

Internet Engineering Task Force (IETF)
Request for Comments: 6870
Updates: 4447
Category: Standards Track
ISSN: 2070-1721

P. Muley, Ed.
M. Aissaoui, Ed.
Alcatel-Lucent
February 2013

Pseudowire Preferential Forwarding Status Bit

Abstract

This document describes a mechanism for signaling the active and standby status of redundant Pseudowires (PWs) between their termination points. A set of Redundant PWs is configured between Provider Edge (PE) nodes in single-segment pseudowire (SS-PW) applications or between Terminating Provider Edge (T-PE) nodes in Multi-Segment Pseudowire (MS-PW) applications.

In order for the PE/T-PE nodes to indicate the preferred PW to use for forwarding PW packets to one another, a new status bit is defined. This bit indicates a Preferential Forwarding status with a value of active or standby for each PW in a redundant set.

In addition, a second status bit is defined to allow peer PE nodes to coordinate a switchover operation of the PW.

Finally, this document updates RFC 4447 by adding details to the handling of the PW status code bits in the PW Status TLV.

Status of This Memo

This is an Internet Standards Track document.

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

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at <http://www.rfc-editor.org/info/rfc6870>.

Copyright Notice

Copyright (c) 2013 IETF Trust and the persons identified as the document authors. All rights reserved.

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

Table of Contents

1. Introduction	4
1.1. Requirements Language	4
2. Motivation and Scope	4
3. Terminology	7
4. PE Architecture	9
5. Modes of Operation	9
5.1. Independent Mode	9
5.2. Master/Slave Mode	12
6. PW State Transition Signaling Procedures	14
6.1. PW Standby Notification Procedures in Independent Mode	14
6.2. PW Standby Notification Procedures in Master/Slave Mode ...	15
6.2.1. PW State Machine	16
6.3. Coordination of PW Switchover	17
6.3.1. Procedures at the Requesting Endpoint	18
6.3.2. Procedures at the Receiving Endpoint	20
7. Status Mapping	20
7.1. AC Defect State Entry/Exit	21
7.2. PW Defect State Entry/Exit	21
8. Applicability and Backward Compatibility	21
9. Security Considerations	22
10. MIB Considerations	22
11. IANA Considerations	22
11.1. Status Code for PW Preferential Forwarding Status	22
11.2. Status Code for PW Request Switchover Status	23
12. Contributors	23
13. Acknowledgments	24
14. References	24
14.1. Normative References	24
14.2. Informative References	24
Appendix A. Applications of PW Redundancy Procedures	26
A.1. One Multi-Homed CE with Single SS-PW Redundancy	26
A.2. Multiple Multi-Homed CEs with SS-PW Redundancy	28
A.3. Multi-Homed CE with MS-PW Redundancy	30
A.4. Multi-Homed CE with MS-PW Redundancy and S-PE Protection ..	31
A.5. Single-Homed CE with MS-PW Redundancy	32
A.6. PW Redundancy between H-VPLS MTU-s and PE-rs	33

1. Introduction

This document provides the extensions to the Pseudowire (PW) control plane to support the protection schemes of the PW redundancy applications described in RFC 6718, "Pseudowire (PW) Redundancy" [8].

It specifies a new PW status bit as well as the procedures Provider Edge (PE) nodes follow to notify one another of the Preferential Forwarding state of each PW in the redundant set, i.e., active or standby. This status bit is different from the PW status bits already defined in RFC 4447, the pseudowire setup and maintenance protocol [2]. In addition, this document specifies a second status bit to allow peer PE nodes to coordinate a switchover operation of the PW from active to standby, or vice versa.

As a result of the introduction of these new status bits, this document updates RFC 4447 by clarifying the rules for processing status bits not originally defined in RFC 4447. It also updates RFC 4447 by defining that a status bit can indicate a status other than a fault or can indicate an instruction to the peer PE. See more details in Section 8.

Section 15 shows in detail how the mechanisms described in this document are used to achieve the desired protection schemes of the applications described in [8].

1.1. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [1].

2. Motivation and Scope

The PW setup and maintenance protocol defines the following status codes in the PW Status TLV to indicate the state for an attachment circuit (AC) and a PW [7]:

0x00000000 - Pseudowire forwarding (clear all failures)

0x00000001 - Pseudowire Not Forwarding

0x00000002 - Local Attachment Circuit (ingress) Receive Fault

0x00000004 - Local Attachment Circuit (egress) Transmit Fault

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

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

The applications defined in [8] allow the provisioning of a primary PW and one or many secondary backup PWs in the same Virtual Private Wire Service (VPWS) or Virtual Private LAN Service (VPLS). The objective of PW redundancy is to maintain end-to-end connectivity for the emulated service by activating the correct PW whenever an AC, a PE, or a PW fails. The correct PW means the one that provides the end-to-end connectivity from Customer Edge (CE) to CE such that packets continue to flow.

A PE node makes a selection of which PW to activate at any given time for the purpose of forwarding user packets. This selection takes into account the local state of the PW and AC, as well as the remote state of the PW and AC as indicated in the PW status bits it received from the peer PE node.

In the absence of faults, all PWs are up both locally and remotely, and a PE node needs to select a single PW to which to forward user packets. This is referred to as the active PW. All other PWs will be in standby and must not be used to forward user packets.

In order for both ends of the service to select the same PW for forwarding user packets, this document defines a new status bit: the Preferential Forwarding status bit. It also defines the procedures the PE nodes follow to indicate the Preferential Forwarding state of a PW to its peer PE node.

In addition, a second status bit is defined to allow peer PE nodes to coordinate a switchover operation of the PW if required by the application. This is known as the Request Switchover status bit.

Together, the mechanisms described in this document achieve the following protection capabilities defined in [8]:

- a. A 1:1 protection in which a specific subset of a path for an emulated service, consisting of a standby PW and/or AC, protects another specific subset of a path for the emulated service, consisting of an active PW and/or AC. An active PW can forward data traffic and control plane traffic, such as Operations, Administration, and Maintenance (OAM) packets. A standby PW does not carry data traffic.
- b. An N:1 protection scheme in which N specific subsets of a path for an emulated service, consisting each of a standby PW and/or AC, protect a specific subset of a path for the emulated service, consisting of an active PW and/or AC.

- c. A mechanism to allow PW endpoints to coordinate the switchover to a given PW by using an explicit request/acknowledgment switchover procedure. This mechanism is complementary to the independent mode of operation and is described in Section 6.3. 6.3. This mechanism can be invoked manually by the user, effectively providing a manual switchover capability. It can also be invoked automatically to resolve a situation where the PW endpoints could not match the two directions of the PW.
- d. A locally configured precedence to govern the selection of a PW when more than one PW qualifies for the active state, as defined in Sections 5.1. and 5.2. The PW with the lowest precedence value has the highest priority. Precedence may be configured via, for example, a local configuration parameter at the PW endpoint.
- e. By configuration, implementations can designate one PW in the 1:1 or N:1 protection as a primary PW and the remaining as secondary PWs. If more than one PW qualifies for the active state, as defined in Sections 5.1 and 5.2, a PE node selects the primary PW in preference to a secondary PW. In other words, the primary PW has implicitly the lowest precedence value. Furthermore, a PE node reverts to the primary PW immediately after it comes back up or after the expiration of a delay effectively achieving revertive protection switching.

1+1 protection (in which one specific subset of a path for an emulated service, consisting of a standby PW and/or AC, protects another specific subset of a path for the emulated service and in which traffic is permanently duplicated at the ingress node on both the currently active and standby subsets of the paths) is not supported.

The above protection schemes are provided using the following operational modes:

1. An independent mode of operation in which each PW endpoint node uses its own local rule to select which PW it intends to activate at any given time, and advertises that PW to the remote endpoints. Only a PW that is up and that indicated active status bit locally and remotely is in the active state and can be used to forward data packets. This is described in Section 5.1.
2. A master/slave mode in which one PW endpoint, the master endpoint, selects and dictates to the other endpoint(s), the slave endpoint(s), which PW to activate. This is described in Section 5.2.

Note that this document specifies the mechanisms to support PW redundancy where a set of redundant PWs terminate on either a PE, in the case of an SS-PW, or on a T-PE, in the case of an MS-PW. PW redundancy scenarios where the redundant set of PW segments terminates on a Switching Provider Edge (S-PE) are for further study.

3. Terminology

Pseudowire (PW): A mechanism that carries the essential elements of an emulated service from one PE to one or more other PEs over a Public Service Network (PSN) [9].

Single-Segment Pseudowire (SS-PW): A PW set up directly between two T-PE devices. The PW label is unchanged between the originating and terminating PEs [6].

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, terminates on a T-PE [6].

Up PW: A PW that has been configured (label mapping exchanged between PEs) and is not showing any of the PW or AC status bits specified in [7]. Such a PW is available for forwarding traffic [8].

Down PW: A PW that either has not been fully configured or has been configured and is showing any of the PW or AC status bits specified in [7]; such a PW is not available for forwarding traffic [8].

Active PW: An up PW used for forwarding user, OAM, and control plane traffic [8].

Standby PW: An up PW that is not used for forwarding user traffic but may forward OAM and specific control plane traffic [8].

Primary PW: The PW that a PW endpoint activates in preference to any other PW when more than one PW qualifies for active state. When the primary PW comes back up after a failure and qualifies for active state, the PW endpoint always reverts to it. The designation of primary is performed by local configuration for the PW at the PE and is only required when revertive protection switching is used [8].

Secondary PW: When it qualifies for active state, a secondary PW is only selected if no primary PW is configured or if the configured primary PW does not qualify for active state (e.g.,

is down). By default, a PW in a redundancy PW set is considered secondary. There is no revertive mechanism among secondary PWs [8].

PW Precedence: This is a configuration local to the PE that dictates the order in which a forwarder chooses to use a PW when multiple PWs all qualify for the active state. Note that a PW that has been configured as primary has, implicitly, the lowest precedence value.

PW Endpoint: A PE where a PW terminates on a point where Native Service Processing is performed, e.g., an SS-PW PE, an MS-PW T-PE, a Hierarchical VPLS (H-VPLS) MTU-s, or PE-rs [8].

Provider Edge (PE): A device that provides PWE3 to a CE [9].

PW Terminating Provider Edge (T-PE): A PE where the customer-facing 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 [6].

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 of the preceding and succeeding segments of the MS-PW. Therefore, it includes 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. An S-PE can exist anywhere a PW must be processed or policy applied. Therefore, it is not limited to the edge of a provider network [6].

MTU-s: A hierarchical virtual private LAN service Multi-Tenant Unit switch, as defined in RFC 4762 [3].

PE-rs: A routing and bridging capable PE as defined in RFC 4762 [3].

FEC: Forwarding Equivalence Class.

OAM: Operations, Administration, and Maintenance.

VCCV: Virtual Connection Connectivity Verification.

This document uses the term 'PE' to be synonymous with both PEs as per RFC 3985 [9] and T-PEs as per RFC 5659 [6].

This document uses the term 'PW' to be synonymous with both PWs as per RFC 3985 [9] and SS-PWs, MS-PWs, and PW segments as per RFC 5659 [6].

4. PE Architecture

Figure 1 shows the PE architecture for PW redundancy, when more than one PW in a redundant set is associated with a single AC. This is based on the architecture in Figure 4b of RFC 3985 [9]. The forwarder selects which of the redundant PWs to use based on the criteria described in this document.

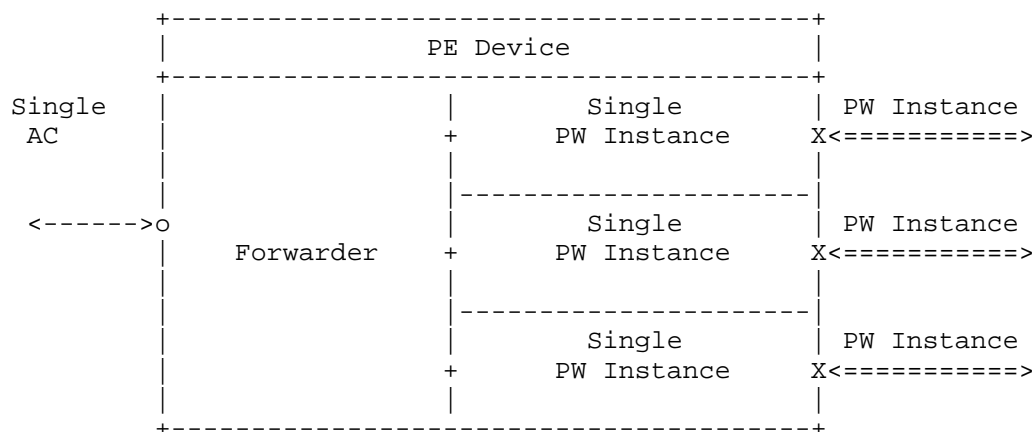


Figure 1. PE Architecture for PW Redundancy

5. Modes of Operation

There are two modes of operation for the use of the PW Preferential Forwarding status bits:

- o independent mode
- o master/slave mode

5.1. Independent Mode

PW endpoint nodes independently select which PWs are eligible to become active and which are not. They advertise the corresponding active or standby Preferential Forwarding status for each PW. Each PW endpoint compares local and remote status bits and uses the PW

that is up at both endpoints and that advertised active Preferential Forwarding status at both the local and remote endpoints.

In this mode of operation, the Preferential Forwarding status indicates the preferred forwarding state of each endpoint but the actual forwarding state of the PW is the result of the comparison of the local and remote forwarding status bits.

If more than one PW qualifies for the active state, each PW endpoint MUST implement a common mechanism to choose the PW for forwarding. The default mechanism MUST be supported by all implementations, and it operates as follows:

1. For a PW using the PWid ID Forwarding Equivalence Class (PWid FEC) [2], the PW with the lowest PWid value is selected.
2. For a PW using the Generalized PWid FEC [2], each PW in a redundant set is uniquely identified at each PE using the following triplet: AGI::SAII::TAII. The unsigned integer form of the concatenated word can be used in the comparison. However, the Source Attachment Individual Identifier (SAII) and Target Attachment Individual Identifier (TAII) values as seen on a PE node are the mirror values of what the peer PE node sees. So that both PE nodes compare the same value, the PE with the lowest system IP address MUST use the unsigned integer form of AGI::SAII::TAII, while the PE with the highest system IP address MUST use the unsigned integer form of AGI::TAII::SAII. This way, both PE nodes will compare the same values. The PW that corresponds to the minimum of the compared values across all PWs in the redundant set is selected.

In the case where the system IP address is not known, it is RECOMMENDED to implement the active PW selection mechanism described next.

In the case of segmented PW, the operator needs to make sure that the PWid or AGI::SAII::TAII of the redundant PWs within the first and last segment are ordered consistently such that the same end-to-end MS-PW gets selected. Otherwise, it is RECOMMENDED to implement the active PW selection mechanism described next.

The PW endpoints MAY also implement the following active PW selection mechanism:

1. If the PW endpoint is configured with the precedence parameter on each PW in the redundant set, it selects the PW with the lowest configured precedence value.

2. If the PW endpoint is configured with one PW as primary and one or more PWs as secondary, it selects the primary PW in preference to all secondary PWs. If a primary PW is not available, it selects the secondary PW with the lowest precedence value. If the primary PW becomes available, a PW endpoint reverts to it immediately or after the expiration of a configurable delay.
3. This active PW selection mechanism assumes the precedence parameter values are configured consistently at both PW endpoints and that unique values are assigned to the PWs in the same redundant set to achieve tiebreaking using this mechanism.

There are scenarios with dual-homing of a CE to PE nodes where each PE node needs to advertise active Preferential Forwarding status on more than one PW in the redundant set. However, a PE MUST always select a single PW for forwarding using the above active PW selection algorithm. An example of such a case is described in 15.2.

There are scenarios where each PE needs to advertise active Preferential Forwarding status on a single PW in the redundant set. In order to ensure that both PE nodes make the same selection, they MUST use the above active PW selection algorithm to determine the PW eligible for active state. An example of such a case is described in 15.5.

In steady state with consistent configuration, a PE will always find an active PW. However, it is possible that such a PW is not found due to a misconfiguration. In the event that an active PW is not found, a management notification SHOULD be generated. If a management notification for failure to find an active PW was generated and an active PW is subsequently found, a management notification SHOULD be generated, so clearing the previous failure indication. Additionally, a PE MAY use the request switchover procedures described in Section 6.3 to have both PE nodes switch to a common PW.

There may also be transient conditions where endpoints do not share a common view of the active/standby state of the PWs. This could be caused by propagation delay of the Targeted Label Distribution Protocol (T-LDP) status messages between endpoints. In this case, the behavior of the receiving endpoint is outside the scope of this document.

Thus, in this mode of operation, the following definition of active and standby PW states apply:

- o Active State

A PW is considered to be in active state when the PW labels are exchanged between its two endpoints and the status bits exchanged between the endpoints indicate the PW is up and its Preferential Forwarding status is active at both endpoints. In this state user traffic can flow over the PW in both directions. As described in Section 5.1, the PE nodes MUST implement a common mechanism to select one PW for forwarding in case multiple PWs qualify for the active state.

- o Standby State

A PW is considered to be in standby state when the PW labels are exchanged between its two endpoints, but the Preferential Forwarding status bits exchanged indicate the PW Preferential Forwarding status is standby at one or both endpoints. In this state, the endpoints MUST NOT forward data traffic over the PW but MAY allow PW OAM packets, e.g., Virtual Connection Connectivity Verification (VCCV) packets [11], to be sent and received in order to test the liveness of standby PWs. The endpoints of the PW MAY also allow the forwarding of specific control plane packets of applications using the PW. The specification of applications and the allowed control plane packets are outside the scope of this document. If the PW is a spoke in H-VPLS, any Media Access Control (MAC) addresses learned via the PW SHOULD be flushed when it transitions to standby state, according to the procedures in RFC 4762 [3] and in [10].

5.2. Master/Slave Mode

One endpoint node of the redundant set of PWs is designated the master and is responsible for selecting which PW both endpoints must use to forward user traffic.

The master indicates the forwarding state in the PW Preferential Forwarding status bit. The other endpoint node, the slave, MUST follow the decision of the master node based on the received status bits. In other words, the Preferential Forwarding status bit sent by the master node indicates the actual forwarding state of the PW at the master node.

There is a single PE master PW endpoint node and one or many PE PW endpoint slave nodes. The assignment of master/slave roles to the PW endpoints is performed by local configuration. Note that the behavior described in this section assumes correct configuration of

the master and slave endpoints. This document does not define a mechanism to detect errors in the configuration, and misconfiguration might lead to protection switchover failing to work correctly. Furthermore, this document does not specify the procedures for a backup master node. In deployments where PE node protection is required, it is recommended to use the independent mode of operation as in the application described in Section 15.2.

One endpoint of the PW, the master, actively selects which PW to activate and uses it for forwarding user traffic. This status is indicated to the slave node by setting the Preferential Forwarding status bit in the status bit TLV to active. It does not forward user traffic to any other of the PW's in the redundant set to the slave node and indicates this by setting the Preferential Forwarding status bit in the status bit TLV to standby for those PWs. The master node MUST ignore any PW Preferential Forwarding status bits received from the slave nodes.

If more than one PW qualifies for the active state, the master PW endpoint node selects one. There is no requirement to specify a default active PW selection mechanism in this case; however, for consistency across implementations, the master PW endpoint SHOULD implement the default active PW selection mechanism described in Section 5.1.

If the master PW endpoint implements the active PW selection mechanism based on primary/secondary and precedence parameters, it MUST comply with the following behavior:

1. If the PW endpoint is configured with the precedence parameter on each PW in the redundant set, it MUST select the PW with the lowest configured precedence value.
2. If the PW endpoint is configured with one PW as primary and one or more PWs as secondary, it MUST select the primary PW in preference to all secondary PWs. If a primary PW is not available, it MUST use the secondary PW with the lowest precedence value. If the primary PW becomes available, a PW endpoint MUST revert to it immediately or after the expiration of a configurable delay.

The slave endpoint(s) are required to act on the status bits received from the master. When the received status bit transitions from active to standby, a slave node MUST stop forwarding over the previously active PW. When the received status bit transitions from standby to active for a given PW, the slave node MUST start forwarding user traffic over this PW.

In this mode of operation, the following definition of active and standby PW states apply:

- o Active State

A PW is considered to be in active state when the PW labels are exchanged between its two endpoints, and the status bits exchanged between the endpoints indicate the PW is up at both endpoints, and the Preferential Forwarding status at the master endpoint is active. In this state, user traffic can flow over the PW in both directions.

- o Standby State

A PW is considered to be in standby state when the PW labels are exchanged between its two endpoints, and the status bits exchanged between the endpoints indicate the Preferential Forwarding status at the master endpoint is standby. In this state, the endpoints **MUST NOT** forward data traffic over the PW but **MAY** allow PW OAM packets, e.g., VCCV, to be sent and received. The endpoints of the PW **MAY** also allow the forwarding of specific control plane packets of applications using the PW. The specification of applications and the allowed control plane packets are outside the scope of this document. If the PW is a spoke in H-VPLS, any MAC addresses learned via the PW **SHOULD** be flushed when it transitions to standby state according to the procedures in RFC 4762 [3] and [10].

6. PW State Transition Signaling Procedures

This section describes the extensions to PW status signaling and the processing rules for these extensions. It defines a new PW Preferential Forwarding status bit that is to be used with the PW Status TLV specified in RFC 4447 [2].

The PW Preferential Forwarding bit, when set, is used to signal either the preferred or actual active/standby forwarding state of the PW by one PE to the far-end PE. The actual semantics of the value being signaled vary according to whether the PW is acting in master/slave or independent mode.

6.1. PW Standby Notification Procedures in Independent Mode

PEs that contain PW endpoints independently select which PW they intend to use for forwarding, depending on the specific application (example applications are described in [8]). They advertise the corresponding preferred active/standby forwarding state for each PW. An active Preferential Forwarding state is indicated by clearing the PW Preferential Forwarding status bit in the PW Status TLV. A standby Preferential Forwarding state is indicated by setting the PW

Preferential Forwarding status bit in the PW Status TLV. This advertisement occurs in both the initial label mapping message and in a subsequent notification message when the forwarding state transitions as a result of a state change in the specific application.

Each PW endpoint compares the updated local and remote status and effectively activates the PW, which is up at both endpoints and which shows both local active and remote active Preferential Forwarding states. The PE nodes MUST implement a common mechanism to select one PW for forwarding in case multiple PWs qualify for the active state, as explained in Section 5.1.

When a PW is in active state, the PEs can forward user packets, OAM packets, and other control plane packets over the PW.

When a PW is in standby state, the PEs MUST NOT forward user packets over the PW but MAY forward PW OAM packets and specific control plane packets.

For MS-PWs, S-PEs MUST relay the PW status notification containing both the existing status bits and the new Preferential Forwarding status bits between ingress and egress PWs as per the procedures defined in [4].

6.2. PW Standby Notification Procedures in Master/Slave Mode

Whenever the master PW endpoint selects or deselects a PW for forwarding user traffic at its end, it explicitly notifies the event to the remote slave endpoint. The slave endpoint carries out the corresponding action on receiving the PW state change notification.

If the PW Preferential Forwarding bit in PW Status TLV received by the slave is set, it indicates that the PW at the master end is not used for forwarding and is thus kept in the standby state. The PW MUST NOT be used for forwarding at slave endpoint. Clearing the PW Preferential Forwarding bit in PW Status TLV indicates that the PW at the master endpoint is used for forwarding and is in active state, and the receiving slave endpoint MUST activate the PW if it was previously not used for forwarding.

When this mechanism is used, a common Group ID in the PWid FEC element or a PW Grouping ID TLV in the Generalized PWid FEC element, as defined in [2], MAY be used to signal PWs in groups in order to minimize the number of LDP status messages that MUST be sent. When PWs are provisioned with such grouping, a termination point sends a single "wildcard" notification message to denote this change in status for all affected PWs. This status message contains either the

PWid FEC TLV with only the Group ID or the Generalized PWid FEC TLV with only the PW Grouping ID TLV. As mentioned in [2], the Group ID field of the PWid FEC element, or the PW Grouping ID TLV in the Generalized PWid FEC element, can be used to send status notification for an arbitrary set of PWs.

For MS-PWs, S-PEs MUST relay the PW status notification containing both the existing and the new Preferential Forwarding status bits between ingress and egress PW segments, as per the procedures defined in [4].

6.2.1. PW State Machine

It is convenient to describe the PW state change behavior in terms of a state machine (Table 1). The PW state machine is explained in detail in the two defined states, and the behavior is presented as a state transition table. The same state machine is applicable to PW groups.

STATE	EVENT	NEW STATE
ACTIVE	PW put in standby (master) Action: Transmit PW Preferential Forwarding bit set	STANDBY
	Receive PW Preferential Forwarding bit set (slave) Action: Stop forwarding over PW	STANDBY
	Receive PW Preferential Forwarding bit set but bit not supported Action: None	ACTIVE
	Receive PW Preferential Forwarding bit clear Action: None.	ACTIVE
STANDBY	PW activated (master) Action: Transmit PW Preferential Forwarding bit clear	ACTIVE
	Receive PW Preferential Forwarding bit clear (slave) Action: Activate PW	ACTIVE
	Receive PW Preferential Forwarding bit clear but bit not supported Action: None	STANDBY
	Receive PW Preferential Forwarding bit set Action: None	STANDBY

Table 1. PW State Transition Table in Master/Slave Mode

6.3. Coordination of PW Switchover

There are PW redundancy applications that require that PE nodes coordinate the switchover to a PW such that both endpoints will forward over the same PW at any given time. One such application for redundant MS-PW is identified in [8]. Multiple MS-PWs are configured between a pair of T-PE nodes. The paths of these MS-PWs are diverse and are switched at different S-PE nodes. Only one of these MS-PWs is active at any given time. The others are put in standby. The endpoints follow the independent mode procedures to use the PW, which is both up and for which both endpoints advertise an active Preferential Forwarding status bit.

The trigger for sending a request to switchover by one endpoint of the MS-PW can be an operational event. For example, a failure that causes the endpoints to be unable to find a common PW for which both endpoints advertise an active Preferential Forwarding status bit. The other trigger is the execution of an administrative maintenance operation by the network operator in order to move the traffic away from the nodes or links currently used by the active PW.

Unlike the case of a master/slave mode of operation, the endpoint requesting the switchover requires explicit acknowledgment from the peer endpoint that the request can be honored before it switches to another PW. Furthermore, any of the endpoints can make the request to switch over.

This document specifies a second status bit that is used by a PE to request that its peer PE switch over to use a different active PW. This bit is referred to as the Request Switchover status bit. The Preferential Forwarding status bit continues to be used by each endpoint to indicate its current local settings of the active/standby state of each PW in the redundant set. In other words, as in the independent mode, it indicates to the far-end which of the PWs is being used to forward packets and which is being put in standby. It can thus be used as a way for the far-end to acknowledge the requested switchover operation.

A PE MAY support the Request Switchover bit. A PE that receives the Request Switchover bit and that does not support it will ignore it.

If the Request Switchover bit is supported by both sending and receiving PEs, the following procedures MUST be followed by both endpoints of a PW to coordinate the switchover of the PW.

S-PEs nodes MUST relay the PW status notification containing the existing status bits, as well as the new Preferential Forwarding and Request Switchover status bits between ingress and egress PW segments as per the procedures defined in [4].

6.3.1. Procedures at the Requesting Endpoint

- a. The requesting endpoint sends a Status TLV in the LDP notification message with the Request Switchover bit set on the PW to which it desires to switch.
- b. The endpoint does not activate, forwarding on that PW at this point in time. It MAY, however, enable receiving on that PW. Thus, the Preferential Forwarding status bit still reflects the currently used PW.

- c. The requesting endpoint starts a timer while waiting for the remote endpoint to acknowledge the request. This timer SHOULD be configurable with a default value of 3 seconds.
- d. If, while waiting for the acknowledgment, the requesting endpoint receives a request from its peer to switch over to the same or a different PW, it MUST perform the following:
 - i. If its address is higher than that of the peer, this endpoint ignores the request and continues to wait for the acknowledgment from its peer.
 - ii. If its system IP address is lower than that of its peer, it aborts the timer and immediately starts the procedures of the receiving endpoint in Section 6.3.2.
- e. If, while waiting for the acknowledgment, the requesting endpoint receives a status notification message from its peer with the Preferential Forwarding status bit cleared in the requested PW, it MUST treat this as an explicit acknowledgment of the request and MUST perform the following:
 - i. Abort the timer.
 - ii. Activate the PW.
 - iii. Send an update status notification message with the Preferential Forwarding status bit and the Request Switchover bit clear on the newly active PW and send an update status notification message with the Preferential Forwarding status bit set in the previously active PW.
- f. If, while waiting for the acknowledgment, the requesting endpoint detects that the requested PW went into down state locally, and could use an alternate PW that is up, it MUST perform the following:
 - i. Abort the timer.
 - ii. Issue a new request to switchover to the alternate PW.
 - iii. Restart the timer.

- g. If, while waiting for the acknowledgment, the requesting endpoint detects that the requested PW went into the down state locally, and could not use an alternate PW that is up, it MUST perform the following:
 - i. Abort the timer.
 - ii. Send an update status notification message with the Preferential Forwarding status bit unchanged and the Request Switchover bit reset for the requested PW.
- h. If, while waiting for the acknowledgment, the timer expires, the requesting endpoint MUST assume that the request was rejected and MAY issue a new request.
- i. If the requesting node receives the acknowledgment after the request expired, it will treat it as if the remote endpoint unilaterally switched between the PWs without issuing a request. In that case, it MAY issue a new request and follow the requesting endpoint procedures to synchronize which PW to use for the transmit and receive directions of the emulated service.

6.3.2. Procedures at the Receiving Endpoint

- a. Upon receiving a status notification message with the Request Switchover bit set on a PW different from the currently active one, and the requested PW is up, the receiving endpoint MUST perform the following:
 - i. Activate the PW.
 - ii. Send an update status notification message with the Preferential Forwarding status bit clear and the Request Switchover bit reset on the newly active PW , and send an update status notification message with the Preferential Forwarding status bit set in the previously active PW.
 - iii. Upon receiving a status notification message with the Request Switchover bit set on a PW, which is different from the currently active PW but is down, the receiving endpoint MUST ignore the request.

7. Status Mapping

The generation and processing of the PW Status TLV MUST follow the procedures in RFC 4447 [2]. The PW Status TLV is sent on the active PW and standby PWs to make sure the remote AC and PW states are always known to the local PE node.

The generation and processing of PW Status TLV by an S-PE node in a MS-PW MUST follow the procedures in [4].

The procedures for determining and mapping PW and AC states MUST follow the rules in [5] with the following modifications.

7.1. AC Defect State Entry/Exit

A PE enters the AC receive (or transmit) defect state for a PW service when one or more of the conditions specified for this PW service in [5] are met.

When a PE enters the AC receive (or transmit) defect state for a PW, it MUST send a forward (reverse) defect indication to the remote peers over all PWs in the redundant set that are associated with this AC.

When a PE exits the AC receive (or transmit) defect state for a PW service, it MUST clear the forward (or reverse) defect indication to the remote peers over all PWs in the redundant set that are associated with this AC.

7.2. PW Defect State Entry/Exit

A PE enters the PW receive (or transmit) defect state for a PW service when one or more of the conditions specified in Section 8.3.1 (Section 8.3.2) in [5] are met for each of the PWs in the redundant set.

When a PE enters the PW receive (or transmit) defect state for a PW service associated with an AC, it MUST send a reverse (or forward) defect indication over one or more of the PWs in the redundant set associated with the same AC if the PW failure was detected by this PE without receiving a forward defect indication from the remote PE [5].

When a PE exits the PW receive (or transmit) defect state for a PW, it MUST clear the reverse (or forward) defect indication over any PW in the redundant associated with the same AC set if applicable.

8. Applicability and Backward Compatibility

The mechanisms defined in this document are to be used in applications where standby state signaling of a PW or PW group is required. Both PWid FEC and Generalized PWid FEC are supported. All PWs that are part of a redundant set MUST use the same FEC type. When the set uses the PWid FEC element, each PW is uniquely identified by its PW ID. When the redundant set uses the Generalized

PWid FEC element, each PW MUST have a unique identifier that consists of the triplet AGI::SAII::TAII.

A PE implementation that uses the mechanisms described in this document MUST negotiate the use of PW Status TLV between its T-LDP peers, as per RFC 4447 [2]. If the PW Status TLV is found to be not supported by either of its endpoints after status negotiation procedures, then the mechanisms specified in this document cannot be used.

A PE implementation that is compliant with RFC 4447 [2] and that does not support the generation or processing of the Preferential Forwarding status bit or of the Request Switchover status bit MUST ignore these status bits if they are set by a peer PE. This document in fact updates RFC 4447 by prescribing the same behavior for any status bit not originally defined in RFC 4447.

Finally, this document updates RFC 4447 by defining that a status bit can indicate a status other than a fault or can indicate an instruction to the peer PE. As a result, a PE implementation compliant to RFC 4447 MUST process each status bit it supports when set according to the rules specific to that status bit.

9. Security Considerations

LDP extensions/options that protect PWs must be implemented because the status bits defined in this document have the same security considerations as the PW setup and maintenance protocol defined in RFC 4447 [2]. It should be noted that the security of a PW redundant set is only as good as the weakest security on any of its members.

10. MIB Considerations

New MIB objects for the support of PW redundancy will be defined in a separate document.

11. IANA Considerations

This document defines the following PW status codes for the PW redundancy application. IANA has allocated these from the "Pseudowire Status Codes Registry".

11.1. Status Code for PW Preferential Forwarding Status

0x00000020 When the bit is set, it indicates PW forwarding standby".

When the bit is cleared, it indicates PW forwarding active".

11.2. Status Code for PW Request Switchover Status

0x00000040 When the bit is set, it represents Request Switchover to this PW.

When the bit is cleared, it represents no specific action.

12. Contributors

The editors would like to thank Matthew Bocci, Pranjali Kumar Dutta, Giles Heron, Marc Lasserre, Luca Martini, Thomas Nadeau, Jonathan Newton, Hamid Ould-Brahim, Olen Stokes, and Daniel Cohn who made a contribution to the development of this document.

Matthew Bocci
Alcatel-Lucent
EMail: matthew.bocci@alcatel-lucent.com

Pranjali Kumar Dutta
Alcatel-Lucent
EMail: pranjali.dutta@alcatel-lucent.com

Giles Heron
Cisco Systems, Inc.
giles.heron@gmail.com

Marc Lasserre
Alcatel-Lucent
EMail: marc.lasserre@alcatel-lucent.com

Luca Martini
Cisco Systems, Inc.
EMail: lmartini@cisco.com

Thomas Nadeau
Juniper Networks
EMail: tnadeau@lucidvision.com

Jonathan Newton
Cable & Wireless Worldwide
EMail: Jonathan.Newton@cw.com

Hamid Ould-Brahim
EMail: ouldh@yahoo.com

Olen Stokes
Extreme Networks
EMail: ostokes@extremenetworks.com

Daniel Cohn
Orckit
daniel.cohn.ietf@gmail.com.

13. Acknowledgments

The authors would like to thank the following individuals for their valuable comments and suggestions, which improved the document both technically and editorially:

Vach Kompella, Kendall Harvey, Tiberiu Grigoriu, John Rigby, Prashanth Ishwar, Neil Hart, Kajal Saha, Florin Balus, Philippe Niger, Dave McDysan, Roman Krzanowski, Italo Busi, Robert Rennison, and Nicolai Leymann.

14. References

14.1. Normative References

- [1] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [2] 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.
- [3] Lasserre, M., Ed., and V. Kompella, Ed., "Virtual Private LAN Service (VPLS) Using Label Distribution Protocol (LDP) Signaling", RFC 4762, January 2007.
- [4] Martini, L., Metz, C., Nadeau, T., Bocci, M., and M. Aissaoui, "Segmented Pseudowire", RFC 6073, January 2011.
- [5] Aissaoui, M., Busschbach, P., Martini, L., Morrow, M., Nadeau, T., and Y(J). Stein, "Pseudowire (PW) Operations, Administration, and Maintenance (OAM) Message Mapping", RFC 6310, July 2011.

14.2. Informative References

- [6] Bocci, M. and S. Bryant, "An Architecture for Multi-Segment Pseudowire Emulation Edge-to-Edge", RFC 5659, October 2009.
- [7] Martini, L., "IANA Allocations for Pseudowire Edge to Edge Emulation (PWE3)", BCP 116, RFC 4446, April 2006.
- [8] Muley, P., Aissaoui, M., and M. Bocci, "Pseudowire Redundancy", RFC 6718, August 2012.

- [9] Bryant, S., Ed., and P. Pate, Ed., "Pseudo Wire Emulation Edge-to-Edge (PWE3) Architecture", RFC 3985, March 2005.
- [10] Dutta, P., Balus, F., Calvignac, G., and O. Stokes "LDP Extensions for Optimized MAC Address Withdrawal in H-VPLS", Work in Progress, October 2011.
- [11] Nadeau, T., Ed., and C. Pignataro, Ed., "Pseudowire Virtual Circuit Connectivity Verification (VCCV): A Control Channel for Pseudowires", RFC 5085, December 2007.

Appendix A. Applications of PW Redundancy Procedures

This section shows how the mechanisms described in this document are used to achieve the desired protection behavior for some of the applications described in "PW Redundancy" [8].

A.1. One Multi-Homed CE with Single SS-PW Redundancy

The following figure illustrates an application of SS-PW redundancy.

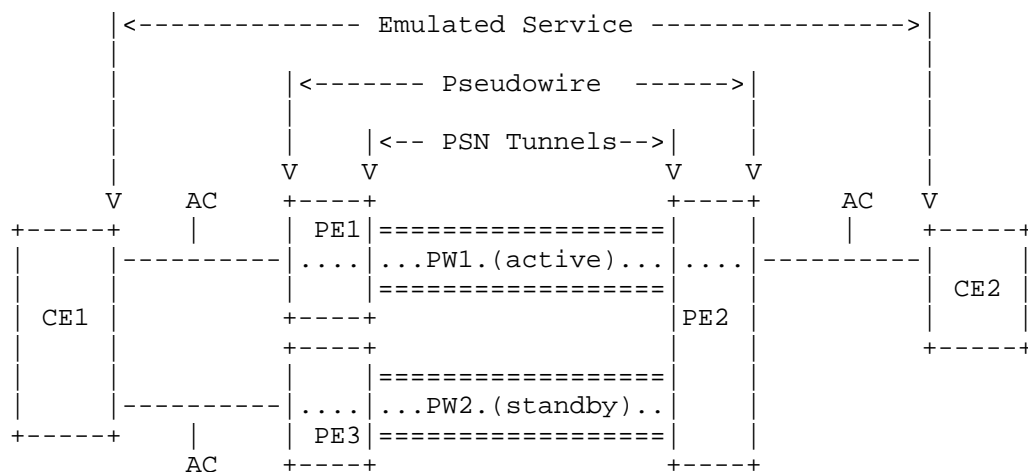


Figure 2. Multi-Homed CE with SS-PW Redundancy

The application in Figure 2 makes use of the independent mode of operation.

CE1 is dual-homed to PE1 and to PE3 by attachment circuits. The method for dual-homing of CE1 to PE1 and to PE3 nodes and the protocols used are outside the scope of this document (see [8]).

In this example, the AC from CE1 to PE1 is active, while the AC from CE1 to PE3 is standby, as determined by the redundancy protocol running on the ACs. Thus, in normal operation, PE1 and PE3 will advertise an active and standby Preferential Forwarding status bit, respectively, to PE2, reflecting the forwarding state of the two ACs to CE1 as determined by the AC dual-homing protocol. PE2 advertises a Preferential Forwarding status bit of active on both PW1 and PW2, since the AC to CE2 is single-homed. As both the local and remote UP/DOWN status and Preferential Forwarding status for PW1 are up and active, traffic is forwarded over PW1 in both directions.

On failure of the AC between CE1 and PE1, the forwarding state of the AC on PE3 transitions to active. PE3 then announces the newly changed Preferential Forwarding status bit of active to PE2. PE1 will advertise a PW status notification message, indicating that the AC between CE1 and PE1 is down. PE2 matches the local and remote Preferential Forwarding status of active and status of "Pseudowire forwarding" and selects PW2 as the new active PW to which to send traffic.

On failure of the PE1 node, PE3 will detect it and will transition the forwarding state of its AC to active. The method by which PE3 detects that PE1 is down is outside the scope of this document. PE3 then announces the newly changed Preferential Forwarding status bit of active to PE2. PE3 and PE2 match the local and remote Preferential Forwarding status of active and UP/DOWN status "Pseudowire forwarding" and select PW2 as the new active PW to which to send traffic. Note that PE2 may have detected that the PW to PE1 went down via T-LDP Hello timeout or via other means. However, it will not be able to forward user traffic until it receives the updated status bit from PE3.

Note that, in this example, the receipt of the AC status on the CE1-PE1 link is normally sufficient for PE2 to switch to PW2. However, the operator may want to trigger the switchover of the PW for administrative reasons, e.g., maintenance; thus, the use of the Preferential Forwarding status bit is required to notify PE2 to trigger the switchover.

Note that the primary/secondary procedures do not apply in this case as the PW Preferential Forwarding status is driven by the AC forwarding state, as determined by the AC dual-homing protocol used.

A.2. Multiple Multi-Homed CEs with SS-PW Redundancy

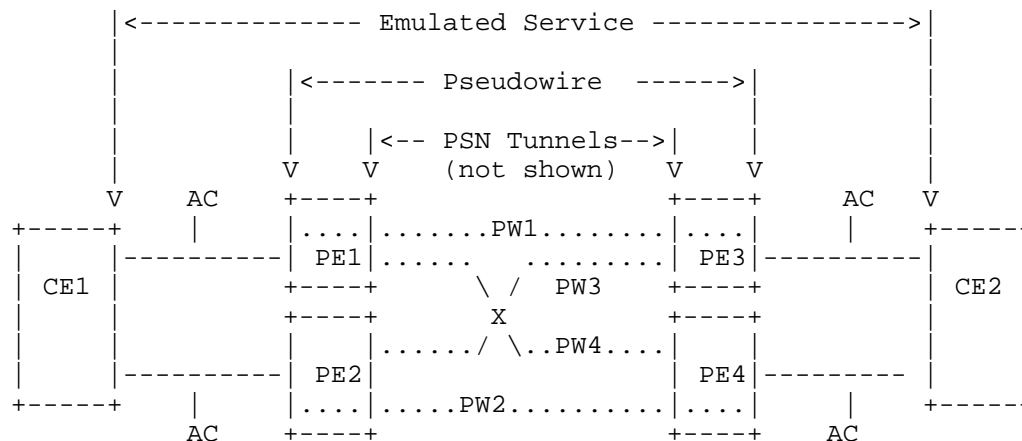


Figure 3. Multiple Multi-Homed CEs with SS-PW Redundancy

The application in Figure 3 makes use of the independent mode of operation.

CE1 is dual-homed to PE1 and PE2. CE2 is dual-homed to PE3 and PE4. The method for dual-homing and the used protocols are outside the scope of this document. Note that the PSN tunnels are not shown in this figure for clarity. However, it can be assumed that each of the PWs shown is encapsulated in a separate PSN tunnel.

Assume that the AC from CE1 to PE1 is active and from CE1 to PE2 it is standby; furthermore, assume that the AC from CE2 to PE3 is standby and from CE2 to PE4 it is active. The method of deriving the active/standby status of the AC is outside the scope of this document.

PE1 advertises the Preferential Forwarding status active and UP/DOWN status "Pseudowire forwarding" for pseudowires PW1 and PW4 connected to PE3 and PE4. This status reflects the forwarding state of the AC attached to PE1. PE2 advertises Preferential Forwarding status standby and UP/DOWN status "Pseudowire forwarding" for pseudowires PW2 and PW3 to PE3 and PE4. PE3 advertises Preferential Forwarding status standby and UP/DOWN status "Pseudowire forwarding" for pseudowires PW1 and PW3 to PE1 and PE2. PE4 advertises the Preferential Forwarding status active and UP/DOWN status "Pseudowire forwarding" for pseudowires PW2 and PW4 to PE2 and PE1, respectively. Thus, by matching the local and remote Preferential Forwarding status of active and UP/DOWN status of

"Pseudowire forwarding" of pseudowires, the PE nodes determine which PW should be in the active state. In this case, it is PW4 that will be selected.

On failure of the AC between CE1 and PE1, the forwarding state of the AC on PE2 is changed to active. PE2 then announces the newly changed Preferential Forwarding status bit of active to PE3 and PE4. PE1 will advertise a PW status notification message, indicating that the AC between CE1 and PE1 is down. PE2 and PE4 match the local and remote Preferential Forwarding status of active and UP/DOWN status "Pseudowire forwarding" and select PW2 as the new active PW to which to send traffic.

On failure of the PE1 node, PE2 will detect the failure and will transition the forwarding state of its AC to active. The method by which PE2 detects that PE1 is down is outside the scope of this document. PE2 then announces the newly changed Preferential Forwarding status bit of active to PE3 and PE4. PE2 and PE4 match the local and remote Preferential Forwarding status of active and UP/DOWN status "Pseudowire forwarding" and select PW2 as the new active PW to which to send traffic. Note that PE3 and PE4 may have detected that the PW to PE1 went down via T-LDP Hello timeout or via other means. However, they will not be able to forward user traffic until they have received the updated status bit from PE2.

Because each dual-homing algorithm running on the two node sets, i.e., {CE1, PE1, PE2} and {CE2, PE3, PE4}, selects the active AC independently, there is a need to signal the active status of the AC such that the PE nodes can select a common active PW for end-to-end forwarding between CE1 and CE2 as per the procedures in the independent mode.

Note that no primary/secondary procedures, as defined in Sections 5.1 and 5.2, apply in this use case as the active/standby status is driven by the AC forwarding state, as determined by the AC dual-homing protocol used.

A.3. Multi-Homed CE with MS-PW Redundancy

The following figure illustrates an application of MS-PW redundancy.

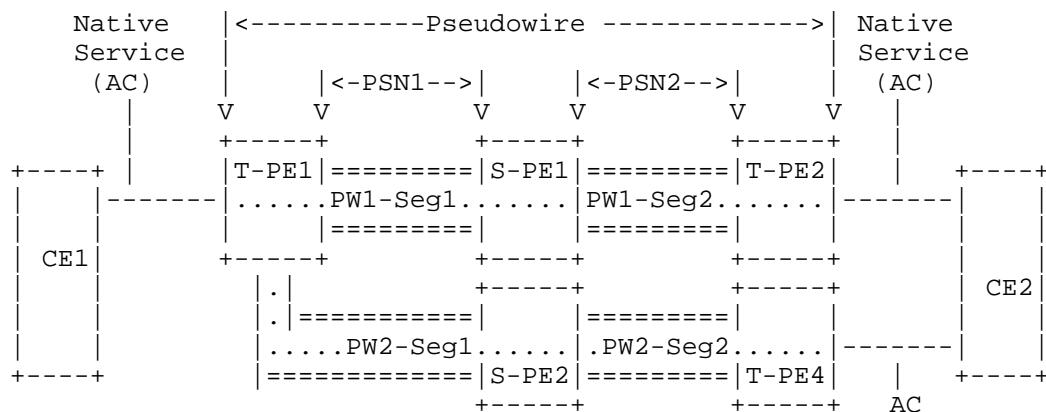


Figure 4. Multi-Homed CE with MS-PW Redundancy

The application in Figure 4 makes use of the independent mode of operation. It extends the application described in Section 15.1. 15.1 of this document and in [8] by adding a pair of S-PE nodes to switch the segments of PW1 and PW2.

CE2 is dual-homed to T-PE2 and T-PE4. PW1 and PW2 are used to extend the resilient connectivity all the way to T-PE1. PW1 has two segments and is an active pseudowire, while PW2 has two segments and is a standby pseudowire. This application requires support for MS-PW with segments of the same type as described in [4].

The operation in this case is the same as in the case of SS-PW, as described in Section 15.1. The only difference is that the S-PE nodes need to relay the PW status notification containing both the UP/DOWN and forwarding status to the T-PE nodes.

A.4. Multi-Homed CE with MS-PW Redundancy and S-PE Protection

The following figure illustrates an application of MS-PW redundancy with 1:1 PW protection.

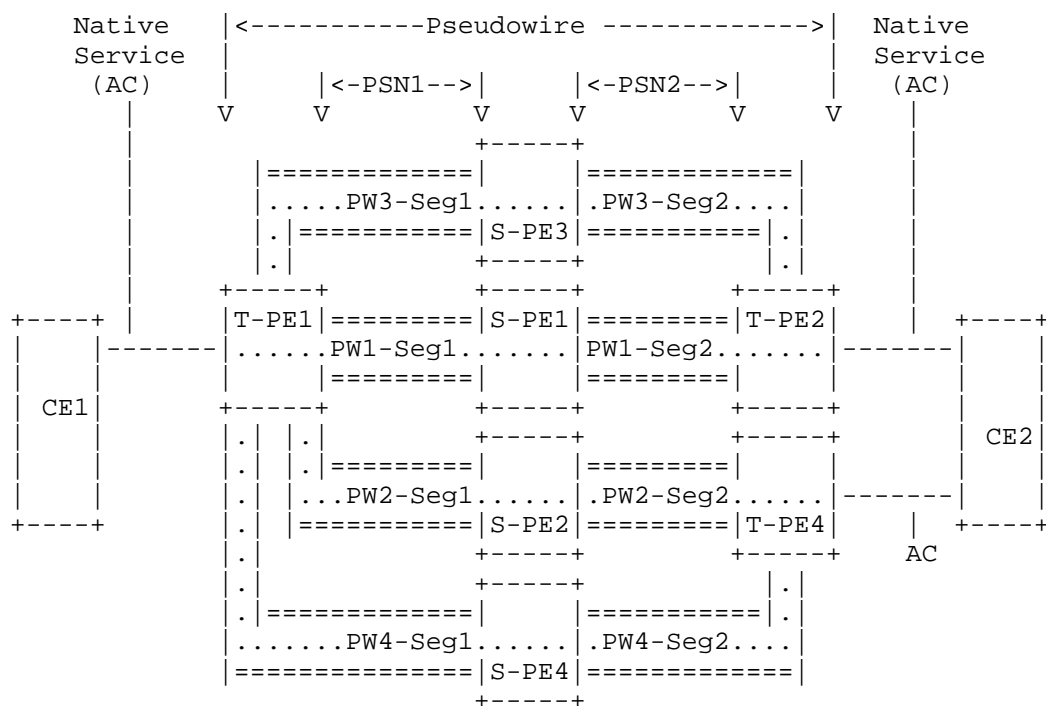


Figure 5. Multi-Homed CE with MS-PW Redundancy and Protection

The application in Figure 5 makes use of the independent mode of operation.

CE2 is dual-homed to T-PE2 and T-PE4. The PW pairs {PW1,PW3} and {PW2,PW4} are used to extend the resilient connectivity all the way to T-PE1, like in the case in Section 15.3, with the addition that this setup provides for S-PE node protection.

CE1 is connected to T-PE1 while CE2 is dual-homed to T-PE2 and T-PE4. There are four segmented PWs. PW1 and PW2 are primary PWs and are used to support CE2 multi-homing. PW3 and PW4 are secondary PWs and are used to support 1:1 PW protection. PW1, PW2, PW3, and PW4 have two segments and they are switched at S-PE1, S-PE2, S-PE3, and S-PE4, respectively.

It is possible that S-PE1 coincides with S-PE4 and/or SP-2 coincides with S-PE3, in particular, where the two PSN domains are interconnected via two nodes. However, Figure 5 shows four separate S-PE nodes for clarity.

The behavior of this setup is exactly the same as the setup in Section 15.3 except that T-PE1 will always see a pair of PWs eligible for the active state, for example, the pair {PW1,PW3} when the AC between CE2 and T-PE2 is in active state. Thus, it is important that both T-PE1 and T-PE2 implement a common mechanism to choose one the two PWs for forwarding, as explained in Section 5.1. Similarly, T-PE1 and T-PE4 must use the same mechanism to select among the pair {PW2,PW4} when the AC between CE2 and T-PE4 is in active state.

A.5. Single-Homed CE with MS-PW Redundancy

The following is an application of the independent mode of operation, along with the request switchover procedures in order to provide N:1 PW protection. A revertive behavior to a primary PW is shown as an example of configuring and using the primary/secondary procedures described in Sections 5.1. and 5.2.

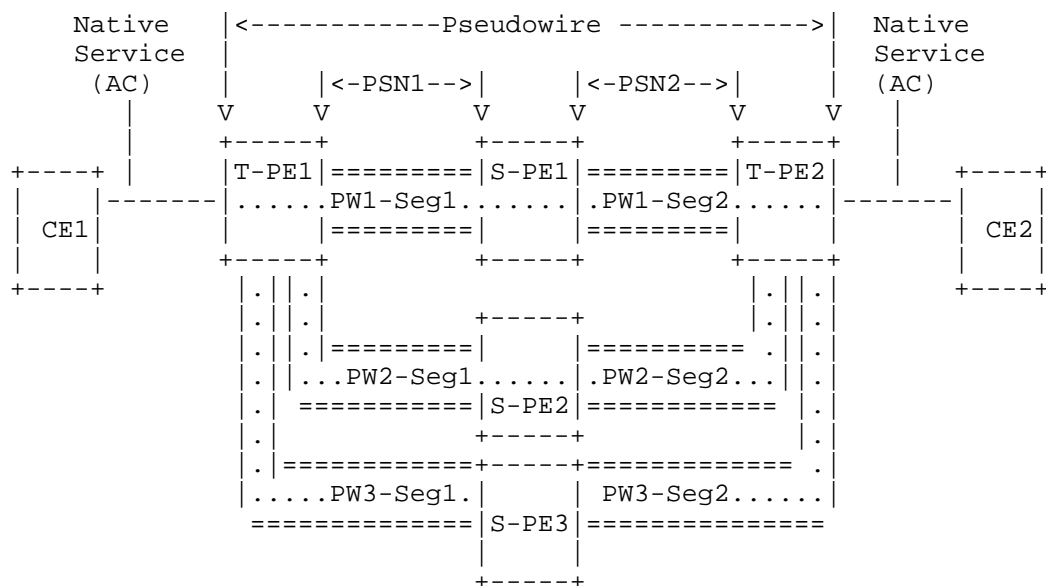


Figure 6. Single-Homed CE with MS-PW Redundancy

CE1 is connected to PE1 in provider edge 1 and CE2 to PE2 in provider edge 2, respectively. There are three segmented PWs: a primary PW, PW1, is switched at S-PE1 and has the lowest precedence value of

zero; a secondary PW, PW2, which is switched at S-PE2 and has a precedence of 1; and another secondary PW, PW3, which is switched at S-PE3 and has a precedence of 2.

The precedence is locally configured at the endpoints of the PW, i.e., T-PE1 and T-PE2. The lower the precedence value, the higher the priority.

T-PE1 and T-PE2 will select the PW they intend to activate based on their local and remote UP/DOWN state, as well as the local precedence configuration. In this case, they will both advertise Preferential Forwarding status bit of active on PW1 and of standby on PW2 and PW3 using priority derived from local precedence configuration. Assuming all PWs are up, T-PE1 and T-PE2 will use PW1 to forward user packets.

If PW1 fails, then the T-PE detecting the failure will send a status notification to the remote T-PE with a Local PSN-facing PW (ingress) Receive Fault bit set, a Local PSN-facing PW (egress) Transmit Fault bit set, or a Pseudowire Not Forwarding bit set. In addition, it will set the Preferential Forwarding status bit on PW1 to standby. It will also advertise the Preferential Forwarding status bit on PW2 as active, as it has the next-lowest precedence value. T-PE2 will also perform the same steps as soon as it is informed of the failure of PW1. Both T-PE nodes will perform a match on the Preferential Forwarding status of active and UP/DOWN status of "Pseudowire forwarding" and will use PW2 to forward user packets.

However, this does not guarantee that the T-PEs will choose the same PW from the redundant set to forward on, for a given emulated service, at all times. This may be due to a mismatch of the configuration of the PW precedence in each T-PE. This may also be due to a failure that caused the endpoints to not be able to match the active Preferential Forwarding status bit and UP/DOWN status bits. In this case, T-PE1 and/or T-PE2 can invoke the request switchover/acknowledgment procedures to synchronize the choice of PW to forward on in both directions.

The trigger for sending a request to switch over can also be the execution of an administrative maintenance operation by the network operator in order to move the traffic away from the T-PE/S-PE nodes/links to be serviced.

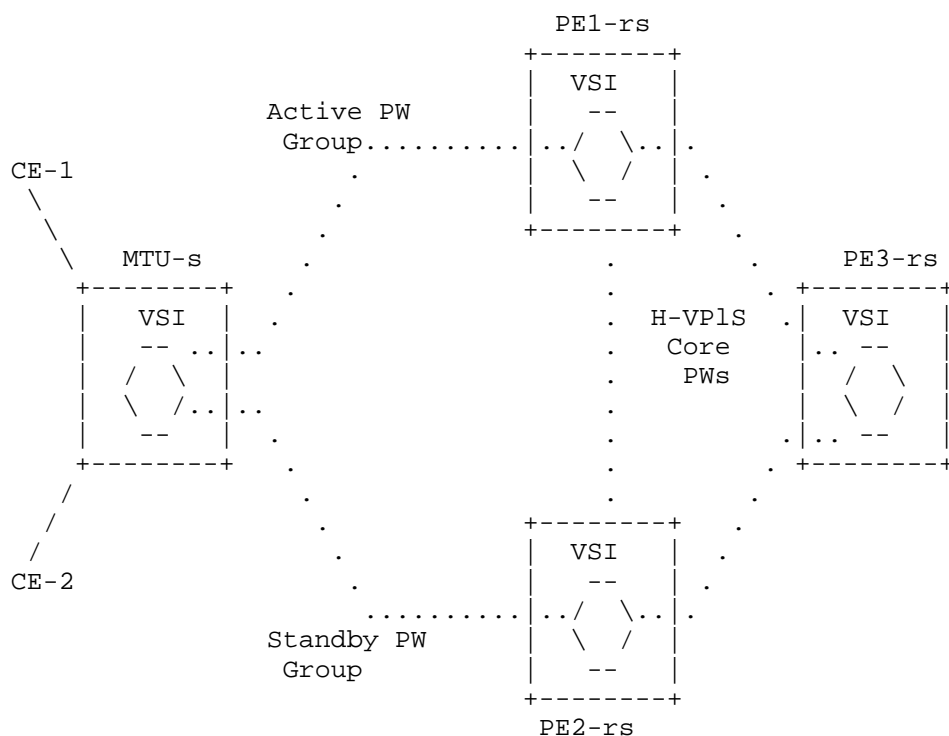
In case the Request Switchover is sent by both endpoints simultaneously, both T-PEs send status notification with the newly selected PW with Request Switchover bit set, waiting for a response from the other endpoint. In such a situation, the T-PE with greater

system address request is given precedence. This helps in synchronizing PWs in the event of mismatch of precedence configuration as well.

On recovery of the primary PW, PW1 is selected to forward traffic and the secondary PW, PW2, is set to standby.

A.6. PW Redundancy between H-VPLS MTU-s and PE-rs

The following figure illustrates the application of use of PW redundancy in H-VPLS for the purpose of dual-homing an MTU-s node to PE nodes using PW spokes. This application makes use of the master/slave mode of operation.



A.6. Multi-Homed MTU-s in H-VPLS Core

MTU-s is dual-homed to PE1-rs and PE2-rs. The primary spoke PWs from MTU-s are connected to PE1-rs, while the secondary PWs are connected to PE2. PE1-rs and PE2-rs are connected to H-VPLS core on the other side of the network. MTU-s communicates to PE1-rs and PE2-rs the forwarding status of its member PWs for a set of Virtual Switch Instances (VSIs) having common status active/standby. It may be

signaled using PW grouping with a common group-id in the Pwid FEC element or Grouping TLV in the Generalized Pwid FEC element, as defined in [2] to scale better. MTU-s derives the status of the PWs based on local policy configuration. In this example, the primary/secondary procedures as defined in Section 5.2 are used, but this can be based on any other policy.

Whenever MTU-s performs a switchover, it sends a wildcard notification message to PE2-rs for the previously standby PW group containing PW Status TLV with PW Preferential Forwarding bit cleared. On receiving the notification, PE-2rs unblocks all member PWs identified by the PW group and the state of the PW group changes from standby to active. All procedures described in Section 6.2 are applicable.

The use of the Preferential Forwarding status bit in master/slave mode is similar to Topology Change Notification in the IEEE Ethernet Bridges controlled by Rapid Spanning Tree Protocol (RSTP) but is restricted over a single hop. When these procedures are implemented, PE-rs devices are aware of switchovers at MTU-s and could generate MAC Withdraw messages to trigger MAC flushing within the H-VPLS full mesh. By default, MTU-s devices should still trigger MAC Withdraw messages, as currently defined in [3], to prevent two copies of MAC Withdraws being sent: one by MTU-s and another one by PE-rs nodes. Mechanisms to disable a MAC Withdraw trigger in certain devices is out of the scope of this document.

Authors' Addresses

Praveen Muley
Alcatel-lucent
701 E. Middlefield Road
Mountain View, CA, 94043, USA

EMail: praveen.muley@alcatel-lucent.com

Mustapha Aissaoui
Alcatel-lucent
600 March Rd
Kanata, ON, Canada K2K 2E6

EMail: mustapha.aissaoui@alcatel-lucent.com

