Network Working Group
Request for Comments: 5254
Category: Informational

N. Bitar, Ed.
Verizon
M. Bocci, Ed.
Alcatel-Lucent
L. Martini, Ed.
Cisco Systems, Inc.

October 2008

Requirements for Multi-Segment Pseudowire Emulation Edge-to-Edge (PWE3)

Status of This Memo

This memo provides information for the Internet community. It does not specify an Internet standard of any kind. Distribution of this memo is unlimited.

Abstract

This document describes the necessary requirements to allow a service provider to extend the reach of pseudowires across multiple domains. These domains can be autonomous systems under one provider administrative control, IGP areas in one autonomous system, different autonomous systems under the administrative control of two or more service providers, or administratively established pseudowire domains.

Table of Contents

1. Introduction

1.1. Scope

This document specifies requirements for extending pseudowires across more than one packet switched network (PSN) domain and/or more than one PSN tunnel. These pseudowires are called multi-segment pseudowires (MS-PWs). Requirements for single-segment pseudowires (SS-PWs) that extend edge to edge across only one PSN domain are specified in [RFC3916]. This document is not intended to invalidate any part of [RFC3985].

This document specifies additional requirements that apply to MS-PWs. These requirements do not apply to PSNs that only support SS-PWs.

1.2. Architecture

The following three figures describe the reference models that are derived from [RFC3985] to support PW emulated services.

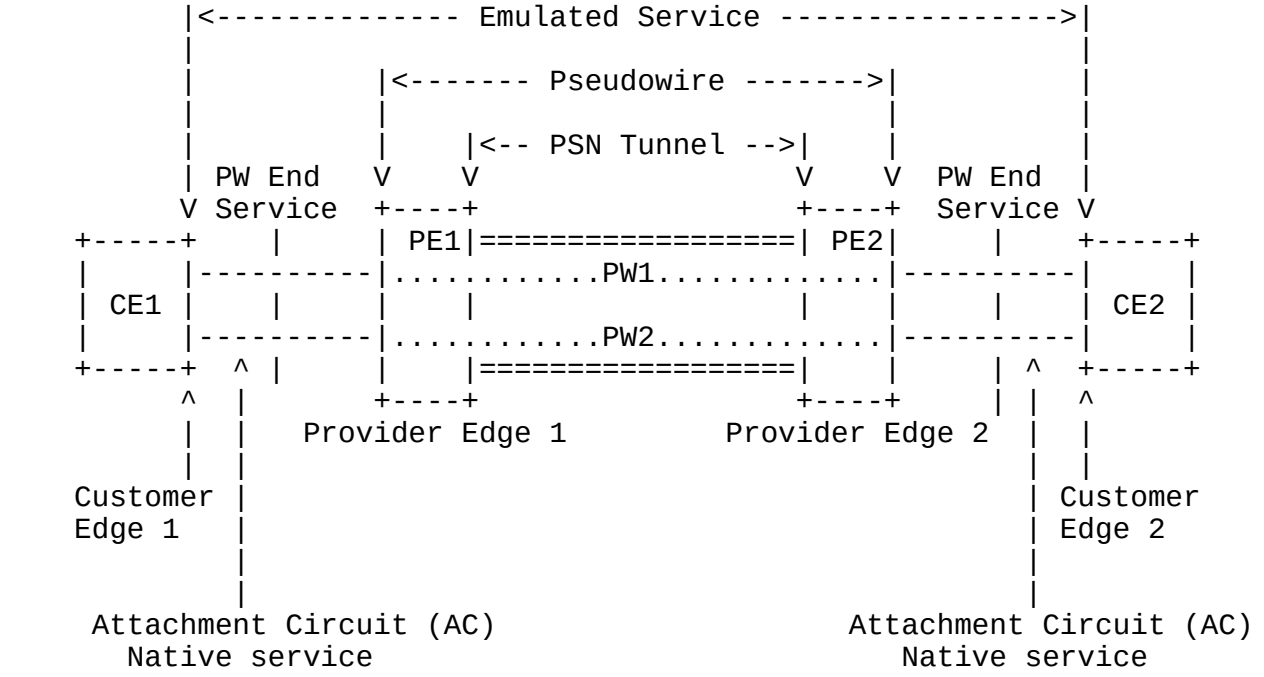

Figure 1: PWE3 Reference Configuration

Figure 1 shows the PWE3 reference architecture [RFC3985]. This architecture applies to the case where a PSN tunnel extends between two edges of a single PSN domain to transport a PW with endpoints at these edges.

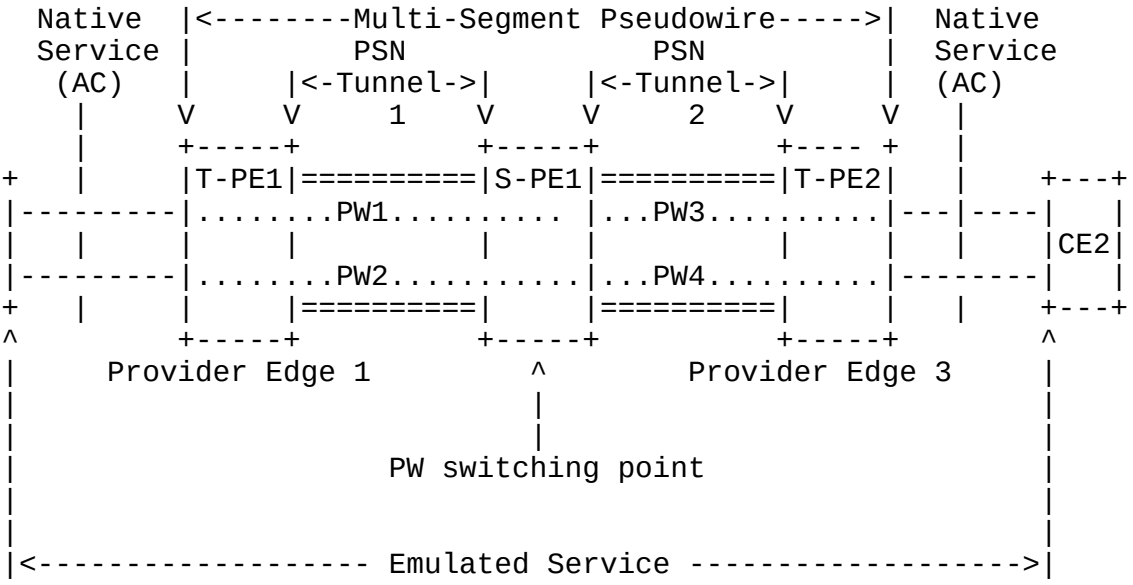

Figure 2: PW Switching Reference Model

Figure 2 extends this architecture to show a multi-segment case. Terminating PE1 (T-PE1) and Terminating PE3 (T-PE3) provide PWE3 service to CE1 and CE2. These PEs terminate different PSN tunnels, PSN Tunnel 1 and PSN Tunnel 2, and may reside in different PSN or pseudowire domains. One PSN tunnel extends from T-PE1 to S-PE1 across PSN1, and a second PSN tunnel extends from S-PE1 to T-PE2 across PSN2.

PWs are used to connect the Attachment circuits (ACs) attached to T-PE1 to the corresponding ACs attached to T-PE2. Each PW on PSN tunnel 1 is switched to a PW in the tunnel across PSN2 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 will be referred to as the PW switching provider edge (S-PE). PW1 and PW3 are segments of the same MS-PW while PW2 and PW4 are segments of another pseudowire. PW segments of the same MS-PW (e.g., PW1 and PW3) MAY be of the same PW type or different types, and PSN tunnels (e.g., PSN Tunnel 1 and PSN Tunnel 2) can be the same or different technology. This document requires support for MS-PWs with segments of the same PW type only.

An S-PE switches an MS-PW from one segment to another based on the PW identifiers (e.g., PW label in case of MPLS PWs). In Figure 2, the domains that PSN Tunnel 1 and PSN Tunnel 2 traverse could be IGP areas in the same IGP network or simply PWE3 domains in a single flat IGP network, for instance.

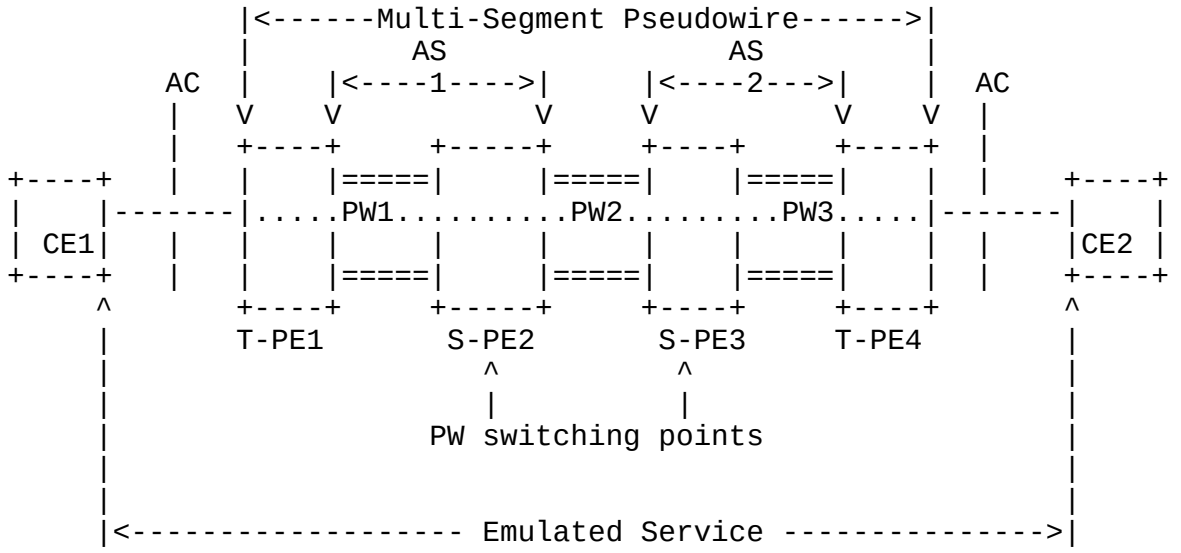

Figure 3: PW Switching Inter-Provider Reference Model

Note that although Figure 2 only shows a single S-PE, a PW may transit more than one S-PEs along its path. For instance, in the multi-AS case shown in Figure 3, there can be an S-PE (S-PE2) at the border of one AS (AS1) and another S-PE (S-PE3) at the border of the other AS (AS2). An MS-PW that extends from the edge of one AS (T-PE1) to the edge of the other AS (T-PE4) is composed of three segments: (1) PW1, a segment in AS1, (2) PW2, a segment between the two border routers (S-PE2 and S-PE3) that are switching PEs, and (3) PWE3, a segment in AS2. AS1 and AS2 could belong to the same provider (e.g., AS1 could be an access network or metro transport network, and AS2 could be an MPLS core network) or to two different providers (e.g., AS1 for Provider 1 and AS2 for Provider 2).

2. Terminology

RFC 3985 [RFC3985] provides terminology for PWE3. The following additional terminology is defined for multi-segment pseudowires:

- 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 an MS-PW. This incorporates the functionality of a PE as defined in RFC 3985.

- Single-Segment Pseudowire (SS-PW). A PW setup directly between two PE devices. Each direction of an SS-PW traverses one PSN tunnel that connects the two 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 an MS-PW by definition MUST terminate on a T-PE.

- PW Segment. A single-segment or a part of a 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 an MS-PW. The S-PE terminates the PSN tunnels transporting the preceding and succeeding segments of the MS-PW. It is therefore a PW switching point for an 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 set up and manage PW segments with other PW switching points and terminating PEs.

2.1. 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].

3. Use Cases

PWE3 defines the signaling and encapsulation techniques for establishing SS-PWs between a pair of terminating PEs (T-PEs), and in the vast majority of cases, this will be sufficient. MS-PWs may be useful in the following situations:

-i. Inter-Provider PWs: An Inter-Provider PW is a PW that extends from a T-PE in one provider domain to a T-PE in another provider domain.

-ii. It may not be possible, desirable, or feasible to establish a direct PW control channel between the T-PEs, residing in different provider networks, to set up and maintain PWs. At a minimum, a direct PW control channel establishment (e.g., targeted LDP session) requires knowledge of and reachability to the remote T-PE IP address. The local T-PE may not have access to this information due to operational or security constraints. Moreover, an SS-PW would require the existence of a PSN tunnel between the local T-PE and the remote T-PE. It may not be feasible or desirable to extend single, contiguous PSN tunnels between T-PEs in one domain and T-PEs in another domain for security and/or scalability reasons or because the two domains may be using different PSN technologies.

-iii. MS-PW setup, maintenance, and forwarding procedures must satisfy requirements placed by the constraints of a multi-provider environment. An example is the inter-AS L2VPN scenario where the T-PEs reside in different provider networks (ASs) and it is the current practice to MD5-key all control traffic exchanged between two networks. An MS-PW allows the providers to confine MD5 key administration for the LDP session to just the PW switching points connecting the two domains.

-iv. PSN Interworking: PWE3 signaling protocols and PSN types may differ in different provider networks. The terminating PEs may be connected to networks employing different PW signaling and/or PSN protocols. In this case, it is not possible to use an SS-PW. An MS-PW with the appropriate interworking performed at the PW switching points can enable PW connectivity between the terminating PEs in this scenario.

-v. Traffic Engineered PSN Tunnels and bandwidth-managed PWs: There is a requirement to deploy PWs edge to edge in large service provider networks. Such networks typically encompass hundreds or thousands of aggregation devices at the edge, each of which would be a PE. Furthermore, there is a requirement that these PWs have explicit bandwidth guarantees. To satisfy these requirements, the PWs will be tunneled over PSN TE-tunnels with bandwidth constraints. A single-segment pseudowire architecture would require that a full mesh of PSN TE-tunnels be provisioned to allow PWs to be established between all PEs. Inter-provider PWs riding traffic engineered tunnels further add to the number of tunnels that would have to be supported by the PEs and the core network as the total number of PEs increases.

In this environment, there is a requirement either to support a sparse mesh of PSN TE-tunnels and PW signaling adjacencies, or to partition the network into a number of smaller PWE3 domains. In either case, a PW would have to pass through more than one PSN tunnel hop along its path. An objective is to reduce the number of tunnels that must be supported, and thus the complexity and scalability problem that may arise.

-vi. Pseudowires in access/metro networks: Service providers wish to extend PW technology to access and metro networks in order to reduce maintenance complexity and operational costs. Today’s access and metro networks are either legacy (Time Division Multiplexed (TDM), Synchronous Optical Network/Synchronous Digital Hierarchy (SONET/SDH), or Frame Relay/Asynchronous Transfer Mode (ATM)), Ethernet, or IP based.

Due to these architectures, circuits (e.g., Ethernet Virtual Circuits (EVCs), ATM VCs, TDM circuits) in the access/metro are traditionally handled as attachment circuits, in their native format, to the edge of the IP-MPLS network where the PW starts. This combination requires multiple separate access networks and complicates end-to-end control, provisioning, and maintenance. In addition, when a TDM or SONET/SDH access network is replaced with a packet-based infrastructure, expenses may be lowered due to moving statistical multiplexing closer to the end-user and converging multiple services onto a single access network.

Access networks have a number of properties that impact the application of PWs. For example, there exist access mechanisms where the PSN is not of an IETF specified type, but uses mechanisms compatible with those of PWE3 at the PW layer.

Here, use case (iv) may apply. In addition, many networks consist of hundreds or thousands of access devices. There is therefore a desire to support a sparse mesh of PW signaling adjacencies and PSN tunnels. Use case (v) may therefore apply. Finally, access networks also tend to differ from core networks in that the access PW setup and maintenance mechanism may only be a subset of that used in the core.

Using the MS-PWs, access and metro network elements need only maintain PW signaling adjacencies with the PEs to which they directly connect. They do not need PW signaling adjacencies with every other access and metro network device. PEs in the PSN backbone, in turn, maintain PW signaling adjacencies among each other. In addition, a PSN tunnel is set up between an access element and the PE to which it connects. Another PSN tunnel needs to be established between every PE pair in the PSN backbone. An MS-PW may be set up from one access network element to another access element with three segments: (1) access-element - PSN-PE, (2) PSN-PE to PSN-PE, and (3) PSN-PE to access element. In this MS-PW setup, access elements are T-PEs while PSN-PEs are S-PEs. It should be noted that the PSN backbone can be also segmented into PWE3 domains resulting in more segments per PW.

3.1. Multi-Segment Pseudowire Setup Mechanisms

This requirements document assumes that the above use cases are realized using one or more of the following mechanisms:

-i. Static Configuration: The switching points (S-PEs), in addition to the T-PEs, are manually provisioned for each segment.

-ii. Pre-Determined Route: The PW is established along an administratively determined route using an end-to-end signaling protocol with automated stitching at the S-PEs.

-iii. Signaled Dynamic Route: The PW is established along a dynamically determined route using an end-to-end signaling protocol with automated stitching at the S-PEs. The route is selected with the aid of one or more dynamic routing protocols.

Note that we define the PW route to be the set of S-PEs through which an MS-PW will pass between a given pair of T-PEs. PSN tunnels along that route can be explicitly specified or locally selected at the S-PEs and T-PEs. The routing of the PSN tunnels themselves is outside the scope of the requirements specified in this document.

4. Multi-Segment Pseudowire Requirements

The following sections detail the requirements that the above use cases put on the MS-PW setup mechanisms.

4.1. All Mechanisms

The following generic requirements apply to the three MS-PW setup mechanisms defined in the previous section.

4.1.1. Architecture

-i. If MS-PWs are tunneled across a PSN that only supports SS-PWs, then only the requirements of [RFC3916] apply to that PSN. The fact that the overlay is carrying MS-PWs MUST be transparent to the routers in the PSN.

-ii. The PWs MUST remain transparent to the P-routers. A P-router is not an S-PE or an T-PE from the MS-PW architecture viewpoint. P-routers provide transparent PSN transport for PWs and MUST not have any knowledge of the PWs traversing them.

-iii. The MS-PWs MUST use the same encapsulation modes specified for SS-PWs.

-iv. The MS-PWs MUST be composed of SS-PWs.

-v. An MS-PW MUST be able to pass across PSNs of all technologies supported by PWE3 for SS-PWs. When crossing from one PSN technology to another, an S-PE must provide the necessary PSN interworking functions in that case.

-vi. Both directions of a PW segment MUST terminate on the same S-PE/T-PE.

-vii. S-PEs MAY only support switching PWs of the same PW type. In this case, the PW type is transparent to the S-PE in the forwarding plane, except for functions needed to provide for interworking between different PSN technologies.

-viii. Solutions MAY provide a way to prioritize the setup and maintenance process among PWs.

4.1.2. Resiliency

Mechanisms to protect an MS-PW when an element on the existing path of an MS-PW fails MUST be provided. These mechanisms will depend on the MS-PW setup. The following are the generic resiliency requirements that apply to all MS-PW setup mechanisms:

-i. Configuration and establishment of a backup PW to a primary PW SHOULD be supported. Mechanisms to perform a switchover from a primary PW to a backup PW upon failure detection SHOULD be provided.

-ii. The ability to configure an end-to-end backup PW path for a primary PW path SHOULD be supported. The primary and backup paths may be statically configured, statically specified for signaling, or dynamically selected via dynamic routing depending on the MS-PW establishment mechanism. Backup and primary paths should have the ability to traverse separate S-PEs. The backup path MAY be signaled at configuration time or after failure.

-iii. The ability to configure a primary PW and a backup PW with a different T-PE from the primary SHOULD be supported.

-iv. Automatic Mechanisms to perform a fast switchover from a primary PW to a backup PW upon failure detection SHOULD be provided.

-v. A mechanism to automatically revert to a primary PW from a backup PW MAY be provided. When provided, it MUST be configurable.

4.1.3. Quality of Service

Pseudowires are intended to support emulated services (e.g., TDM and ATM) that may have strict per-connection quality-of-service (QoS) requirements. This may include either absolute or relative guarantees on packet loss, delay, and jitter. These guarantees are, in part, delivered by reserving sufficient network resources (e.g., bandwidth), and by providing appropriate per-packet treatment (e.g., scheduling priority and drop precedence) throughout the network.

For SS-PWs, a traffic engineered PSN tunnel (i.e., MPLS-TE) may be used to ensure that sufficient resources are reserved in the P-routers to provide QoS to PWs on the tunnel. In this case, T-PEs MUST have the ability to automatically request the PSN tunnel resources in the direction of traffic (e.g., admission control of PWs onto the PSN tunnel and accounting for reserved bandwidth and

available bandwidth on the tunnel). In cases where the tunnel supports multiple classes of service (CoS) (e.g., E-LSP), bandwidth management is required per CoS.

For MS-PWs, each S-PE maps a PW segment to a PSN tunnel. Solutions MUST enable S-PEs and T-PEs to automatically bind a PW segment to a PSN tunnel based on CoS and bandwidth requirements when these attributes are specified for a PW. Solutions SHOULD also provide the capability of binding a PW segment to a tunnel as a matter of policy configuration. S-PEs and T-PEs must be capable of automatically requesting PSN tunnel resources per CoS.

S-PEs and T-PEs MUST be able to associate a CoS marking (e.g., EXP field value for MPLS PWs) with PW PDUs. CoS marking in the PW PDUs affects packet treatment. The CoS marking depends on the PSN technology. Thus, solutions must enable the configuration of necessary mapping for CoS marking when the MS-PW crosses from one PSN technology to another. Similarly, different administrative domains may use different CoS values to imply the same CoS treatment. Solutions MUST provide the ability to define CoS marking maps on S-PEs at administrative domain boundaries to translate from one CoS value to another as a PW PDU crosses from one domain to the next.

[RFC3985] requires PWs to respond to path congestion by reducing their transmission rate. Alternatively, RFC 3985 permits PWs that do not have a congestion control mechanism to transmit using explicitly reserved capacity along a provisioned path. Because MS-PWs are a type of PW, this requirement extends to them as well. RFC 3985 applied to MS-PWs consequently requires that MS-PWs employ a congestion control mechanism that is effective across an MS path, or requires an explicit provisioning action that reserves sufficient capacity in all domains along the MS path before the MS-PW begins transmission. S-PEs are therefore REQUIRED to reject attempts to establish MS-PW segments for PW types that either do not utilize an appropriate congestion control scheme or when resources that are sufficient to support the transmission rate of the PW cannot be reserved along the path.

4.1.4. Congestion Control

[RFC3985] requires all PWs to respond to congestion, in order to conform to [RFC2914]. In the absence of a well-defined congestion control mechanism, [RFC3985] permits PWs to be carried across paths that have been provisioned such that the traffic caused by PWs has no harmful effect on concurrent traffic that shares the path, even under congestion. These requirements extend to the MS-PWs defined in this document.

Path provisioning is frequently performed through QoS reservation protocols or network management protocols. In the case of SS-PWs, which remain within a single administrative domain, a number of existing protocols can provide this provisioning functionality. MS-PWs, however, may transmit across network domains that are under the control of multiple entities. QoS provisioning across such paths is inherently more difficult, due to the required inter-domain interactions. It is important to note that these difficulties do not invalidate the requirement to provision path capacity for MS-PW use. Each domain MUST individually implement a method to control congestion. This can be by QoS reservation, or other congestion control method. MS-PWs MUST NOT transmit across unprovisioned, best effort, paths in the absence of other congestion control schemes, as required by [RFC3985].

Solutions MUST enable S-PEs and T-PEs on the path of an MS-PW to notify other S-PEs and T-PEs on that path of congestion, when it occurs. Congestion may be indicated by queue length, packet loss rate, or bandwidth measurement (among others) crossing a respective threshold. The action taken by a T-PE that receives a notification of congestion along the path of one of its PWs could be to re-route the MS-PW to an alternative path, including an alternative T-PE if available. If a PE, or an S-PE has knowledge that a particular link or tunnel is experiencing congestion, it MUST not set up any new MS-PW that utilize that link or tunnel. Some PW types, such as TDM PWs, are more sensitive to congestion than others. The reaction to a congestion notification MAY vary per PW type.

4.1.5. Additional Generic Requirements for MS-PW Setup Mechanisms

The MS-PW setup mechanisms MUST accommodate the service provider's practices, especially in relation to security, confidentiality of SP information, and traffic engineering. Security and confidentiality are especially important when the MS-PWs are set up across autonomous systems in different administrative domains. The following are generic requirements that apply to the three MS-PW setup mechanisms defined earlier:

-i. The ability to statically select S-PEs and PSN tunnels on a PW path MUST be provided. Static selection of S-PEs is by definition a requirement for the static configuration and signaled/static route setup mechanisms. This requirement satisfies the need for forcing an MS-PW to traverse specific S-PEs to enforce service provider security and administrative policies.

-ii. Solutions SHOULD minimize the amount of configuration needed to set up an MS-PW.

-iii. Solutions should support different PW setup mechanisms on the same T-PE, S-PE, and PSN network.

-iv. Solutions MUST allow T-PEs to simultaneously support use of SS-PW signaling mechanisms as specified in [RFC4447], as well as MS-PW signaling mechanisms.

-v. Solutions MUST ensure that an MS-PW will be set up when a path that satisfies the PW constraints for bandwidth, CoS, and other possible attributes does exist in the network.

-vi. Solutions must clearly define the setup procedures for each mechanism so that an MS-PW setup on T-PEs can be interpreted as successful only when all PW segments are successfully set up.

-vii. Admission control to the PSN tunnel needs to be performed against available resources, when applicable. This process MUST be performed at each PW segment comprising the MS-PW. PW admission control into a PSN tunnel MUST be configurable.

-viii. In case the PSN tunnel lacks the resources necessary to accommodate the new PW, an attempt to signal a new PSN tunnel, or increase the capacity of the existing PSN tunnel MAY be made. If the expanded PSN tunnel fails to set up, the PW MUST fail to set up.

-ix. The setup mechanisms must allow the setup of a PW segment between two directly connected S-PEs without the existence of a PSN tunnel. This requirement allows a PW segment to be set up between two (Autonomous System Border Routers (ASBRs) when the MS-PW crosses AS boundaries without the need for configuring and setting up a PSN tunnel. In this case, admission control must be done, when enabled, on the link between the S-PEs.

4.1.6. Routing

An objective of MS-PWs is to provide support for the following connectivity:

-i. MS-PWs MUST be able to traverse multiple service provider administrative domains.

-ii. MS-PWs MUST be able to traverse multiple autonomous systems within the same administrative domain.

-iii. MS-PWs MUST be able to traverse multiple autonomous systems belonging to different administrative domains.

-iv. MS-PWs MUST be able to support any hybrid combination of the aforementioned connectivity scenarios, including both PW transit and termination in a domain.

4.2. Statically Configured MS-PWs

When the MS-PW segments are statically configured, the following requirements apply in addition to the generic requirements previously defined.

4.2.1. Architecture

There are no additional requirements on the architecture.

4.2.2. MPLS-PWs

Solutions should allow for the static configuration of MPLS labels for MPLS-PW segments and the cross-connection of these labels to preceding and succeeding segments. This is especially useful when an MS-PW crosses provider boundaries and two providers do not want to run any PW signaling protocol between them. A T-PE or S-PE that allows the configuration of static labels for MS-PW segments should also simultaneously allow for dynamic label assignments for other MS-PW segments. It should be noted that when two interconnected S-PEs do not have signaling peering for the purpose of setting up MS-PW segments, they should have in-band PW Operations and Maintenance (OAM) capabilities that relay PW or attachment circuit defect notifications between the adjacent S-PEs.

4.2.3. Resiliency

The solution should allow for the protection of a PW segment, a contiguous set of PW segments, as well as the end-to-end path. The primary and protection segments must share the same segment endpoints. Solutions should allow for having the backup paths set up prior to the failure or as a result of failure. The choice should be made by configuration. When resources are limited and cannot satisfy all PWs, the PWs with the higher setup priorities should be given preference when compared with the setup priorities of other PWs being set up or the holding priorities of existing PWs.

Solutions should strive to minimize traffic loss between T-PEs.

4.2.4. Quality of Service

The CoS and bandwidth of the MS-PW must be configurable at T-PEs and S-PEs.

4.3. Signaled PW Segments

When the MS-PW segments are dynamically signaled, the following requirements apply in addition to the generic requirements previously defined. The signaled MS-PW segments can be on the path of a statically configured MS-PW, signaled/statically routed MS-PW, or signaled/dynamically routed MS-PW.

There are four different mechanisms that are defined to setup SS-PWs:

-i. Static set up of the SS-PW (MPLS or L2TPv3 forwarding)

-ii. LDP using PWid Forwarding Equivalence Class (FEC) 128

-iii. LDP using the generalized PW FEC 129

-iv. L2TPv3

The MS-PW setup mechanism MUST be able to support PW segments signaled with any of the above protocols; however, the specification of which combinations of SS-PW signaling protocols are supported by a specific implementation is outside the scope of this document.

For the signaled/statically routed and signaled/dynamically routed MS-PW setup mechanisms, the following requirements apply in addition to the generic requirements previously defined.

4.3.1. Architecture

There are no additional requirements on the architecture.

4.3.2. Resiliency

Solutions should allow for the signaling of a protection path for a PW segment, sequence of segments, or end-to-end path. The protection and primary paths for the protected segment(s) share the same respective segments endpoints. When admission control is enabled, systems must be careful not to double account for bandwidth allocation at merged points (e.g., tunnels). Solutions should allow for having the backup paths set up prior to the failure or as a result of failure. The choice should be made by configuration at the endpoints of the protected path. When resources are limited and cannot satisfy all PWs, the PWs with the higher setup priorities

should be given preference when compared with the setup priorities of other PWs being set up or the holding priorities of existing PWs. Procedures must allow for the primary and backup paths to be diverse.

4.3.3. Quality of Service

When the T-PE attempts to signal an MS-PW, the following capability is required:

- -i. Signaling must be able to identify the CoS associated with an MS-PW.
- -ii. Signaling must be able to carry the traffic parameters for an MS-PW per CoS. Traffic parameters should be based on existing INTSERV definitions and must be used for admission control when admission control is enabled.
- -iii. The PW signaling MUST enable separate traffic parameter values to be specified for the forward and reverse directions of the PW.
- -iv. PW traffic parameter representations MUST be the same for all types of MS-PWs.
- -v. The signaling protocol must be able to accommodate a method to prioritize the PW setup and maintenance operation among PWs.

4.3.4. Routing

See the requirements for "Resiliency" above.

4.3.5. Additional Requirements on Signaled MS-PW Setup Mechanisms

The following are further requirements on signaled MS-PW setup mechanisms:

- -i. The signaling procedures MUST be defined such that the setup of an MS-PW is considered successful if all segments of the MS-PW are successfully set up.
- -ii. The MS-PW path MUST have the ability to be dynamically set up between the T-PEs by provisioning only the T-PEs.

-iii. Dynamic MS-PW setup requires that a unique identifier be associated with a PW and be carried in the signaling message. That identifier must contain sufficient information to determine the path to the remote T-PE through intermediate S-PEs.

-iv. In a single-provider domain, it is natural to have the T-PE identified by one of its IP addresses. This may also apply when an MS-PW is set up across multiple domains operated by the same provider. However, some service providers have security and confidentiality policies that prevent them from advertising reachability to routers in their networks to other providers (reachability to an ASBR is an exception). Thus, procedures MUST be provided to allow dynamic set up of MS-PWs under these conditions.

4.4. Signaled PW / Dynamic Route

The following requirements apply, in addition to those in Sections 4.1 and 4.3, when both dynamic signaling and dynamic routing are used.

4.4.1. Architecture

There are no additional architectural requirements.

4.4.2. Resiliency

The PW routing function MUST support dynamic re-routing around failure points when segments are set up using the dynamic setup method.

4.4.3. Quality of Service

There are no additional QoS requirements.

4.4.4. Routing

The following are requirements associated with dynamic route selection for an MS-PW:

-i. Routing must enable S-PEs and T-PEs to discover S-PEs on the path to a destination T-PE.

-ii. The MS-PW routing function MUST have the ability to automatically select the S-PEs along the MS-PW path. Some of the S-PEs MAY be statically selected and carried in the signaling to constrain the route selection process.

-iii. The PW routing function MUST support re-routing around failures that occur between the statically configured segment endpoints. This may be done by choosing another PSN tunnel between the two segment endpoints or setting up an alternative tunnel.

-iv. Routing protocols must be able to advertise reachability information of attachment circuit (AC) endpoints. This reachability information must be consistent with the AC identifiers carried in signaling.

5. Operations and Maintenance (OAM)

OAM mechanisms for the attachment circuits are defined in the specifications for PW emulated specific technologies (e.g., ITU-T I.610 [i610] for ATM). These mechanisms enable, among other things, defects in the network to be detected, localized, and diagnosed. They also enable communication of PW defect states on the PW attachment circuit. Note that this document uses the term OAM as Operations and Management as per ITU-T I.610.

The interworking of OAM mechanisms for SS-PWs between ACs and PWs is defined in [PWE3-OAM]. These enable defect states to be propagated across a PWE3 network following the failure and recovery from faults.

OAM mechanisms for MS-PWs MUST provide at least the same capabilities as those for SS-PWs. In addition, it should be possible to support both segment and end-to-end OAM mechanisms for both defect notifications and connectivity verification in order to allow defects to be localized in a multi-segment network. That is, PW OAM segments can be T-PE to T-PE, T-PE to S-PE, or S-PE to S-PE.

The following requirements apply to OAM for MS-PWs:

-i. Mechanisms for PW segment failure detection and notification to other segments of an MS-PW MUST be provided.

-ii. MS-PW OAM SHOULD be supported end-to-end across the network.

-iii. Single ended monitoring SHOULD be supported for both directions of the MS-PW.

-iv. SS-PW OAM mechanisms (e.g., [RFC5085]) SHOULD be extended to support MS-PWs on both an end-to-end basis and segment basis.

-v. All PE routers along the MS-PW MUST agree on a common PW OAM mechanism to use for the MS-PW.

-vi. At the S-PE, defects on an PSN tunnel MUST be propagated to all PWs that utilize that particular PSN tunnel.

-vii. The directionality of defect notifications MUST be maintained across the S-PE.

-viii. The S-PE SHOULD be able to behave as a segment endpoint for PW OAM mechanisms.

-ix. The S-PE MUST be able to pass T-PE to T-PE PW OAM messages transparently.

-x. Performance OAM is required for both MS-PWs and SS-PWs to measure round-trip delay, one-way delay, jitter, and packet loss ratio.

6. Management of Multi-Segment Pseudowires

Each PWE3 approach that uses MS-PWs SHOULD provide some mechanisms for network operators to manage the emulated service. Management mechanisms for MS-PWs MUST provide at least the same capabilities as those for SS-PWs, as defined in [RFC3916].

It SHOULD also be possible to manage the additional attributes for MS-PWs. Since the operator that initiates the establishment of an MS-PW may reside in a different PSN domain from the S-PEs and one of the T-PEs along the path of the MS-PW, mechanisms for the remote management of the MS-PW SHOULD be provided.

The following additional requirements apply:

6.1. MIB Requirements

-i. MIB Tables MUST be designed to facilitate configuration and provisioning of the MS-PW at the S-PEs and T-PEs.

-ii. The MIB(s) MUST facilitate inter-PSN configuration and monitoring of the ACs.

6.2. Management Interface Requirements

-i. Mechanisms MUST be provided to enable remote management of an MS-PW at an S-PE or T-PE. It SHOULD be possible for these mechanisms to operate across PSN domains. An example of a commonly available mechanism is the command line interface (CLI) over a telnet session.

-ii. For security or other reasons, it SHOULD be possible to disable the remote management of an MS-PW.

7. Security Considerations

This document specifies the requirements both for MS-PWs that can be set up across domain boundaries administered by one or more service providers (inter-provider MS-PWs), and for MS-PWs that are only set up across one provider (intra-provider MS-PWs).

7.1. Inter-Provider MS-PWs

The security requirements for MS-PW setup across domains administered by one service provider are the same as those described under security considerations in [RFC4447] and [RFC3916]. These requirements also apply to inter-provider MS-PWs.

In addition, [RFC4111] identifies user and provider requirements for L2 VPNs that apply to MS-PWs described in this document. In this section, the focus is on the additional security requirements for inter-provider operation of MS-PWs in both the control plane and data plane, and some of these requirements may overlap with those in [RFC4111].

7.1.1. Data-Plane Security Requirements

By security in the "data plane", we mean protection against the following possibilities:

-i. Packets from within an MS-PW traveling to a PE or an AC to which the PW is not intended to be connected, other than in a manner consistent with the policies of the MS-PW.

-ii. Packets from outside an MS-PW entering the MS-PW, other than in a manner consistent with the policies of the MS-PW.

MS-PWs that cross service provider (SP) domain boundaries may connect one T-PE in a SP domain to a T-PE in another provider domain. They may also transit other provider domains even if the two T-PEs are under the control of one SP. Under these scenarios, there is a

chance that one or more PDUs could be falsely inserted into an MS-PW at any of the originating, terminating, or transit domains. Such false injection can be the result of a malicious attack or fault in the S-PE. Solutions MAY provide mechanisms for ensuring the end-to-end authenticity of MS-PW PDUs.

The data plane security requirements at a service provider border for MS-PWs are similar to those for inter-provider BGP/MPLS IP Virtual Private Networks [RFC4364]. In particular, an S-PE or T-PE SHOULD discard a packet received from a particular neighbor over the service provider border unless one of the following two conditions holds:

-i. Any MPLS label processed at the receiving S-PE or T-PE, such the PSN tunnel label or the PW label has a label value that the receiving system has distributed to that neighbor; or

-ii. Any MPLS label processed at the receiving S-PE or T-PE, such as the PSN tunnel label or the PW label has a label value that the receiving S-PE or T-PE has previously distributed to the peer S-PE or T-PE beyond that neighbor (i.e., when it is known that the path from the system to which the label was distributed to the receiving system is via that neighbor).

One of the domains crossed by an MS-PW may decide to selectively mirror the PDUs of an MS-PW for eavesdropping purposes. It may also decide to selectively hijack the PDUs of an MS-PW by directing the PDUs away from their destination. In either case, the privacy of an MS-PW can be violated.

Some types of PWs make assumptions about the security of the underlying PSN. The minimal security provided by an MPLS PSN might not be sufficient to meet the security requirements expected by the applications using the MS-PW. This document does not place any requirements on protecting the privacy of an MS-PW PDU via encryption. However, encryption may be required at a higher layer in the protocol stack, based on the application or network requirements.

The data plane of an S-PE at a domain boundary MUST be able to police incoming MS-PW traffic to the MS-PW traffic parameters or to an administratively configured profile. The option to enable/disable policing MUST be provided to the network administrator. This is to ensure that an MS-PW or a group of MS-PWs do not grab more resources than they are allocated. In addition, the data plane of an S-PE MUST be able to police OAM messages to a pre-configured traffic profile or to filter out these messages upon administrative configuration.

An ingress S-PE MUST ensure that an MS-PW receives the CoS treatment configured or signaled for that MS-PW at the S-PE. Specifically, an S-PE MUST guard against packets marked in the exp bits or IP-header Differentiated Services (DS) field (depending on the PSN) for a better CoS than they should receive.

An ingress S-PE MUST be able to define per-interface or interface-group (a group may correspond to interfaces to a peer-provider) label space for MPLS-PWs. An S-PE MUST be configurable not to accept labeled packets from another provider unless the bottom label is a PW-label assigned by the S-PE on the interface on which the packet arrived.

Data plane security considerations for SS-PWs specified in [RFC3985] also apply to MS-PWs.

7.1.2. Control-Plane Security Requirements

An MS-PW connects two attachment circuits. It is important to make sure that PW connections are not arbitrarily accepted from anywhere, or else a local attachment circuit might get connected to an arbitrary remote attachment circuit. The fault in the connection can start at a remote T-PE or an S-PE.

Where a PW segment crosses a border between one provider and another provider, the PW segment endpoints (S-PEs) SHOULD be on ASBRs interconnecting the two providers. Directly interconnecting the S-PEs using a physically secure link, and enabling signaling and routing authentication between the S-PEs, eliminates the possibility of receiving an MS-PW signaling message or packet from an untrusted peer. Other configurations are possible. For example, P routers for the PSN tunnel between the adjacent S-PEs/T-PEs may reside on the ASBRs. In which case, the S-PEs/T-PEs MUST satisfy themselves of the security and privacy of the path.

The configuration and maintenance protocol MUST provide a strong authentication and control protocol data protection mechanism. This option MUST be implemented, but it should be deployed according to the specific PSN environment requirements. Furthermore, authentication using a signature for each individual MS-PW setup message MUST be available, in addition to an overall control protocol session authentication and message validation.

Since S-PEs in different provider networks SHOULD reside at each end of a physically secure link, or be interconnected by a limited number of trusted PSN tunnels, each S-PE will have a trust relationship with only a limited number of S-PEs in other ASs. Thus, it is expected that current security mechanisms based on manual key management will

be sufficient. If deployment situations arise that require large scale connection to S-PEs in other ASs, then a mechanism based on RFC 4107 [RFC4107] MUST be developed.

Peer authentication protects against IP address spoofing but does not prevent one peer (S-PE or T-PE) from connecting to the wrong attachment circuit. Under a single administrative authority, this may be the result of a misconfiguration. When the MS-PW crosses multiple provider domains, this may be the result of a malicious act by a service provider or a security hole in that provider network. Static manual configuration of MS-PWs at S-PEs and T-PEs provides a greater degree of security. If an identification of both ends of an MS-PW is configured and carried in the signaling message, an S-PE can verify the signaling message against the configuration. To support dynamic signaling of MS-PWs, whereby only endpoints are provisioned and S-PEs are dynamically discovered, mechanisms SHOULD be provided to configure such information on a server and to use that information during a connection attempt for validation.

An incoming MS-PW request/reply MUST NOT be accepted unless its IP source address is known to be the source of an "eligible" peer. An eligible peer is an S-PE or a T-PE with which the originating S-PE or T-PE has a trust relationship. The number of such trusted T-PEs or S-PEs is bounded and not anticipated to create a scaling issue for the control plane authentication mechanisms.

If a peering adjacency has to be established prior to exchanging setup requests/responses, peering MUST only be done with eligible peers. 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 automatically generated from the local PW configuration information.

Furthermore, the restriction of peering sessions to specific interfaces MUST also be provided. The S-PE and T-PE MUST drop the unaccepted signaling messages in the data path to avoid a Denial-of-Service (DoS) attack on the control plane.

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

S-PEs that connect one provider domain to another provider domain MUST have the capability to rate-limit signaling traffic in order to prevent DoS attacks on the control plane. Furthermore, detection and disposition of malformed packets and defense against various forms of attacks that can be protocol-specific MUST be provided.

7.2. Intra-Provider MS-PWs

Security requirements for pseudowires are provided in [RFC3916]. These requirements also apply to MS-PWs.

MS-PWs are intended to enable many more PEs to provide PWE3 services in a given service provider network than traditional SS-PWs, particularly in access and metro environments where the PE may be situated closer to the ultimate endpoint of the service. In order to limit the impact of a compromise of one T-PE in a service provider network, the data path security requirements for inter-provider MS-PWs also apply to intra-provider MS-PWs in such cases.

8. Acknowledgments

The editors gratefully acknowledge the following contributors: Dimitri Papadimitriou (Alcatel-Lucent), Peter Busschbach (Alcatel-Lucent), Sasha Vainshtein (Axerra), Richard Spencer (British Telecom), Simon Delord (France Telecom), Deborah Brungard (AT&T), David McDysan (Verizon), Rahul Aggarwal (Juniper), Du Ke (ZTE), Cagatay Buyukkoc (ZTE), and Stewart Bryant (Cisco).

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.

[RFC3916] Xiao, X., Ed., McPherson, D., Ed., and P. Pate, Ed., "Requirements for Pseudo-Wire Emulation Edge-to-Edge (PWE3)", RFC 3916, September 2004.

[RFC3985] Bryant, S., Ed., and P. Pate, Ed., "Pseudo Wire Emulation Edge-to-Edge (PWE3) Architecture", RFC 3985, March 2005.

9.2. Informative References

[i610] Recommendation I.610 "B-ISDN operation and maintenance principles and functions", February 1999.

[RFC5085] Nadeau, T., Ed., and C. Pignataro, Ed., "Pseudowire Virtual Circuit Connectivity Verification (VCCV): A Control Channel for Pseudowires", RFC 5085, December 2007.

[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.

[RFC4111] Fang, L., Ed., "Security Framework for Provider-Provisioned Virtual Private Networks (PPVPNs)", RFC 4111, July 2005.

[PWE3-OAM] Nadeau, T., Ed., Morrow, M., Ed., Busschbach, P., Ed., Alissaoui, M.,Ed., D. Allen, Ed., "Pseudo Wire (PW) OAM Message Mapping", Work in Progress, March 2005.

[RFC2914] Floyd, S., "Congestion Control Principles", BCP 41, RFC 2914, September 2000.

[RFC4364] Rosen, E. and Y. Rekhter, "BGP/MPLS IP Virtual Private Networks (VPNs)", RFC 4364, February 2006.

[RFC4107] Bellovin, S. and R. Housley, "Guidelines for Cryptographic Key Management", BCP 107, RFC 4107, June 2005.

Authors' Addresses

Nabil Bitar
Verizon
117 West Street
Waltham, MA 02145
EMail: nabil.n.bitar@verizon.com

Matthew Bocci
Alcatel-Lucent Telecom Ltd,
Voyager Place
Shoppenhangers Road
Maidenhead
Berks, UK
EMail: matthew.bocci@alcatel-lucent.co.uk

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

Full Copyright Statement

Intellectual Property