Internet Engineering Task Force (IETF)
Request for Comments: 6478
Updates: 5885
Category: Standards Track
ISSN: 2070-1721

L. Martini
G. Swallow
G. Heron
Cisco
M. Bocci
Alcatel-Lucent

May 2012

Pseudowire Status for Static Pseudowires

Abstract

This document specifies a mechanism to signal Pseudowire (PW) status messages using a PW associated channel (ACh). Such a mechanism is suitable for use where no PW dynamic control plane exits, known as static PWs, or where a Terminating Provider Edge (T-PE) needs to send a PW status message directly to a far-end T-PE. The mechanism allows PW Operations, Administration, and Maintenance (OAM) message mapping and PW redundancy to operate on static PWs. This document also updates RFC 5885 in the case when Bi-directional Forwarding Detection (BFD) is used to convey PW status-signaling information.

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/rfc6478.

Table of Contents

1. Introduction

The default control plane for Pseudowire (PW) technology, as defined in [RFC4447], is based on the Label Distribution Protocol (LDP). However, that document also describes a static provisioning mode without a control plane. When a static PW is used, there is no method to transmit the status of the PW or attachment circuit (AC) between the two Provider Edge (PE) devices at each end of the PW. This document defines a method to transport the PW status codes defined in Section 5.4.2 of [RFC4447] and elsewhere [REDUNDANCY] in-band with the PW data using a generic associated channel [RFC5586].

2. Specification of Requirements

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

3. Terminology

ACh: Associated Channel

ACH: Associated Channel Header

FEC: Forwarding Equivalence Class

LDP: Label Distribution Protocol

LSP: Label Switching Path

MS-PW: Multi-Segment Pseudowire

PE: Provider Edge

PW: Pseudowire

SS-PW: Single-Segment Pseudowire

S-PE: Switching Provider Edge Node of MS-PW

T-PE: Terminating Provider Edge Node of MS-PW

4. Applicability

As described in [RFC4447] and [RFC6310], a PE that establishes an MPLS PW using means other than LDP, e.g., by static configuration, MUST support some alternative method of status reporting. The

procedures described in this document are for use when PWs are statically configured and an LDP control plane is not available.

As defined in [RFC4447], a PE that establishes a PW using LDP MUST use the PW status TLV mechanism for AC and PW status and defect notification on that PW. In order to avoid duplicate notifications and potentially conflicting notifications, such PEs MUST NOT use the mechanisms described in this document for those PWs, except that the S-PE bypass mode described in Section 5.5 MAY be used when both T-PEs at each end of the PW use LDP to establish the PW.

In order to protect against duplicate notifications and potentially conflicting notifications, when the Pseudowire Status protocol for Static Pseudowires described in this document is used, the BFD VCCV (Virtual Circuit Connectivity Verification) status-signaling mechanisms described in [RFC5885] (CV Types 0x08 and 0x20) MUST NOT be used. BFD VCCV for fault detection (CV types 0x04 and 0x10) MAY still be used.

5. Pseudowire Status Operation

5.1. PW OAM Message

The PW status TLV as defined in Section 5.4.2 of [RFC4447] is transported in a PW OAM message using the PW ACH.

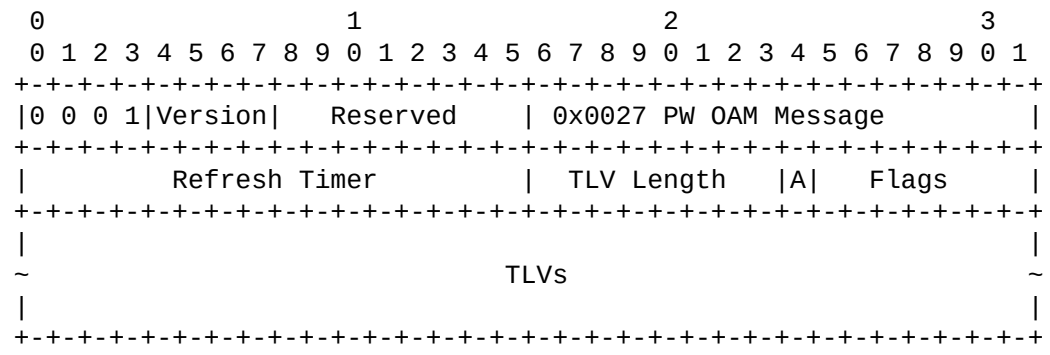

Figure 1: ACh PW OAM Message Packet Header

The first 32 bits are the standard ACH header construct as defined in [RFC5586].

The first nibble (0001b) indicates the ACH instead of PW data. The version and the reserved values are both set to 0 as specified in [RFC4385].

The refresh timer is an unsigned integer and specifies refresh time in seconds with a range from 1 to 65535. The value 0 means that the refresh timer is set to infinity, and the PW OAM message will never be refreshed, and will never timeout.

The TLV length field indicates the length of all TLVs only. This document defines only the transport of the PW status TLV, as defined in Section 5.4.2, [RFC4447], in the TLV field. In the future, additional TLVs may be defined to be used in this field with code points allocated from the IANA registry called "LDP TLV Type Name Space".

The A flag bit is used to indicate an acknowledgment of the PW status TLV included. The rest of the flag bits are reserved and they MUST be set to 0 on transmit, and ignored upon receipt. When the A bit is set, the refresh timer value is a requested timer value.

The PW OAM Message code point value is 0x0027.

5.2. Sending a PW Status Message

The PW Status messages are sent in-band using the PW OAM message containing the PW Status TLV for a particular PW, as defined in [RFC4447]. The PW Status TLV format is almost as defined in [RFC4447] and is repeated here for the reader's convenience:

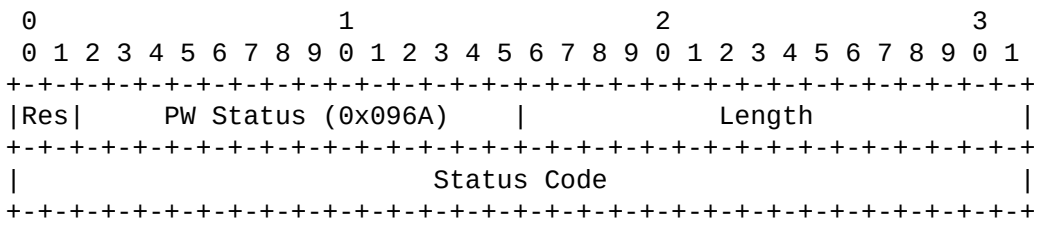

Figure 2: PW Status TLV Format

Unlike the case in [RFC4447], here, the first 2 bits are reserved, and MUST be set to zero on transmit and ignored on receipt.

The PW Status TLV is prepended with a PW OAM message header and sent on the ACh of the PW to which the status update applies.

To clear a particular status indication, the PE needs to send a new PW OAM message containing a PW Status TLV with the corresponding bit cleared as defined in [RFC4447].

The procedures described in [RFC6073] that apply to an S-PE and PW using an LDP control plane also apply when sending PW status using

the PW OAM channel. The OPTIONAL procedures using the SP-PE TLV described in [RFC6073] can also be applied when sending PW status using the PW OAM channel.

The detailed message transmit and message receive procedures are specified in the next section. PW OAM status messages MUST NOT be used as a connectivity verification method.

5.3. PW OAM Status Message Transmit and Receive

Unlike the PW status procedures defined in [RFC4447], with this method there is no TCP/IP session or session management. Therefore, unlike the TCP/IP case, where each message is sent only once, the PW OAM message containing the PW status TLV needs to be transmitted repeatedly to ensure reliable message delivery. If a malformed TLV or an unknown TLV is received in a PW OAM status message, the TLV MUST be ignored, and the PE SHOULD report the event to the operator.

A PW OAM message containing a PW status TLV with a new status bit set or reset will be transmitted immediately by the PE. Unless the message is acknowledged within a second, the PW OAM message will then be repeated twice more at an initial interval of one second. Subsequently, the PW OAM message will be transmitted with an interval specified by the refresh timer value in the packet. Note that this value MAY be updated in the new PW OAM message packet, in which case the new refresh timer value becomes the new packet transmit interval.

The suggested default value for the refresh timer is 600 seconds. This default is adequate for typical deployments, and PEs are designed to take into account processing these messages at the required rate.

When a PW OAM message containing a status TLV is received, a timer is started according to the refresh rate specified in the packet. If another non-zero PW status message is not received within 3.5 times the specified timer value, the status condition will timeout in 3.5 times the last refresh timer value received, and the default status of zero is assumed on the PW. It is also a good practice to introduce some jitter in the delay between refresh transmissions, as long as the maximum jitter delay is within the prescribed maximum refresh time of 3.5 times the specified timer value for 3 consecutive refresh packets.

To clear a particular status fault, the PE need only send an updated message with the corresponding bit cleared. If the PW status code is zero, the PW OAM message will be sent like any other PW OAM status message using the procedures described above; however, transmission will cease after 3 PW status messages have been sent at one second

intervals and before the refresh timer expires. A PW status message of zero MAY be acknowledged using the procedures described in Section 5.3.1. If it is acknowledged, then a timer value of zero MUST be used. This SHOULD cause the PE sending the PW status notification message with a PW status code equal to zero to stop sending and to continue normal operation.

5.3.1. Acknowledgment of PW Status

A PE receiving a PW OAM message containing a PW status message MAY acknowledge the PW status message by simply building a reply packet with the same format and status code as the received PW OAM message, but with the A bit set, and transmitting it on the PW ACh back to the source of the PW OAM message. The receiving PE MAY use the refresh timer field in the acknowledgement packet to request a new refresh interval from the originator of the PW OAM message. The timer value set in the reply packet SHOULD then be used by the originator of the PW OAM message as the new transmit interval. If the requested refresh timer value is to be used, then, when the the current timer expires, the PW OAM message transmission interval is set to the new value and the new value is sent in the PW OAM message. If the transmitting PE does not want to use the new timer value (for local policy reasons, or because it simply cannot support it), it MUST refresh the PW OAM message with the timer value it desires. The receiving PE will then set its timeout timer according to the new refresh timer value that is in the packet received, regardless of what timer value it requested. The receiving PE MUST NOT request a new refresh timer value more than once per refresh interval.

The suggested default value for the refresh timer value in the acknowledgment packet is 600 seconds.

If the sender PE receives an acknowledgment message that does not match the current active PW status message being sent, it simply ignores the acknowledgment packet.

If a PE that has received a non-zero status code for a PW detects by any means that the far end PE has become unreachable, it will follow the standard defect entry procedures of [RFC6310], Section 6.2.

5.4. MPLS Label Stack

With one exception, all PW OAM status messages are sent to the adjacent PE across the PSN tunnel. In many cases, the transmitting PE has no way to determine whether the adjacent PE is an S-PE or a T-PE. This is a necessary behavior to preserve backward compatibility with PEs that do not understand MS-PWs. In the procedures described in this document, there are two possible

destinations for the PW OAM status messages: the adjacent PE or the T-PE. Sending a PW status message directly to the T-PE is an enhanced method that is only applicable using PW OAM status messages sent in the PW ACH.

5.4.1. Label Stack for a Message Destined to the Next PE

A PE that needs to forward a PW OAM status message to the adjacent PE across the PSN tunnel MUST set the PW label TTL field to 1. Furthermore, if the control word is not in use on the particular PW, the PE MUST place the GAL reserved label [RFC5586] below the PW label with the TTL field set to 1.

5.4.2. Label Stack for a Message Destined to the Egress PE

This is also known as "S-PE bypass mode"; see below. A T-PE that requires sending a PW OAM status message directly to the corresponding T-PE at the other end of the PW MUST set the TTL of the PW label to a value that is sufficient to reach the corresponding T-PE. This value will be greater than one, but will be set according to the local policy on the transmitting T-PE. Furthermore, if the control word is not in use on the particular PW, the PE MUST also place the GAL reserved label [RFC5586] below the PW label with the TTL field set to 1.

5.5. S-PE Bypass Mode

S-PE bypass mode enables a T-PE that uses LDP as the PW setup and control protocol to bypass all S-PEs that might be present along the MS-PW and to send a message directly to the remote T-PE. This is used for very fast message transmission in-band with the PW PDUs. This mode is OPTIONAL and MUST be supported by both T-PEs to be enabled. This mode MUST NOT be used if the first PW segment connected to each T-PE is not using LDP.

Note that this method MUST NOT be used to send messages that are permitted to originate at an S-PE. Otherwise, race conditions could occur between messages sent via the control plane by S-PEs and messages sent via the data plane by T-PEs.

Status codes, except for those listed below, MUST NOT be sent using the S-PE bypass procedure and MUST be ignored on reception.

0x00000002 - Local Attachment Circuit (ingress) Receive Fault

0x00000004 - Local Attachment Circuit (egress) Transmit Fault

0x00000020 - PW forwarding standby

0x00000040 - Request switchover to this PW

Note that since "clear all failures" may be sent by an S-PE, it MUST NOT be sent using the S-PE bypass mode.

When S-PE bypass mode is enabled, all PW Status TLVs received using this method have priority over PW Status TLVs sent via control protocols such as LDP [RFC4447]. However, the same PW Status TLVs MUST also be sent in LDP to keep the S-PEs state updated.

5.5.1. S-PE Bypass Mode LDP Flag Bit

When a PW Segment along an MS-PW is using the LDP control protocol and wishes to request the use of the S-PE bypass status message mode, it sets the B bit in the generic protocol flags interface parameters sub-TLV as shown in Figure 3. This flag can only be set by a T-PE using LDP as the PW configuration and management protocol. If the S-PE bypass mode LDP flag bit in the generic protocol flags interface parameter does not match in the FEC advertisement for directions of a specific PW, that PW MUST NOT be enabled.

The interface parameter is defined as follows:

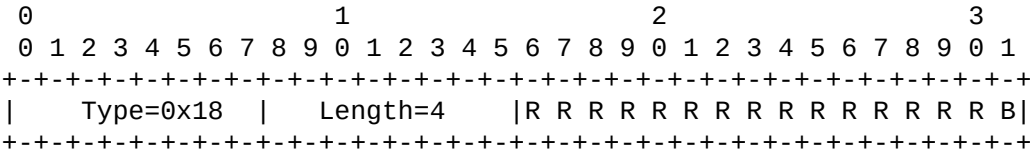

Figure 3: PW Generic Protocol Flags Sub-TLV

TLV Type

Type 0x18 - PW Generic Protocol Flags.

Length

TLV length is always 4 octets.

Flags

Bit B, in position 31 above, is set to request the S-PE bypass mode. R bits are to be allocated by IANA as described in the IANA section. If they are not allocated, they are to be considered as reserved for future use and MUST be zero on transmission and ignored on reception of this TLV.

If the T-PE receives an LDP label mapping message containing a generic protocol flags interface parameter TLV with the bit B set, then the T-PE receiving the label mapping message MAY send S-PE bypass status messages in the PW ACh. If bit B of said TLV is not set, or the TLV is not present, then the T-PE receiving the label mapping message MUST NOT send S-PE bypass status messages in the PW ACh.

6. S-PE Operation

The S-PE will operate according to the procedures defined in [RFC6073]. The following additional procedures apply to the case where a static PW segment is switched to a dynamic PW segment that uses LDP, and the case where a static PW segment is switched to another static PW segment.

6.1. Static PW to Another Static PW

The procedures that are described in [RFC6073] Section 10 also apply to the case of a static PW switched to another static PW. The LDP header is simply replaced by the PW OAM header; otherwise, the packet format will be identical. The information that is necessary to form an SP-PE TLV MUST be configured in the S-PE, or no SP-PE TLV will be sent. [RFC6073] defines the IANA "Pseudowire Switching Point PE TLV Type" registry. In order to support the static PW configuration and addressing scheme, the following new code point has been assigned:

Type	Length	Description
0x07	24	Static PW/MPLS-TP PW segment ID of last PW segment traversed

The format of this TLV is that of the "Static Pseudowire Sub-TLV" defined in [RFC6426].

6.2. Dynamic PW to Static PW or Vice Versa

The procedures that are described in Section 10 of [RFC6073] also apply to this situation. However, if the PW label of the LDP-controlled PW segment is withdrawn by the adjacent PE, the S-PE will set the PW status code "0x00000001 - Pseudowire Not Forwarding" to the adjacent PW on the static PW segment.

The S-PE will only withdraw its label for the dynamic, LDP-controlled PW segment if the S-PE is not provisioned.

7. Security Considerations

The security measures described in [RFC4447], [RFC5085], and [RFC6073] are adequate for the proposed mechanism.

8. IANA Considerations

IANA has set up the registry of "PW Generic Protocol Flags". These are bit strings of length 16. Bit 0 is defined in this document. Bits 1 through 15 are to be assigned by IANA using the "IETF Review" policy defined in [RFC5226].

Any requests for allocation from this registry require a description of up to 65 characters.

Initial PW Generic Protocol Flags value allocations are as follows:

Bit Mask		Description	
0x0001	-	S-PE bypass mode	[RFC6478]

This document uses a new Associated Channel Type. IANA already maintains the "Pseudowire Associated Channel Types" registry. The value 0x0027 has been assigned with the description "PW OAM Message".

This document uses a new Pseudowire Switching Point PE TLV Type. IANA already maintains the "Pseudowire Switching Point PE sub-TLV Type" registry. A value of 0x07 has been assigned with the description "Static PW/MPLS-TP PW segment ID of last PW segment traversed".

This document uses a new interface parameter type. IANA already maintains the "Pseudowire Interface Parameters Sub-TLV type Registry". A value of 0x18 has been assigned with the description "PW Generic Protocol Flags".

9. References

9.1. Normative References

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

[RFC4385] Bryant, S., Swallow, G., Martini, L., and D. McPherson, "Pseudowire Emulation Edge-to-Edge (PWE3) Control Word for Use over an MPLS PSN", RFC 4385, February 2006.

[RFC4447] Martini, L., Ed., Rosen, E., El-Aawar, N., Smith, T., and G. Heron, "Pseudowire Setup and Maintenance Using the Label Distribution Protocol (LDP)", RFC 4447, April 2006.

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

[RFC5226] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 5226, May 2008.

[RFC6073] Martini, L., Metz, C., Nadeau, T., Bocci, M., and M. Aissaoui, "Segmented Pseudowire", RFC 6073, January 2011.

[RFC6310] 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.

[RFC6426] Gray, E., Bahadur, N., Boutros, S., and R. Aggarwal, "MPLS On-Demand Connectivity Verification and Route Tracing", RFC 6426, November 2011.

9.2. Informative References

[REDUNDANCY] Muley, P., Ed., and M. Aissaoui, Ed., "Pseudowire Preferential Forwarding Status Bit", Work in Progress, September 2011.

[RFC5885] Nadeau, T., Ed., and C. Pignataro, Ed., "Bidirectional Forwarding Detection (BFD) for the Pseudowire Virtual Circuit Connectivity Verification (VCCV)", RFC 5885, June 2010.

[RFC5586] Bocci, M., Ed., Vigoureux, M., Ed., and S. Bryant, Ed., "MPLS Generic Associated Channel", RFC 5586, June 2009.

Authors' Addresses

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

George Swallow
Cisco Systems, Inc.
300 Beaver Brook Road
Boxborough, Massachusetts 01719
United States
EMail: swallow@cisco.com

Giles Heron
Cisco Systems
9-11 New Square
Bedfont Lakes
Feltham
Middlesex
TW14 8HA
United Kingdom
EMail: giheron@cisco.com

Matthew Bocci
Alcatel-Lucent
Voyager Place
Shoppenhangers Road
Maidenhead
Berks
SL6 2PJ
United Kingdom
EMail: matthew.bocci@alcatel-lucent.com