Network Working Group
Request for Comments: 3443
Updates: 3032
Category: Standards Track

P. Agarwal
Brocade
B. Akyol
Cisco Systems

January 2003

Time To Live (TTL) Processing in Multi-Protocol Label Switching (MPLS) Networks

Status of this Memo

Copyright Notice

Abstract

This document describes Time To Live (TTL) processing in hierarchical Multi-Protocol Label Switching (MPLS) networks and is motivated by the need to formalize a TTL-transparent mode of operation for an MPLS label-switched path. It updates RFC 3032, "MPLS Label Stack Encoding". TTL processing in both Pipe and Uniform Model hierarchical tunnels are specified with examples for both "push" and "pop" cases. The document also complements RFC 3270, "MPLS Support of Differentiated Services" and ties together the terminology introduced in that document with TTL processing in hierarchical MPLS networks.

Conventions used in this document

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. Introduction and Motivation

This document describes Time To Live (TTL) processing in hierarchical MPLS networks. We believe that this document adds details that have not been addressed in [MPLS-ARCH, MPLS-ENCAPS], and that the methods presented in this document complement [MPLS-DS].

In particular, a new mode of operation (referred to as the Pipe Model) is introduced to support the practice of configuring MPLS LSPs such that packets transiting the LSP see the tunnel as a single hop regardless of the number of intermediary label switch routers (LSR). The Pipe Model for TTL is currently being used in multiple networks and is provided as an option configurable by the network operator by several vendors.

This document formalizes the TTL processing in MPLS networks and ties it with the terminology introduced in [MPLS-DS].

2. TTL Processing in MPLS Networks

2.1. Changes to RFC 3032 [MPLS-ENCAPS]

a) [MPLS-ENCAPS] only covers the Uniform Model and does NOT address the Pipe Model or the Short Pipe Model. This document addresses these two models and for completeness will also address the Uniform Model.

b) [MPLS-ENCAPS] does not cover hierarchical LSPs. This document addresses this issue.

c) [MPLS-ENCAPS] does not define TTL processing in the presence of Penultimate Hop Popping (PHP). This document addresses this issue.

2.2. Terminology and Background

As defined in [MPLS-ENCAPS], MPLS packets use a MPLS shim header that indicates the following information about a packet:

a) MPLS Label (20 bits)
b) TTL (8 bits)
c) Bottom of stack (1 bit)
d) Experimental bits (3 bits)

The experimental bits were later redefined in [MPLS-DS] to indicate the scheduling and shaping behavior that could be associated with an MPLS packet.

[MPLS-DS] also defined two models for MPLS tunnel operation: Pipe and Uniform Models. In the Pipe Model, a MPLS network acts like a circuit when MPLS packets traverse the network such that only the LSP ingress and egress points are visible to nodes that are outside the tunnel. A Short variation of the Pipe Model is also defined in [MPLS-DS] to differentiate between different egress forwarding and QoS treatments. On the other hand, the Uniform Model makes all the

nodes that a LSP traverses visible to nodes outside the tunnel. We will extend the Pipe and Uniform Models to include TTL processing in the following sections. Furthermore, TTL processing, when performing PHP, is also described in this document. For a detailed description of Pipe and Uniform Models, please see [MPLS-DS].

TTL processing in MPLS networks can be broken down into two logical blocks: (i) the incoming TTL determination to take into account any tunnel egress due to MPLS Pop operations; (ii) packet processing of (possibly) exposed packets and outgoing TTLs.

We also note here that signaling the LSP type (Pipe, Short Pipe or Uniform Model) is out of the scope of this document, and that is also not addressed in the current versions of the label distribution protocols, e.g. LDP [MPLS-LDP] and RSVP-TE [MPLS-RSVP]. Currently, the LSP type is configured by the network operator manually by means of either a command line or network management interface.

2.3. New Terminology

iTTL: The TTL value to use as the incoming TTL. No checks are performed on the iTTL.

oTTL: This is the TTL value used as the outgoing TTL value (see section 3.5 for exception). It is always (iTTL - 1) unless otherwise stated.

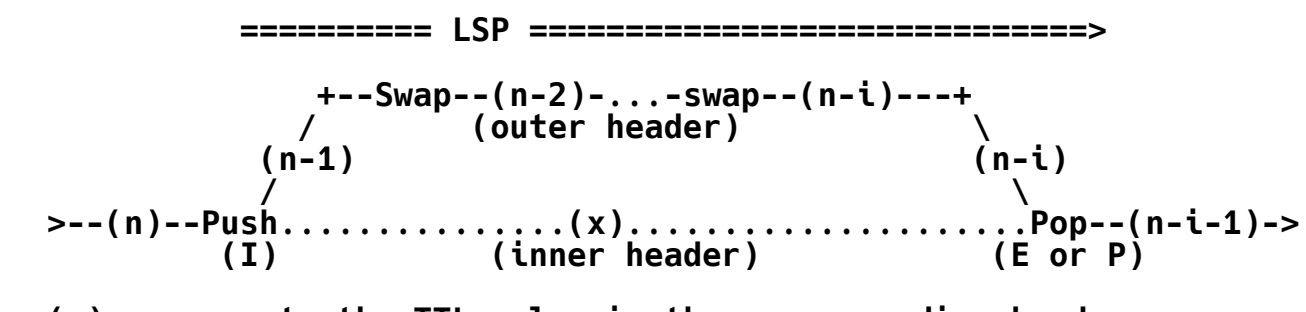
oTTL Check: Check if oTTL is greater than 0. If the oTTL Check is false, then the packet is not forwarded. Note that the oTTL check is performed only if any outgoing TTL (either IP or MPLS) is set to oTTL (see section 3.5 for exception).

3. TTL Processing in different Models

This section describes the TTL processing for LSPs conforming to each of the 3 models (Uniform, Short Pipe and Pipe) in the presence/absence of PHP (where applicable).

3.1. TTL Processing for Uniform Model LSPs (with or without PHP)

(consistent with [MPLS-ENCAPS]):

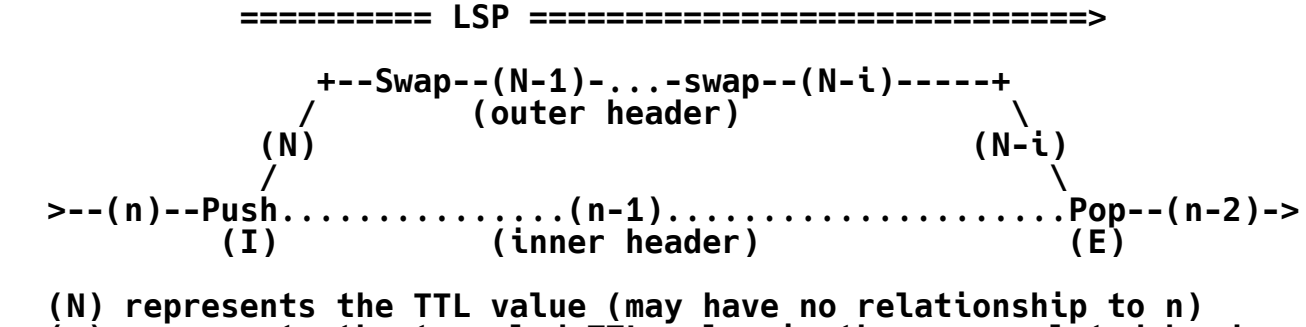


(n) represents the TTL value in the corresponding header
(x) represents non-meaningful TTL information
(I) represents the LSP ingress node
(P) represents the LSP penultimate node
(E) represents the LSP Egress node

This picture shows TTL processing for a Uniform Model MPLS LSP. Note that the inner and outer TTLs of the packets are synchronized at tunnel ingress and egress.

3.2. TTL Processing for Short Pipe Model LSPs

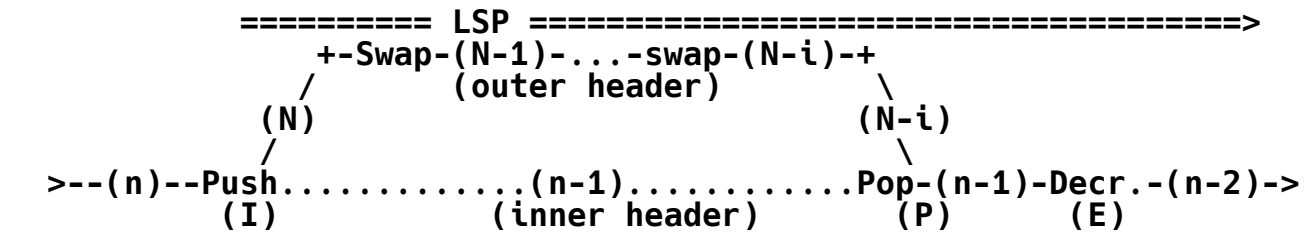
3.2.1. TTL Processing for Short Pipe Model LSPs without PHP



(N) represents the TTL value (may have no relationship to n)
(n) represents the tunneled TTL value in the encapsulated header
(I) represents the LSP ingress node
(E) represents the LSP Egress node

The Short Pipe Model was introduced in [MPLS-DS]. In the Short Pipe Model, the forwarding treatment at the egress LSR is based on the tunneled packet, as opposed to the encapsulating packet.

3.2.2. TTL Processing for Short Pipe Model with PHP:

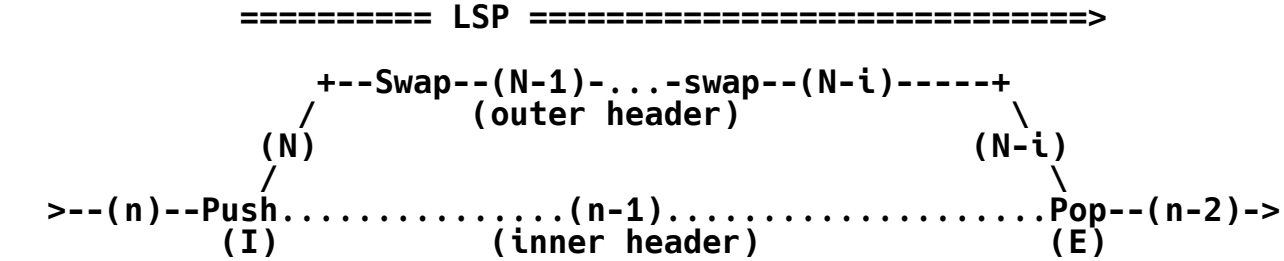


(N) represents the TTL value (may have no relationship to n)
(n) represents the tunneled TTL value in the encapsulated header
(I) represents the LSP ingress node
(P) represents the LSP penultimate node
(E) represents the LSP egress node.

Since the label has already been popped by the LSP's penultimate node, the LSP egress node just decrements the header TTL.

Also note that at the end of the Short Pipe Model LSP, the TTL of the tunneled packet has been decremented by two, with or without PHP.

3.3. TTL Processing for Pipe Model LSPs (without PHP only):



(N) represents the TTL value (may have no relationship to n)
(n) represents the tunneled TTL value in the encapsulated header
(I) represents the LSP ingress node
(E) represents the LSP Egress node

From the TTL perspective, the treatment for a Pipe Model LSP is identical to the Short Pipe Model without PHP.

3.4. Incoming TTL (iTTL) determination

If the incoming packet is an IP packet, then the iTTL is the TTL value of the incoming IP packet.

If the incoming packet is an MPLS packet and we are performing a Push/Swap/PHP, then the iTTL is the TTL of the topmost incoming label.

If the incoming packet is an MPLS packet and we are performing a Pop (tunnel termination), the iTTL is based on the tunnel type (Pipe or Uniform) of the LSP that was popped. If the popped label belonged to a Pipe Model LSP, then the iTTL is the value of the TTL field of the header, exposed after the label was popped (note that for the purpose of this document, the exposed header may be either an IP header or an MPLS label). If the popped label belonged to a Uniform Model LSP, then the iTTL is equal to the TTL of the popped label. If multiple Pop operations are performed sequentially, then the procedure given above is repeated with one exception: the iTTL computed during the previous Pop is used as the TTL of subsequent labels being popped; i.e. the TTL contained in the subsequent label is essentially ignored and replaced with the iTTL computed during the previous pop.

3.5. Outgoing TTL Determination and Packet Processing

After the iTTL computation is performed, the oTTL check is performed. If the oTTL check succeeds, then the outgoing TTL of the (labeled/unlabeled) packet is calculated and packet headers are updated as defined below.

If the packet was routed as an IP packet, the TTL value of the IP packet is set to oTTL (iTTL - 1). The TTL value(s) for any pushed label(s) is determined as described in section 3.6.

For packets that are routed as MPLS, we have four cases:

1) Swap-only: The routed label is swapped with another label and the TTL field of the outgoing label is set to oTTL.

2) Swap followed by a Push: The swapped operation is performed as described in (1). The TTL value(s) of any pushed label(s) is determined as described in section 3.6.

3) Penultimate Hop Pop (PHP): The routed label is popped. The oTTL check should be performed irrespective of whether the oTTL is used to update the TTL field of the outgoing header. If the PHPed label belonged to a Short Pipe Model LSP, then the TTL field of the PHP exposed header is neither checked nor updated. If the

PHPed label was a Uniform Model LSP, then the TTL field of the PHP exposed header is set to the oTTL. The TTL value(s) of additional labels are determined as described in section 3.6

4) Pop: The pop operation happens before routing and hence it is not considered here.

3.6. Tunnel Ingress Processing (Push)

For each pushed Uniform Model label, the TTL is copied from the label/IP-packet immediately underneath it.

For each pushed Pipe Model or Short Pipe Model label, the TTL field is set to a value configured by the network operator. In most implementations, this value is set to 255 by default.

3.7. Implementation Remarks

1) Although iTTL can be decremented by a value larger than 1 while it is being updated or oTTL is being determined, this feature should be only used for compensating for network nodes that are not capable of decrementing TTL values.

2) Whenever iTTL is decremented, the implementer must make sure that the value does not become negative.

3) In the Short Pipe Model with PHP enabled, the TTL of the tunneled packet is unchanged after the PHP operation.

4. Conclusion

This Internet Document describes how the TTL field can be processed in an MPLS network. We clarified the various methods that are applied in the presence of hierarchical tunnels and completed the integration of Pipe and Uniform Models with TTL processing.

5. Security Considerations

This document does not add any new security issues other than the ones defined in [MPLS-ENCAPS, MPLS-DS]. In particular, the document does not define a new protocol or expand an existing one and does not introduce security problems into the existing protocols. The authors believe that clarification of TTL handling in MPLS networks benefits service providers and their customers since troubleshooting is simplified.

6. References

6.1. Normative References

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

[MPLS-ARCH] Rosen, E., Viswanathan, A. and R. Callon, "Multiprotocol Label Switching Architecture", RFC 3031, January 2001.

[MPLS-ENCAPS] Rosen, E., Tappan, D., Fedorkow, G., Rekhter, Y., Farinacci, D., Li, T. and A. Conta, "MPLS Label Stack Encoding", RFC 3032, January 2001.

[MPLS-DS] Le Faucheur, F., Wu, L., Davie, B., Davari, S., Vaananen, P., Krishnan, R., Cheval, P. and J. Heinanen, "Multi-Protocol Label Switching (MPLS) Support of Differentiated Services", RFC 3270, May 2002.

6.2. Informative References

[MPLS-LDP] Andersson, L., Doolan, P., Feldman, N., Fredette, A. and B. Thomas, "LDP Specification", RFC 3036, January 2001.

[MPLS-RSVP] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V. and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, December 2001.

7. Acknowledgements

The authors would like to thank the members of the MPLS working group for their feedback. We would especially like to thank Shahram Davari and Loa Andersson for their careful review of the document and their comments.

8. Author's Addresses

Puneet Agarwal
Brocade Communications Systems, Inc.
1745 Technology Drive
San Jose, CA 95110

EMail: puneet@acm.org

Bora Akyol
Cisco Systems
170 W. Tasman Drive
San Jose, CA 95134

EMail: bora@cisco.com

9. Full Copyright Statement

Acknowledgement

Funding for the RFC Editor function is currently provided by the Internet Society.