Internet Engineering Task Force

Internet Draft

Intended status: Standards Track

Expires: August 2017

Luca Martini Monoski LLC George Swallow Cisco Elisa Bellagamba Ericsson

February 2017

MPLS LSP PW status refresh reduction for Static Pseudowires

draft-ietf-pals-status-reduction-02.txt

Status of this Memo

This Internet-Draft is submitted to IETF in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at http://www.ietf.org/ietf/lid-abstracts.txt.

The list of Internet-Draft Shadow Directories can be accessed at http://www.ietf.org/shadow.html.

This Internet-Draft will expire on August 13, 2010

Abstract

This document describes a method for generating an aggregated pseudowire status message transmitted on a Multi-Protocol Label Switching (MPLS) Label Switched Path (LSP) to indicate the status of one or more pseudowires carried on the LSP.

The method for transmitting the pseudowire (PW) status information is not new, however this protocol extension allows a Service Provider (SP) to reliably monitor the individual PW status while not

overwhelming the network with multiple periodic status messages. This is achieved by sending a single cumulative summary status verification message for all the PWs grouped in the same LSP.

Table of Contents

| <u>1</u> | Introduction | <u>3</u> |
|--------------|--|-----------------|
| <u>1.1</u> | Requirements Language | <u>3</u> |
| <u>1.2</u> | Terminology | <u>3</u> |
| <u>1.3</u> | Notational Conventions in Backus-Naur Form | 3 3 4 |
| <u>2</u> | PW status refresh reduction protocol | 4 4 5 5 5 6 6 9 |
| <u>2.1</u> | Protocol states | <u>4</u> |
| <u>2.1.1</u> | INACTIVE | <u>5</u> |
| 2.1.2 | STARTUP | <u>5</u> |
| 2.1.3 | ACTIVE | <u>5</u> |
| <u>2.2</u> | Timer value change transition procedure | <u>5</u> |
| <u>3</u> | PW status refresh reduction procedure | <u>6</u> |
| <u>4</u> | PW status refresh reduction Message Encoding | <u>6</u> |
| <u>5</u> | PW status refresh reduction Control Messages | 9 |
| <u>5.1</u> | Notification message | <u>10</u> |
| <u>5.2</u> | PW Configuration Message | 11 |
| <u>5.2.1</u> | MPLS-TP Tunnel ID | 12 |
| <u>5.2.2</u> | PW ID configured List | 12 |
| <u>5.2.3</u> | PW ID unconfigured List | <u>13</u> |
| <u>6</u> | PW provisioning verification procedure | <u>13</u> |
| <u>6.0.4</u> | PW ID List advertising and processing | 14 |
| <u>7</u> | Security Considerations | <u>14</u> |
| <u>8</u> | IANA Considerations | <u>15</u> |
| <u>8.1</u> | PW Status Refresh Reduction Message Types | <u>15</u> |
| 8.2 | PW Configuration Message Sub-TLVs | <u>15</u> |
| <u>8.3</u> | PW Status Refresh Reduction Notification Codes | <u>16</u> |
| <u>9</u> | References | <u>16</u> |
| <u>9.1</u> | Normative References | <u>16</u> |
| 9.2 | Informative References | <u>17</u> |
| 10 | Authors' Addresses | 17 |

Martini, et al. [Page 2]

1. Introduction

When PWs use a Multi Protocol Label Switched (MPLS) network as the Packet Switched Network (PSN), they are setup according to [RFC8077] static configuration mode and the PW status information is propagated using the method described in [RFC6478]. There are 2 basic modes of operation described in [RFC6478] section 5.3: Periodic retransmission of non-zero status messages, and a simple acknowledge of PW status (sec 5.3.1 of [RFC6478]). The LSP level protocol described below applies to the case when PW status is acknowledged immediately with a requested refresh value of zero (no refresh). In this case the PW status refresh reduction protocol is necessary for several reasons, such as:

- -i. Greatly increase the scalability of the PW status protocol by reducing the amount of messages that a PE needs to periodically send to it's neighbors.
- -ii. Detect a remote PE restart.
- -iii. If the local state is lost for some reason, the PE needs to be able to request a status refresh reduction from the remote PE
- -iv. Optionally detect a remote PE provisioning change.

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 [RFC2119].

1.2. Terminology

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

1.3. Notational Conventions in Backus-Naur Form

All multiple-word atomic identifiers use underscores () between the words to join the words. Many of the identifiers are composed of a concatenation of other identifiers. These are expressed using Backus-Naur Form (using double-colon - "::" - notation).

Where the same identifier type is used multiple times in a concatenation, they are qualified by a prefix joined to the identifier by a dash (-). For example Src-Node ID is the Node ID of a node referred to as Src (where "Src" is short for "source" in this example).

The notation does not define an implicit ordering of the information elements involved in a concatenated identifier.

2. PW status refresh reduction protocol

PW status refresh reduction protocol consists of a simple message that is sent at the LSP level using the MPLS Generic Associated Channel. [RFC5586]

A PE using the PW status refresh reduction protocol MUST send the PW status refresh reduction Message as soon as a PW is configured on a particular LSP. The message is then re-transmitted at a locally configured interval indicated in the refresh timer field. If no acknowledgment is received, the protocol does not reach active state, and the PE SHOULD NOT send any PW status messages with a refresh timer of zero as described in [RFC6478] section 5.3.1.

It is worth noting that no relationship is existing between the locally configured timer for the refresh reduction protocol and the PW individual status refresh timers.

2.1. Protocol states

The protocol can be in 3 possible states: INACTIVE, STARTUP, and ACTIVE.

2.1.1. INACTIVE

This state is entered when the protocol is turned off. This state is also entered if all PW on a specific LSP are unprovisioned, or the feature is unprovisioned.

2.1.2. STARTUP

In this state the PE transmits periodic PW status refresh reduction messages, with the Ack Session ID set to 0. The PE remains in this state until a PW status refresh message is received with the correct local session ID in the Ack Session ID Field. This state can be exited to the ACTIVE or INACTIVE state.

2.1.3. ACTIVE

This state is entered once the PE receives a PW status refresh reduction message with the correct local session ID in the Ack Session ID Field within 3.5 times the refresh timer field value of the last PW status refresh reduction message transmitted. This state is immediately exited as follows:

- -i. A valid PW status refresh reduction message is not received within 3.5 times the current refresh timer field value. (assuming a timer transition procedure is not in progress) New state: STARTUP
- -ii. A PW status refresh reduction message is received with the wrong, or a zero, Ack Session ID field value. New state: STARTUP
- -iii. All PWs using the particular LSP are unprovisioned, or the protocol is disabled. New state: INACTIVE

2.2. Timer value change transition procedure

If a PE needs to change the refresh timer value field while the PW refresh reduction protocol is in the ACTIVE state, the following procedure must be followed:

- -i. A PW status refresh reduction message is transmitted with the new timer value.
- -ii. If the new value is greater then the original one the PE will operate on the new timer value immediately.
- -iii. If the new value is smaller then the original one, the PE will operate according to the original timer value for a period 3.5 times the original timer value, or until the first valid PW status refresh reduction message is received.

Martini, et al. [Page 5]

A PE receiving a PW status refresh reduction message with a new timer value, will immediately transmit an acknowledge PW status refresh reduction message, and start operating according to the new timer value.

3. PW status refresh reduction procedure

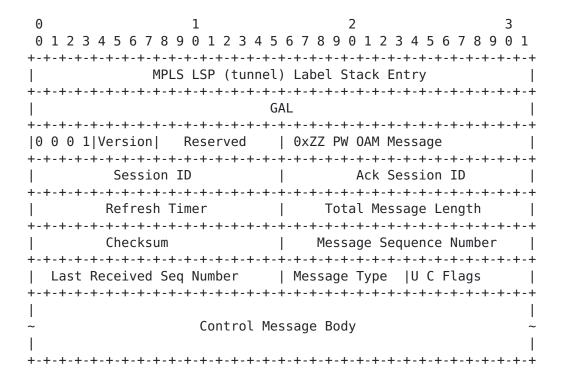
When the refresh reduction protocol, on a particular LSP, is in the ACTIVE state, the PE can send all PW status messages, for PWs on that LSP, with a refresh timer value of zero. This greatly decreases the amount of messages that the PE needs to transmit to the remote PE because once the PW status message for a particular PW is acknowledged, further repetitions of that message are no longer necessary.

To further mitigate the amount of possible messages when an LSP starts forwarding traffic, care should be taken to permit the PW refresh reduction protocol to reach the ACTIVE state quickly, and before the the first PW status refresh timer expires. This can be achieved by using a PW status refresh reduction Message refresh timer value that is much smaller then the PW status message refresh timer value in use. (sec 5.3.1 of [RFC6478])

If the refresh reduction protocol session is terminated by entering the INACTIVE or STARTUP states, the PE MUST immediately re-send all the previously sent PW status messages for that particular LSP for which the session terminated. In this case the refresh timer value MUST NOT be set to zero, and MUST be set according to the local policy of the PE router.

4. PW status refresh reduction Message Encoding

The packet containing the refresh reduction message is encoded as follows: (omitting link layer information)



This message contains the following fields:

* PW OAM Message.

This field indicates the Generic Associated Channel type in the GACH header as defined in [RFC5586].

Note: Channel type 0xZZ pending IANA allocation.

* Session ID

A non-zero, locally selected session number that is not preserved if the local PE restarts.

In order to get a locally unique session ID, the recommended choice is to perform a CRC-16 giving as input the following data

|Y|Y|M|M|D|D|H|H|M|M|S|S|L|L|L|

Where: YY: are the decimal two last digit of the current year MM: are the decimal two digit of the current month DD: are the decimal two digit of the current day HHMMSSLLL: are the decimal digits of the current time expressed in (hour, minutes, seconds, milliseconds) If the calculation results in an already existing Session ID, a unique Session ID can be generated by adding 1 to

Martini, et al. [Page 7]

the result until the Session ID is unique. Any other method to generate a locally unique session ID is also acceptable.

* Ack Session ID

The Acknowledgment Session ID received from the remote PE.

* Refresh Timer.

A non zero unsigned 16 bit integer value greater or equal to 10, in milliseconds, that indicates the desired refresh interval. The default value of 30000 is RECOMENDED.

* Total Message Length

Total length in octets of the Checksum, Message Type, Flags, Message Sequence Number, and control message body. A value of zero means that no control message is present, and therefore that no Checksum, and following fields are present either.

* Checksum

A 16 bit field containing the one's complement of the one's complement sum of the entire message (including the GACH header), with the checksum field replaced by zero for the purpose of computing the checksum. An all-zero value means that no checksum was transmitted. Note that when the checksum is not computed, the header of the bundle message will not be covered by any checksum.

* Message Sequence Number

An unsigned 16 bit integer number that is started from 1 when the protocol enters ACTIVE state. The sequence numbers wraps back to 1 when the maximum value is reached. The value of zero is reserved and MUST NOT be used.

* Last Received Message Sequence Number

The sequence number of the last message received. In no message has yet been received during this session, this field is set to zero.

* Message Type

The Type of the control message that follows. Control message types are allocated in this document, and by IANA.

Martini, et al. [Page 8]

* (U) Unknown flag bit.

Upon receipt of an unknown messageor TLV, if U is clear (=0), the keepalive session MUST be terminated by entering STARTUP state; if U is set (=1), the unknown message, or message contining a unknown TLV, MUST be acknowledged and silently ignored and the following messages, or TLVs, if any, processed as if the unknown message, or TLV did not exist.

- * (C) Configuration flag bit. The C Bit is used to signal the end of PW configuration transmission. If it is set, the sending PE has finished sending all it's current configuration information.
- * Flags (Reserved)

6 bits of flags reserved for future use, they MUST be set to 0 on transmission, and ignored on reception.

* Control Message Body

The Control Message body is defined in a section below, and is specific to the type of message.

It should be noted that the Checksum, Message Sequence Number, Last Received Message Sequence Number, Message Type, Flags, and control message body are OPTIONAL. The length field is used to parse how many optional fields are included. Hence all optional fields that precede a specific field that needs to be included in a specific implementation MUST be included if that optional field is also included.

If any of the above vaules are outside the specified range, a notification message is returned with a code "PW configuration not supported.", and the message is ignored.

PW status refresh reduction Control Messages

PW status refresh reduction Control messages consist of the Checksum, Message Sequence Number, Last Received Message Sequence Number, Message Type, Flags, and control message body.

When there is the need to send a PW status refresh reduction Control Messages, the system can attach it to a scheduled PW status refresh reduction or send one ahead of time. In any case PW status refresh reduction Control Messages always piggy back on normal messages.

A PW refresh reduction message is also called a PW status refresh

Martini, et al. [Page 9]

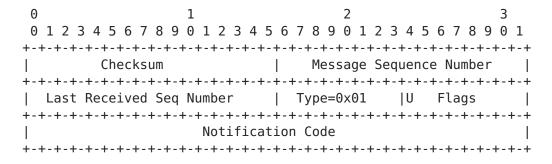
reduction Control Message if it contains a control message construct.

There can only be one control message construct per PW status refresh reduction Message. If the U bit is set, and a PE receiving the PW status refresh reduction Message does not understand the control message, the control message MUST be silently ignored. However the message sequence number MUST still be acknowledged by sending a null message back with the appropriate value in the Last Message Received Field. If a control message is not acknowledged, after 3.5 times the value of the Refresh Timer, a fatal notification "unacknowledged control message" MUST be sent, and the PW refresh reduction session MUST be terminated.

If a PE does not want or need to send a control message, the Checksum, and all following fields MUST NOT be sent, and the Total Message Length field is then set to zero.

<u>5.1</u>. Notification message

The most common use of the Notification Message is to acknowledge the reception of a message by indicating the received message sequence number in the "Last Received Sequence Number" field. The notification message is encoded as follows:



The message type is set to 0x01, and the U bit is treated as described in the above section. The Notification Codes are a 32 bit quantity assigned by IANA. (see IANA consideration section) Notification codes are either considered "Error codes" or simple notifications. If the Notification code is an Error code as indicated in the IANA allocation registry, the keepalive session MUST be terminated by entering STARTUP state. The 7 bits of flags are reserved for future use, they MUST be set to 0 on transmission, and ignored on reception.

Martini, et al. [Page 10]

5.2. PW Configuration Message

The PW status refresh reduction TLVs are informational TLVs, that allow the remote PE to verify certain provisioning information. This message contain a series of sub-TLVs in no particular order, that contain PW and LSP configuration information. The message has no preset length limit, however its total length will be limited by the transport network Maximum Transmit Unit (MTU). PW refresh reduction messages MUST NOT be fragmented. If a sender has more configuration information to send than will fit into one PW Configuration Message it may send further messages carrying further TLVs.

| 0 1 | 2 | 3 | | | | | |
|--|---------------------|-----------------|--|--|--|--|--|
| 0 1 2 3 4 5 6 7 8 9 0 1 2 3 | 4 5 6 7 8 9 0 1 2 3 | 4 5 6 7 8 9 0 1 | | | | | |
| +-+-+-+-+-+- | +-+-+-+-+-+-+-+- | +-+-+-+-+-+-+ | | | | | |
| Checksum | Message Seq | uence Number | | | | | |
| +- | | | | | | | |
| Last Received Seq Number | Type=0x02 | U C (Flags) | | | | | |
| +- | | | | | | | |
| ~ | | ~ | | | | | |
| PW Configuration Message Sub-TLVs | | | | | | | |
| ~ | | ~ | | | | | |
| +- | | | | | | | |

The PW Configuration Message type is set to 0x02. For this message the U-bit is set to 1 as processing of these messages is OPTIONAL.

The C Bit is used to signal the end of PW configuration transmission. If it is set, the sending PE has finished sending all its current configuration information. The PE transmitting the configuration MUST set the C bit on the last PW configuration message when all current PW configuration has been sent.

PW Configuration Message Sub-TLVs have the following generic format:

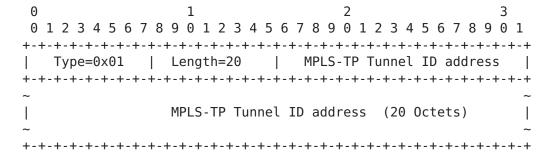
```
2
        1
0\ 1\ 2\ 3\ 4\ 5\ 6\ 7\ 8\ 9\ 0\ 1\ 2\ 3\ 4\ 5\ 6\ 7\ 8\ 9\ 0\ 1\ 2\ 3\ 4\ 5\ 6\ 7\ 8\ 9\ 0\ 1
| Length
                Value
            Value Continued
```

Martini, et al. [Page 11]

5.2.1. MPLS-TP Tunnel ID

This TLV contains the address of the MPLS-TP tunnel ID. When the configuration message is used for a particular keepalive session the MPLS-TP Tunnel ID sub-TLV MUST be sent at least once.

The MPLS Tunnel ID address is encoded as follows:



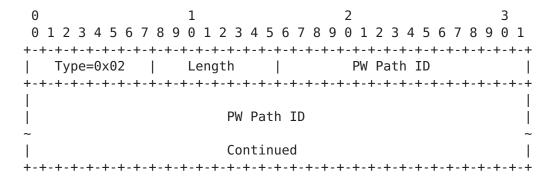
The MPLS point to point tunnel ID is defined in [RFC6370] as follows:

Src-Global Node ID::Src-Tunnel Num::Dst-Global Node ID::Dst-Tunnel Num

Note that a single address is enough to identify the tunnel, and the source end of the message.

5.2.2. PW ID configured List

This OPTIONAL TLV contains a list of the provisioned PWs on the LSP.



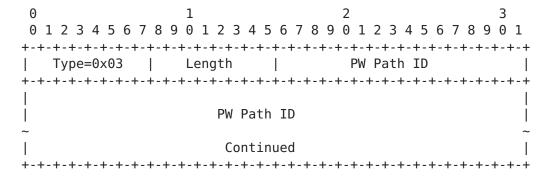
The PW Path ID is a 32 octet pseudowire path identifier specified in [RFC6370] as follows: AGI::Src-Global ID::Src-Node ID::Src-AC ID:: Dst-Global_ID::Dst-Node_ID::Dst-AC_ID

Martini, et al. [Page 12]

The number of PW Path IDs in the TLV will be inferred by the length of the TLV up to a maximum of 8. The procedure for processing this TLV will be described in a section below.

5.2.3. PW ID unconfigured List

This OPTIONAL TLV contains a list of the PWs that have been unprovisioned on the LSP. Note that it is a fatal session error to send the same PW address in both the configured list TLV , and the unconfigured list TLV in the same configuration message. If the this error occurs, an error notification message is returned with the error code of "PW Configuration TLV conflict" and the session is terminated by entering STARTUP state.



The PW Path ID is a 32 octet pseudowire path identifier specified in [RFC6370] as follows: AGI::Src-Global ID::Src-Node ID::Src-AC ID:: Dst-Global ID::Dst-Node ID::Dst-AC ID

The number of PW Path IDs in the TLV will be inferred by the length of the TLV up to a maximum of 8.

6. PW provisioning verification procedure

The advertisement of the PW configuration message is OPTIONAL.

A PE that desires to use the PW configuration message to verify the configuration of PWs on a particular LSP, should advertise its PW configuration to the remote PE on LSPs that have active keepalive sessions. When a PE receives PW configuration information using this protocol and it is not supporting or is not willing to use the information, it MUST acknowledge all the PW configuration messages with a notification of "PW configuration not supported". In this case, the information in the control messages is silently ignored. If a PE receives such a notification it SHOULD stop sending PW

Martini, et al. [Page 13]

configuration control messages for the duration of the PW refresh reduction keepalive session.

If PW configuration information is received, it is used to verify the accuracy of the local configuration information against the remote PE's configuration information. If a configuration mismatch is detected, where a particular PW is configured locally but not on the remote PE, the following action SHOULD be taken:

- -i. The local PW MUST be considered in "Not Forwarding" State.
- -ii. The PW Attachment Circuit status is set to reflect the PW fault.
- -iii. An Alarm SHOULD be raised to a network management system.

6.0.4. PW ID List advertising and processing

When configuration messages are advertised along a particular LSP, the PE sending the messages needs to check point the configuration information sent by setting the C bit when all currently known configuration information has been sent. This process allows the receiving PE to immediately proceed to verify all the currently configured PWs on that LSP, eliminating the need for a long waiting period.

If a new PW is added to a particular LSP, the PE MUST place the configuration verification of this PW on hold for a period of at least 30 seconds. This is necessary to minimize false positive events of mis-configuration due to the ends of the PW being slightly out of sync.

7. Security Considerations

The security considerations of [RFC6478] are adequate for the proposed mechanism since the operating environment is almost identical to the one where this protocol would be deployed. It should also be noted that since this protocol is designed to be deployed between two adjacent PEs connected by a physical link, it is not possible to misdirect or inject traffic without compromising the PW transport link itself. All these situations are covered in the security considerations of [RFC6478].

Martini, et al. [Page 14]

8. IANA Considerations

All the registries in this section are to be created or updated as appropriate in the Pseudowire Name Spaces (PWE3).

8.1. PW Status Refresh Reduction Message Types

IANA needs to set up a registry of "PW status refresh reduction Control Messages". These are 8-bit values. Type value 1 through 2 are defined in this document. Type values 3 through 64 are to be assigned by IANA using the "Expert Review" policy defined in RFC5226. Type values 65 through 127, 0 and 255 are to be allocated using the IETF review policy defined in [RFC5226]. Type values 128 through 254 are reserved for vendor proprietary extensions and are to be assigned by IANA, using the "First Come First Served" policy defined in RFC5226.

The Type Values are assigned as follows:

Type Message Description ---------0x01 Notification message 0x02 PW Configuration Message

8.2. PW Configuration Message Sub-TLVs

IANA needs to set up a registry of "PW status refresh reduction Configuration Message Sub-TLVs". These are 8-bit values. Type value 1 through 2 are defined in this document. Type values 3 through 64 are to be assigned by IANA using the "Expert Review" policy defined in RFC5226. Type values 65 through 127, 0 and 255 are to be allocated using the IETF review policy defined in [RFC5226]. Type values 128 through 254 are reserved for vendor proprietary extensions and are to be assigned by IANA, using the "First Come First Served" policy defined in RFC5226.

The Type Values are assigned as follows:

| sub-TLV type | Description |
|--------------|----------------------------|
| | |
| 0×01 | MPLS-TP Tunnel ID address. |
| 0x02 | PW ID configured List. |
| 0x03 | PW ID unconfigured List. |

8.3. PW Status Refresh Reduction Notification Codes

IANA needs to set up a registry of "PW status refresh reduction Notification Codes". These are 32-bit values. Type value 0 through 7 are defined in this document. Type values 8 through 65536 are to be assigned by IANA using the "Expert Review" policy defined in RFC5226. Type values 65536 through 134,217,728, 0 and 4,294,967,295 are to be allocated using the IETF review policy defined in [RFC5226]. Type values 134,217,729 through 4,294,967,294 are reserved for vendor proprietary extensions and are to be assigned by IANA, using the "First Come First Served" policy defined in RFC5226.

For each value assigned IANA should also track whether the value constitutes an error as described in <u>Section 5.1</u>. When values are assigned by IETF Review, the setting of this column must be documented in the RFC that requests the allocation. For Expert Review and FCFS assignments, the setting of this column must be made clear by the requester at the time of assignment.

The Type Values are assigned as follows:

| Error? | Description |
|--------|--|
| | |
| No | Null Notification. |
| No | PW configuration rejected. |
| Yes | PW Configuration TLV conflict. |
| No | Unknown TLV (U-bit=1) |
| Yes | Unknown TLV (U-bit=0) |
| No | Unknown Message Type |
| No | PW configuration not supported. |
| Yes | Unacknowledged control message. |
| | No No Yes No Yes No No |

9. References

9.1. Normative References

- [RFC2119] Bradner. S, "Key words for use in RFCs to Indicate Requirement Levels", <u>RFC 2119</u>, March, 1997.
- [RFC8077] "Pseudowire Setup and Maintenance Using the Label Distribution Protocol (LDP)", L. Martini, G. Heron, RFC8077, february 2017.
- [RFC6478] L. Martini, G. Swallow, G. Heron, M. Bocci "Pseudowire Status for Static Pseudowires", <u>RFC6478</u>, May 2012

Martini, et al. [Page 16]

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

9.2. Informative References

10. Authors' Addresses

Luca Martini Monoski LLC.

e-mail: lmartini@monoski.com

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

Elisa Bellagamba Ericsson EAB Torshamnsgatan 48 16480, Stockholm Sweden e-mail: elisa.bellagamba@gmail.com

Copyright Notice

Copyright (c) 2017 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.

Expiration Date: August 2017