Network Working Group
INTERNET-DRAFT
Intended Status: Standards Track

Sami Boutros
Siva Sivabalan
George Swallow
Shaleen Saxena
Cisco Systems

Vishwas Manral
Hewlett Packard Co.

Sam Aldrin
Huawei Technologies, Inc.

Expires: September 25, 2014 March 24, 2014

Definition of Time-to-Live TLV for LSP-Ping Mechanisms
draft-ietf-mpls-lsp-ping-ttl-tlv-07.txt

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) from any node on the path of the MS-PW. This document defines a TLV to address this shortcoming.

Table of Contents

1. Introduction

A MS-PW may span across multiple service provider networks. In order to allow Service Providers (SP) to verify segments of such MS-PW 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 LSP-Ping echo request packet to any another node along the path of the MS-PW and receive the corresponding echo reply. If the originator of the 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 echo request MAY need to know how many hops away the originator of the echo request is so that it can set the TTL value on the MPLS header for the 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 (LSR) on the LSP to the any other LSR on the LSP the receiver may need to know the TTL to send the 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 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 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 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

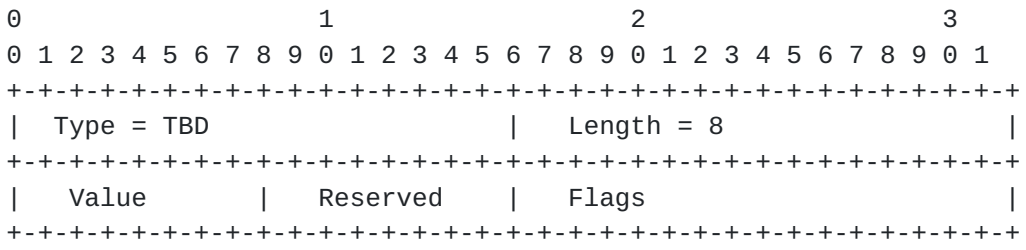


Figure 1: Time To Live TLV format

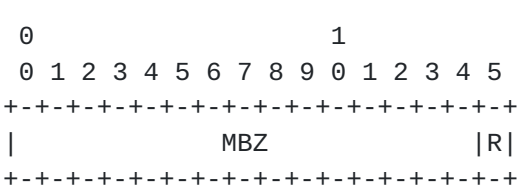
The TTL TLV has the format shown in Figure 1.

Value

The value of the TTL as specified by this TLV

Flags

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



One flag is defined for now, the R flag; the rest of the flags are currently undefined and 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

This TLV shall be included in the echo request by the originator of request. The use of this TLV is optional. If a receiver does not understand the TTL TLV, it will simply ignore the TLV (Type value of

TLV is assumed to be in the range of optional TLV's which SHOULD be ignored if an implementation does not support or understand them). In the absence of TTL TLV or if TTL TLV is ignored by a receiver, the determination of the TTL value used in the MPLS label on the echo reply is beyond the scope of this document.

If a receiver understands the TTL TLV, and the TTL TLV is present in the echo request, and if the value field is zero, the LSP Ping Echo request packet SHOULD be dropped.

If a receiver understands the TTL TLV, and the TTL TLV is present in the echo request, the receiver MUST use the TTL value specified in TLV in the MPLS header of the echo reply. In other words, if the value of the TTL provided by this TLV does not match the TTL determined by other means, such as Switching Point TLV in MS-PW, then TTL TLV must be used. This will aid the originator of the 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.

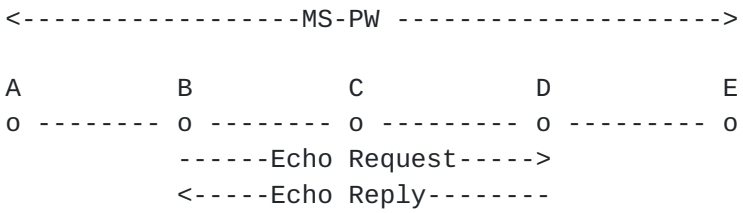


Figure 2: Use-case with MS-PWs

Let us assume a 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 LSP-Ping request with the TTL TLV is originated from B and sent towards D. The 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 echo request or from the optional LDP switching point TLV. The echo request is intercepted at D because of TTL expiry. D detects the TTL TLV in the request, and use the TTL value (i.e., 2) specified in the TLV on the MPLS label of the echo reply. The echo reply will be intercepted by B because of TTL expiry.

The same operation will apply in the case a co-routed bidirectional LSP and we want to check connectivity from an intermediate LSR B to another LSR D, from B.

4.1. Traceroute mode

In the traceroute mode TTL value in the TLV is successively set to 1, 2, 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 echo request packet was intercepted before the intended destination for reason other than label TTL expiry. This could be due 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 Echo Reply packet must be modified by deducting the incoming label TTL on the received packet from TTL TLV value. If the echo request packet is punted to the CPU before the incoming label TTL is deducted, then another 1 must be deducted. In other words:

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

5. Security Considerations

This draft allows the setting of the TTL value in the MPLS Label of an 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 which 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 [RFC4379] security section 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 is requested to assign TLV type value to the following TLV from the "Multiprotocol Label Switching Architecture (MPLS) Label Switched Paths (LSPs) Parameters - TLVs" registry, "TLVs and sub-TLVs" sub-registry.

Time To Live TLV (See Section 3). The value must be assigned from the range (32768-49161) of optional TLVs.

7. Acknowledgements

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

8. References

8.1 Normative References

[1] K. Kompella, G. Swallow, "Detecting Multi-Protocol Label Switched (MPLS) Data Plane Failures", RFC 4379, February 2006.

[2] T. Nadeau, et. al, "Pseudowire Virtual Circuit Connectivity Verification (VCCV): A Control Channel for Pseudowires ", RFC 5085, December 2007.

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

Authors' Addresses

Sami Boutros
Cisco Systems, Inc.
3750 Cisco Way
San Jose, California 95134
USA
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 , MASSACHUSETTS 01719
United States
Email: swallow@cisco.com

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

Vishwas Manral
Hewlett Packard Co.
19111 Pruneridge Ave,
Cupertino, CA 95014 USA
United States
EMail: vishwas.manral@hp.com

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

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