            Definition of Time to Live TLV for LSP-Ping Mechanisms

Abstract

   LSP-Ping is a widely deployed Operation, Administration, and
   Maintenance (OAM) mechanism in MPLS networks. However, in the
   present form, this mechanism is inadequate to verify connectivity of
   a segment of a Multi-Segment Pseudowire (MS-PW) and/or bidirectional
   co-routed Label Switched Path (LSP) from any node on the path of the
   MS-PW and/or bidirectional co-routed LSP. This document defines a
   TLV to address this shortcoming.

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/rfc7394.

Copyright Notice

Table of Contents

1. Introduction

   An MS-PW may span across multiple service provider networks. In
   order to allow Service Providers (SPs) to verify segments of such
   MS-PWs from any node on the path of the MS-PW, any node along the
   path of the MS-PW, should be able to originate an MPLS Echo Request
   packet to any other node along the path of the MS-PW and receive the
   corresponding MPLS Echo Reply. If the originator of the MPLS Echo
   Request is at the end of a MS-PW, the receiver of the request can
   send the reply back to the sender without knowing the hop-count
   distance of the originator. The reply will be intercepted by the
   originator regardless of the TTL value on the reply packet. But, if
   the originator is not at the end of the MS-PW, the receiver of the
   MPLS Echo Request may need to know how many hops away the originator

of the MPLS Echo Request is so that it can set the TTL value on the
MPLS header for the MPLS Echo Reply to be intercepted at the
originator node.

In MPLS networks, for bidirectional co-routed LSPs, if it is desired
to verify connectivity from any intermediate node Label Switching
Router (LSR) on the LSP to the any other LSR on the LSP the receiver
may need to know the TTL to send the MPLS Echo Reply with, so as the
packet is intercepted by the originator node.

A new optional TTL TLV is defined in this document. This TLV will be
added by the originator of the MPLS Echo Request to inform the
receiver how many hops away the originator is on the path of the
MS-PW or bidirectional LSP.

This mechanism only works if the MPLS Echo Reply is sent down the
co-routed LSP; hence, the scope of this TTL TLV is currently limited
to MS-PW or bidirectional co-routed MPLS LSPs. The presence of the
TLV implies the use of the return path of the co-routed LSP, if the
return path is any other mechanism, then the TLV in the MPLS Echo
Request MUST be ignored.

2. Terminology

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 [RFC2119].

LSR: Label Switching Router

MPLS-TP: MPLS Transport Profile

MS-PW: Multi-Segment Pseudowire

PW: Pseudowire

TLV: Type Length Value

TTL: Time To Live

3. Time To Live TLV

3.1. TTL TLV Format

```
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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  Type = 32769                 |      Length = 8               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Value     |    Reserved     |   Flags                       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
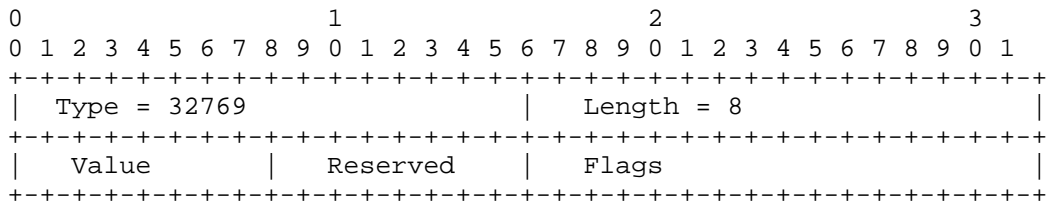
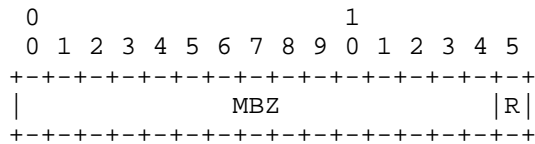   Figure 1: Time To Live TLV Format

   The TTL TLV has the format shown in Figure 1.

   Value

      The value of the TTL as specified by this TLV

   Flags

      The Flags field is a bit vector with the following format:

```
 0                   1
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|            MBZ              |R|
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

      One flag is defined for now, the R flag. The rest of the
      flags are Reserved - MUST be zero (MBZ) when sending and
      ignored on receipt.

      The R flag (Reply TTL) is set signify that the value is
      meant to be used as the TTL for the reply packet. Other bits
      may be defined later to enhance the scope of this TLV.

3.2. Usage

   The TTL TLV MAY be included in the MPLS Echo Request by the
   originator of the request.

   If the TTL TLV is present and the receiver does not understand TTL
   TLVs, it will simply ignore the TLV, as is the case for all optional
   TLVs. If the TTL TLV is not present or is not processed by the
   receiver, any determination of the TTL value used in the MPLS label
   on the LSP-Ping echo reply is beyond the scope of this document.

If the TTL TLV is present and the receiver understands TTL TLVs, one
of the following two conditions apply:

o If the TTL TLV value field is zero, the LSP-Ping echo request
   packet SHOULD be dropped.

o Otherwise, the receiver MUST use the TTL value specified in the
   TTL TLV when it creates the MPLS header of the MPLS Echo Reply.
   The TTL value in the TTL TLV takes precedence over any TTL value
   determined by other means, such as from the Switching Point TLV in
   the MS-PW. This precedence will aid the originator of the LSP-
   Ping echo request in analyzing the return path.

4. Operation

   In this section, we explain a use case for the TTL TLV with an MPLS
   MS-PW.

```
            <------------------MS-PW -------------------->


            A          B          C          D          E
            o -------- o -------- o --------- o --------- o
                       ---MPLS Echo Request--->
                       <--MPLS Echo Reply------
```
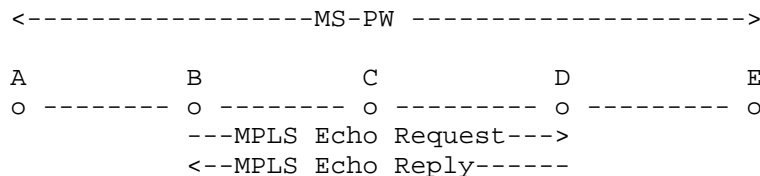
            Figure 2: Use-Case with MS-PWs

   Let us assume an MS-PW going through LSRs A, B, C, D, and E.
   Furthermore, assume that an operator wants to perform a connectivity
   check between B and D, from B. Thus, an MPLS Echo Request with the
   TTL TLV is originated from B and sent towards D. The MPLS Echo
   Request packet contains the FEC of the PW Segment between C and D.
   The value field of the TTL TLV and the TTL field of the MPLS label
   are set to 2, the choice of the value 2 will be based on the operator
   input requesting the MPLS Echo Request or from the optional LDP
   switching point TLV. The MPLS Echo Request is intercepted at D
   because of TTL expiry. D detects the TTL TLV in the request and uses
   the TTL value (i.e., 2) specified in the TLV on the MPLS label of the
   MPLS Echo Reply. The MPLS Echo Reply will be intercepted by B
   because of TTL expiry.

   The same operation will apply when we have a co-routed bidirectional
   LSP and we want to check connectivity from an intermediate LSR "B" to
   another LSR "D".

4.1. Traceroute Mode

   In traceroute mode, the TTL value in the TLV is set to 1 for the
   first Echo Request, then to 2 for the next, and so on. This is
   similar to the TTL values used for the label set on the packet.

4.2. Error Scenario

   It is possible that the MPLS Echo Request packet was intercepted
   before the intended destination for reasons other than label TTL
   expiry. This could be due to network faults, misconfiguration, or
   other reasons. In such cases, if the return TTL is set to the value
   specified in the TTL TLV, then the echo response packet will continue
   beyond the originating node. This becomes a security issue.

   To prevent this, the label TTL value used in the MPLS Echo Reply
   packet MUST be modified by deducting the incoming label TTL on the
   received packet from TTL TLV value. If the MPLS Echo Request packet
   is punted to the CPU before the incoming label TTL is deducted, then
   another 1 MUST be added. In other words:

   Return TTL Value on the MPLS Echo Reply packet = (TTL TLV Value) -
   (Incoming Label TTL) + 1

5. Security Considerations

   This document allows the setting of the TTL value in the MPLS Label
   of an MPLS Echo Reply, so that it can be intercepted by an
   intermediate device. This can cause a device to get a lot of LSP-
   Ping packets that get redirected to the CPU.

   However, the same is possible even without the changes mentioned in
   this document. A device should rate limit the LSP-Ping packets
   redirected to the CPU so that the CPU is not overwhelmed.

   The recommendation in the Security Considerations of [RFC4379]
   applies, to check the source address of the MPLS Echo Request;
   however, the source address can now be any node along the LSP path.

   A faulty transit node changing the TTL TLV value could make the wrong
   node reply to the MPLS Echo Request, and/or the wrong node to receive
   the MPLS Echo Reply. An LSP trace may help identify the faulty
   transit node.

6. IANA Considerations

   IANA has assigned a TLV type value to the following TLV from the
   "Multi-Protocol Label Switching (MPLS) Label Switched Paths (LSPs)
   Ping Parameters" registry in the "TLVs" subregistry.

      Time To Live TLV (see Section 3).

   IANA has allocated the value 32769.

7. References

7.1 Normative References

   [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate
             Requirement Levels", BCP 14, RFC 2119, March 1997,
             <http://www.rfc-editor.org/info/rfc2119>.

   [RFC4379] Kompella, K. and G. Swallow, "Detecting Multi-Protocol
             Label Switched (MPLS) Data Plane Failures", RFC 4379,
             February 2006, <http://www.rfc-editor.org/info/rfc4379>.

   [RFC5085] Nadeau, T., Ed., and C. Pignataro, Ed., "Pseudowire Virtual
             Circuit Connectivity Verification (VCCV): A Control Channel
             for Pseudowires", RFC 5085, December 2007,
             <http://www.rfc-editor.org/info/rfc5085>.

Acknowledgements

   The authors would like to thank Greg Mirsky for his comments.

Contributors

   Michael Wildt
   Cisco Systems, Inc.
   1414 Massachusetts Avenue
   Boxborough, MA 01719
   United States
   EMail: mwildt@cisco.com

Authors' Addresses

   Sami Boutros
   Cisco Systems, Inc.
   3750 Cisco Way
   San Jose, CA 95134
   United States
   EMail: sboutros@cisco.com

   Siva Sivabalan
   Cisco Systems, Inc.
   2000 Innovation Drive
   Kanata, Ontario, K2K 3E8
   Canada
   EMail: msiva@cisco.com

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

   Shaleen Saxena
   Cisco Systems, Inc.
   1414 Massachusetts Avenue
   Boxborough, MA 01719
   United States
   EMail: ssaxena@cisco.com

   Vishwas Manral
   Ionos Networks
   4100 Moorpark Ave, Suite 122
   San Jose, CA 95117
   United States
   EMail: vishwas@ionosnetworks.com

   Sam Aldrin
   Huawei Technologies, Inc.
   1188 Central Express Way,
   Santa Clara, CA 95051
   United States
   EMail: aldrin.ietf@gmail.com