Internet Engineering Task Force (IETF)
Request for Comments: 7267
Updates: 6073
Category: Standards Track
ISSN: 2070-1721

L. Martini, Ed.
Cisco Systems, Inc.
M. Bocci, Ed.
F. Balus, Ed.
Alcatel-Lucent

June 2014

Dynamic Placement of Multi-Segment Pseudowires

Abstract

RFC 5254 describes the service provider requirements for extending the reach of pseudowires (PWs) across multiple Packet Switched Network domains. A multi-segment PW is defined as a set of two or more contiguous PW segments that behave and function as a single point-to-point PW. This document describes extensions to the PW control protocol to dynamically place the segments of the multi-segment pseudowire among a set of Provider Edge (PE) routers. This document also updates RFC 6073 by updating the value of the Length field of the PW Switching Point PE Sub-TLV Type 0x06 to 14.

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/rfc7267.

Table of Contents

1. Introduction

1.1. Scope

[RFC5254] describes the service provider requirements for extending the reach of pseudowires across multiple Packet Switched Network (PSN) domains. This is achieved using a multi-segment pseudowire (MS-PW). An MS-PW is defined as a set of two or more contiguous pseudowire (PW) segments that behave and function as a single point-to-point PW. This architecture is described in [RFC5659].

The procedures for establishing PWs that extend across a single PSN domain are described in [RFC4447], while procedures for setting up PWs across multiple PSN domains or control plane domains are described in [RFC6073].

The purpose of this document is to specify extensions to the pseudowire control protocol [RFC4447], and [RFC6073] procedures, to enable multi-segment PWs to be dynamically placed. The procedures follow the guidelines defined in [RFC5036] and enable the reuse of existing TLVs, and procedures defined for Single-Segment Pseudowires (SS-PWs) in [RFC4447]. Dynamic placement of point-to-multipoint (P2MP) PWs is for further study and outside the scope of this document.

1.2. 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 [RFC2119].

1.3. Terminology

[RFC5659] provides terminology for multi-segment pseudowires.

This document defines the following additional terms:

- Source Terminating Provider Edge (ST-PE): A Terminating Provider Edge (T-PE) that assumes the active signaling role and initiates the signaling for multi-segment PWs.

- Target Terminating Provider Edge (TT-PE): A Terminating Provider Edge (T-PE) that assumes the passive signaling role. It waits and responds to the multi-segment PW signaling message in the reverse direction.

- Forward Direction: ST-PE to TT-PE.

- Reverse Direction: TT-PE to ST-PE.

- Pseudowire Routing (PW routing): The dynamic placement of the segments that compose an MS-PW, as well as the automatic selection of Switching PEs (S-PEs).

1.4. Architecture Overview

The following figure shows the reference model, derived from [RFC5659], to support PW emulated services using multi-segment PWs.

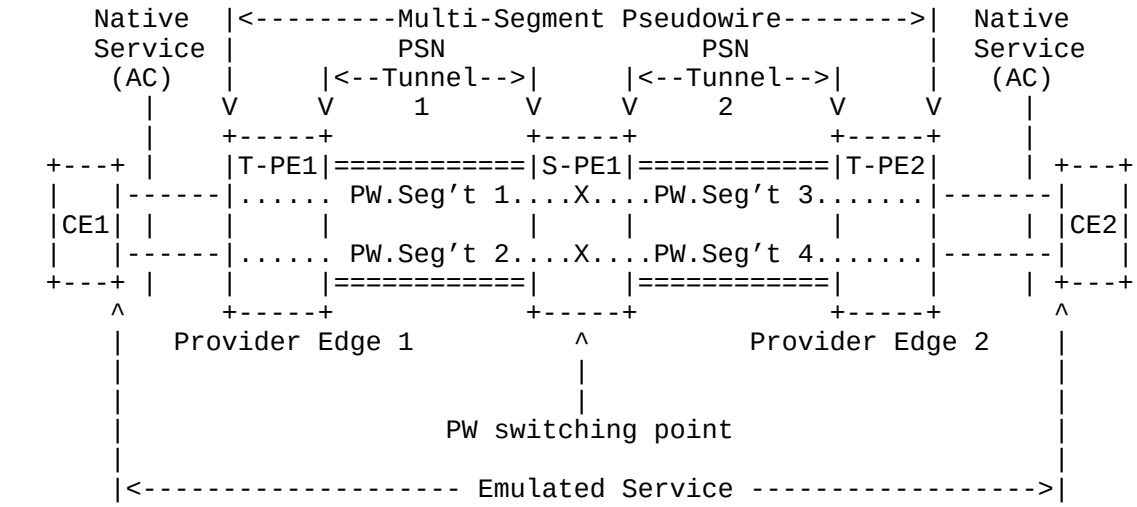

Figure 1: MS-PW Reference Model

The PEs that provide services to CE1 and CE2 are Terminating PE1 (T-PE1) and Terminating PE2 (T-PE2), respectively. A PSN tunnel extends from T-PE1 to Switching PE1 (S-PE1), and a second PSN tunnel extends from S-PE1 to T-PE2 . PWs are used to connect the attachment circuits (ACs) attached to PE1 to the corresponding ACs attached to T-PE2.

A PW segment on PSN Tunnel 1 is connected to a PW segment on PSN Tunnel 2 at S-PE1 to complete the multi-segment PW (MS-PW) between T-PE1 and T-PE2. S-PE1 is therefore the PW switching point and is referred to as the switching provider edge (S-PE). PW Segment 1 and PW Segment 3 are segments of the same MS-PW, while PW Segment 2 and PW Segment 4 are segments of another MS-PW. PW segments of the same MS-PW (e.g., PW Segment 1 and PW Segment 3) MUST be of the same PW type, and PSN tunnels can be of the same or a different technology. An S-PE switches an MS-PW from one segment to another based on the PW identifiers (PWid, or Attachment Individual Identifier (AII)). How

the PW protocol data units (PDUs) are switched at the S-PE depends on the PSN tunnel technology: in the case of a Multiprotocol Label Switching (MPLS) PSN to another MPLS PSN, PW switching involves a standard MPLS label swap operation.

Note that although Figure 1 only shows a single S-PE, a PW may transit more than one S-PE along its path. Although [RFC5659] describes MS-PWs that span more than one PSN, this document does not specify how the Label Distribution Protocol (LDP) is used for PW control [RFC4447] in an inter-AS (Autonomous System) environment.

2. Applicability

This document describes the case where the PSNs carrying the MS-PW are only MPLS PSNs using the Generalized Pseudowire Identifier (PWid) Forwarding Equivalence Class (FEC) element (also known as FEC 129).

Interactions with an IP PSN using the Layer 2 Tunneling Protocol version 3 (L2TPv3) as described in Section 8 of [RFC6073] are left for further study.

2.1. Changes to Existing PW Signaling

The procedures described in this document make use of existing LDP TLVs and related PW signaling procedures described in [RFC4447] and [RFC6073]. The following optional TLV is also defined:

- A Bandwidth TLV to address QoS Signaling requirements (see Section 4.2).

This document also updates the value of the Length field of the PW Switching Point PE Sub-TLV Type 0x06 to 14.

3. PW Layer 2 Addressing

Single-segment pseudowires on an MPLS PSN can use attachment circuit identifiers for a PW using FEC 129. In the case of a dynamically placed MS-PW, there is a requirement for the attachment circuit identifiers to be globally unique, for the purposes of reachability and manageability of the PW. Referencing Figure 1 above, individual globally unique addresses MUST be allocated to all the ACs and S-PEs of an MS-PW.

3.1. Attachment Circuit Addressing

The attachment circuit addressing is derived from AII Type 2 [RFC5003], as shown here:

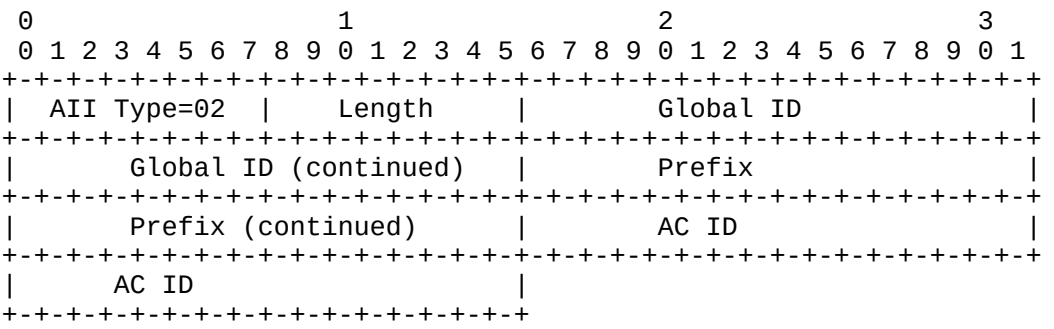

Figure 2: AII Type 2 TLV Structure

The fields are defined in Section 3.2 of [RFC5003].

Addressing schemes based on AII Type 2 permit varying levels of AII summarization, thus reducing the scaling burden on PW routing. PW addressing based on AII Type 2 is suitable for point-to-point provisioning models where auto-discovery of the address at the TT-PE is not required. That is, it is known a priori by provisioning.

Implementations of the following procedure MUST interpret the AII type to determine the meaning of the address format of the AII, irrespective of the number of segments in the MS-PW. All segments of the PW MUST be signaled with the same AII type.

A unique combination of Global ID, Prefix, and AC ID parts of the AII Type 2 are assigned to each AC. In general, the same Global ID and Prefix are be assigned for all ACs belonging to the same T-PE. This is not a strict requirement, however. A particular T-PE might have more than one Prefix assigned to it, and likewise a fully qualified AII with the same Global ID/Prefix but different AC IDs might belong to different T-PEs.

For the purpose of MS-PWs, the AII MUST be globally unique across all PSNs that are spanned by the MS-PW.

The AII for a local attachment circuit of a given T-PE of an MS-PW and the AII of the corresponding attachment circuit on a far-end T-PE (with respect to the LDP signaling) are known as the Source Attachment Individual Identifier (SAII) and Target Attachment Individual Identifier (TAII) as per [RFC6074].

3.2. S-PE Addressing

Each S-PE MUST be assigned an address that uniquely identifies it from a pseudowire perspective, in order to populate the PW Switching Point PE (SP-PE) TLV specified in [RFC6073]. For this purpose, at least one Attachment Identifier (AI) address of the format similar to AII Type 2 [RFC5003] composed of the Global ID, and Prefix part, only, MUST be assigned to each S-PE.

If an S-PE is capable of dynamic MS-PW signaling but is not assigned with an S-PE address, then on receiving a dynamic MS-PW Label Mapping message the S-PE MUST return a Label Release with the "Resources Unavailable" (0x38) status code.

4. Dynamic Placement of MS-PWs

[RFC6073] describes a procedure for concatenating multiple pseudowires together. This procedure requires each S-PE to be manually configured with the information required for each segment of the MS-PW. The procedures in the following sections describe a method to extend [RFC6073] by allowing the automatic selection of predefined S-PEs and dynamically establishing an MS-PW between two T-PEs.

4.1. Pseudowire Routing Procedures

The AII Type 2 described above contains a Global ID, Prefix, and AC ID. The TAII is used by S-PEs to determine the next SS-PW destination for LDP signaling.

Once an S-PE receives an MS-PW Label Mapping message containing a TAII with an AII that is not locally present, the S-PE performs a lookup in a PW AII routing table. If this lookup results in an IP address for the next-hop PE with reachability information for the AII in question, then the S-PE will initiate the necessary LDP messaging procedure to set up the next PW segment. If the PW AII routing table lookup does not result in an IP address for a next-hop PE, the destination AII has become unreachable, and the PW setup MUST fail. In this case, the next PW segment is considered unprovisioned, and a Label Release MUST be returned to the T-PE with a status message of "AII Unreachable".

If the TAII of an MS-PW Label Mapping message received by a PE contains the Prefix matching the locally provisioned prefix on that PE but an AC ID that is not provisioned, then the LDP liberal label retention procedures apply, and the Label Mapping message is retained.

To allow for dynamic end-to-end signaling of MS-PWs, information MUST be present in S-PEs to support the determination of the next PW signaling hop. Such information can be provisioned (equivalent to a static route) on each S-PE, or disseminated via regular routing protocols (e.g., BGP).

4.1.1. AII PW Routing Table Lookup Aggregation Rules

All PEs capable of dynamic MS-PW path selection MUST build a PW AII routing table to be used for PW next-hop selection.

The PW addressing scheme (AII Type 2 as defined in [RFC5003]) consists of a Global ID, a 32-bit Prefix, and a 32-bit Attachment Circuit ID.

An aggregation scheme similar to that used for classless IPv4 addresses can be employed. A length mask (8 bits) is specified as a number ranging from 0 to 96 that indicates which Most Significant Bits (MSBs) are relevant in the address field when performing the PW address-matching algorithm.

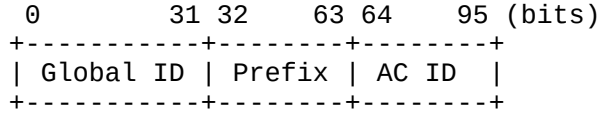

Figure 3: PW Addressing Scheme

During the signaling phase, the content of the (fully qualified) TAII Type 2 field from the FEC 129 TLV is compared against routes from the PW routing table. Similar to the IPv4 case, the route with the longest match is selected, determining the next signaling hop and implicitly the next PW segment to be signaled.

4.1.2. PW Static Route

For the purpose of determining the next signaling hop for a segment of the pseudowire, the PEs MAY be provisioned with fixed-route entries in the PW next-hop routing table. The static PW entries will follow all the addressing rules and aggregation rules described in the previous sections. The most common use of PW static provisioned routes is this example of the "default" route entry as follows:

Global ID = 0 Prefix = 0 AC ID = 0, Prefix Length = 0
Next Signaling Hop = {IP Address of next-hop S-PE or T-PE}

4.1.3. Dynamic Advertisement with BGP

Any suitable routing protocol capable of carrying external routing information MAY be used to propagate MS-PW path information among S-PEs and T-PEs. However, T-PEs and S-PEs MAY choose to use the Border Gateway Protocol (BGP) [RFC4271] with the Multiprotocol Extensions as defined in [RFC4760] to propagate PW address information throughout the PSN. PW address information is only propagated by PEs that are capable of PW switching. Therefore, the multiprotocol BGP neighbor topology MUST coincide with the topology of T-PEs and S-PEs.

Contrary to Layer 2 VPN signaling methods that use BGP for auto-discovery [RFC6074], in the case of the dynamically placed MS-PW, the source T-PE knows a priori (by provisioning) the AC ID on the terminating T-PE that signaling should use. Hence, there is no need to advertise a "fully qualified" 96-bit address on a per-PW attachment circuit basis. Only the T-PE Global ID, Prefix, and prefix length need to be advertised as part of well-known BGP procedures; see [RFC4760].

Since PW Endpoints are provisioned in the T-PEs, the ST-PE will use this information to obtain the first S-PE hop (i.e., first BGP next hop) to where the first PW segment will be established. Any subsequent S-PEs will use the same information (i.e., the next BGP next hop(s)) to obtain the next signaling hop(s) on the path to the TT-PE.

The PW dynamic path Network Layer Reachability Information (NLRI) is advertised in BGP UPDATE messages using the MP_REACH_NLRI and MP_UNREACH_NLRI attributes [RFC4760]. The {AFI, SAFI} value pair used to identify this NLRI is (AFI=25, SAFI=6). A route target MAY also be advertised along with the NLRI.

The Next Hop field of the MP_REACH_NLRI attribute SHALL be interpreted as an IPv4 address whenever the length of the NextHop address is 4 octets, and as an IPv6 address whenever the length of the NextHop address is 16 octets.

The NLRI field in the MP_REACH_NLRI and MP_UNREACH_NLRI is a prefix comprising an 8-octet Route Distinguisher, the Global ID, the Prefix, and the AC ID, and encoded as defined in Section 4 of [RFC4760].

This NLRI is structured as follows:

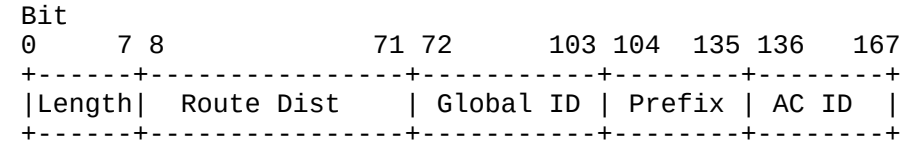

Figure 4: NLRI Field Structure

The Length field is the prefix length of the Route Distinguisher + Global ID + Prefix + AC ID in bits.

Except for the default PW route, which is encoded as a 0-length Prefix, the minimum value of the Length field is 96 bits. Lengths of 128 bits to 159 bits are invalid, as the AC ID field cannot be aggregated. The maximum value of the Length field is 160 bits. BGP advertisements received with invalid Prefix lengths MUST be rejected as having a bad packet format.

4.2. LDP Signaling

The LDP signaling procedures are described in [RFC4447] and expanded in [RFC6073]. No new LDP signaling components are required for setting up a dynamically placed MS-PW. However, some optional signaling extensions are described below.

One of the requirements that MUST be met in order to achieve the QoS objectives for a PW on a segment is that a PSN tunnel MUST be selected that can support at least the required class of service and that has sufficient bandwidth available.

Such PSN tunnel selection can be achieved where the next hop for a PW segment is explicitly configured at each PE, whether the PE is a T-PE or an S-PE in the case of a segmented PW, without dynamic path selection (as per [RFC6073]). In these cases, it is possible to explicitly configure the bandwidth required for a PW so that the T-PE or S-PE can reserve that bandwidth on the PSN tunnel.

Where dynamic path selection is used and the next hop is therefore not explicitly configured by the operator at the S-PE, a mechanism to signal the bandwidth for the PW from the T-PE to the S-PEs is required. This is accomplished by including an optional PW Bandwidth TLV. The PW Bandwidth TLV is specified as follows:

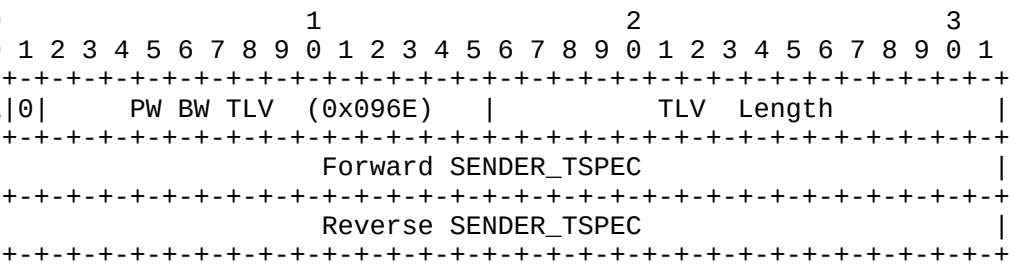

Figure 5: PW Bandwidth TLV Structure

The PW Bandwidth TLV fields are as follows:

- TLV Length: The length of the value fields in octets. Value = 64.

- Forward SENDER_TSPEC = the SENDER_TSPEC for the forward direction of the PW, as defined in Section 3.1 of [RFC2210].

- Reverse SENDER_TSPEC = the SENDER_TSPEC for the reverse direction of the PW, as defined in Section 3.1 of [RFC2210].

The complete definitions of the content of the SENDER_TSPEC objects are found in Section 3.1 of [RFC2210]. The forward SENDER_TSPEC refers to the data path in the direction ST-PE to TT-PE. The reverse SENDER_TSPEC refers to the data path in the direction TT-PE to ST-PE.

In the forward direction, after a next-hop selection is determined, a T/S-PE SHOULD reference the forward SENDER_TSPEC object to determine an appropriate PSN tunnel towards the next signaling hop. If such a tunnel exists, the MS-PW signaling procedures are invoked with the inclusion of the PW Bandwidth TLV. When the PE searches for a PSN tunnel, any tunnel that points to a next hop equivalent to the next hop selected will be included in the search (the LDP address TLV is used to determine the next-hop equivalence).

When an S/T-PE receives a PW Bandwidth TLV, once the PW next hop is selected, the S/T-PE MUST request the appropriate resources from the PSN. The resources described in the reverse SENDER_TSPEC are allocated from the PSN toward the originator of the message or

previous hop. When resources are allocated from the PSN for a specific PW, the allocation SHOULD account for the resource usage of the PW.

In the case where PSN resources towards the previous hop are not available, the following procedure MUST be followed:

i. The PSN MAY allocate more QoS resources, e.g., bandwidth, to the PSN tunnel.

ii. The S-PE MAY attempt to set up another PSN tunnel to accommodate the new PW QoS requirements.

iii. If the S-PE cannot get enough resources to set up the segment in the MS-PW, a Label Release MUST be returned to the previous hop with a status message of "Bandwidth resources unavailable".

In the latter case, the T-PE receiving the status message MUST also withdraw the corresponding PW Label Mapping message for the opposite direction if it has already been successfully set up.

If an ST-PE receives a Label Mapping message, the following procedure MUST be followed:

If the ST-PE has already sent a Label Mapping message for this PW, then the ST-PE MUST check to see if this Label Mapping message originated from the same LDP peer to which the corresponding Label Mapping message for this particular PW was sent. If it is the same peer, the PW is established. If it is a different peer, then the ST-PE MUST send a Label Release message with a status code of "PW Loop Detected" to the PE that originated the LDP Label Mapping message.

If the PE has not yet sent a Label Mapping message for this particular PW, then it MUST send the Label Mapping message to this LDP peer, regardless of what the PW TAII routing lookup result is.

4.2.1. Multiple Alternative Paths in PW Routing

A next-hop selection for a specific PW may find a match with a PW route that has multiple next hops associated with it. Multiple next hops may be either configured explicitly as static routes or learned through BGP routing procedures. Implementations at an S-PE or T-PE MAY use selection algorithms, such as CRC32 on the FEC TLV or flow-aware transport of PWs [RFC6391], for load balancing of PWs across multiple next hops, so that each PW has a single next hop. The details of such selection algorithms are outside the scope of this document.

4.2.2. Active/Passive T-PE Election Procedure

When an MS-PW is signaled, each T-PE might independently initiate signaling the MS-PW. This could result in a different path being used by each direction of the PW. To avoid this situation, one T-PE MUST initiate PW signaling (i.e., take an active role), while the other T-PE waits to receive the LDP Label Mapping message before sending the LDP Label Mapping message for the reverse direction of the PW (i.e., take a passive role). The active T-PE (the ST-PE) and the passive T-PE (the TT-PE) MUST be identified before signaling begins for a given MS-PW. Both T-PEs MUST use the same method for identifying which is active and which is passive.

A T-PE SHOULD determine whether it assumes the active role or the passive role using procedures similar to those of [RFC5036], Section 2.5.2, Bullet 2. The T-PE compares the Source Attachment Individual Identifier (SAII) [RFC6074] with the Target Attachment Individual Identifier (TAII) [RFC6074] as unsigned integers, and if the SAII > TAII, the T-PE assumes the active role. Otherwise, it assumes the passive role.

The following procedure for comparing the SAII and TAII as unsigned integers SHOULD be used:

- If the SAII Global ID > TAII Global ID, then the T-PE is active
  - else if the SAII Global ID < TAII Global ID, then the T-PE is passive
    - else if the SAII Prefix > TAII Prefix, then the T-PE is active
      - else if the SAII Prefix < TAII Prefix, then the T-PE is passive
        - else if the SAII AC ID > TAII AC ID, then the T-PE is active
          - else if the SAII AC ID < TAII AC ID, then the T-PE is passive
            - else there is a configuration error

4.2.3. Detailed Signaling Procedures

On receiving a Label Mapping message, the S-PE MUST inspect the FEC TLV. If the receiving node has no local AII matching the TAII for that Label Mapping message, then the Label Mapping message SHOULD be forwarded on to another S-PE or T-PE. The S-PE will check to see if the FEC is already installed for the forward direction:

- If the FEC is already installed and the received Label Mapping was received from the same LDP peer to which the forward LDP Label Mapping was sent, then this Label Mapping represents signaling in the reverse direction for this MS-PW segment.

- If the FEC is already installed and the received Label Mapping was received from a different LDP peer to which the forward LDP Label Mapping was sent, then the received Label Mapping MUST be released with a status code of "PW Loop Detected".

- If the FEC is not already installed, then this represents signaling in the forward direction.

The following procedures are then executed, depending on whether the Label Mapping was determined to be for the forward or the reverse direction of the MS-PW.

For the forward direction:

i. Determine the next-hop S-PE or T-PE according to the procedures above. If next-hop reachability is not found in the S-PE's PW AII routing table, then a Label Release MUST be sent with status code "AII Unreachable". If the next-hop S-PE or T-PE is found and is the same LDP peer that sent the Label Mapping message, then a Label Release MUST be returned with status code "PW Loop Detected". If the SAII in the received Label Mapping is local to the S-PE, then a Label Release MUST be returned with status code "PW Loop Detected".

ii. Check to see if a PSN tunnel exists to the next-hop S-PE or T-PE. If no tunnel exists to the next-hop S-PE or T-PE, the S-PE MAY attempt to set up a PSN tunnel.

iii. Check to see if a PSN tunnel exists to the previous hop. If no tunnel exists to the previous-hop S-PE or T-PE, the S-PE MAY attempt to set up a PSN tunnel.

iv. If the S-PE cannot get enough PSN resources to set up the segment to the next-hop or previous-hop S-PE or T-PE, a Label Release MUST be returned to the T-PE with a status message of "Resources Unavailable".

v. If the Label Mapping message contains a Bandwidth TLV, allocate the required resources on the PSN tunnels in the forward and reverse directions according to the procedures above.

vi. Allocate a new PW label for the forward direction.

vii. Install the FEC for the forward direction.

viii. Send the Label Mapping message with the new forward label and the FEC to the next-hop S-PE/T-PE.

For the reverse direction:

i. Install the FEC received in the Label Mapping message for the reverse direction.

ii. Determine the next signaling hop by referencing the LDP sessions used to set up the PW in the forward direction.

iii. Allocate a new PW label for the next hop in the reverse direction.

iv. Install the FEC for the next hop in the reverse direction.

v. Send the Label Mapping message with a new label and the FEC to the next-hop S-PE/ST-PE.

5. Procedures for Failure Handling

5.1. PSN Failures

Failures of the PSN tunnel MUST be handled by PSN mechanisms. An example of such a PSN mechanism is MPLS fast reroute [RFC4090]. If the PSN is unable to re-establish the PSN tunnel, then the S-PE SHOULD follow the procedures defined in Section 10 of [RFC6073].

5.2. S-PE Failures

For defects in an S-PE, the procedures defined in [RFC6073] SHOULD be followed. A T-PE or S-PE may receive an unsolicited Label Release message from another S-PE or T-PE with various failure codes, such as "Loop Detected", "PW Loop Detected", "Resources Unavailable", "Bad Strict Node Error", or "AII Unreachable". All these failure codes indicate a generic class of PW failures at an S-PE or T-PE.

If an unsolicited Label Release message with such a failure status code is received at a T-PE, then it is RECOMMENDED that the T-PE attempt to re-establish the PW immediately. However, the T-PE MUST throttle its PW setup message retry attempts with an exponential backoff in situations where PW setup messages are being constantly released. It is also RECOMMENDED that a T-PE detecting such a situation take action to notify an operator.

S-PEs that receive an unsolicited Label Release message with a failure status code SHOULD follow this procedure:

i. If the Label Release is received from an S-PE or T-PE in the forward or reverse signaling direction, then the S-PE MUST tear down both segments of the PW. The status code received in the Label Release message SHOULD be propagated when sending the Label Release for the next segment.

5.3. PW Reachability Changes

In general, an established MS-PW will not be affected by next-hop changes in AII reachability information.

If there is a next-hop change in AII reachability information in the forward direction, the T-PE MAY elect to tear down the MS-PW by sending a Label Withdraw message to the downstream S-PE or T-PE. The teardown MUST also be accompanied by an unsolicited Label Release message and will be followed by an attempt by the T-PE to re-establish the MS-PW.

If there is a change in the AII reachability information in the forward direction at an S-PE, the S-PE MAY elect to tear down the MS-PW in both directions. A label withdrawal is sent in each direction followed by an unsolicited Label Release. The unsolicited Label Release messages MUST be accompanied by the status code "AII Unreachable". This procedure is OPTIONAL. Note that this procedure is likely to be disruptive to the emulated service. PW Redundancy [RFC6718] MAY be used to maintain the connectivity used by the emulated service in the case of a failure of the PSN or S-PE.

A change in AII reachability information in the reverse direction has no effect on an MS-PW.

6. Operations, Administration, and Maintenance (OAM)

The OAM procedures defined in [RFC6073] may also be used for dynamically placed MS-PWs. A PW Switching Point PE TLV [RFC6073] is used to record the switching points that the PW traverses.

In the case of an MS-PW where the PW Endpoints are identified by using globally unique AII addresses based on FEC 129, there is no pseudowire identifier (PWid) defined on a per-segment basis. Each individual PW segment is identified by the address of the adjacent S-PE(s) in conjunction with the SAII and TAII.

In this case, the following TLV type (0x06) MUST be used in place of type 0x01 in the PW Switching Point PE TLV:

Type	Length	Description
0x06	14	L2 PW address of PW Switching Point

The above sub-TLV MUST be included in the PW Switching Point PE TLV once per individual PW switching point, following the same rules and procedures as those described in [RFC6073]. A more detailed description of this sub-TLV is also given in Section 7.4.1 of [RFC6073]. However, the length value MUST be set to 14 ([RFC6073] states that the length value is 12, but this does not correctly represent the actual length of the TLV).

7. Security Considerations

This document specifies extensions to the protocols already defined in [RFC4447] and [RFC6073]. The extensions defined in this document do not affect the security considerations for those protocols, but [RFC4447] and [RFC6073] do impose a set of security considerations that are applicable to the protocol extensions specified in this document.

It should be noted that the dynamic 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.

Note that the PW control protocol may be used to establish and maintain an MS-PW across administrative boundaries. Section 13 of [RFC6073] specifies security considerations applicable to LDP used in this manner, including considerations for establishing the integrity of, and authenticating, LDP control messages. These considerations also apply to the protocol extensions specified in this document.

Note that the protocols for dynamically distributing AII reachability information may have their own security considerations. However, those protocol specifications are outside the scope of this document.

8. IANA Considerations

8.1. Correction

IANA has corrected a minor error in the "Pseudowire Switching Point PE sub-TLV Type" registry. The entry 0x06 "L2 PW address of the PW Switching Point" has been corrected to Length 14 and the reference changed to [RFC6073] and this document as follows:

Type	Length	Description	Reference
0x06	14	L2 PW Address of PW Switching Point	[RFC6073][RFC7267]

8.2. LDP TLV Type Name Space

This document defines one new LDP TLV type. IANA already maintains a registry for LDP TLV types, called the "TLV Type Name Space" registry, within the "Label Distribution Protocol (LDP) Parameters" registry as defined by [RFC5036]. IANA has assigned the following value.

Value	Description	Reference	Notes/Registration Date
0x096E	Bandwidth TLV	This document	

8.3. LDP Status Codes

This document defines three new LDP status codes. IANA maintains a registry of these codes, called the "Status Code Name Space" registry, in the "Label Distribution Protocol (LDP) Parameters" registry as defined by [RFC5036]. The IANA has assigned the following values.

Range/Value	E	Description	Reference
0x00000037	0	Bandwidth resources unavailable	This document
0x00000038	0	Resources Unavailable	This document
0x00000039	0	AII Unreachable	This document

8.4. BGP SAFI

IANA has allocated a new BGP SAFI for "Network Layer Reachability Information used for Dynamic Placement of Multi-Segment Pseudowires" in the IANA "SAFI Values" registry [RFC4760] within the "Subsequent Address Family Identifiers (SAFI) Parameters" registry. The IANA has assigned the following value.

Value	Description	Reference
6	Network Layer Reachability Information used for Dynamic Placement of Multi-Segment Pseudowires	This document

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.

[RFC2210] Wroclawski, J., "The Use of RSVP with IETF Integrated Services", RFC 2210, September 1997.

[RFC4447] Martini, L., Ed., Rosen, E., El-Aawar, N., Smith, T., and G. Heron, "Pseudowire Setup and Maintenance Using the Label Distribution Protocol (LDP)", RFC 4447, April 2006.

[RFC5003] Metz, C., Martini, L., Balus, F., and J. Sugimoto, "Attachment Individual Identifier (AII) Types for Aggregation", RFC 5003, September 2007.

[RFC5036] Andersson, L., Ed., Minei, I., Ed., and B. Thomas, Ed., "LDP Specification", RFC 5036, October 2007.

[RFC6073] Martini, L., Metz, C., Nadeau, T., Bocci, M., and M. Aissaoui, "Segmented Pseudowire", RFC 6073, January 2011.

9.2. Informative References

[RFC4090] Pan, P., Ed., Swallow, G., Ed., and A. Atlas, Ed., "Fast Reroute Extensions to RSVP-TE for LSP Tunnels", RFC 4090, May 2005.

[RFC4271] Rekhter, Y., Ed., Li, T., Ed., and S. Hares, Ed., "A Border Gateway Protocol 4 (BGP-4)", RFC 4271, January 2006.

[RFC4760] Bates, T., Chandra, R., Katz, D., and Y. Rekhter, "Multiprotocol Extensions for BGP-4", RFC 4760, January 2007.

[RFC5254] Bitar, N., Ed., Bocci, M., Ed., and L. Martini, Ed., "Requirements for Multi-Segment Pseudowire Emulation Edge-to-Edge (PWE3)", RFC 5254, October 2008.

[RFC5659] Bocci, M. and S. Bryant, "An Architecture for Multi-Segment Pseudowire Emulation Edge-to-Edge", RFC 5659, October 2009.

[RFC6074] Rosen, E., Davie, B., Radoaca, V., and W. Luo, "Provisioning, Auto-Discovery, and Signaling in Layer 2 Virtual Private Networks (L2VPNs)", RFC 6074, January 2011.

[RFC6391] Bryant, S., Ed., Filsfils, C., Drafz, U., Kompella, V., Regan, J., and S. Amante, "Flow-Aware Transport of Pseudowires over an MPLS Packet Switched Network", RFC 6391, November 2011.

[RFC6718] Muley, P., Aissaoui, M., and M. Bocci, "Pseudowire Redundancy", RFC 6718, August 2012.

10. Contributors

The editors gratefully acknowledge the following people for their contributions to this document:

Nabil Bitar
Verizon
40 Sylvan Road
Waltham, MA 02145
US

EMail: nabil.bitar@verizon.com

Himanshu Shah
Ciena Corp.
35 Nagog Park
Acton, MA 01720
US

EMail: hshah@ciena.com

Mustapha Aissaoui
Alcatel-Lucent
600 March Road
Kanata
ON, Canada

EMail: mustapha.aissaoui@alcatel-lucent.com

Jason Rusmisel
Alcatel-Lucent
600 March Road
Kanata
ON, Canada

EMail: Jason.rusmisel@alcatel-lucent.com

Andrew G. Malis
Huawei
2330 Central Expressway
Santa Clara, CA 95050
US

EMail: agmalis@gmail.com

Chris Metz
Cisco Systems, Inc.
3700 Cisco Way
San Jose, CA 95134
US

EMail: chmetz@cisco.com

David McDysan
Verizon
22001 Loudoun County Pkwy.
Ashburn, VA 20147
US

EMail: dave.mcdysan@verizon.com

Jeff Sugimoto
Alcatel-Lucent
701 E. Middlefield Rd.
Mountain View, CA 94043
US

EMail: jeffery.sugimoto@alcatel-lucent.com

Mike Loomis
Alcatel-Lucent
701 E. Middlefield Rd.
Mountain View, CA 94043
US

EMail: mike.loomis@alcatel-lucent.com

11. Acknowledgements

The editors also gratefully acknowledge the input of the following people: Paul Doolan, Mike Duckett, Pranjal Dutta, Ping Pan, Prayson Pate, Vasile Radoaca, Yeongil Seo, Yetik Serbest, and Yuichiro Wada.

Authors' Addresses

Luca Martini (editor)
Cisco Systems, Inc.
9155 East Nichols Avenue, Suite 400
Englewood, CO 80112
US

EMail: lmartini@cisco.com

Matthew Bocci (editor)
Alcatel-Lucent
Voyager Place
Shoppenhangers Road
Maidenhead
Berks, UK

EMail: matthew.bocci@alcatel-lucent.com

Florin Balus (editor)
Alcatel-Lucent
701 E. Middlefield Rd.
Mountain View, CA 94043
US

EMail: florin@nuagenetworks.net