Network Working Group
Internet Draft
Expiration Date: January 2009
Intended status: Standards Track

Luca Martini
Chris Metz
Cisco Systems Inc.

Thomas D. Nadeau
BT

Matthew Bocci
Florin Balus
Mustapha Aissaoui
Alcatel-Lucent

Mike Duckett
Bellsouth

July 2008

Segmented Pseudo Wire

draft-ietf-pwe3-segmented-pw-09.txt

Abstract

This document describes how to connect pseudowires (PW) between two distinct PW control planes or PSN domains. The PW control planes may belong to independent autonomous systems, or the PSN technology is heterogeneous, or a PW might need to be aggregated at a specific PSN point. The PW packet data units are simply switched from one PW to

another without changing the PW payload.

Table of Contents

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. Terminology

- PW Terminating Provider Edge (T-PE). A PE where the customer-facing attachment circuits (ACs) are bound to a PW forwarder. A Terminating PE is present in the first and last segments of a MS-PW. This incorporates the functionality of a PE as defined in [RFC3985].

- Single-Segment Pseudowire (SS-PW). A PW setup directly between two T-PE devices. Each PW in one direction of a SS-PW traverses one PSN tunnel that connects the two T-PEs.

- Multi-Segment Pseudowire (MS-PW). A static or dynamically configured set of two or more contiguous PW segments that behave and function as a single point-to-point PW. Each end of a MS-PW by definition MUST terminate on a T-PE.

- PW Segment. A part of a single-segment or multi-segment PW, which is set up between two PE devices, T-PEs and/or S-PEs.

- PW Switching Provider Edge (S-PE). A PE capable of switching the control and data planes of the preceding and succeeding PW segments in a MS-PW. The S-PE terminates the PSN tunnels of the preceding and succeeding segments of the MS-PW. It is therefore a

- PW switching point for a MS-PW. A PW Switching Point is never the S-PE and the T-PE for the same MS-PW. A PW switching point runs necessary protocols to setup and manage PW segments with other PW switching points and terminating PEs.

3. Introduction

PWE3 defines the signaling and encapsulation techniques for establishing SS-PWs between a pair of ultimate PEs and in the vast majority of cases this will be sufficient. MS-PWs come into play in two general cases:

-i. When it is not possible, desirable or feasible to establish a PW control channel between the ultimate source and destination PEs. At a minimum PW control channel establishment requires knowledge of and reachability to the remote (ultimate) PE IP address. The local (ultimate) PE may not have access to this information related to topology, operational or security constraints.

An example is the inter-AS L2VPN scenario where the ultimate PEs reside in different provider networks (ASes) and it is

the practice to MD5-key all control traffic exchanged between two networks. Technically a SS-PW could be used but this would require MD5-keying on ALL ultimate source and destination PE nodes. An MS-PW allows the providers to confine MD5 key administration to just the PW switching points connecting the two domains.

A second example might involve a single AS where the PW setup path between the ultimate PEs is computed by an external entity (i.e. client-layer routing protocol). Assume a full mesh of PWE3 control channels established between PE-A, PE-B and PE-C. A client-layer L2 connection tunneled through a PW is required between ultimate PE-A and PE-C. The external entity computes a PW setup path that passes through PE-B. This results in two discrete PW segments being built: one between PE-A and PE-B and one between PE-B and PE-C. The successful client-layer L2 connection between ultimate PE-A and ultimate PE-C requires that PE-B performs the PWE3 switching process.

A third example involves the use of PWs in hierarchical IP/MPLS networks. Access networks connected to a backbone use PWs to transport customer payloads between customer sites serviced by the same access network and up to the edge of the backbone where they can be terminated or switched onto a succeeding PW segment crossing the backbone. The use of PWE3 switching between the access and backbone networks can potentially reduce the PWE3 control channels and routing information processed by the access network T-PEs.

It should be noted that PWE3 switching does not help in any way to reduce the amount of PW state supported by each access network T-PE.

-ii. PWE3 signaling and encapsulation protocols are different. The ultimate PEs are connected to networks employing different PW signaling and encapsulation protocols. In this case it is not possible to use a SS-PW. A MS-PW with the appropriate interworking performed at the PW switching points can enable PW connectivity between the ultimate PEs in this scenario.

There are four different signaling protocols that are defined to signal PWs:

-i. Static configuration of the PW (MPLS or L2TPv3).
-ii. LDP using FEC 128
-iii. LDP using the generalized FEC 129
-iv. L2TPv3

4. General Description

A pseudowire (PW) is a tunnel established between two provider edge (PE) nodes to transport L2 PDUs across a packet switched network (PSN) as described in Figure 1 and in [PWE3-ARCH]. Many providers are looking at PWs as a means of migrating existing (or building new) L2VPN services (i.e. Frame-Relay, ATM, Ethernet) on top of a PSN by using PWs. PWs may span multiple autonomous systems of the same or different provider networks. In these scenarios PW control channels (i.e. targeted LDP, L2TPv3) and PWs will cross AS boundaries.

Inter-AS L2VPN functionality is currently supported and several techniques employing MPLS encapsulation and LDP signaling have been documented [2547BIS]. It is also straightforward to support the same inter-AS L2VPN functionality employing L2TPv3. In this document we define methodology to switch a PW between two PW control planes.

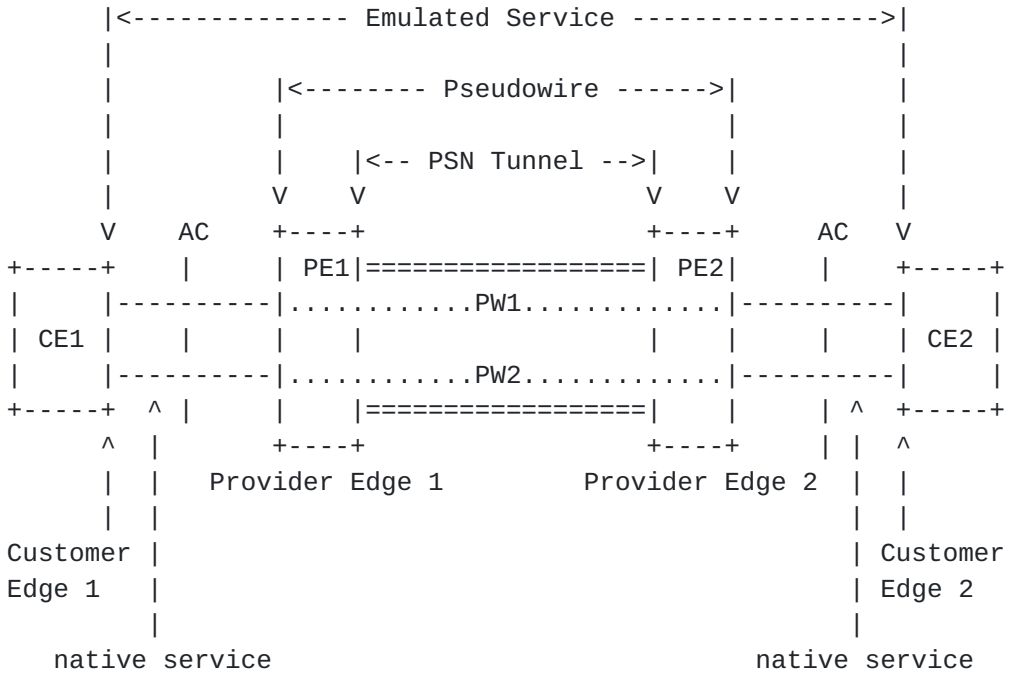


Figure 1: PWE3 Reference Model

There are two methods for switching a PW between two PW control planes. In the first method (Figure 2), the two control planes

terminate on different PEs.

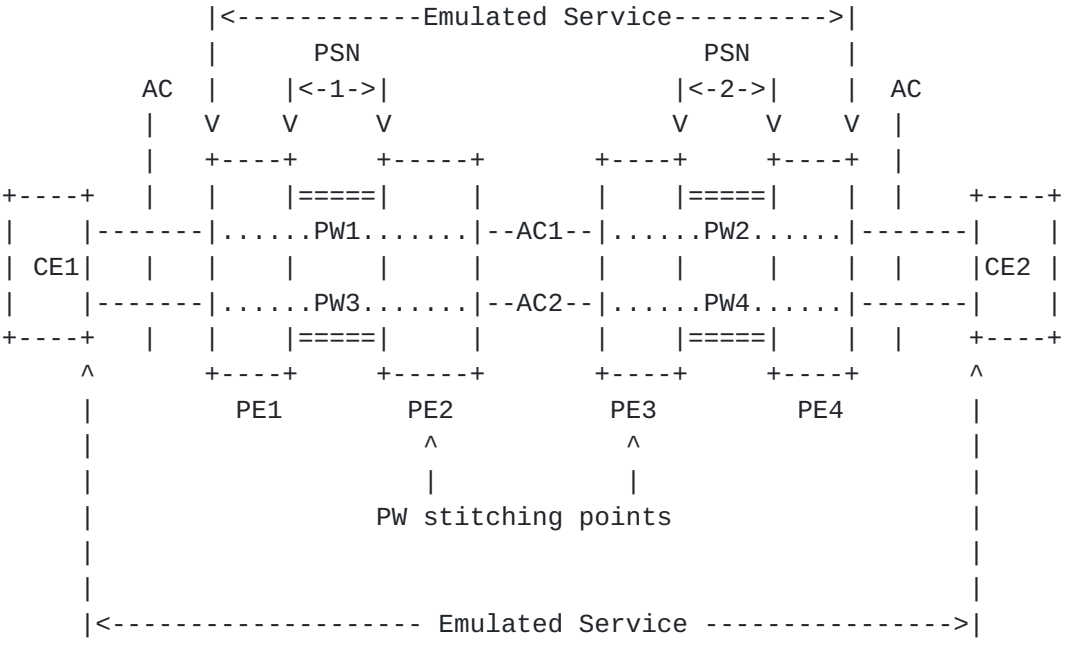


Figure 2: PW Switching using ACs Reference Model

In Figure 2, pseudowires in two separate PSNs are stitched together using native service attachment circuits. PE2 and PE3 only run the control plane for the PSN to which they are directly attached. At PE2 and PE3, PW1 and PW2 are connected using attachment circuit AC1, while PW3 and PW4 are connected using attachment circuit AC2.

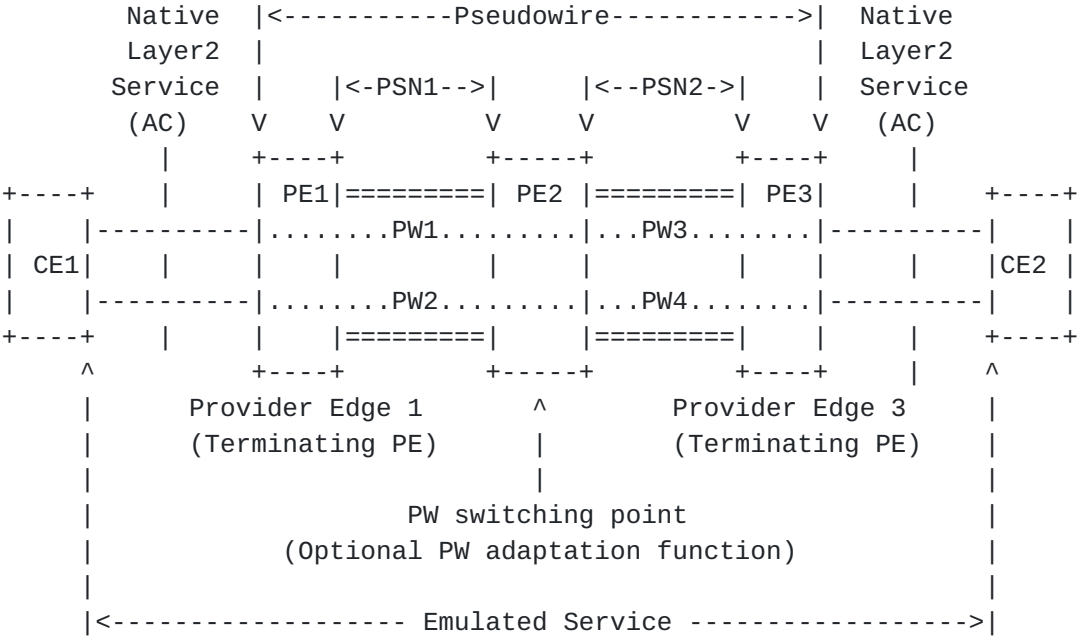


Figure 3: PW Control Plane Switching Reference Model

In Figure 3 PE2 runs two separate control planes: one toward PE1, and one Toward PE3. The PW switching point is at PE2 which is configured to connect PW1 and PW3 together to complete the multi-hop PW between PE1 and PE3. PW1 and PW3 MUST be of the same PW type, but PSN1 and PSN2 need not be the same technology. In the latter case, if the PW is switched to a different technology, the PEs must adapt the PDU encapsulation between the different PSN technologies. In the case where PSN1 and PSN2 are the same technology the PW PDU does not need to be modified, and PDUs are then switched between the pseudowires at the PW label level.

It should be noted that it is possible to adapt one PSN technology to a different one, for example MPLS over an IP or GRE [RFC4023] encapsulation, but this is outside the scope of this document. Further, one could perform an interworking function on the PWs themselves at the PW switching point, allowing conversion from one PW type to another, but this is also outside the scope of this document.

This document describes procedures for building multi-segment pseudowires using manual configuration of the switching point PE2. Other documents may build on this base specification to automate the configuration and selection of PE2. It should also be noted that a PW can traverse multiple PW switching points along it's path, and the edge PEs will not require any specific knowledge of how many PW switching points the PW has traversed (though this may be reported for troubleshooting purposes).

In general the approach taken is to connect the individual control planes by passing along any signaling information immediately upon reception. First the PW switching point is configured to switch a SS-PW from a specific peer to another SS-PW destined for a different peer. No control messages are exchanged yet as the PW switching point PE does not have enough information to actually initiate the PW setup messages. However, if a session does not already exist, a control protocol (LDP/L2TP) session is setup. In this model the MS-PW setup is starting from the T-PE devices. Next once the T-PE is configured it sends the PW control setup messages. These messages are received, and immediately used to form the PW setup messages for the next SS-PW of the MS-PW. If one of the Switching PEs doesn't accept an LDP Label Mapping message then a Label Release message may be sent back to the originator T-PE depending on the cause of the error. LDP liberal label retention mode still applies, hence if a PE is simply not configured yet , the label mapping is stored for future use. A MS-PW is declared UP when all the constituent SS-PWs are UP.

5. PW Switching and Attachment Circuit Type

The PWs in each PSN are established independently, with each PSN being treated as a separate PWE3 domain. For example, in Figure 2 for case of MPLS PSNs, PW1 is setup between PE1 and PE2 using the LDP targeted session as described in [RFC4447], and at the same time a separate pseudowire, PW2, is setup between PE3 and PE4. The ACs for PW1 and PW2 at PE2 and PE3 MUST be configured such that they are the same PW type e.g. ATM VCC, Ethernet VLAN, etc.

6. Applicability

The general applicability of MS-PWs and their relationship to L2VPNs is described in [MS-PW-ARCH]. The applicability of a PW type, as specified in the relevant RFC for that encapsulation (e.g. [RFC4717] for ATM), applies to each segment. This section describes further applicability considerations.

In comon with SS-PWs, the performance of any segment of a MS-PW is equal to the performance of the PSN plus any impairments introduced by the PW layer itself. Therefore it is not possible for the MS-PW to provide better performance than the PSN over which it is transported. Furthermore, the overall performance of an MS-PW is no better than the worst performing segment of that MS-PW.

Since different PSN types may be able to achieve different maximum performance objectives, it is necessary to carefully consider which PSN types are used along the path of a MS-PW.

7. PW-MPLS to PW-MPLS Control Plane Switching

Referencing Figure 3, PE2 sets up a PW1 using the LDP targeted session as described in [RFC4447], at the same time a separate pseudowire PW3 is setup to PE3. Each PW is configured independently on the PEs, but on PE2 pseudowire PW1 is connected to pseudowire PW3. PDUs are then switched between the pseudowires at the PW label level. Hence the data plane does not need any special knowledge of the specific pseudowire type. A simple standard MPLS label swap operation is sufficient to connect the two PWs, and in this case the PW adaptation function is not used. However when pushing a new PSN label the TTL SHOULD be set to 255, or some other locally configured fixed value.

This process can be repeated as many times as necessary, the only limitation to the number of PW switching points traversed is imposed by the TTL field of the PW MPLS Label. The setting of the TTL is a matter of local policy on the originating PE, but SHOULD be set to 255.

There are three MPLS to MPLS PW control planes:
-i. Static configuration of the PW.
-ii. LDP using FEC 128
-iii. LDP using the generalized FEC 129
This results in four distinct PW switching situations that are significantly different, and must be considered in detail:
-i. PW Switching between two static control planes.
-ii. Static Control plane switching to LDP dynamic control plane.
-iii. Two LDP control planes using the same FEC type
-iv. LDP using FEC 128, to LDP using the generalized FEC 129

7.1. Static Control plane switching

In the case of two static control planes the PW switching point MUST be configured to direct the MPLS packets from one PW into the other. There is no control protocol involved in this case. When one of the control planes is a simple static PW configuration and the other control plane is a dynamic LDP FEC 128 or generalized PW FEC, then the static control plane should be considered identical to an attachment circuit (AC) in the reference model of Figure 1. The switching point PE SHOULD signal the proper PW status if it detects a failure in sending or receiving packets over the static PW. Because the PW is statically configured, the status communicated to the dynamic LDP PW will be limited to local interface failures. In this case, the PW switching point PE behaves in a very similar manner to a T-PE, assuming an active role. This means that the S-PE will immediately send the LDP Label Mapping message if the static PW is deemed to be UP.

7.2. Two LDP control planes using the same FEC type

As stated in a section above, the PW switching point PE should assume an initial passive role. This means that once independent PWs are configured on the switching point, the LSR does not advertise the LDP PW FEC mapping until it has received at least one of the two PW LDP FECs from a remote PE. This is necessary because the switching point LSR does not know a priori what the interface parameter field in the initial FEC advertisement will contain.

The PWID is a unique number between each pair of PEs. Hence Each SS-

PW that forms an MS-PW may have a different PWID. In the case of The Generalized PW FEC, the AGI/SAI/TAI may have to also be different for some, or sometimes all, SS-PWs.

7.2.1. FEC 129 Active/Passive T-PE Election Procedure

When a MS-PW is signaled using FEC 129, each T-PE might independently start signaling the MS-PW. If the MS-PW path is not statically configured, in certain cases the signaling procedure could result in an attempt to setup each direction of the MS-PW through different paths. To avoid this situation one of the T-PE MUST start the PW signaling (active role), while the other waits to receive the LDP label mapping before sending the respective PW LDP label mapping message. (passive role). When the MS-PW path not statically configured, the Active T-PE (the ST-PE) and the passive T-PE (the TT-PE) MUST be identified before signaling is initiated for a given MS-PW.

The determination of which T-PE assume the active role SHOULD be done as follows:

the SAII and TAII are compared as unsigned integers, if the SAII is bigger then the T-PE assumes the active role.

The selection process to determine which T-PE assumes the active role MAY be superseded by manual provisioning.

7.3. LDP FEC 128 to LDP using the generalized FEC 129

When a PE is using the generalized FEC 129, there are two distinct roles that a PE can assume: active and passive. A PE that assumes the active role will send the LDP PW setup message, while a passive role PE will simply reply to an incoming LDP PW setup message. The PW switching point PE, will always remain passive until a PWID FEC 128 LDP message is received, which will cause the corresponding generalized PW FEC LDP message to be formed and sent. If a generalized FEC PW LDP message is received while the switching point PE is in a passive role, the corresponding PW FEC 128 LDP message will be formed and sent.

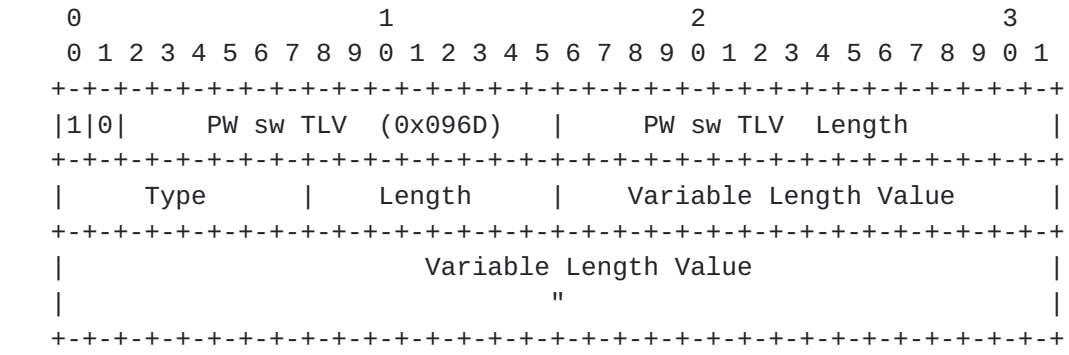
PW IDs need to be mapped to the corresponding AGI/TAI/SAI and vice versa. This can be accomplished by local PW switching point configuration, or by some other means, such as some form of auto discovery. Such other means are outside the scope of this document.

7.4. LDP PW switching point TLV

The edge to edge PW might traverse several switching points, in separate administrative domains. For management and troubleshooting reasons it is useful to record information about the switching points that the PW traverses. This is accomplished by using a PW switching point TLV.

Note that sending the PW switching point TLV is OPTIONAL, however the PE or S-PE MUST process the TLV upon reception. The PW switching point TLV is appended to the PW FEC at each switching point. Each S-PE can append a PW switching point TLV, and the order of the PW switching point TLVs MUST be preserved. The S-PE TLV MUST be sent if VCCV operation is required beyond the first MS-PW segment from a T-PE.

The PW switching point TLV encoded as follows:



[note] LDP TLV type is pending IANA approval.

- PW sw TLV Length

  Specifies the total length of all the following PW switching point TLV fields in octets

- Type

  Encodes how the Value field is to be interpreted.

- Length

  Specifies the length of the Value field in octets.

- Value

  Octet string of Length octets that encodes information to be interpreted as specified by the Type field.

PW Switching point TLV Types are assigned by IANA according the the process defined in the "IANA Allocations" section below.

The PW switching Point TLV is an OPTIONAL TLV that should appear only once for each switching point traversed. If the S-PE TLV is not present with the required sub-TLVs, VCCV operation will not function.

For local policy reasons, a particular S-PE can filter out all S-PE TLVs in a label mapping message that traverses it and not include it's own S-PE TLV. In this case, from any upstream PE, it will appear as if this particular S-PE is the T-PE. This might be necessary , depending on local policy if the S-PE is at the Service provider administrative boundary.

7.4.1. PW Switching Point Sub-TLVs

Below are details specific to PW Switching Point Sub-TLVs described in this document:

- PW ID of last PW segment traversed. This is only applicable if the last PW segment traversed used LDP FEC 128 to signal the PW.

  This sub-TLV type contains a PW ID in the format of the PWID described in [RFC4447]. This is just a 32 bit unsigned integer number.

- PW Switching Point description string.

  An optional description string of text up to 80 characters long.

- Local IP address of PW Switching Point.

  The Local IP V4 or V6 address of the PW Switching Point. This is an OPTIONAL Sub-TLV. In most cases this will be the local LDP session IP address of the S-PE.

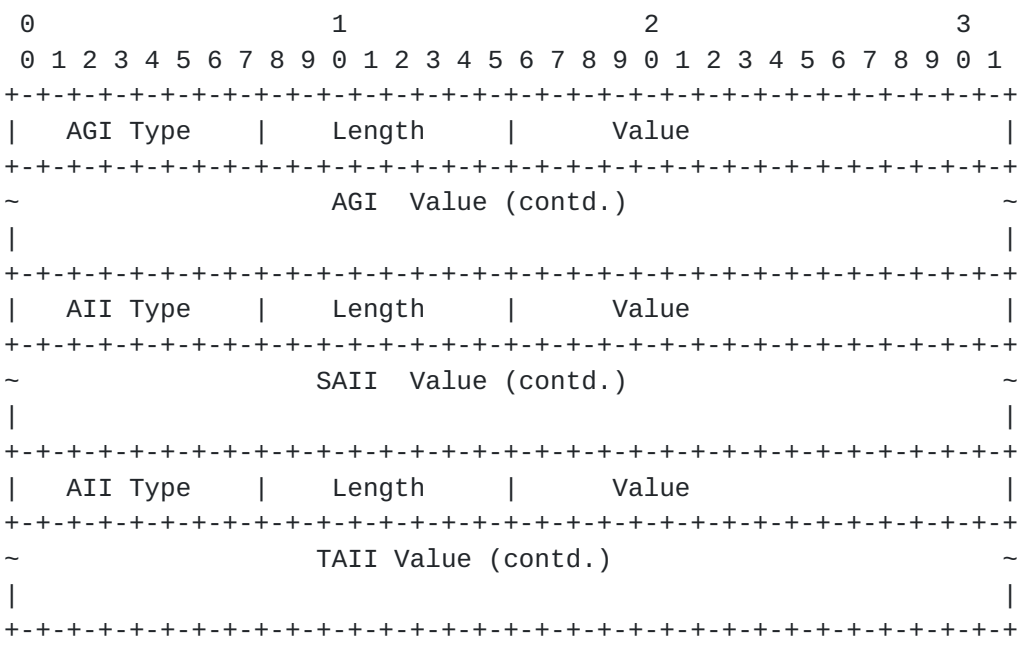
- Remote IP address of the last PW Switching Point traversed or of the T-PE

  The IP V4 or V6 address of the last PW Switching Point traversed or of the T-PE. This is an OPTIONAL Sub-TLV. In most cases this will be the remote IP address of the LDP session. This Sub-TLV SHOULD only be included if there are no other S-PE TLV present from other S-PEs, or if the remote ip address of the LDP session does not correspond to the "Local IP address of PW Switching Point" TLV value contained in the last S-PE TLV.

- The FEC of last PW segment traversed.

  This is only applicable if the last PW segment traversed used LDP FEC 129 to signal the PW.

  The Attachment Identifier of the last PW segment traversed. This is coded in the following format:



- L2 PW address of PW Switching Point (recommended format).

This sub-TLV type contains a L2 PW address of PW Switching Point in the format described in [RFC5003]. This includes the AII type field , and length, as well as the L2 PW address.

7.4.2. Adaptation of Interface Parameters

[RFC4447] defines several interface parameters, which are used by the Network Service Processing (NSP) to adapt the PW to the Attachment Circuit (AC). The interface parameters are only used at the end points, and MUST be passed unchanged across the PW switching point. However the following interface parameters MAY be modified as follows:

- 0x03 Optional Interface Description string This Interface parameter MAY be modified, or altogether removed from the FEC element depending on local configuration policies.

- 0x09 Fragmentation indicator This parameter MAY be inserted in the FEC by the switching point if it is capable of re-assembly of fragmented PW frames according to [PWE3-FRAG].

- 0x0C VCCV parameter This Parameter contains the CC type , and CV type bit fields. The CV type bit field MUST be reset to reflect the CV type supported by the S-PE. CC type bit field MUST have bit 1 "Type 2: MPLS Router Alert Label " set to 0. The other bit fields MUST be reset to reflect the CC type supported by the S-PE.

7.5. Group ID

The Group ID (GR ID) is used to reduce the number of status messages that need to be sent by the PE advertising the PW FEC. The GR ID has local significance only, and therefore MUST be mapped to a unique GR ID allocated by the PW switching point PE.

7.6. PW Loop Detection

A switching point PE SHOULD check the OPTIONAL PW switching Point TLV, to verify if it's own IP address appears in it. If it's IP address appears in a received PW switching Point TLV, the PE SHOULD break the loop, and send a label release message with the following error code:

Assignment	E	Description
0x0000003A	0	"PW Loop Detected"

[note: error code pending IANA allocation]

8. PW-MPLS to PW-L2TPv3 Control Plane Switching

Both MPLS and L2TPv3 PWs may be static or dynamic. This results in four possibilities when switching between L2TPv3 and MPLS.

-i. Switching between MPLS and L2TPv3 static control planes.
-ii. Switching between a static MPLS PW and a dynamic L2TPv3 PW.
-iii. Switching between a static L2TPv3 PW and a dynamic MPLS PW.
-iv. Switching between a dynamic MPLS PW and a dynamic L2TPv3 PW.

8.1. Static MPLS and L2TPv3 PWs

In the case of two static control planes, the PW switching point MUST be configured to direct packets from one PW into the other. There is no control protocol involved in this case. The configuration MUST include which MPLS VC Label maps to which L2TPv3 Session ID (and associated Cookie, if present) as well as which MPLS Tunnel Label maps to which PE destination IP address.

8.2. Static MPLS PW and Dynamic L2TPv3 PW

When a statically configured MPLS PW is switched to a dynamic L2TPv3 PW, the static control plane should be considered identical to an attachment circuit (AC) in the reference model of Figure 1. The switching point PE SHOULD signal the proper PW status if it detects a failure in

sending or receiving packets over the static PW. Because the PW is statically configured, the status communicated to the dynamic L2TPv3 PW will be limited to local interface failures. In this case, the PW switching point PE behaves in a very similar manner to a T-PE, assuming an active role.

8.3. Static L2TPv3 PW and Dynamic LDP/MPLS PW

When a statically configured L2TPv3 PW is switched to a dynamic LDP/MPLS PW, then the static control plane should be considered identical to an attachment circuit (AC) in the reference model of Figure 1. The switching point PE SHOULD signal the proper PW status (via an L2TPv3 SLI message) if it detects a failure in sending or receiving packets over the static PW. Because the PW is statically configured, the status communicated to the dynamic LDP/MPLS PW will be limited to local interface failures. In this case, the PW switching point PE behaves in a very similar manner to a T-PE, assuming an active role.

8.4. Dynamic LDP/MPLS and L2TPv3 PWs

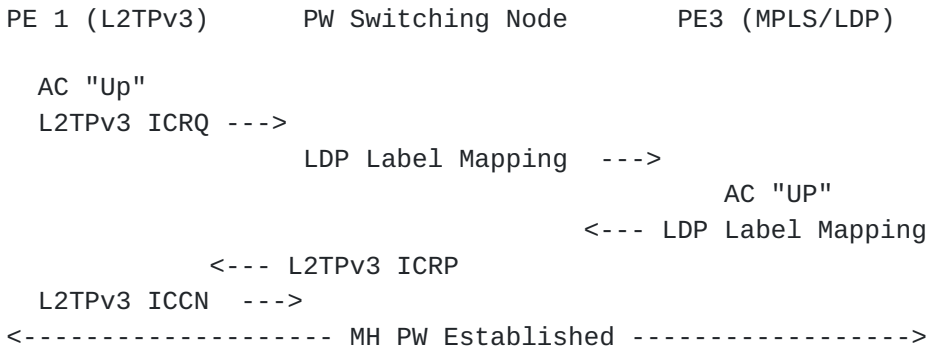
When switching between dynamic PWs, the switching point always assumes an initial passive role. Thus, it does not initiate an LDP/MPLS or L2TPv3 PW until it has received a connection request (Label Mapping or ICRQ) from one side of the node. Note that while MPLS PWs are made up of two unidirectional LSPs bonded together by FEC identifiers, L2TPv3 PWs are bidirectional in nature, setup via a 3-message exchange (ICRQ, ICRP and ICCN). Details of Session

Establishment, Tear Down, and PW Status signaling are detailed below.

8.4.1. Session Establishment

When the PW switching point receives an L2TPv3 ICRQ message, the identifying AVPs included in the message are mapped to FEC identifiers and sent in an LDP label mapping message. Conversely, if an LDP Label Mapping message is received, it is either mapped to an ICRP message or causes an L2TPv3 session to be initiated by sending an ICRQ.

Following are two example exchanges of messages between LDP and L2TPv3. The first is a case where an L2TPv3 T-PE initiates an MS-PW, the second is a case where an MPLS T-PE initiates an MS-PW.

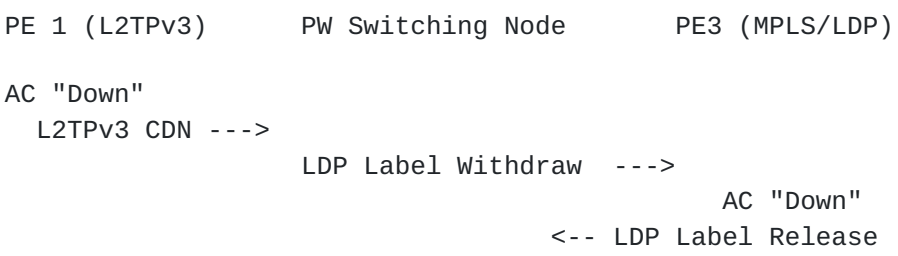


8.4.2. Adaptation of PW Status message

L2TPv3 uses the SLI message to indicate a interface status change (such as the interface transitioning from "Up" or "Down"). MPLS/LDP PWs either signal this via an LDP Label Withdraw or the PW Status Notification message defined in section 4.4 of [RFC4447].

8.4.3. Session Tear Down

L2TPv3 uses a single message, CDN, to tear down a pseudowire. The CDN message translates to a Label Withdraw message in LDP. Following are two example exchanges of messages between LDP and L2TPv3. The first is a case where an L2TPv3 T-PE initiates the termination of an MS-PW, the second is a case where an MPLS T-PE initiates the termination of an MS-PW.



8.5. Adaptation of L2TPv3 AVPs to Interface Parameters

[RFC4447] defines several interface parameters which MUST be mapped to the equivalent AVPs in L2TPv3 setup messages.

* Interface MTU

  The Interface MTU parameter is mapped directly to the L2TP Interface MTU AVP defined in [RFC4667]

* Max Number of Concatenated ATM cells

  This interface parameter is mapped directly to the L2TP "ATM Maximum Concatenated Cells AVP" described in section 6 of [RFC4454].

* Optional Interface Description String

  This string may be carried as the "Call-Information AVP" described in section 2.2 of [L2TP-INFOMSG]

* PW Type

  The PW Type defined in [RFC4446] is mapped to the L2TPv3 "PW Type" AVP defined in [L2TPv3].

* PW ID (FEC 128)

  For FEC 128, the PW ID is mapped directly to the L2TPv3 "Remote End ID" AVP defined in [L2TPv3].

* Generalized FEC 129 SAI/TAI

  Section 4.3 of [RFC4667] defines how to encode the SAI and TAI parameters. These can be mapped directly.

Other interface parameter mappings will either be defined in a future version of this document, or are unsupported when switching between LDP/MPLS and L2TPv3 PWs.

8.6. Switching Point TLV in L2TPv3

When translating between LDP and L2TPv3 control messages, the PW Switching Point TLV described earlier in this document is carried in a single variable length L2TP AVP present in the ICRQ, ICRP messages, and optionally in the ICCN message.

The L2TP "Switching Point AVP" is Attribute Type TBA-L2TP-AVP-1. The AVP MAY be hidden (the L2TP AVP H-bit may be 0 or 1), the length of the AVP is 6 plus the length of the series of Switching Point sub-TLVs included in the AVP, and the AVP MUST NOT be marked Mandatory (the L2TP AVP M-bit MUST be 0).

8.7. L2TPv3 and MPLS PW Data Plane

When switching between an MPLS and L2TP PW, packets are sent in their entirety from one PW to the other, replacing the MPLS label stack with the L2TPv3 and IP header or vice versa. There are some situations where an additional amount of interworking must be provided between the two data planes at the PW switching node.

8.7.1. PWE3 Payload Convergence and Sequencing

Section 5.4 of [PWE3-ARCH] discusses the purpose of the various shim headers necessary for enabling a pseudowire over an IP or MPLS PSN. For L2TPv3, the Payload Convergence and Sequencing function is carried out via the Default L2-Specific Sublayer defined in [L2TPv3]. For MPLS, these two functions (together with PSN Convergence) are carried out via the MPLS Control Word. Since these functions are different between MPLS and L2TPv3, interworking between the two may be necessary.

The L2TP L2-Specific Sublayer and MPLS Control Word are shim headers which in some cases are not necessary to be present at all. For example, an Ethernet PW with sequencing disabled will generally not require an MPLS Control Word or L2TP Default L2-Specific Sublayer to be present at all. In this case, Ethernet frames are simply sent from one PW to the other without any modification beyond the MPLS and L2TP/IP encapsulation and decapsulation.

The following section offers guidelines for how to interwork between L2TP and MPLS for those cases where the Payload Convergence, Sequencing, or PSN Convergence functions are necessary on one or both sides of the switching node.

8.7.2. Mapping

The MPLS Control Word consists of (from left to right):

-i. These bits are always zero in MPLS are not necessary to be mapped to L2TP.

-ii. These six bits may be used for Payload Convergence depending on the PW type. For ATM, the first four of these bits are defined in [RFC4717]. These map directly to the bits defined in [RFC4454]. For Frame Relay, these bits indicate how to set the bits in the Frame Relay header which must be regenerated for L2TP as it carries the Frame Relay header intact.

-iii. L2TP determines its payload length from IP. Thus, this Length field need not be carried directly to L2TP. This Length field will have to be calculated and inserted for MPLS when necessary.

-iv. The Default L2-Specific Sublayer has a sequence number with different semantics than that of the MPLS Control Word. This difference eliminates the possibility of supporting sequencing across the MS-PW by simply carrying the sequence number through the switching point transparently. As such, sequence numbers MAY be supported by checking the sequence numbers of packets arriving at the switching point and regenerating a new sequence number in the proper format for the PW on egress. If this type of sequence interworking at the switching node is not supported, and a T-PE requests sequencing of all packets via the L2TP control channel during session setup, the switching node SHOULD NOT allow the session to be established by sending a CDN message with Result Code set to 17 "sequencing not supported" (subject to IANA Assignment).

9. Operation And Management

9.1. Extensions to VCCV to Support Switched PWs

Single-hop pseudowires are signaled using the Virtual Circuit Connectivity Verification (VCCV) parameter included in the interface parameter field of the PW ID FEC TLV or the sub-TLV interface parameter of the Generalized PW ID FEC TLV as described in [RFC5085]. When a switching point exist between PE nodes, it is required to be able to continue operating VCCV end-to-end across a switching point and to provide the ability to trace the path of the MS-PW over any number of segments.

This document provides a method for achieving these two objectives. This method is based on re-using the existing VCCV CW and decrementing the TTL of the PW label at each hop in the path of the MS-PW.

9.2. PW-MPLS to PW-MPLS OAM Data Plane Indication

9.2.1. Decreasing the PW Label TTL

As stated above the S-PE MUST perform a standard MPLS label swap operation on the MPLS PW label. By the rules defined in [RFC3032] the PW label TTL MUST be decreased at every S-PE. Once the PW label TTL reaches the value of 0, the packet is sent to the control plane to be processed. Hence, by controlling the PW TTL value of the PW label it is possible to select exactly which hop will respond to the VCCV packet.

9.3. Signaling OAM Capabilities for Switched Pseudowires

Like in SS-PW, MS-PW VCCV capabilities are signaled using the VCCV parameter included in the interface parameter field of the PW ID FEC TLV or the sub-TLV interface parameter of the Generalized PW ID FEC TLV as described in [RFC5085].

In Figure 3 T-PE1 uses the VCCV parameter included in the interface parameter field of the PW ID FEC TLV or the sub-TLV interface parameter of the Generalized PW ID FEC TLV to indicate to the far end T-PE2 what VCCV capabilities T-PE1 supports. This is the same VCCV parameter as would be used if T-PE1 and T-PE2 were connected directly. S-PE2, which is a PW switching point, as part of the adaptation function for interface parameters, processes locally the VCCV parameter then passes it to T-PE2. If there were multiple S-PEs on the path between T-PE1 and T-PE2, each would carry out the same processing, passing along the VCCV parameter. The local processing of the VCCV parameter removes CC Types specified by the originating T-PE that are not supported on the S-PE. For example, if T-PE1 indicates supports CC Types 1,2,3 and the Then the S-PE removes the Router Alert CC Type=2, leaving the rest of the TLV unchanged, and passes the modified VCCV parameter to the next S-PE along the path.

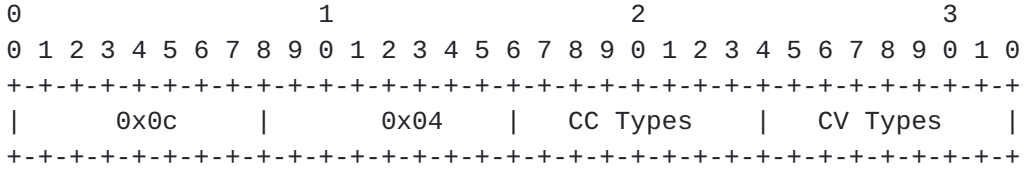
The far end T-PE (T-PE2) receives the VCCV parameter indicating only the CC types that are supported by the initial T-PE (T-PE1) and all S-PEs along the PW path.

9.4. OAM Capability for MS-PWs Demultiplexed using MPLS

The VCCV parameter ID is defined as follows in [RFC4446]:

Parameter ID	Length	Description
0x0c	4	VCCV

The format of the VCCV parameter field is as follows:



0x01 Type 1: PWE3 control word with 0001b as first nibble as defined in [RFC4385].
0x02 Type 2: MPLS Router Alert Label.
0x04 Type 3: MPLS PW De-multiplexor Label TTL = 1 (Type 3).

9.4.1. MS-PW and VCCV CC Type 1

VCCV CC type 1 is normally supported between T-PEs, and MAY be removed by an S-PE as a matter of local security policy. When using CC type 1 for MS-PWs the PE transmitting the VCCV packet MUST set the TTL to the appropriate value to reach the destination S-PE. However if the packet is destined for the T-PE, the TTL can be set to any value that is sufficient for the packet to reach the T-PE.

9.4.2. MS-PW and VCCV CC type 2

VCCV CC type 2 is not supported for MS-PWs and MUST be removed form a VCCV parameter field by the S-PE.

9.4.3. MS-PW and VCCV CC type 3

VCCV CC type 3 can be used for MS-PWs, however if the CW is enabled VCCV type 1 is preferred according to the rules in [RFC5085]. Note that for using the VCCV type 3, TTL method, the PE will set the PW label TTL to the appropriate value necessary to reach the target PE, otherwise the VCCV packet might be forwarded over the AC to the CPE.

9.4.4. Detailed VCCV Procedures

In order to test the end-to-end connectivity of the multi-segment PW, a S-PE must include the FEC used in the last segment to the destination T-PE. This information is either configured at the sending T-PE or is obtained by processing the corresponding sub-TLVs of the PW switching point TLV. The necessary S-PE sub-TLVs are :

Type Description
0x01 PW ID of last PW segment traversed
0x03 Local IP address of PW Switching Point
0x04 Remote IP address of last PW Switching Point traversed or
of the T-PE

9.4.4.1. End to End verification between T-PEs

In Figure 3, if T-PE1, S-PE and T-PE2 support Control Word , the PW control plane will automatically negotiate the use of the CW. VCCV CC type 3 will function correctly whether the CW is enable or not on the PW. However VCCV type 1 for (which can be use for end to end

verification only), is only supported if the CW is enabled.

At the S-PE the data path operations include an outer label pop, inner label swap and new outer label push. Note that there is no requirement for the S-PE to inspect the CW. Thus, the end-to-end connectivity of the multi-segment pseudowire can be verified by performing all of the following steps:

-i. T-PE forms a VCCV-ping echo request message with the FEC matching that of the last segment PW to the destination T-PE.

-ii. T-PE sets the inner PW label TTL to the exact value to allow the packet to reach the far end T-PE. (the value is determined by counting the number of S-PEs from the control plane information) Alternatively, if CC type 1 is supported the packet can be encapsulated according to CC type 1 in [RFC5085]

-iii. T-PE sends a VCCV packet that will follow the exact same data path at each S-PE as that taken by data packets.

-iv. S-PE may performs an outer label pop, if PHP is disabled, and will perform an inner label swap with TTL decrement, and new outer label push.

-v. There is no requirement for the S-PE to inspect the CW.

-vi. The VCCV packet is diverted to VCCV control processing at the destination T-PE.

-vii. Destination T-PE replies using the specified reply mode, i.e., reverse PW path or IP path.

9.4.4.2. Partial verification from T-PE

In order to trace part of the multi-segment pseudowire, the TTL of the PW label may be used to force the VCCV message to 'pop out' at an intermediate node. When the TTL expires, the S-PE can determine that the packet is a VCCV packet by either checking the control word (CW) , or if the CW is not in use by checking for a valid IP header with UDP destination port 3503. The packet should then be diverted to VCCV processing.

In Figure 2, if T-PE1 sends a VCCV message with the TTL of the PW label equal to 1, the TTL will expire at the S-PE. T-PE1 can thus verify the first segment of the pseudowire.

The VCCV packet is built according to [RFC4379] section 3.2.9 for FEC 128, or 3.2.10 for a FEC 129 PW. All the information necessary to build the VCCV LSP ping packet is collected by inspecting the S-PE TLVs.

Note that this use of the TTL is subject to the caution expressed in [RFC5085]. If a penultimate LSR between S-PEs or between an S-PE and a T-PE manipulates the PW label TTL, the VCCV message may not emerge from the MS-PW at the correct S-PE.

9.4.4.3. Partial verification between S-PEs

Assuming that all nodes along an MS-PW support the Control Word CC Type 3, VCCV between S-PEs may be accomplished using the PW label TTL as described above. In Figure 3, the S-PE may verify the path between it and T-PE2 by sending a VCCV message with the PW label TTL set to 1. Given a more complex network with multiple S-PEs, an S-PE may verify the connectivity between it and an S-PE two segments away by sending a VCCV message with the PW label TTL set to 2. Thus, an S-PE can diagnose connectivity problems by successively increasing the TTL. All the information needed to build the proper VCCV echo request packet as described in [RFC4379] section 3.2.9 or 3.2.10 is obtained automatically from the LDP label mapping that contains S-PE TLVs.

9.4.5. Optional FEC Reply in VCCV LSP Ping packet

When A S-PE along the PW path receives an VCCV LSP Ping echo request packet the following OPTIONAL procedure can be followed in addition to the procedure described below:

-i. S-PE validates the echo request with the FEC.
-ii. The S-PE build the standard LSP ping reply packet to be sent back.
-iii. The S-PE appends the FEC128 information for the next segment along the MS-PW to the LSP PING reply packet.

This FEC information can then be compared to the S-PE TLV information received from the control plane when the PW was first signalled. This FEC information MUST not be sent in the reply if the S-PE is not sending an S-PE TLV for administrative reasons in the same situation as explained previously.

9.4.6. Processing of an VCCV Echo Message in a MS-PW

The challenge for the control plane is to be able to build the VCCV echo request packet with the necessary information such as the target FEC 128 PW sub-TLV (FEC128) of the downstream PW segment which the packet is destined for. This could be even more difficult in situations in which the MS-PW spans different providers and Autonomous Systems.

9.4.6.1. Sending a VCCV Echo Request

When in the "ping" mode of operation, the sender of the echo request message requires the FEC of the last segment to the target S-PE/T-PE node. This information can either be configured manually or be obtained by inspecting the corresponding sub-TLVs of the PW switching point TLV. However, the PW switching point TLV is optional and there is no guarantee that all S-PE nodes will populate it with their system address, the PWid of the last PW segment traversed, and the last system address of of the last PE traversed by the label mapping message. If all information is not available, VCCV LSP ping mode will not function.

9.4.6.2. Receiving an VCCV Echo Request

Upon receiving a VCCV echo request the control plane on S-PEs (or the target node of each segment of the MS-PW) validates the request and responds to the request with an echo reply consisting of a return code of 8 (label switched at stack-depth) indicating that it is an S-PE and not the egress router for the MS-PW.

If the node is the T-PE or the egress node of the MS-PW, it responds to the echo request with an echo reply with a return code of 3 (egress router).

9.4.7. VCCV Trace Operations

As an example, in Figure 3, VCCV trace can be performed on the MS-PW originating from T-PE1 by a single operational command. The following process ensues:

-i. T-PE1 sends a VCCV echo request with TTL set to 1 and a FEC containing the pseudowire information of the first segment (PW1 between T-PE1 and S-PE) to S-PE for validation. If FEC Stack Validation is enabled, the request may also include additional sub-TLV such as LDP Prefix and/or RSVP LSP dependent on the type of transport tunnel the segmented PW

is riding on.

-ii. S-PE validates the echo request with the FEC. Since it is a switching point between the first and second segment it builds an echo reply with a return code of 8 and sends the echo reply back to T-PE1.

-iii. T-PE1 builds a second VCCV echo request based on the infomation obtained from the control plane (S-PE TLV). It then increments the TTL and sends it out to T-PE2. Note that the VCCV echo request packet is switched at the S-PE datapath and forwarded to the next downstream segment without any involvement from the control plane.

-iv. T-PE2 receives and validates the echo request with the FEC. Since T-PE2 is the destination node or the egress node of the MS-PW it replies to T-PE1 with an echo reply with a return code of 3 (Egress Router).

-v. T-PE1 receives the echo reply from T-PE2. T-PE1 is made aware that T-PE2 is the destination of the MS-PW because the echo reply has a return code of is 3. The trace process is completed.

If no echo reply is received, or an error code is received from a particular PE, the trace process MUST stop immediately, and no packets MUST be sent further along the MS-PW.

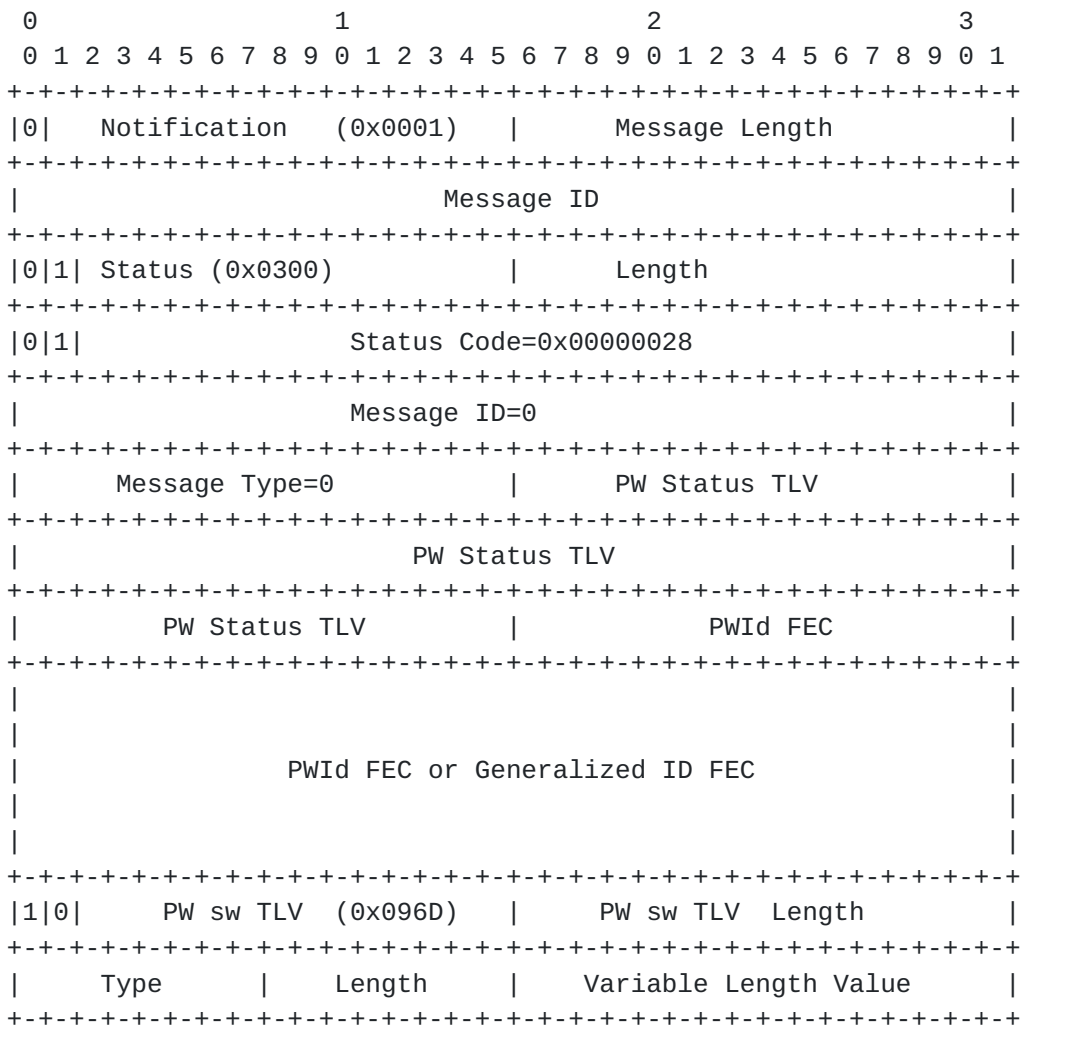
For more detail on the format of the VCCV echo packet, refer to [RFC5085] and [RFC4379]. The TTL here refers to that of the inner (PW) label TTL.

9.5. Mapping Switched Pseudowire Status

In the PW switching with attachment circuits case (Figure 2), PW status messages indicating PW or attachment circuit faults SHOULD be mapped to fault indications or OAM messages on the connecting AC as defined in [PW-MSG-MAP]. If the AC connecting two PWs crosses an administrative boundary, then the manner in which those OAM messages are treated at the boundary is out of scope of this draft.

In the PW control plane switching case (Figure 3), there is no attachment circuit at the PW switching point, but the two PWs are connected together. Similarly, the status of the PWs are forwarded unchanged from one PW to the other by the control plane switching function. However, it may sometimes be necessary to communicate status of one of the locally attached SS-PW at a PW switching point.

For LDP this can be accomplished by sending an LDP notification message containing the PW status TLV, as well as an OPTIONAL PW switching point TLV as follows:



Only one PW switching point TLV can be present in this message. This message is then relayed by each PW switching point unchanged. The T-PE decodes the status message and the included PW switching point TLV to detect exactly where the fault occurred. At the T-PE if there is no PW switching point TLV included in the LDP status notification then the status message can be assumed to have originated at the remote T-PE.

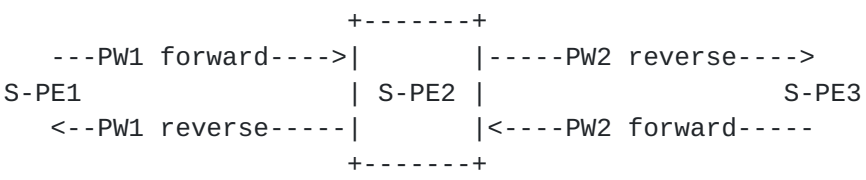
The merging of the received T-LDP status and the local status for the PW segments at an S-PE can be summarized as follows:

-i. When the local status for both PW segments is UP, the S-PE passes any received AC or PW status bits unchanged, i.e., the status notification TLV is unchanged but the VCid in the case of a FEC 128 TLV is set to value of the PW segment to the next hop.

-ii. When the local status for any of the PW segments is Down, the S-PE always sends "PW Down" status bits regardless if the received status bits from the remote node indicated "PW UP/Down". AC status bit are passed along unchanged.

9.5.1. S-PE initiated PW status messages

The PW fault directions are defined as follows:



When a local fault is detected by the S-PE, a PW status message is sent in both directions along the PW. Since there are no attachment circuits on an S-PE, only the following status messages are relevant:

0x00000008 - Local PSN-facing PW (ingress) Receive Fault
0x00000010 - Local PSN-facing PW (egress) Transmit Fault

Each S-PE needs to store only two 32-bit PW status words for each SS-PW: One for local failures , and one for remote failures (normally received from another PE). The first failure will set the appropriate bit in the 32-bit status word, and each subsequent failure will be ORed to the appropriate PW status word. In the case of the PW status word storing remote failures, this rule has the effect of a logical OR operation with the first failure received on the particular SS-PW.

It should be noted that remote failures received on an S-PE are just passed along the MS-PW unchanged while local failures detected an an S-PE are signalled on both SS-PWs.

A T-PE can receive multiple failures from S-PEs along the MH-PW, however only the failure from the remote closest S-PE will be stored (last pw status message received). The PW status word received are

just ORed to any existing remote PW status already stored on the T-PE.

Given that there are two SS-PW at a particular S-PE for a particular MH-PW, there are for possible failure cases as follows:

-i. PW2 reverse direction fault
-ii. PW1 reverse direction fault
-iii. PW2 forward direction fault
-iv. PW1 forward direction fault

It should also be noted that once a PW status notification message is initiated at a PW switching point for a particular PW status bit any further status message, for the same status bit, received from an upstream neighbor is processed locally and not forwarded until the PW switching point original status error state is cleared.

Each S-PE along the MS-PW MUST store any PW status messages transiting it. If more then one status message with the same PW status bit set is received by a T-PE only the last PW status message is stored.

9.5.1.1. Local PW2 reverse direction fault

When this failure occurs the S-PE will take the following actions:

* Send a PW status message to S-PE3 containing "0x00000010 - Local PSN-facing PW (egress) Transmit Fault"
* Send a PW status message to S-PE1 containing "0x00000008 - Local PSN-facing PW (ingress) Receive Fault"
* Store 0x00000010 in the local PW status word for the SS-PW toward S-PE3.

9.5.1.2. Local PW1 reverse direction fault

When this failure occurs the S-PE will take the following actions:

* Send a PW status message to S-PE1 containing "0x00000010 - Local PSN-facing PW (egress) Transmit Fault"
* Send a PW status message to S-PE3 containing "0x00000008 - Local PSN-facing PW (ingress) Receive Fault"
* Store 0x00000010 in the local PW status word for the SS-PW toward S-PE1.

9.5.1.3. Local PW2 forward direction fault

When this failure occurs the S-PE will take the following actions:

* Send a PW status message to S-PE3 containing "0x00000008 - Local PSN-facing PW (ingress) Receive Fault"
* Send a PW status message to S-PE1 containing "0x00000010 - Local PSN-facing PW (egress) Transmit Fault"
* Store 0x00000008 in the local PW status word for the SS-PW toward S-PE3.

9.5.1.4. Local PW1 forward direction fault

When this failure occurs the S-PE will take the following actions:

* Send a PW status message to S-PE1 containing "0x00000008 - Local PSN-facing PW (ingress) Receive Fault"
* Send a PW status message to S-PE3 containing "0x00000010 - Local PSN-facing PW (egress) Transmit Fault"
* Store 0x00000008 in the local PW status word for the SS-PW toward S-PE1.

9.5.1.5. Clearing Faults

Remote PW status fault clearing messages received by an S-PE will only be forwarded if there are no corresponding local faults on the S-PE. (local faults always supersede remote faults)

Once the local fault has cleared, and there is no corresponding (same PW status bit set) remote fault, a PW status messages is sent out to the adjacent PEs clearing the fault.

When a PW status fault clearing message is forwarded, the S-PE will always send the S-PE TLV associated with the PE which cleared the fault.

9.5.2. PW status messages and S-PE TLV processing

When a PW status message is received that includes a S-PE TLV, the S-PE TLV information MAY be stored, along with the contents of the PW status Word according to the procedures described above. The S-PE TLV stored is always the S-PE TLV that is associated with the PE that set that particular last fault. If subsequent PW status message for the same PW status bit are received the S-PE TLV will overwrite the previously stored S-PE TLV.

9.5.3. T-PE processing of PW status messages

The PW switching architecture is based on the concept that the T-PE should process the PW LDP messages in the same manner as if it was participating in the setup of a SS-PW. However T-PE participating a MS-PW, SHOULD be able to process the PW switching point TLV. Otherwise the processing of PW status messages , and other PW setup messages is exactly as described in [RFC4447].

9.6. Pseudowire Status Negotiation Procedures

Pseudowire Status signaling methodology, defined in [RFC4447], SHOULD be transparent to the switching point.

9.7. Status Dampening

When the PW control plane switching methodology is used to cross an administrative boundary it might be necessary to prevent excessive status signaling changes from being propagated across the administrative boundary. This can be achieved by using a similar method as commonly employed for the BGP protocol route advertisement dampening. The details of this OPTIONAL algorithm are a matter of implementation, and are outside the scope of this document.

10. Peering Between Autonomous Systems

The procedures outlined in this document can be employed to provision and manage MS-PWs crossing AS boundaries.

The use of more advanced mechanisms involving auto-discovery and ordered PWE3 MS-PW signaling will be covered in a separate document.

11. Security Considerations

This document specifies the LDP and L2TPv3 extensions that are needed for setting up and maintaining pseudowires. The purpose of setting up pseudowires is to enable layer 2 frames to be encapsulated and transmitted from one end of a pseudowire to the other. Therefore we treat the security considerations for both the data plane and the control plane.

11.1. Data Plane Security

Data plane security consideration as discussed in [RFC4447], [L2TPv3], and [PWE3-ARCH] apply to this extension without any changes.

11.1.1. VCCV Security considerations

The VCCV technology for MS-PW offers a method for the service provider to verify the data path of a specific PW. This involves sending a packet to a specific PE and receiving an answer which either confirms , or indicates that the information contained in the packet is incorrect. This is a very similar process to the commonly used IP ICMP ping , and TTL expired methods for IP networks. It should be noted that when using VCCV Type 3 for PW when the CW is not enabled, if a packet is crafted with a TTL greater then the number of hops along the MS-PW path, or an S-PE along the path mis-processes the TTL, the packet could mistakenly be forwarded out the attachment circuit as a native PW packet. This packet would most likely be treated as an error packet by the CE. However if this possibility is not acceptable, the CW should be enabled to guarantee that a VCCV packet will never be mistakenly forwarded to the AC.

11.2. Control Protocol Security

General security considerations with regard to the use of LDP are specified in section 5 of RFC 3036. Security considerations with regard to the L2TPv3 control plane are specified in [L2TPv3]. These considerations apply as well to the case where LDP or L2TPv3 is used to set up PWs.

A Pseudowire connects two attachment circuits. It is important to make sure that LDP connections are not arbitrarily accepted from anywhere, or else a local attachment circuit might get connected to an arbitrary remote attachment circuit. Therefore an incoming session request MUST NOT be accepted unless its IP source address is known to be the source of an "eligible" peer. The set of eligible peers could be pre-configured (either as a list of IP addresses, or as a list of address/mask combinations), or it could be discovered dynamically via an auto-discovery protocol which is itself trusted. (Obviously if the auto-discovery protocol were not trusted, the set of "eligible peers" it produces could not be trusted.)

Even if a connection request appears to come from an eligible peer, its source address may have been spoofed. So some means of preventing source address spoofing must be in place. For example, if

all the eligible peers are in the same network, source address filtering at the border routers of that network could eliminate the possibility of source address spoofing.

For a greater degree of security, the LDP MD5 authentication key option, as described in section 2.9 of RFC 3036, or the Control Message Authentication option of [L2TPv3] MAY be used. This provides integrity and authentication for the control messages, and eliminates the possibility of source address spoofing. Use of the message authentication option does not provide privacy, but privacy of control messages are not usually considered to be highly urgent. Both the LDP and L2TPv3 message authentication options rely on the configuration of pre-shared keys, making it difficult to deploy when the set of eligible neighbors is determined by an auto-configuration protocol.

When the Generalized ID FEC Element is used, it is possible that a particular peer may be one of the eligible peers, but may not be the right one to connect to the particular attachment circuit identified by the particular instance of the Generalized ID FEC element. However, given that the peer is known to be one of the eligible peers (as discussed above), this would be the result of a configuration error, rather than a security problem. Nevertheless, it may be advisable for a PE to associate each of its local attachment circuits with a set of eligible peers, rather than having just a single set of eligible peers associated with the PE as a whole.

12. IANA Considerations

12.1. L2TPv3 AVP

This document uses a ne L2TP parameter, IANA already maintains a registry of name "Control Message Attribute Value Pair" defined by [RFC3438]. The following new values are required:

TBA-L2TP-AVP-1 - PW Switching Point AVP

12.2. LDP TLV TYPE

This document uses several new LDP TLV types, IANA already maintains a registry of name "TLV TYPE NAME SPACE" defined by RFC3036. The following value is suggested for assignment:

TLV type	Description
0x096D	Pseudowire Switching TLV

12.3. LDP Status Codes

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

Assignment	E	Description
0x0000003A	0	"PW Loop Detected"

12.4. L2TPv3 Result Codes

This document uses several new L2TPv3 status codes, IANA already maintains a registry of name "L2TPv3 Result Codes" defined by RFCxxxx. The following value is suggested for assignment:

Assignment	Description
17	"sequencing not supported"

12.5. New IANA Registries

IANA needs to set up a registry of "PW Switching Point TLV Type". These are 8-bit values. Types value 1 through 3 are defined in this document. Type values 4 through 64 are to be assigned by IANA using the "Expert Review" policy defined in RFC2434. Type values 65 through 127, 0 and 255 are to be allocated using the IETF consensus policy defined in [RFC2434]. Types values 128 through 254 are reserved for vendor proprietary extensions and are to be assigned by IANA, using the "First Come First Served" policy defined in RFC2434.

The Type Values are assigned as follows:

Type	Length	Description
0x01	4	PW ID of last PW segment traversed
0x02	variable	PW Switching Point description string
0x03	4/16	Local IP address of PW Switching Point
0x04	4/16	Remote IP address of last PW Switching Point traversed or of the T-PE
0x05	variable	AI of last PW segment traversed
0x06	10	L2 PW address of PW Switching Point

13. 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.

14. Full Copyright Statement

15. Acknowledgments

The authors wish to acknowledge the contributions of Satoru Matsushima, Wei Luo, Neil Mcgill, Skip Booth, Neil Hart, Michael Hua, and Tiberiu Grigoriu.

16. Normative References

[RFC4385] " Pseudowire Emulation Edge-to-Edge (PWE3) Control Word for Use over an MPLS PSN", S. Bryant, et al., RFC4385, February 2006.

[RFC4446] "IANA Allocations for Pseudowire Edge to Edge mulation (PWE3)", L. Martini, RFC4446, April 2006.

[RFC4447] "Transport of Layer 2 Frames Over MPLS", Martini, L., et al., rfc4447 April 2006.

[RFC3985] Stewart Bryant, et al., PWE3 Architecture, RFC3985

[2547BIS] "BGP/MPLS IP VPNs", Rosen, E, Rekhter, Y. draft-ietf-l3vpn-rfc2547bis-03.txt (work in progress), October 2004.

[L2TPv3] "Layer Two Tunneling Protocol (Version 3)", J. Lau, M. Townsley, I. Goyret, RFC3931

[RFC5085] Nadeau, T., et al."Pseudo Wire Virtual Circuit Connection Verification (VCCV), A Control Channel for Pseudowires", RFC5085 December 2007.

[RFC2434] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations section in RFCs", BCP 26, RFC 2434, October 1998.

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

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

17. Informative References

[RFC4023] "Encapsulating MPLS in IP or Generic Routing Encapsulation (GRE)", Rosen, E, Rekhter, Y. RFC4023, March 2005.

[PWE3-ARCH] "PWE3 Architecture" Bryant, et al., draft-ietf-pwe3-arch-07.txt (work in progress), June 2003.

[PWE3-FRAG] "PWE3 Fragmentation and Reassembly", A. Malis, W. M. Townsley, draft-ietf-pwe3-fragmentation-05.txt (work in progress) February 2004

[RFC4667] "Layer 2 Virtual Private Network (L2VPN) Extensions for Layer 2 Tunneling Protocol (L2TP)", Luo, Wei, RFC4667, W. Luo, September 2006

[L2TP-INFOMSG] "L2TP Call Information Messages", Mistretta, Goyret, McGill, Townsley, draft-mistretta-l2tp-infomsg-02.txt, (work in progress), July 2004

[RFC4454] "Asynchronous Transfer Mode (ATM) over Layer 2 Tunneling Protocol Version 3 (L2TPv3)", Singh, Townsley, Pignataro, RFC4454, May 2006 (work in progress), March 2004.

[RFC4717] "Encapsulation Methods for Transport of (ATM) MPLS Networks", Martini et al., RFC4717, December 2006

[RFC3438] W. M. Townsley, "Layer Two Tunneling Protocol (L2TP) Internet"

[PW-MSG-MAP] "Pseudo Wire (PW) OAM Message Mapping", Nadeau et al, draft-ietf-pwe3-oam-msg-map-02.txt, (work in progress), February 2005

[RFC4379] "Detecting Multi-Protocol Label Switched (MPLS) Data Plane Failures", RFC4379, February 2006.

[RFC3032] "MPLS Label Stack Encoding", RFC3032, January 2001

[MS-PW-ARCH] "An Architecture for Multi-Segment Pseudo Wire Emulation Edge-to-Edge", Bocci et al, draft-ietf-pwe3-ms-pw-arch-03.txt June 2007

18. Author Information

Luca Martini
Cisco Systems, Inc.
9155 East Nichols Avenue, Suite 400
Englewood, CO, 80112
e-mail: lmartini@cisco.com

Thomas D. Nadeau
BT
BT Centre
81 Newgate Street
London, EC1A 7AJ
United Kingdom
e-mail: tom.nadeau@bt.com

Chris Metz
Cisco Systems, Inc.
e-mail: chmetz@cisco.com

Mike Duckett
Bellsouth
Lindbergh Center
D481
575 Morosgo Dr
Atlanta, GA 30324
e-mail: mduckett@bellsouth.net

Matthew Bocci
Alcatel-Lucent
Grove House, Waltham Road Rd
White Waltham, Berks, UK. SL6 3TN
e-mail: matthew.bocci@alcatel-lucent.co.uk

Florin Balus
Alcatel-Lucent
701 East Middlefield Rd.
Mountain View, CA 94043
e-mail: florin.balus@alcatel-lucent.com

Mustapha Aissaoui
Alcatel-Lucent
600, March Road,
Kanata, ON, Canada
e-mail: mustapha.aissaoui@alcatel-lucent.com