigned a further code point from the IETF consensus portion of this registry as follows:

TLV Type	Value	Reference
-----	-----	-----
L2 PW Address of Switching Point	0x0805	This Document

6. Security Considerations

This document introduces no new security considerations beyond those discussed in [RFC5036], [RFC4447], and [RFC7267]. The security considerations detailed in those documents apply to the protocol extensions described in this RFC.

As with [RFC7267], it should be noted that the path selection mechanisms specified in this document enable the network to automatically select the S-PEs that are used to forward packets on the MS-PW. Appropriate tools, such as the Virtual Circuit Connectivity Verification (VCCV) trace mechanisms specified in [RFC6073], can be used by an operator of the network to verify the path taken by the MS-PW and therefore be satisfied that the path does not represent an additional security risk.

7. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997, <<http://www.rfc-editor.org/info/rfc2119>>.
- [RFC3212] Jamoussi, B., Andersson, L., Callon, R., Dantu, R., Wu, L., Doolan, P., Worster, T., Feldman, N., Fredette, A., Girish, M., Gray, E., Heenanen, J., Kilty, T., and A. Malis, "Constraint-Based LSP Setup using LDP", RFC 3212, January 2002, <<http://www.rfc-editor.org/info/rfc3212>>.
- [RFC4447] Martini, L., Rosen, E., El-Aawar, N., Smith, T., and G. Heron, "Pseudowire Setup and Maintenance Using the Label Distribution Protocol (LDP)", RFC 4447, April 2006, <<http://www.rfc-editor.org/info/rfc4447>>.
- [RFC5003] Metz, C., Martini, L., Balus, F., and J. Sugimoto, "Attachment Individual Identifier (AII) Types for Aggregation", RFC 5003, September 2007, <<http://www.rfc-editor.org/info/rfc5003>>.
- [RFC5036] Andersson, L., Minei, I., and B. Thomas, "LDP Specification", RFC 5036, October 2007, <<http://www.rfc-editor.org/info/rfc5036>>.
- [RFC6073] Martini, L., Metz, C., Nadeau, T., Bocci, M., and M. Aissaoui, "Segmented Pseudowire", RFC 6073, January 2011, <<http://www.rfc-editor.org/info/rfc6073>>.
- [RFC7267] Martini, L., Bocci, M., and F. Balus, "Dynamic Placement of Multi-Segment Pseudowires", RFC 7267, June 2014, <<http://www.rfc-editor.org/info/rfc7267>>.

Acknowledgements

The authors gratefully acknowledge the contribution of the RFC 3212 [RFC3212] authors for the specification of the ER TLV and the ER-Hop TLV, which are reused by this document. The authors also gratefully acknowledge the input of Lizhong Jin.

Authors' Addresses

Pranjal Kumar Dutta
Alcatel-Lucent
701 E. Middlefield Road
Mountain View, California 94043
United States

EMail: pranjal.dutta@alcatel-lucent.com

Matthew Bocci
Alcatel-Lucent
Voyager Place, Shoppenhangers Road
Maidenhead, Berks SL6 2PJ
United Kingdom

EMail: matthew.bocci@alcatel-lucent.com

Luca Martini
Cisco Systems
9155 East Nichols Avenue, Suite 400
Englewood, Colorado 80112
United States

EMail: lmartini@cisco.com

