              Pseudowire Preferential Forwarding Status Bit
                   draft-ietf-pwe3-redundancy-bit-07.txt

Abstract

   This document describes a mechanism for standby status signaling 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 needed
   to indicate a preferential forwarding status 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 the PW operational status textual
   convention defined in RFC 5542 [9].

time. It is inappropriate to use Internet-Drafts as reference
material or to cite them other than as "work in progress."

This Internet-Draft will expire on August 1, 2012.

Copyright Notice

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].

Table of Contents

## 1. Introduction

   In Virtual Private Wire Services (VPWS) or Virtual Private Local Area
   network Services (VPLS) that use SS-PWs, protection for the PW is
   provided by the packet switched network (PSN) layer. This may be a
   Resource Reservation Protocol with Traffic Engineering (RSVP-TE)
   label switched path (LSP) with a fast reroute (FRR) backup or an end-
   to-end backup LSP. There are, however, applications where PSN
   protection is insufficient to fully protect the PW-based service.
   These include the following:

   In a VPWS service where the Customer Edge (CE) node is dual homed to
   a pair of PE nodes, PW redundancy mechanisms are required to ensure
   that the correct PW is used for forwarding when attachment circuit
   (AC) redundancy is used. PW redundancy mechanisms are also required
   when multiple redundant MS-PWs are used between T-PEs, to ensure that
   both T-PEs use the same MS-PW to forward to one another.

   In a hierarchical VPLS (H-VPLS) service, PW redundancy mechanisms are
   required to enable a multi-tenant unit switch (MTU-s) to be dual-
   homed to two PE-rs devices.

In these cases, pseudowire redundancy mechanisms are required. These
scenarios are described in the PW redundancy and framework document
[6].

Scenarios, such as those above, therefore rely on a set of two or
more pseudowires to protect a given VPWS or VPLS . Only one of these
pseudowires is used by the PEs to forward user traffic on at any
given time. This is the active PW. The other PWs in the set are
considered standby and are not used for forwarding unless they become
active. This provides a 1:1 or N:1 PW protection with the possibility
of multi-homing between the CE and the PEs.

In order to support AC or spoke PW redundancy, at least one of the
PEs on which a PW terminates MUST be different from that on which the
primary PW terminates, as described in [6]. Figure 1-1 illustrates an
application of active and standby PWs.

```
        |<-------------- Emulated Service ---------------->|
         |                                                |
         |           |<------- Pseudowire ------->|        |
         |           |                           |        |
         |           |    |<-- PSN Tunnels-->|    |        |
         |           V    V                  V    V        |
         V    AC    +----+                  +----+   AC   V
   +-----+    |    | PE1|================|    |    |   +-----+
   |     |----------|....|...PW1.(active)...|....|----------|     |
   |     |    |    |    |================|    |        | CE2 |
   | CE1 |         +----+                  |PE2 |        |    |
   |     |    |    +----+                  |    |        +-----+
   |     |    |    |    |================|    |
   |     |----------|....|...PW2.(standby)..|    |
   +-----+    |    | PE3|================|    |
           AC    +----+                  +----+
```
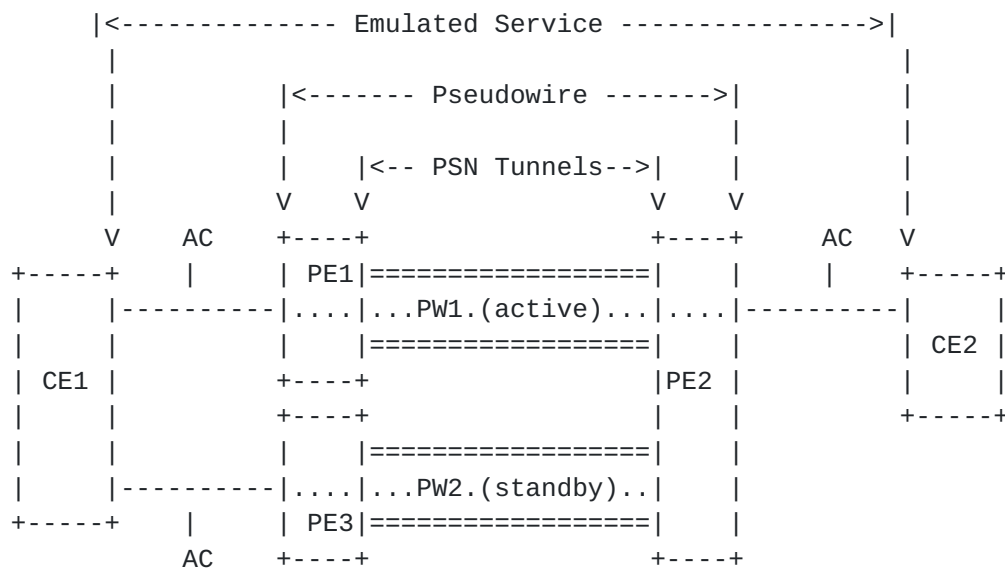
Figure 1-1: Reference Model for PW Redundancy

In MS-PW applications, PW redundancy is also required to protect the
service against failures of the switching PEs, which cannot be
protected by PSN mechanisms. In addition, PW redundancy is also
required if CEs are dual-homed to the PEs, as described above. In
this case, multiple MS-PWs are configured between a pair of T-PE

nodes, as described in Figure 2 of [6]. The paths of these MS-PWs are diverse in that they are switched at different Switching PE (S-PE) nodes. Only one of these MS-PWs is active at any given time, while the others are standby.

This document specifies a new PW status bit to indicate the preferential forwarding status of the PW for the purpose of notifying the remote PE of the preferential forwarding state of each PW in the redundancy set i.e. Active or Standby. This status bit is different from the PW status bits already defined in the PWE3 control protocol [2]. In addition, a second status bit is defined to allow peer PE nodes to coordinate a switchover operation of the PW from Active to standby, or vice versa.

## [2]. Motivation and Scope

The PWE3 control protocol [2] defines the following status codes in the PW status TLV to indicate the state for an AC and a PW:

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 scenarios defined in [6] allow the provisioning of a primary PW and one or many secondary PWs in the same VPWS or VPLS service.

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 forward user packets to. 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, to allow a PE node 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 PW protection capabilities:

a. A MANDATORY 1:1 PW protection with a single active PW and one
   standby PW. 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 OPTIONAL N:1 PW protection scheme with a single active PW
   and N standby PWs.

c. An OPTIONAL 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. . 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. An OPTIONAL, locally configured precedence to govern the
   selection of a PW when more than one PW qualify 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. OPTIONALLY, implementations can designate by configuration one
   PW in the 1:1 or N:1 set as a primary PW and the remaining as
   secondary PWs. If more than one PW qualify 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. The PE node can use the PW precedence
   to select a secondary PW among many that qualify for active
   state.

These 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 it to the
      remote endpoints. Only a PW which is UP and which 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.

The above mechanisms and operational modes allow the following:

   a. Multi-homing of a CE device to two or more PE nodes.

   b. Multi-homing of a PE node to two or more PE nodes.

More details of how these schemes are used can be found in
Informative Appendix A.

Note that this document specifies the mechanisms to support PW
redundancy where a set of redundant PWs terminate on either a PE (for
SS-PW) or a T-PE (for MS-PW). PW redundancy scenarios where the
redundant set of PW segments terminates on an 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 PSN [7].

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 T-PEs [4].

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 [4].

UP PW:    A PW which has been configured (label mapping exchanged
          between PEs) and is not in any of the PW or AC defect
          states specified in [2]. Such a PW is available for
          forwarding traffic.

DOWN PW:  A PW that has either not been fully configured, or has been
          configured and is in any of the PW or AC defect states
          specified in [2], such a PW is not available for forwarding
          traffic.

Active PW:  An UP PW used for forwarding user, OAM and control plane
          traffic.

Standby PW: An UP PW that is not used for forwarding user traffic,
          but MAY forward OAM and specific control plane traffic.

Primary PW: The PW which 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 behavior is used.

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.

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 which 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., A SS-PW PE, an MS-PW
          T-PE, or an H-VPLS MTU-s or PE-rs.

Provider Edge (PE): A device that provides PWE3 to a CE [7].

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 [4].

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.  It
therefore 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.  It is therefore not
limited to the edge of a provider network [4].

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

PE-rs: A hierarchical virtual private LAN service switch, as
defined in RFC 4762 [3].

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 [7] and T-PEs as per RFC 5659 [4].

This document uses the term 'PW' to be synonymous with both PWs as
per RFC 3985 [7] and SS-PWs and MS-PWs as per RFC 5659 [4].

## 4. PE Architecture

Figure 4-1 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 [7]. The
forwarder selects which of the redundant PWs to using the criteria
described in this document.

```
            +----------------------------------------+
            |                 PE Device              |
            +----------------------------------------+
   Single   |                  |      Single    | PW Instance
     AC     |                  +   PW Instance   X<===========>
            |                  |                 |
            |                  |--------------------|
   <------>o |                  |      Single    | PW Instance
            |     Forwarder    +   PW Instance   X<===========>
            |                  |                 |
            |                  |--------------------|
            |                  |      Single    | PW Instance
            |                  +   PW Instance   X<===========>
            |                  |                 |
            +----------------------------------------+
```
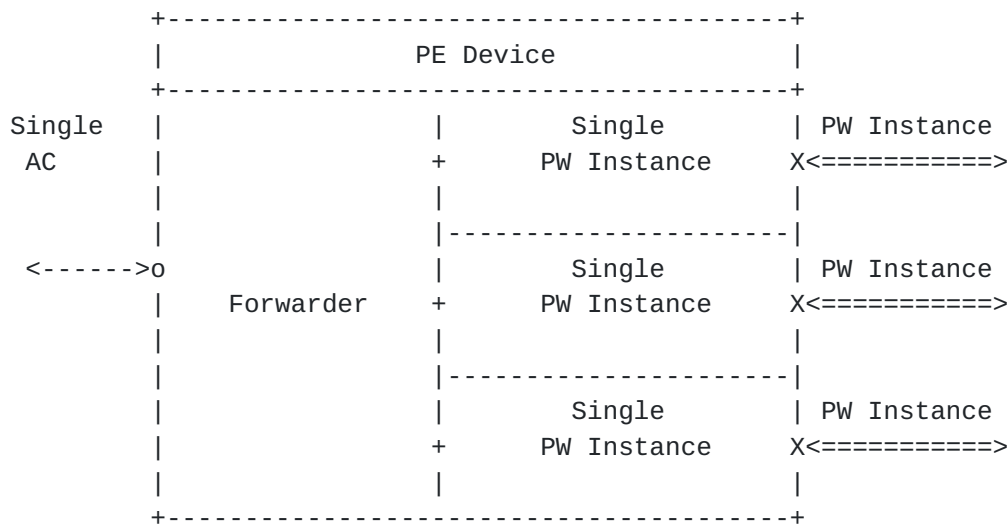
                Figure 4-1 PE Architecure for PW redundancy


## 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
   operates as follows:

1. For FEC128 PW, the PW with the lowest pw-id value is selected.

2. For FEC129 PW, 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 SAII and TAII values as seen on a PE
   node are the mirror values of what the peer PE node sees. To have
   both PE nodes compare the same value we propose that the PE with
   the lowest system IP address use the unsigned integer form of
   AGI::SAII::TAII while the PE with the highest system IP address
   use the unsigned integer form of AGI::TAII::SAII. This way, both
   PEs will compare the same values. The PW which corresponds to the
   minimum of the compared values across all PWs in the redundant is
   selected.

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

   In the case of segmented PW, the operator needs to make sure that
   the pw-id 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 optional active PW selection mechanism described
   next.

The PW endpoints MAY also implement the following optional active PW
selection mechanism.

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.

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
   redundancy set to achieve tie-breaking 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 redundancy 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 advertize Active
preferential forwarding status on a single PW in the redundancy 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 atate. 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 mis-configuration. In the event that an active PW is not found,
a management indication SHOULD be generated. If a management
indication for failure to find an active PW was generated and an
active PW is subsequently found, a management indication SHOULD be
generated, so clearing the previous failure indication. Additionally,
a PE MAY use the optional 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 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 liveliness 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
is 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 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.

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 redundancy 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 qualify 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 but 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 optional active PW selection
mechanism based on primay/secondary and precedence parameters, it
MUST follow the following behaviour:

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 is 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" bit Status Code 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 a
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 [6]). 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 [5].

## 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 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 PW FEC TLV with only the Group ID, or else it contains the PW Generalized 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 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 [5].

## 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)         STANDBY
                        Action: Transmit PW preferential
                                forwarding bit set

                        Receive PW Preferential Forwarding   STANDBY
                           bit set   (slave)
                        Action: Stop forwarding over PW

                        Receive PW Preferential Forwarding   ACTIVE
                           bit set but bit not supported
                        Action: None

                        Receive PW Preferential Forwarding   ACTIVE
                           bit clear
                        Action: None.


        STANDBY         PW activated (master)              ACTIVE
                        Action: Transmit PW preferential
                          forwarding bit clear

                        Receive PW Preferential Forwarding   ACTIVE
                           bit clear (slave)
                        Action: Activate PW

                        Receive PW Preferential Forwarding   STANDBY
                           bit clear but bit not supported
                        Action: None

                        Receive PW Preferential Forwarding   STANDBY
                           bit set
                        Action: None
```

Table 1  PW State Transition Table in Master/Slave Mode


## 6.3. Coordination of PW Switchover

There are PW redundancy applications which 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 [6]. 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 of the MS-PW by one endpoint can be an operational event, for example a failure, which 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 acknowledgement 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 switchover.

This document specifies a second status bit that is used by a PE to request that its peer PE switchover to use a different active PW. This bit is referred to as the 'request PW 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 redundancy 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.

The request switchover bit is OPTIONAL and, if received by a PE, is ignored if not understood.

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 [5].

**[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 it desires to switch to.

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 the remote
   endpoint to acknowledge the request.

d. If while waiting for the acknowledgment, the requesting endpoint
   receives a request from its peer to switchover 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 acknowledgement 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 which is UP, it MUST perform the
   following:

        i. Abort the timer.

       ii. Issue a new request to switchover to the alternate PW.

      iii. Re-start 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 which 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 different from the

currently active one, and the requested PW 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 [5].

The procedures for determining and mapping PW and AC states MUST
follow the rules in [8] with the following modifications.

### 7.1. AC Defect State Entry/Exit

A PE enters the AC receive (or transmit) defect state for a PW when
one or more of the conditions specified for this PW type in [8] 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 redundancy set when required by the PW type
in [8] .

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 redundancy set when required by
the PW type in [8] .

### 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 8.2.1
(Section 8.2.2) in [8] are met for all PWs in the redundancy set.

When a PE enters the PW receive (or transmit) defect state for a PW
service, it MUST send a reverse (or forward) defect indication over
one or more of the PWs in the redundancy set if the PW failure was
detected by this PE without receiving a forward defect indication
from the remote PE [8].

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 redundancy set if applicable.

**8**. **Applicability and Backward Compatibility**

The mechanisms defined in this document are applicable to
applications where standby state signaling of a PW or PW group is
required. Both PW FEC 128 and 129 are supported. All PWs which are
part of a redundant set MUST use the same FEC type. When the set uses
FEC 128 PWs, each PW is uniquely identified by its PW-ID. When the
redundant set uses FEC 129 PWs, each PW MUST have a unique identifier
which 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 PW Status TLV is found to be not
supported by either of its endpoint after status negotiation
procedures, then the mechanisms specified in this document cannot be
used.

A PE implementation compliant to RFC 4447 [2], and which does not
support the generation or processing of the 'Preferential Forwarding'
status bit or of the 'request switchover' status bit, will ignore
these status bits if they are received from a peer PE.

**9**. **Security Considerations**

This document uses the LDP extensions that are needed for protecting
pseudowires. It will have the same security properties as in the PWE3
control protocol [2].

**10**. **MIB Considerations**

This document makes the following update to the PwOperStatusTC
textual convention in RFC 5542 [9]:

```
PwOperStatusTC ::= TEXTUAL-CONVENTION
      STATUS      current
      DESCRIPTION
        "Indicates the operational status of the PW.



   - up(1):           Ready to pass packets.

   - down(2):         PW signaling is not yet finished, or
                      indications available at the service
                      level indicate that the PW is not
                      passing packets.
```

           - testing(3):          AdminStatus at the PW level is set to
                                   test.



           - dormant(4):    The PW is not in a condition to pass
                            packets but is in a 'pending' state,
                            waiting for some external event. For example, the
     PW Preferential Forwarding status state machine as defined in [RFC
     XXXX (this document)] is in state "STANDBY".

           - notPresent(5): Some component is missing to accomplish
                            the setup of the PW.  It can be
                            configuration error, incomplete
                            configuration, or a missing H/W component.

           - lowerLayerDown(6): One or more of the lower-layer interfaces
                                responsible for running the underlying PSN
                                is not in OperStatus 'up' state."

        SYNTAX    INTEGER {
            up(1),
            down(2),
            testing(3),
            dormant(4),
            notPresent(5),
            lowerLayerDown(6)
            }



## 11. IANA Considerations


   This document defines the following PW status codes for the PW
   redundancy application. IANA is requested to allocate these from the
   PW 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. Major Contributing Authors

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

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

Pranjal Kumar Dutta
Alcatel-Lucent
Email: pranjal.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
CA Technologies
thomas.nadeau@ca.com

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

Hamid Ould-Brahim
Nortel
Email: hbrahim@nortel.com

Olen Stokes
Extreme Networks
Email: ostokes@extremenetworks.com

Daniel Cohn
Orckit
DanielC@orckit.com


## 13. Acknowledgments

## 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., et al., "Pseudowire Setup and Maintenance using
       LDP", RFC 4447, April 2006.

[3]    Kompella,V., Lasserrre, M. , et al., ''Virtual Private LAN
       Service (VPLS) Using LDP Signalling'', RFC 4762, January 2007.


### 14.2. Informative References

[4]    Bocci, M., Bryant, S., et al., ''An Architecture for Multi-
       Segment Pseudowire Emulation Edge-to-Edge'', RFC 5659, October
       2009.

[5]    Martini, L., et al., "Segmented Pseudo Wire", RFC 6073, January
       2011.

[6]     Muley, P., et al., ''Pseudowire (PW) Redundancy'', draft-ietf-pwe3-redundancy-07.txt'', April 2012.

[7]     Bryant, S., et al., "Pseudo Wire Emulation Edge-to-Edge (PWE3) Architecture", RFC 3985, March 2005

[8]     Aissaoui, M., et al., "Pseudo Wire (PW) OAM Message Mapping", RFC 6310, July 2011.

[9]     Nadeau, T., Zelig, D., Nicklass, O., ''Definitions of Textual Conventions for Pseudowire (PW) Management'', RFC 5542, May 2009

[10]    Dutta, P., Lasserre, M., Stokes, O., ''LDP Extensions for Optimized MAC Address Withdrawal in H-VPLS'', draft-ietf-l2vpn-vpls-ldp-mac-opt-05.txt, October 2011

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

## 15. 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 scenarios described in the PW redundancy requirements and framework document [6].

### 15.1. One Multi-homed CE with single SS-PW redundancy

The following figure illustrates an application of single segment pseudowire redundancy.

```
      |<-------------- Emulated Service ---------------->|
      |                                                  |
      |              |<------- Pseudo Wire ------>|      |
      |              |                            |      |
      |              |    |<-- PSN Tunnels-->|    |      |
      |              V    V                  V    V      |
      V    AC    +----+                  +----+   AC   V
   +-----+   |    | PE1|================|    |    |   +-----+
   |     |---------|....|...PW1.(active)...|....|----------|     |
   |     |   |    |    |================|    |    |   | CE2 |
   | CE1 |       +----+                  |PE2 |       |     |
   |     |       +----+                  |    |       +-----+
   |     |   |    |    |================|    |    |
   |     |---------|....|...PW2.(standby)..|    |
   +-----+   |    | PE3|================|    |
         AC    +----+                  +----+
```

                 Figure 15-1 Multi-homed CE with single SS-PW redundancy

   The application in Figure 15-1 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 [6]).

   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 "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
   '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 select PW2 as the new active pseudowire to send
   traffic to.

On failure of 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 pseudowire to send
traffic to. 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 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, and 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 PW Preferential Forwarding status is driven by the AC
forwarding state as determined by the AC dual-homing protocol used.

## 15.2. Multiple Multi-homed CEs with single SS-PW redundancy

```
          |<-------------- Emulated Service --------------->|
          |                                                 |
          |          |<------- Pseudo Wire ------>|         |
          |          |                            |         |
          |          |    |<-- PSN Tunnels-->|     |        |
          |          V    V    (not shown)  V     V         |
          V    AC    +----+                 +----+    AC   V
      +-----+   |    |....|.......PW1........|....|    |   +-----+
      |     |---------| PE1|......    .........| PE3|----------|    |
      | CE1 |        +----+     \ /  PW3    +----+        | CE2 |
      |     |        +----+      X          +----+        |     |
      |     |        |    |......./ \..PW4....|     |      |     |
      |     |---------| PE2|                | PE4|--------- |     |
      +-----+   |    |....|.....PW2..........|....|    |   +-----+
          AC    +----+                 +----+    AC
```

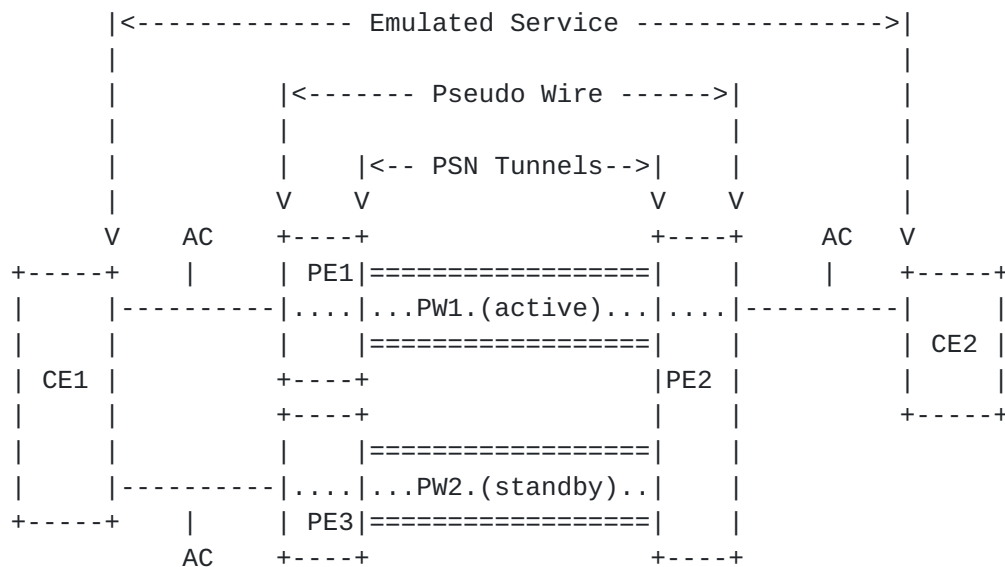                Figure 15-2 Multiple Multi-homed CEs with single SS-PW redundancy

The application in Figure 15-2 makes use of the Independent mode of
operation.

CE1 is dual-homed to PE1 and PE2. CE2 is dual-homed 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, from CE1 to PE2 is Standby; furthermore, assume that the AC from CE2 to PE3 is Standby and from CE2 to PE4 is Active. The method of deriving Active/Standby status of the AC is outside the scope of this document.

PE1 advertises the preferential 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 status "Standby" and UP/DOWN status "Pseudowire forwarding" for pseudowires PW2 and PW3 to PE3 and PE4. PE3 advertises preferential status "Standby" and UP/DOWN status "Pseudowire forwarding" for pseudowires PW1 and PW3 to PE1 and PE2. PE4 advertise the preferential 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 pseudowire to send traffic to.


On failure of PE1 node, PE2 will detect it 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 pseudowire to send traffic to. 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
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 any primary/secondary procedures, as defined in sections
5.1.  and 5.2. , do not 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.
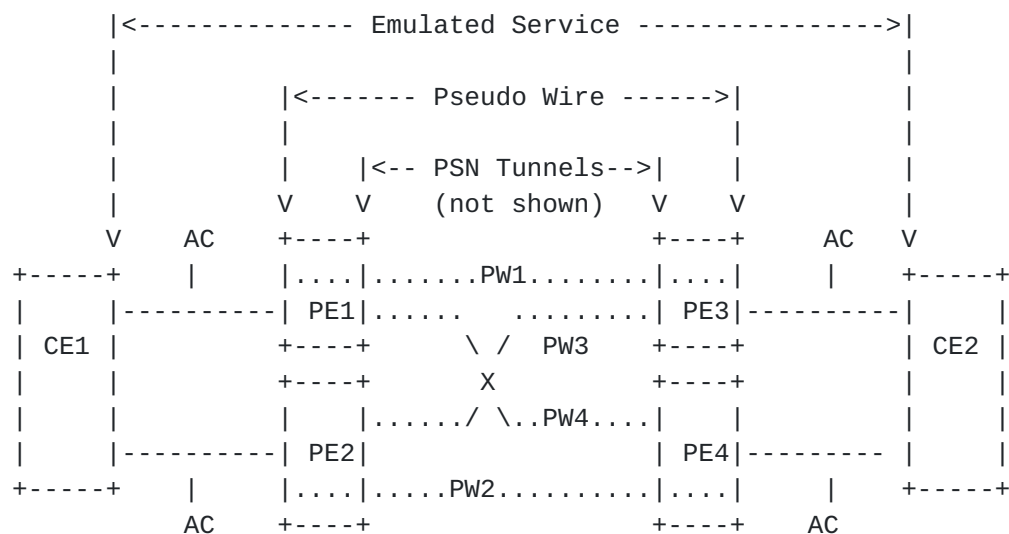
### 15.3. Multi-homed CE with MS-PW redundancy

The following figure illustrates an application of multi-segment
pseudowire redundancy.

```
        Native    |<-----------Pseudo Wire----------->|  Native
        Service   |                                   |  Service
         (AC)     |    |<-PSN1-->|      |<-PSN2-->|    |   (AC)
          |       V    V         V      V         V    V    |
          |      +-----+         +-----+         +-----+
  +----+  |      |T-PE1|=========|S-PE1|=========|T-PE2|   |  +----+
  |     |-------|......PW1-Seg1.......|PW1-Seg2.......|-------|  |
  |     |       |      |=========|     |=========|     |       |  |
  | CE1|       +-----+         +-----+         +-----+       |  |
  |     |         |.|           +-----+         +-----+       | CE2|
  |     |         |.|=========|      |=========|     |       |  |
  |     |         |.....PW2-Seg1......|.PW2-Seg2......|-------|  |
  +----+         |============|S-PE2|=========|T-PE4|   |  +----+
                  +-----+         +-----+   AC
```

              Figure 15-3 Multi-homed CE with MS-PW redundancy

The application in Figure 15-3 makes use of the Independent mode of
operation. It extends the application described in Section 15.1. of
this document and in [6] 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 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 [5].

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.

## 15.4. Multi-homed CE with MS-PW redundancy and S-PE protection

The following figure illustrates an application of multi-segment
pseudowire redundancy with 1:1 PW protection.

```
          Native   |<-----------Pseudo Wire------------->|  Native
          Service  |                                     |  Service
           (AC)    |      |<-PSN1-->|     |<-PSN2-->|     |   (AC)
             |     V      V         V     V         V     V    |
             |                      +-----+                    |
             |        |============|     |============|        |
             |        |.....PW3-Seg1......|.PW3-Seg2....|      |
             |        |.|==========|S-PE3|==========|.|        |
             |        |.|           +-----+          |.|        |
             |      +-----+         +-----+         +-----+    |
         +----+ |    |T-PE1|=========|S-PE1|=========|T-PE2|    | +----+
         |    |-------|......PW1-Seg1.......|PW1-Seg2.......|-------|    |
         |    |    |      |=========|     |=========|     |    |    |
         | CE1|    +-----+         +-----+         +-----+    |    |
         |    |    |.| |.|          +-----+         +-----+    | CE2|
         |    |    |.| |.|=========|     |=========|     |    |    |
         |    |    |.| |...PW2-Seg1......|.PW2-Seg2......|-------|    |
         +----+    |.| |==========|S-PE2|=========|T-PE4|    | +----+
                   |.|           +-----+         +-----+    AC
                   |.|           +-----+          |.|
                   |.|============|     |==========|.|
                   |.......PW4-Seg1......|.PW4-Seg2....|
                   |==============|S-PE4|============|
                                 +-----+
```

   Figure 15-4  Multi-homed CE with MS-PW redundancy and protection
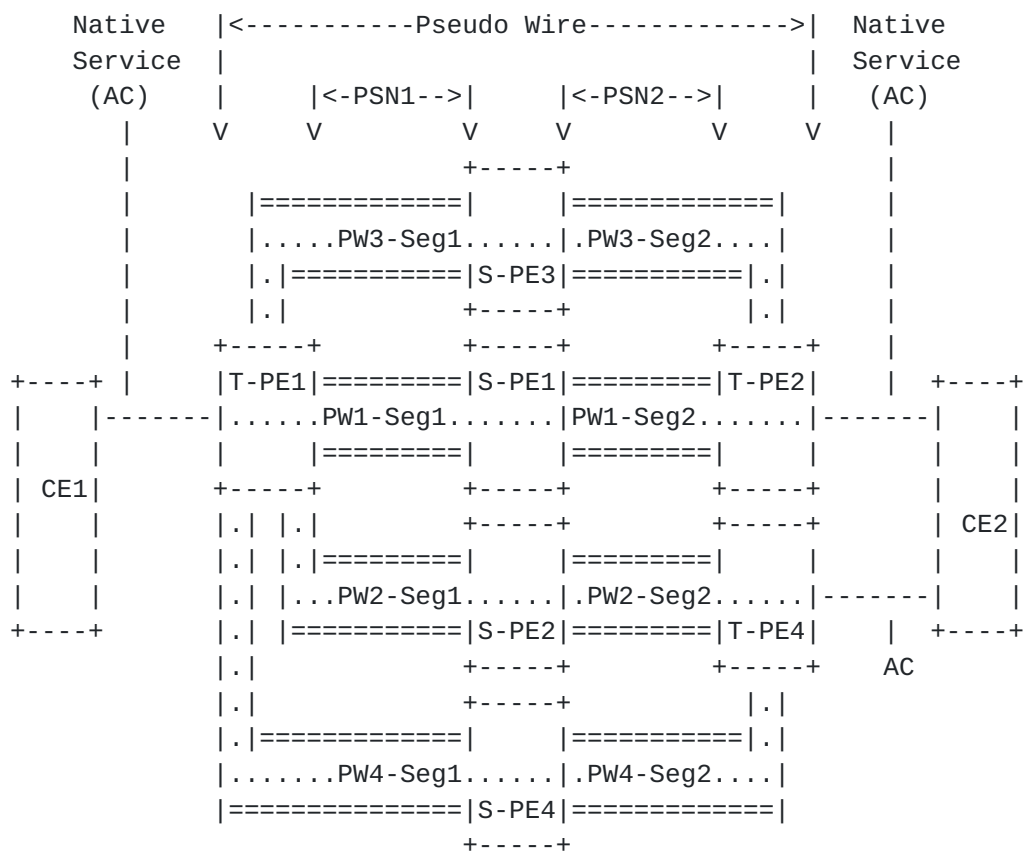
The application in Figure Figure 15-4 makes use of the
Independent mode of operation.

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 15-4 shows four
separate S-PE nodes for clarity.

The behavior of this setup is exactly the same as in 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.

**15.5. Single Homed CE with MS-PW redundancy**

   The following is an application of the independent mode of operation
   along with the optional 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. .

```
          Native    |<------------Pseudo Wire------------>|  Native
         Service  |                                     |  Service
          (AC)    |      |<-PSN1-->|     |<-PSN2-->|     |  (AC)
             |     V     V         V     V         V     V   |
             |     +-----+         +-----+         +-----+   |
    +----+ |     |T-PE1|=========|S-PE1|=========|T-PE2|   |   +----+
    |    |-------|......PW1-Seg1.......|.PW1-Seg2......|-------|   |
    | CE1|       |     |=========|     |=========|     |       | CE2|
    |    |       +-----+         +-----+         +-----+       |    |
    +----+        |.||.|                         |.||.|        +----+
                  |.||.|         +-----+         |.||.|
                  |.||.|=========|     |=========|.||.|
                  |.||...PW2-Seg1......|.PW2-Seg2...||.|
                  |.| ==========|S-PE2|=========== |.|
                  |.|           +-----+            |.|
                  |.|===========+-----+============ .|
                  |.....PW3-Seg1.|     | PW3-Seg2......|
                   =============|S-PE3|===============
                                |     |
                                +-----+
```
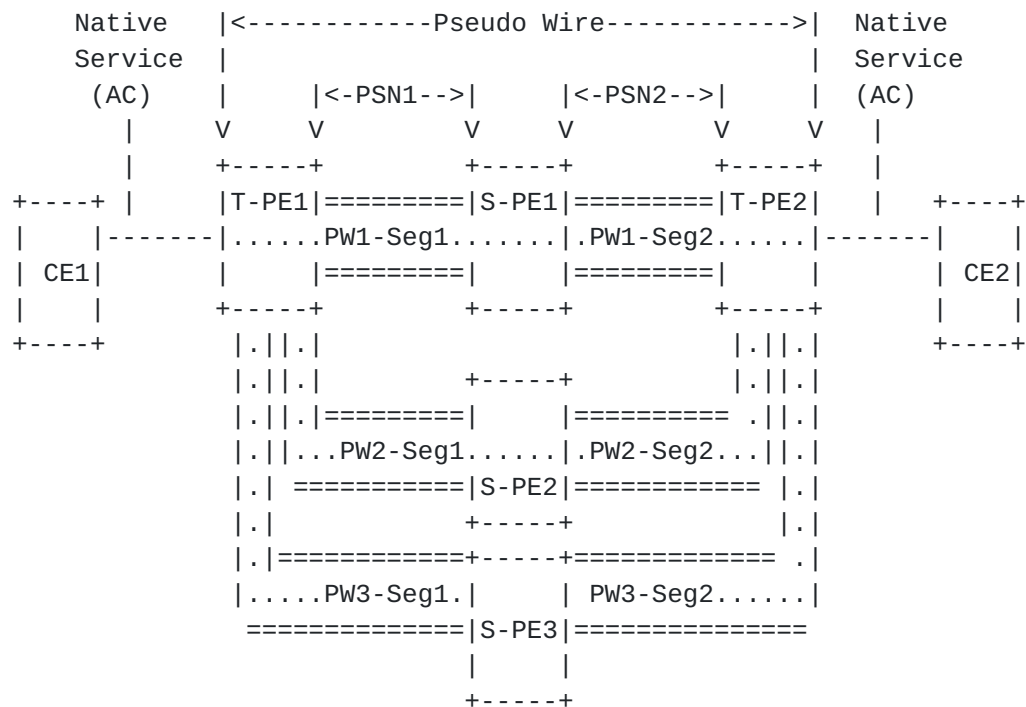
   Figure 15-5 Single homed CE with multi-segment pseudowire redundancy

   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. A primary PW, PW1 has the lowest
   precedence value of zero. A secondary PW, PW2, which is switched at
   S-PE2 and has a precedence of 1. Finally, another secondary PW, PW3,
   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.
   Lower the precedence value, 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, or 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 which 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 optional request switchover/acknowledgement procedures to synchronize the choice of PW to forward on in both directions.

The trigger for sending a request to switchover 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 response from the other endpoint. In such situation, the T-PE with greater system address request is given precedence. This helps in synchronizing PWs in event of mismatch of precedence configuration as well.

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

## 15.6. PW redundancy between H-VPLS MTU-s and PE-rs

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.

```
                                          PE1-rs
                                        +--------+
                                        | VSI    |
                     Active PW          |   --   |
                     Group..........|../  \..|.
        CE-1                       .  |  \  /  | .
          \                       .  |   --   | .
           \                    .    +--------+   .
            \   MTU-s        .              .       .   PE3-rs
             +--------+   .                 .     . +--------+
             |  VSI   | .                   . H-VPlS .| VSI    |
             |  -- ..|..                    .  Core   |.. --   |
             | /  \  |                      .  PWs    | /  \   |
             | \  /..|..                    .         | \  /   |
             |  --   | .                    .         .|.. --   |
             +--------+   .                  .          . +--------+
            /              .                  .          .
           /               .         +--------+   .
          /                .         | VSI    | .
        CE-2               .         |   --   | .
                ..........|../  \..|.
               Standby PW          | \  /   |
               Group               |   --   |
                                   +--------+
                                     PE2-rs
```
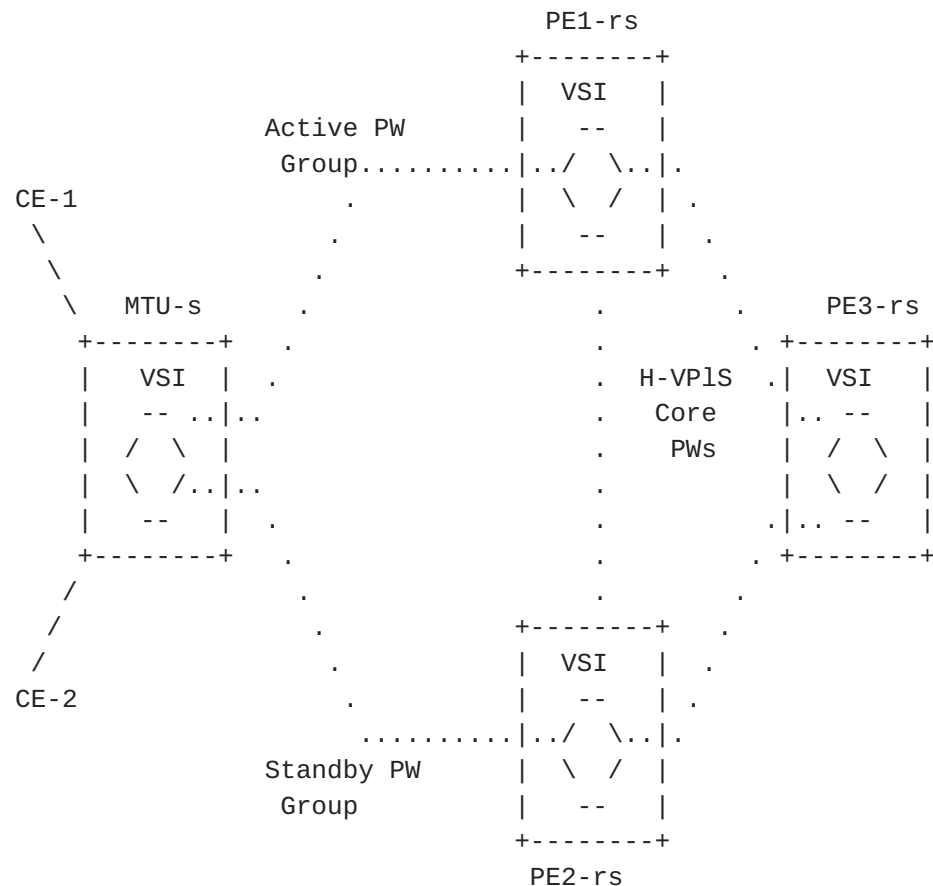
                Figure 15-6 Multi-homed MTU-s in H-VPLS core

   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 network. MTU-s communicates to PE1-rs and PE2-rs the
   forwarding status of its member PWs for a set of VSIs having common
   status Active/Standby. It may be signaled using PW grouping with
   common group-id in PWid FEC Element or Grouping TLV in 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 state of 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 RSTP controlled
IEEE Ethernet Bridges 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 to be sent, one by MTU-s
and another one by PE-rs nodes. Mechanisms to disable MAC Withdraw
trigger in certain devices is out of the scope of this document

Author's Addresses

Praveen Muley
Alcatel-lucent
701 E. Middlefiled Road
Mountain View, CA, 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