Requirements and Framework for Point-to-Multipoint Pseudowires over MPLS Packet Switched Networks

Abstract

This document presents a set of requirements and a framework for providing a point-to-multipoint pseudowire (PW) over MPLS Packet Switched Networks. The requirements identified in this document are related to architecture, signaling, and maintenance aspects of point-to-multipoint PW operation. They are proposed as guidelines for the standardization of such mechanisms. Among other potential applications, point-to-multipoint PWs can be used to optimize the support of multicast Layer 2 services (Virtual Private LAN Service and Virtual Private Multicast Service).

Status of This Memo

This document is not an Internet Standards Track specification; it is published for informational purposes.

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). Not all documents approved by the IESG are a candidate for any level of Internet Standard; see 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/rfc7338.
Copyright Notice

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

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

This document may contain material from IETF Documents or IETF Contributions published or made publicly available before November 10, 2008. The person(s) controlling the copyright in some of this material may not have granted the IETF Trust the right to allow modifications of such material outside the IETF Standards Process. Without obtaining an adequate license from the person(s) controlling the copyright in such materials, this document may not be modified outside the IETF Standards Process, and derivative works of it may not be created outside the IETF Standards Process, except to format it for publication as an RFC or to translate it into languages other than English.
Table of Contents

1. Introduction ....................................................3
  1.1. Problem Statement ..........................................3
  1.2. Scope of This Document .....................................4
  1.3. Conventions Used in This Document ..........................4
2. Definitions .....................................................5
  2.1. Acronyms ...................................................5
  2.2. Terminology ................................................5
3. P2MP PW Requirements ............................................6
  3.1. Reference Model ............................................6
  3.2. P2MP PW and Underlying Layer ...............................7
  3.3. P2MP PW Construction .......................................9
  3.4. P2MP PW Signaling Requirements ............................10
    3.4.1. P2MP PW Identifier ..................................10
    3.4.2. PW Type Mismatch ....................................10
    3.4.3. Interface Parameters Sub-TLV ........................10
    3.4.4. Leaf Grafting/Pruning .................................10
    3.4.5. Failure Detection and Reporting ....................11
    3.4.6. Protection and Restoration ..........................11
    3.4.7. Scalability ........................................13
4. Backward Compatibility .........................................13
5. Security Considerations ........................................13
6. References .....................................................14
  6.1. Normative References ......................................14
  6.2. Informative References ....................................14
7. Acknowledgments ................................................15
8. Contributors ...................................................16

1. Introduction

1.1. Problem Statement

As defined in the pseudowire architecture [RFC3985], a pseudowire
(PW) is a mechanism that emulates the essential attributes of a
telemcomunications service (such as a T1 leased line or Frame Relay)
over an IP or MPLS Packet Switched Network (PSN). It provides a
single service that is perceived by its user as an unshared link or
circuit of the chosen service. A pseudowire is used to transport
Layer 1 or Layer 2 traffic (e.g., Ethernet, Time-Division
Multiplexing (TDM), ATM, and Frame Relay) over a Layer 3 PSN.
Pseudowire Emulation Edge-to-Edge (PWE3) operates "edge to edge" to
provide the required connectivity between the two endpoints of
the PW.

The point-to-multipoint (P2MP) topology described in [VPMS-REQS] and
required to provide P2MP Layer 2 VPN service can be achieved using
one or more P2MP PWs. The use of PW encapsulation enables P2MP
services to transport Layer 1 or Layer 2 data. This could be achieved using a set of point-to-point PWs, with traffic replication at the Root Provider Edge (PE), but at the cost of bandwidth efficiency, as duplicate traffic would be carried multiple times on shared links.

This document defines the requirements for a point-to-multipoint PW (P2MP PW). A P2MP PW is a mechanism that emulates the essential attributes of a P2MP telecommunications service such as a P2MP ATM Virtual Circuit over a Packet Switched Network.

The required functions of P2MP PWs include encapsulating service-specific Protocol Data Units (PDUs) arriving at an ingress Attachment Circuit (AC), carrying them across a tunnel to one or more egress ACs, managing their timing and order, and any other operations required to emulate the behavior and characteristics of the service as faithfully as possible.

1.2. Scope of This Document

The document describes the general architecture of P2MP PW with a reference model, mentions the notion of data encapsulation, and outlines specific requirements for the setup and maintenance of a P2MP PW. In this document, the requirements focus on the Single-Segment PW model. The requirements for realizing P2MP PW in the Multi-Segment PW model [RFC5254] are left for further study. This document refers to [RFC3916] for other aspects of P2MP PW implementation, such as "Packet Processing" (Section 4 of that document) and "Faithfulness of Emulated Services" (Section 7 of that document).

1.3. Conventions Used in This Document

Although this is a requirements specification not a protocol specification, the key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted to apply to protocol solutions designed to meet these requirements as described in [RFC2119].
2. Definitions

2.1. Acronyms

- P2P: Point-to-Point
- P2MP: Point-to-Multipoint
- PW: Pseudowire
- PSN: Packet Switched Network
- SS-PW: Single-Segment Pseudowire

2.2. Terminology

This document uses terminology described in [RFC5659]. It also introduces additional terms needed in the context of P2MP PW.

**P2MP PW (also referred to as PW tree):**
Point-to-Multipoint Pseudowire. A PW attached to a source Customer Edge (CE) used to distribute Layer 1 or Layer 2 traffic to a set of one or more receiver CEs. The P2MP PW is unidirectional (i.e., carrying traffic from Root PE to Leaf PEs) and optionally supports a return path.

**P2MP SS-PW:**
Point-to-Multipoint Single-Segment Pseudowire. A single-segment P2MP PW set up between the Root PE attached to the source CE and the Leaf PEs attached to the receiver CEs. The P2MP SS-PW uses P2MP Label Switched Paths (LSPs) as PSN tunnels.

**Root PE:**
P2MP PW Root Provider Edge. The PE attached to the traffic source CE for the P2MP PW via an Attachment Circuit (AC).

**Leaf PE:**
P2MP PW Leaf Provider Edge. A PE attached to a set of one or more traffic receiver CEs, via ACs. The Leaf PE replicates traffic to the CEs based on its Forwarder function [RFC3985].

**P2MP PSN Tunnel:**
In the P2MP SS-PW topology, the PSN tunnel is a general term indicating a virtual P2MP connection between the Root PE and the Leaf PEs. A P2MP tunnel may potentially carry multiple P2MP PWs inside (aggregation). This document uses terminology from the document describing the MPLS multicast architecture [RFC5332] for MPLS PSN.
3. P2MP PW Requirements

3.1. Reference Model

As per the definition in [RFC3985], a pseudowire (PW) both originates and terminates on the edge of the same packet switched network (PSN). The PW label is unchanged between the originating and terminating Provider Edges (PEs). This is also known as a single-segment pseudowire (SS-PW) -- the most fundamental network model of PWE3.

A P2MP PW can be defined as point-to-multipoint connectivity from a Root PE connected to a traffic source CE to one or more Leaf PEs connected to traffic receiver CEs. It is considered to be an extended architecture of the existing P2P SS-PW technology.

Figure 1 describes the P2MP PW reference model that is derived from [RFC3985] to support P2MP emulated services.

```
     +----+         +-----+         +----+     |
     |PE1 |---------|  P  |=========|PE2 |AC2  |     +----+
     |    |         |   ......PW1.......> |---------->|CE2  |
     |    |=========|   . |=========|    |     |     +----+
     |    |         |   . |=========|    |     |     +----+
     +----+         +-----+         +----+     |     +----+
    Native Service |<-----P2MP PSN tunnel --->| Service LEAF
ROOT V (AC) V V V V V V V V V V
V     +----+         +-----+         +----+     |
     |PE1 |---------|  P  |=========|PE2 |AC2  |     +----+
     |    |         |   ......PW1.......> |---------->|CE2  |
     |    |=========|   . |=========|    |     |     +----+
     |    |---------|  PE4 |=========|    |     |     +----+
     |    |         |   ......PW1.......> |---------->|CE3  |
     |     +-----+         +-----+         +-----+     |
    Native Service |<-----P2MP PSN tunnel --->| Service LEAF
V     +----+         +-----+         +----+     |
     |CE1 |---------|  AC1 |---------|CW1  |AC3  |     +----+
     |    |         |   . |=========|    |     |     +----+
     |    |---------|  AC5 |---------|    |     |     +----+
     |     +-----+         +-----+         +-----+     |
```

Figure 1: P2MP PW Reference Model

This architecture applies to the case where a P2MP PSN tunnel extends between edge nodes of a single PSN domain to transport a unidirectional P2MP PW with endpoints at these edge nodes. In this model, a single copy of each PW packet is sent over the PW on the P2MP PSN tunnel and is received by all Leaf PEs due to the P2MP...
nature of the PSN tunnel. The P2MP PW SHOULD be traffic optimized, i.e., only one copy of a P2MP PW packet or PSN tunnel (underlying layer) packet is sent on any single link along the P2MP path. P routers participate in P2MP PSN tunnel operation but not in the signaling of P2MP PWs.

The Reference Model outlines the basic pieces of a P2MP PW. However, several levels of replication need to be considered when designing a P2MP PW solution:

- Ingress PE replication to CEs: traffic is replicated to a set of local receiver CEs
- P router replication in the core: traffic is replicated by means of a P2MP PSN tunnel (P2MP LSP)
- Egress PE replication to CEs: traffic is replicated to local receiver CEs

Theoretically, it is also possible to consider Ingress PE replication in the core; that is, all traffic is replicated to a set of P2P PSN transport tunnels at ingress, not using P router replication at all.

However, this approach may lead to duplicate copies of each PW packet being sent over the same physical link, specifically in the case where multiple PSN tunnels transit that physical link. Hence, this approach is not preferred.

Specific operations that MUST be performed at the PE on the native data units are not described here since the required pre-processing (Forwarder (FWRD) and Native Service Processing (NSP)) defined in Section 4.2 of [RFC3985] is also applicable to P2MP PW.

P2MP PWs are generally unidirectional, but a Root PE may need to receive unidirectional P2P return traffic from any Leaf PE. For that purpose, the P2MP PW solution MAY support an optional return path from each Leaf PE to the Root PE.

3.2. P2MP PW and Underlying Layer

The definition of MPLS multicast encapsulation [RFC5332] specifies the procedure to carry MPLS packets that are to be replicated and a copy of the packet sent to each of the specified next hops. This notion is also applicable to a P2MP PW packet carried by a P2MP PSN tunnel.

To be more precise, a P2MP PSN tunnel corresponds to a "point-to-multipoint data link or tunnel" described in Section 3 of [RFC5332].
Similarly, P2MP PW labels correspond to "the top labels (before applying the data link or tunnel encapsulation) of all MPLS packets that are transmitted on a particular point-to-multipoint data link or tunnel".

In the P2MP PW architecture using the SS-PW network model, the PW-PDU [RFC3985] is replicated by the underlying P2MP PSN tunnel layer. Note that the PW label is unchanged, and hidden in switching, by the transit P routers.

In a solution, a P2MP PW MUST be supported over a single P2MP PSN tunnel as the underlying layer of traffic distribution. Figure 2 gives an example of P2MP PW topology relying on a single P2MP LSP. The PW tree is composed of one Root PE (i1) and several Leaf PEs (e1, e2, e3, e4).

The mechanisms for establishing the PSN tunnel are outside the scope of this document, as long as they enable the essential attributes of the service to be emulated.

\[
\begin{array}{c}
\text{i1} \\
/ \\
/ \ \\
/ \ \\
/ \ \\
e1 e2 e3 e4
\end{array}
\]

Figure 2: Example of P2MP Underlying Layer for P2MP PW

A single P2MP PSN tunnel MUST be able to serve the traffic from more than one P2MP PW in an aggregated way, i.e., multiplexing.

A P2MP PW solution MAY support different P2MP PSN tunneling technology (e.g., MPLS over GRE [RFC4023] or P2MP MPLS LSP) or different setup protocols (e.g., multipoint extensions for LDP (mLDP) [RFC6388] and P2MP RSVP-TE [RFC4875]).

The P2MP LSP associated to the P2MP PW can be selected either by user configuration or by dynamically using a multiplexing/demultiplexing mechanism.

The P2MP PW multiplexing SHOULD be used based on the overlap rate between P2MP LSP and P2MP PW. As an example, an existing P2MP LSP may attach more leaves than the ones defined as Leaf PEs for a given
P2MP PW. It may be attractive to reuse it to minimize new
configuration, but using this P2MP LSP would cause non-Leaf PEs
(i.e., not part of the P2MP PW) to receive unwanted traffic.

Note: no special configuration is needed for non-Leaf PEs to drop
that unwanted traffic because they do not have forwarding information
entries unless they process the setup operation for corresponding
P2MP PWs (e.g., signaling).

The operator SHOULD determine whether it is acceptable to partially
multiplex the P2MP PW onto a P2MP LSP, and a minimum congruency rate
may be defined to enable the Root PE to make this determination. The
congruency rate SHOULD take into account several items, including:

- the amount of overlap between the Leaf PEs of the P2MP PW and the
  existing egress PE routers of the P2MP LSP. If there is a
  complete overlap, the congruency is perfect and the rate is 100%.

- the impact on other traffic (e.g., from other VPNs) supported over
  the P2MP LSP.

With this procedure, a P2MP PW is nested within a P2MP LSP. This
allows multiplexing several PWs over a common P2MP LSP. Prior to the
P2MP PW signaling phase, the Root PE determines which P2MP LSP will
be used for this P2MP PW. The PSN tunnel can be an existing PSN
tunnel or the Root PE can create a new P2MP PSN tunnel. Note that
the ingress PE may modify or re-create an existing P2MP PSN tunnel in
order to add one or more leaf PEs to enable it to transport the P2MP
PW.

3.3. P2MP PW Construction

[RFC5332] introduces two approaches to assigning MPLS labels (meaning
PW labels in the P2MP PW context): Upstream-Assigned [RFC5331] and
Downstream-Assigned. However, it is out of scope of this document
which one should be used in PW construction. It is left to the
specification of the solution.

The following requirements apply to the establishment of P2MP PWs:

- PE nodes MUST be configurable with the P2MP PW identifiers and
  ACs.

- A discovery mechanism SHOULD allow the Root PE to discover the
  Leaf PEs, or vice versa.
- Solutions SHOULD allow single-sided operation at the Root PE for the selection of some AC(s) at the Leaf PE(s) to be attached to the PW tree so that the Root PE controls the leaf attachment.

- The Root PE SHOULD support a method to be informed about whether a Leaf PE has successfully attached to the PW tree.

3.4. P2MP PW Signaling Requirements

3.4.1. P2MP PW Identifier

The P2MP PW MUST be uniquely identified. This unique P2MP PW identifier MUST be used for all signaling procedures related to this PW (PW setup, monitoring, etc.).

3.4.2. PW Type Mismatch

The Root PE and Leaf PEs of a P2MP PW MUST be configured with the same PW type as defined in [RFC4446] for P2P PW. In case of a type mismatch, a PE SHOULD abort attempts to attach the Leaf PE to the P2MP PW.

3.4.3. Interface Parameters Sub-TLV

Some interface parameters [RFC4446] related to the AC capability have been defined according to the PW type and are signaled during the PW setup.

Where applicable, a solution is REQUIRED to ascertain whether the AC at the Leaf PE is capable of supporting traffic coming from the AC at the Root PE.

In case of a mismatch, the passive PE (Root or Leaf PE, depending on the signaling process) SHOULD support mechanisms to reject attempts to attach the Leaf PE to the P2MP PW.

3.4.4. Leaf Grafting/Pruning

Once the PW tree is established, the solution MUST allow the addition or removal of a Leaf PE, or a subset of leaves to/from the existing tree, without any impact on the PW tree (data and control planes) for the remaining Leaf PEs.

The addition or removal of a Leaf PE MUST also allow the P2MP PSN tunnel to be updated accordingly. This may cause the P2MP PSN tunnel to add or remove the corresponding Leaf PE.
3.4.5. Failure Detection and Reporting

Since the underlying layer has an end-to-end P2MP topology between the Root PE and the Leaf PEs, the failure reporting and processing procedures are implemented only on the edge nodes.

Failure events may cause one or more Leaf PEs to become detached from the PW tree. These events MUST be reported to the Root PE, using appropriate out-of-band or in-band Operations, Administration, and Maintenance (OAM) messages for monitoring.

It MUST be possible for the operator to choose the out-of-band or in-band monitoring tools or both to monitor the Leaf PE status. For management purposes, the solution SHOULD allow the Root PE to be informed of Leaf PEs’ failure.

Based on these failure notifications, solutions MUST allow the Root PE to update the remaining leaves of the PW tree.

- A solution MUST support an in-band status notification mechanism to detect failures: unidirectional point-to-multipoint traffic failure. This MUST be realized by enhancing existing unicast PW methods, such as Virtual Circuit Connectivity Verification (VCCV) for seamless and familiar operation as defined in [RFC5085].

- In case of failure, it MUST correctly report which Leaf PEs are affected. This MUST be realized by enhancing existing PW methods, such as LDP Status Notification. The notification message SHOULD include the type of fault (P2MP PW, AC, or PSN tunnel).

- A Leaf PE MAY be notified of the status of the Root PE’s AC.

- A solution MUST support OAM message mapping [RFC6310] at the Root PE and Leaf PE if a failure is detected on the source CE.

3.4.6. Protection and Restoration

It is assumed that if recovery procedures are required, the P2MP PSN tunnel will support standard MPLS-based recovery techniques. In that case, a mechanism SHOULD be implemented to avoid race conditions between recovery at the PSN level and recovery at the PW level.

An alternative protection scheme MAY rely on the PW layer.

Leaf PEs MAY be protected via a P2MP PW redundancy mechanism. In the example depicted below, a standby P2MP PW is used to protect the active P2MP PW. In that protection scheme, the AC at the Root PE MUST serve both P2MP PWs. In this scenario, the criteria for
switching over SHOULD be defined, e.g., failure of one or all leaves of the active P2MP PW will trigger switchover of the whole P2MP PW.

![Diagram of P2MP PW Redundancy for Protecting Leaf PEs]

Note that some of the nodes/links in this figure can be physically shared; this depends on the service provider policy of network redundancy.

The Root PE MAY be protected via a P2MP PW redundancy mechanism. In the example depicted below, a standby P2MP PW is used to protect the active P2MP. A single AC at the Leaf PE MUST be used to attach the CE to the primary and the standby P2MP PW. The Leaf PE MUST support protection mechanisms in order to select the active P2MP PW.

![Diagram of P2MP PW Redundancy for Protecting Root PEs]
3.4.7. Scalability

The solution SHOULD scale at worst linearly for message size, memory requirements, and processing requirements, with the number of Leaf PEs.

Increasing the number of P2MP PWs between a Root PE and a given set of Leaf PEs SHOULD NOT cause the P router to increase the number of entries in its forwarding table by the same or greater proportion. Multiplexing P2MP PWs to P2MP PSN tunnels achieves this.

4. Backward Compatibility

Solutions MUST be backward compatible with current PW standards. Solutions SHOULD utilize existing capability advertisement and negotiation procedures for the PEs implementing P2MP PW endpoints.

The implementation of OAM mechanisms also implies the advertisement of PE capabilities to support specific OAM features. The solution MAY allow advertising P2MP PW OAM capabilities. A solution MUST NOT allow a P2MP PW to be established to PEs that do not support P2MP PW functionality. It MUST have a mechanism to report an error for incompatible PEs.

In some cases, upstream traffic is needed from downstream CEs to upstream CEs. The P2MP PW solution SHOULD allow a return path (i.e., from the Leaf PE to the Root PE) that provides upstream connectivity.

In particular, the same ACs MAY be shared between the downstream and upstream directions. For downstream, a CE receives traffic originated by the Root PE over its AC. For upstream, the CE MAY also send traffic destined to the same Root PE over the same AC.

5. Security Considerations

The security requirements common to PW are raised in Section 11 of [RFC3916]. P2MP PW is a variant of the initial P2P PW definition, and those requirements (and the security considerations from [RFC3985]) also apply. The security considerations from [RFC5920] and [RFC6941] also apply to the IP/MPLS and MPLS-TP deployment scenarios, respectively.

Some issues specifically due to P2MP topology need to be addressed in the definition of the solution:

- The solution SHOULD provide means to protect the traffic delivered to receivers (Integrity, Confidentiality, Endpoint Authentication).
The solution SHOULD support means to protect the P2MP PW as a whole against attacks that would lead to any kind of denial of service.

Specifically, safeguard mechanisms should be considered to avoid any negative impact on the whole PW tree when any one receiver or any group of receivers is attacked. Safeguard mechanisms for both the data plane and the control plane need to be considered.

6. References

6.1. Normative References


6.2. Informative References


7. Acknowledgments

The authors thank the following people: the authors of [RFC4461] since the structure and content of this document were, for some sections, largely inspired by [RFC4461]; JL. Le Roux and A. Cauvin for the discussions, comments, and support; Adrian Farrel for his Routing Area Director review; and IESG reviewers.
8. Contributors

Philippe Niger
France Telecom
2, avenue Pierre-Marzin
22307 Lannion Cedex
France
EMail: philippe.niger@orange-ftgroup.com

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

Lei Wang
Telenor
Snaroyveien 30
Fornebu 1331
Norway
EMail: lei.wang@telenor.com

Rahul Aggarwal
Juniper Networks
1194 North Mathilda Ave.
Sunnyvale, CA  94089
US
EMail: rahul@juniper.net

Simon Delord
Telstra
380 Flinders Lane
Melbourne
Australia
EMail: simon.delord@gmail.com
Martin Vigoureux  
Alcatel-Lucent France  
Route de Villejust  
91620 Nozay  
France  

EMail: martin.vigoureux@alcatel-lucent.fr

Lizhong Jin  
ZTE Corporation  
889, Bibo Road  
Shanghai, 201203  
China  

EMail: lizho.jin@gmail.com
Authors’ Addresses

Frederic Jounay (editor)
Orange CH
4 rue caudray 1020 Renens
Switzerland
EMail: frederic.jounay@orange.ch

Yuji Kamite (editor)
NTT Communications Corporation
1-1-6 Uchisaiwai-cho, Chiyoda-ku
Tokyo 100-8019
Japan
EMail: y.kamite@ntt.com

Giles Heron
Cisco Systems, Inc.
9 New Square
Bedfont Lakes
Feltham
Middlesex
TW14 8HA
United Kingdom
EMail: giheron@cisco.com

Matthew Bocci
Alcatel-Lucent Telecom Ltd
Voyager Place
Shoppenhangers Road
Maidenhead
Berks
United Kingdom
EMail: Matthew.Bocci@alcatel-lucent.com